Characterization of *Larva64* prim-pol activity and endogenous protein-protein interactions

A thesis presented to the faculty of the Graduate School of
Western Carolina University in partial fulfillment of the
requirements for the degree of Master of Science in Chemistry.

By

Hunter Chase Bishop

Committee Members

Jamie Wallen, PhD (Committee Chair)

Department of Chemistry and Physics

Maria Gainey, PhD (Committee Member)

Department of Chemistry and Physics

Brian Byrd, PhD, MSPH (Committee Member)

School of Health Sciences

November 2024

ACKNOWLEDGEMENTS

I would like to thank the members of my committee, Dr. Jamie Wallen, Dr. Maria Gainey, and Dr. Brian Byrd for their fostering of my scientific journey at WCU since freshman year. Their wisdom, concern, and most importantly, tough love, pushed me to find what I loved in science once again. Without them, I would not be here. I would like to take this opportunity to apologize to them, and all the members of our lab, for the litany of unprompted dad-jokes and puns.

I would like to express gratitude to Dr. Scott Huffman and Dr. Curtis Beimborn for their encouragement, humor, and coffee when I wanted to quit, the Western Carolina University Chemistry and Physics Department for a top-notch affordable education, and my college family at The Bridge Church who loved me and supported me in all my highs and lows. To my HP Elitebook 8460p that got me through college; thank you for somehow surviving all this time without catastrophically dying. Lastly, I would like to thank my beloved Cassidy for being the love of my life and just being awesome.

"Count it all joy, my brothers, when you meet trials of various kinds, for you know that the testing of your faith produces steadfastness. And let steadfastness have its full effect, that you may be perfect and complete, lacking in nothing."

James 1:2-4 (ESV)

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF ABBREVIATIONS

prim-pol: DNA primase-polymerase enzyme AEP

dsDNA: double-stranded DNA

ssDNA: single-stranded DNA

rNTP: ribonucleoside triphosphate

dNTP: deoxyribonucleoside triphosphate

Larva64: prim-pol gene 64 from mycobacteriophage Larva

kb: DNA kilobase

SD: strand-displacement

dsM13: double-stranded M13 DNA

ssM13: single stranded M13 DNA

bp: DNA base pair

WT: wild-type

PG: *PicoGreen*^{TM} dye

ABSTRACT

Characterization of *Larva64* prim-pol activity and endogenous protein-protein interactions

Hunter Chase Bishop

Western Carolina University (November 2024)

Director: Dr. Jamie Wallen

Primase-polymerases (prim-pols) are enzymes that exhibit both primase and polymerase activities in a single enzyme. Many of these enzymes possess additional DNA binding domains, such as helicase domains, fused to the prim-pol domain, which suggests that these enzymes may have all the required functions necessary to replicate a genome. Actinobacteriophage Larva contains a prim-pol protein (*Larva64*) that has an N-terminal prim-pol domain fused to a C-terminal predicted helicase domain, and it has been shown that *Larva64* is essential for viral survival. Previous work has demonstrated that *Larva64* can polymerize dsDNA *de novo*; however, our current data suggest that the enzyme works at a speed too slow to fully replicate the roughly 62 kb viral genome in a timely manner, indicating other interactions or proteins assist in Larva's genome synthesis. Using a qualitative agarose gel electrophoresis-based assay, we demonstrated that *Larva64* is unable to synthesize a dsDNA product using rNTPs rather than dNTPs, and the addition of rNTPs with dNTPs does not appear to increase the overall polymerization rate. Moreover, while *Larva64* can readily synthesize DNA in the presence of Mg^{2+}, it fails to use other divalent cations for catalysis. Prior research on *Larva64* identified the N-terminus as a prim-pol

domain and the C-terminus as a helicase/ATPase domain. D81A and K415A mutations in *Larva64* were produced, with the D81A mutation rendering polymerase activity inactive, and the K415A mutation being comparable to wild-type activity. To determine a quantitative rate of dsDNA synthesis, *PicoGreen*^{TM} is utilized as a fluorescent dye for quenched polymerization reactions at various time points. The results from the *PicoGreen*^{TM} assay reveal an observable increase in fluorescence with time, as well as *Larva64's* ability to hand-off synthesis to faster bacterial and viral polymerases, Klenow, and T7 polymerase, respectively. Alpha Fold 3 models were generated for *Larva64* in different stoichiometric orientations, suggesting that *Larva64* exists either in a monomeric or hexameric orientation. Finally, a bacterial two-hybrid assay was utilized to identify an endogenous protein-protein interaction between *Larva64* and *Larva48*, a sliding clamp protein. The interaction between a sliding clamp protein and a prim-pol is novel to the literature. The interactions between bacterial host polymerases, the viral sliding clamp, or a combination of proteins could potentially assist in the genome replication of phage Larva.

CHAPTER 1: INTRODUCTION

Primase-polymerases, more commonly known as prim-pols, are enzymes that possess the properties of both a primase and polymerase needed for DNA replication. DNA replication is an essential biochemical component for all cellular life, as well as for viruses. DNA replication usually requires three separate enzymes: a helicase, a primase, and a polymerase. Each of these enzymes plays an essential role, with the helicase "unzipping" the double-stranded DNA (dsDNA), primase forming a small strand of RNA on both pieces of the now unzipped DNA, to which the DNA polymerase will start at the 3'-end of the RNA primers to synthesize complementary DNA. This trifecta of enzymes works together to form two identical strands of dsDNA (1).

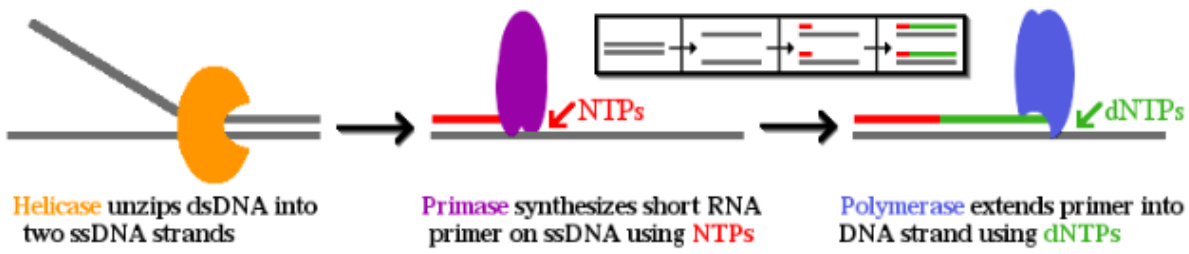

Figure 1: Schematic of DNA replication (Credit: Folse, 2020)

While many prim-pols require a helicase to "unzip" the DNA, they have the ability to both prime and polymerize DNA during DNA synthesis. The two-for-one property of prim-pols is intriguing as it may offer insight into the early origins of DNA replication (1). There have been recent studies that have suggested that certain prim-pols can function as a helicase as well and perform strand displacement synthesis (2). Strand displacement (SD) synthesis refers to DNA synthesis that is able to move through a double-stranded junction and replicate one side of the DNA strand and release a piece of single-stranded DNA (ssDNA) which then becomes double-stranded in random fashion (3). The ability for a singular enzyme to serve multiple roles in DNA synthesis provides impetus

for research into prim-pols. Notable prim-pols are the human PrimPol, a human DNA repair protein, and the NrS-1 prim-pol, a deep-sea phage prim-pol responsible for DNA replication and de novo primer synthesis for phage NrS-1 (4,5). Depending on the nucleotide being incorporated, human PrimPol has a DNA synthesis rate of 4.2 base/second to 0.307 base/second, for an average of 2.3 base/second (6). The human PrimPol is not responsible for DNA replication of the genome, unlike NrS-1 prim-pol which is responsible for replicating its entire 37,159 bp phage genome (5). This shows that prim-pols can serve as both main and supporting roles in genome replication, all the more prompting for study.

Prim-pols are found in a number of bacteriophages, which are viruses that infect solely bacteria and have potential as an alternative for the treatment of antibiotic-resistant bacterial infections. Bacteriophages are the most abundant biological entities on Earth, making up an incredibly rich genomic and proteomic landscape to explore. The emerging public health crisis of antibiotic-resistant bacteria has resulted in increased interest and research into bacteriophages as clinical therapies (7). Studies of viral genome replication and host-virus interactions can help tailor and adapt bacteriophages for eventual pharmaceutical applications. Actinophages are typically linear dsDNA viruses that infect bacteria belonging to phylum Actinobacteria. Mycobacteriophages are of particular interest in the realm of actinophages as they infect certain *Mycobacterium* hosts. Notable members of genus *Mycobacterium* include *Mycobacterium tuberculosis* and *Mycobacterium lepra*e, the bacteria responsible for tuberculosis and leprosy respectively (8). Further investigation into mycobacteriophage genomic replication, where prim-pols are a subject of interest, can potentially lead to more effective phage treatments against antibiotic resistant infections. Many bacteriophages contain genes annotated as prim-pols based on domain conservation software, and when compared to the number of bacteriophages being discovered and

annotated every year, prim-pols in bacteriophages are largely understudied. Mycobacteriophage Larva was originally isolated from *Mycobacterium smegmatis* mc^2155 in Williamsburg, VA and is classified as a temperate phage of the Siphoviridae morphotype belonging to cluster K5 (9). Larva's gene 64 (*Larva64*) encodes a prim-pol, and it is the only encoded DNA polymerase identified in the phage. Thus, *Larva64* is presumably the sole source of Larva's endogenous genetic replication abilities, and prime for study.

Prim-pols, including *Larva64,* are able to perform *de novo* DNA synthesis due to their primase and polymerase activities. These activities can be demonstrated using a rolling circle assay. Rolling circle assays are most often utilized to measure a polymerase's ability to extend a DNA primer annealed to circular ssDNA, such as ssM13, polymerizing a complementary strand of DNA (10). This DNA comes from M13 bacteriophage which possesses a single-stranded, circular viral genome 6.4 kb long (11). Since prim-pols, including *Larva64,* can synthesize *de novo*, there is no need to anneal a DNA primer to the ssM13 for rolling circle polymerization assays involving them. The progression of DNA synthesis can be measured by quenching the reaction at various time points via the addition of a quench buffer and measuring the size of the dsDNA products. These quench buffers often contain EDTA, which chelates the Mg^{2+} necessary for enzymatic function, thus stopping the synthesis reaction (12). The size of the resulting dsDNA products can be measured using agarose gels for qualitative results or using dsDNA binding fluorescent dyes to measure the size quantitatively. Rolling circle assays would be used exclusively in this study to characterize the activity of *Larva64* via qualitative and quantitative assays.

Previous research at Western Carolina University by Nathan Folse in 2020 expressed and purified *Larva64* and characterized *de novo* enzymatic activity, as well as protein-protein interactions (PPi) between *Larva64* and *M. smegmatis* proteins via a bacterial two-hybrid assay

(B2H) (13). Folse demonstrated an interaction between *Larva64* and transketolase, 2-isopropylmalate synthase (2-IPM), transcription termination/antitermination protein NusA, chaperonin GroEL, and a fifth protein of unknown function from *M. smegmatis*. The enzymatic activity of *Larva64* however was not quantified, rather only shown to be present via gel-based assays and that the enzyme appeared to be Mg^{2+} dependent. The two-hybrid assay shed light on potential interactions between *Larva64* and *M. smegmatis* proteins but did not explore protein-protein interactions within the virus itself. Moreover, using CRISPRi, Folse was able to demonstrate that when *Larva64* is silenced in the viral genome, the virus is unable to survive, whereas when *Larva65*, a helicase, is silenced, the virus is able to survive. This shows that the enzyme is essential for the function of phage Larva and given that it is the only polymerase present in the viral genome, it supports the assertion that *Larva64* is the endogenous source of viral genome replication. This also shows the massive importance of studying prim-pols in bacteriophage, as damage to or alterations to bacteriophage prim-pols could result in phage therapies becoming ineffective against antibiotic resistant bacteria.

Results of gel-based assays done by Folse in 2020 have shown that the enzyme works at a speed too slow to fully replicate the roughly 62 kb viral genome in a timely manner, however no quantitative rate was determined. Enzymatic rates for dsDNA synthesis can be assessed using various methods, namely using radioactively marked nucleotides or fluorescent dyes that bind to dsDNA (14, 15). *PicoGreen*^{TM} (PG) is an intercalating fluorescent dye which binds nearly exclusively to double-stranded DNA. It has an emission and excitation of 485 nm and 528 nm, respectively (16), and is commonly used to assess dsDNA concentration and enzyme rate (17, 18, 19, 20, 21, 22). The institution where this research was conducted discouraged the usage of

radioactivity for this project, so PG would be utilized to determine the rate of synthesis for *Larva64.*

Given that *Larva64* is essential for phage Larva's survival, and is the only encoded polymerase for the virus, the lack of polymerization haste calls into question the means of Larva's genomic replication. This indicates that other interactions or proteins may assist in Larva's genome synthesis. Two-hybrid systems are a genetic approach for detecting protein-protein interactions (PPi) *in vivo* (23). This system allows for a wide library of proteins to be tested against one another to determine interactions that would potentially serve an important role in an organism or virus. If PPi occurs, a phenotypic change in bacterial colonies will occur, such as color change or growth on antibiotic doped media. Folse in 2020 previously conducted a two-hybrid assay to determine interactions between *Larva64* and *M. smegmatis* proteins, but did not investigate endogenous PPi. These endogenous PPi may reveal assistance *Larva64* is receiving to synthesize the viral genome.

Protein structure determines the function of the protein. As such, knowing and understanding protein structures are very important in bacteriophage DNA replication research as improper folding or damaged structure can lead to ineffective phage therapies. One key component of protein structure is whether an individual protein is functional on its own or a subunit of a larger complex. The subunit stoichiometry of protein complexes often plays an important role in determining and regulating a protein's function (24). For example, sliding clamp proteins often form a homodimer complex in the shape of a ring which allows them to envelop DNA and facilitate more efficient operation of DNA replication enzymes (25). The stoichiometry of the protein and resulting shape of the protein complex determine the function of the enzyme, however determining these structures can be difficult and time consuming (26). One alternative to traditional methods of protein structure characterization is AI modeling, namely using AlphaFold 3. Alpha Fold 3 is a

valuable tool to observe potential protein structures *in silico* (27) requiring just a protein sequence to produce a model. This tool can be used to determine potential structures of *Larva64* as well as its stoichiometry, thus giving insight to the slow function of the enzyme.

The aims of this study are firstly to qualitatively characterize the activity of the *Larva64* prim-pol through identifying variables that could potentially increase the speed of *Larva64*. Secondly, we aim to quantitatively characterize the DNA synthesis activity of *Larva64* via a quenched fluorescence assay. Thirdly, we seek to determine in what stoichiometric orientation *Larva64* exists based off of Alpha Fold 3 modeling and comparisons to proteins of similar sequence using HHPred. Lastly, by determining endogenous protein-protein interactions we seek to determine by what mechanism phage Larva's genome is replicated, and what role *Larva64* plays in that replication.

CHAPTER TWO: QUALITATIVE CHARACTERIZATION OF LARVA64 ENZYMATIC ACTIVITY

In Vitro Qualitative Characterization of *Larva64*

To demonstrate *de novo* DNA synthesis activities and show the polymerization of dsDNA, a rolling circle assay was employed using ssM13 DNA. Biochemical qualitative analysis from this study has confirmed multiple prim-pol activity characteristics observed by Folse, 2020 and revealed novel characteristics of *Larva64*. *Larva64* was confirmed to have robust *de novo* synthesis activity, is unable to use $MnCl_2$, and a slow synthesis speed. *Larva64* is unable to utilize any divalent cation tested besides $MgCl_2$ for DNA synthesis, is unable to synthesize a double-stranded product given just rNTPs, nor see an increase in speed given the addition of rNTPs with dNTPs. The enzyme did not see an increase in speed with a K415A mutation, although a D81A mutation did eliminate all *de novo* DNA synthesis activity. Optimization of enzyme buffer pH and NaCl conditions were conducted, leading to an adjustment to pH of 8.0 and NaCl concentration of 25 mM. Lastly, *Larva64* was shown to be able to hand off synthesis to faster DNA polymerases incapable of *de novo* DNA synthesis and was confirmed to perform strand displacement synthesis. All qualitative assays included in this study were performed in triplicate. .

Long Run Gel:

To show the DNA synthesis activity of *Larva64* over a longer period of time and at more intervals than Folse, 2020, an assay was performed to quench and measure dsDNA products at 10-minute intervals from 10-110 minutes. dsDNA products are measured throughout the qualitative characterization via agarose gels. Figure 2 shows an agarose gel with DNA products of a “Long Run” *Larva64* assay.

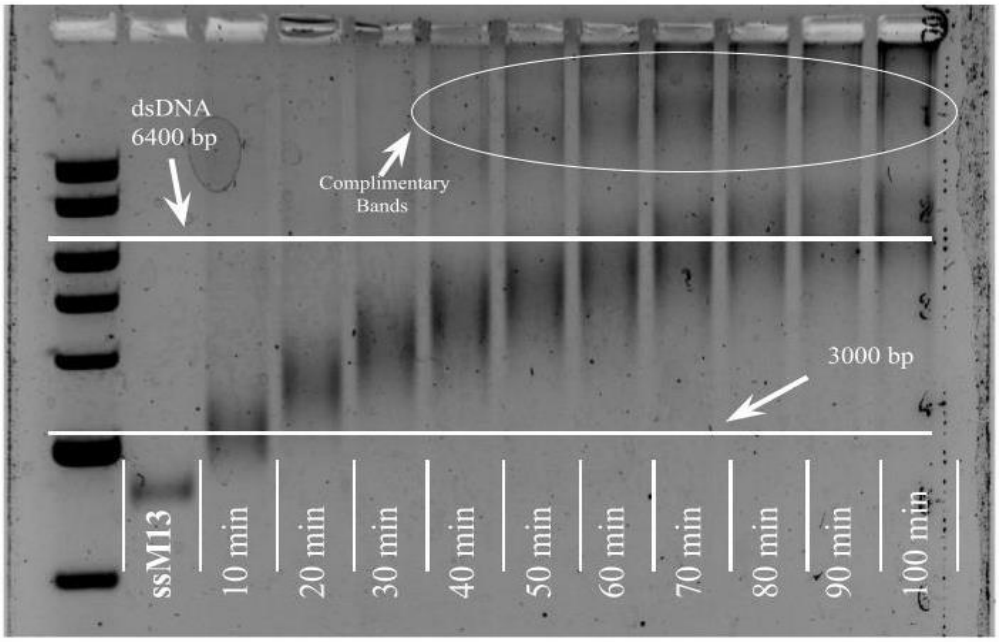

Figure 2: Gel-based assay results of "Long Run". The first lane is a 1 kb ladder, followed by ssM13, and 10-minute increment quenches of Larva64 and accompanying dsDNA product. White horizontal lines indicate where 3000 and 6400 bp dsDNA products should be.

A sample which lacks the addition of *Larva64* serves as a 0-minute time point for the assay, and for all following assays. Quenched at 10-minute intervals, this assay shows the synthesis of a DNA complement over time. By roughly the 60-minute time point there is a band at the 6.4 kb marker, indicating the completion of synthesis around the entirety of the circular M13 DNA stand. For an enzyme to make a product 6.4 kb in 60 minutes, that would translate to a rate of roughly 1.8 bp per second based on this qualitative assay. There is a dark line of bands from dsDNA products that run complementary above the prominent bands, well above the 6.4 kb marker, which indicates strand-displacement activity of *Larva64*. The observation that synthesis is not complete until ~60 minutes confirms that *Larva64*'s activity is slow. Multiple alterations to enzymatic conditions were conducted to discern if *Larva64* activity could be enhanced so as to better accept it as the sole protein responsible for genome replication in phage Larva.

$MgCl_2$ Dependence:

Prior research by Folse (2020) demonstrated that *Larva64* was Mg^{2+} dependent and unable to use Mn^{2+} in lieu of Mg^{2+}. To confirm this dependence and observe if *Larva64* was able to utilize any other divalent metals, a gel-based assay was conducted with identical concentrations of different metal salts replacing $MgCl_2$. These metal salts were $MnCl_2$, $CaCl_2$, $NiCl_2$, $CoCl_2$, and $ZnSO_4$. A lane with no metal salt served as a negative control. The quench assay consisted of two time points (30 minutes and 2 hours), with a sample containing $MgCl_2$ serving as the positive control (Figure 3).

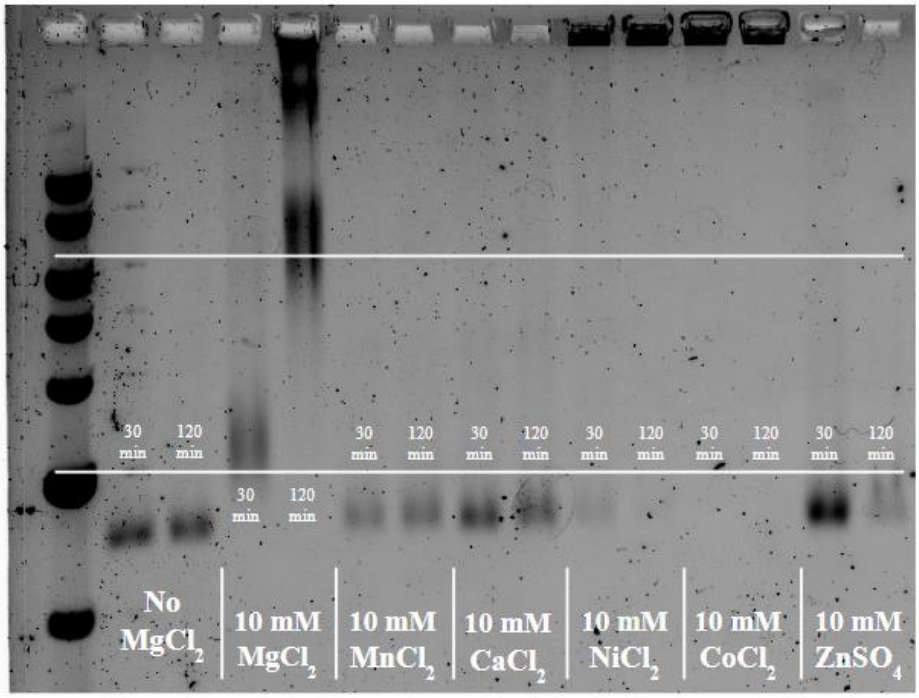

Figure 3: Divalent cation (M^{2+}) dependence of Larva64 prim-pol activity, various metals compared to $MgCl_2$. Quenched at 30- and 120-minute intervals, we see that the only metal salt that allows activity is $MgCl_2$. White horizontal lines indicate where 3000 and 6400 bp dsDNA products should be.

The results of the assay showed that no substitute metal salt allowed for activity. Ni and Co have dark bands within the well resulting from the protein immediately crashing upon the addition of the metal salt. Upon addition of the $NiCl_2$ and $CoCl_2$ salts to the sample, the solution went

cloudy as a precipitate formed. $ZnSO_4$ appeared to have slight activity; however, it was dismissed with later assays that showed no product being produced beyond the ssM13 line at 30- and 60-minute increments. Many DNA polymerases are capable of utilizing different metal ions in lieu of Mg^{2+} including all of the metal ions tested (28); however, it appears that *Larva64* is not one of those proteins. With the testing of alternate divalent metal salts, none allowed synthesis except for the control Mg^{2+}. With this information, incorrect metal salts were ruled out as a cause of the slow speed of *Larva64.*

dNTP and rNTP Synthesis:

Given that some prim-pols, such as the human PrimPol, are capable of synthesizing nucleic acids using both rNTPs and dNTPs, we investigated if *Larva64* would be able to synthesize a double-stranded product using only rNTPs (29). Moreover, we also investigated whether the presence of rNTPs would cause an increase of priming activity and possibly increase polymerization speed. A gel-based assay was conducted with enzyme conditions that included either 1 mM dNTPs, 1 mM dNTPs + 1 mM rNTPs, or 1 mM rNTPs to see if a double-stranded product could be formed or if there would be an increase in speed. The timepoints for the assay were 10, 30, 60, and 120 minutes (Figure 4).

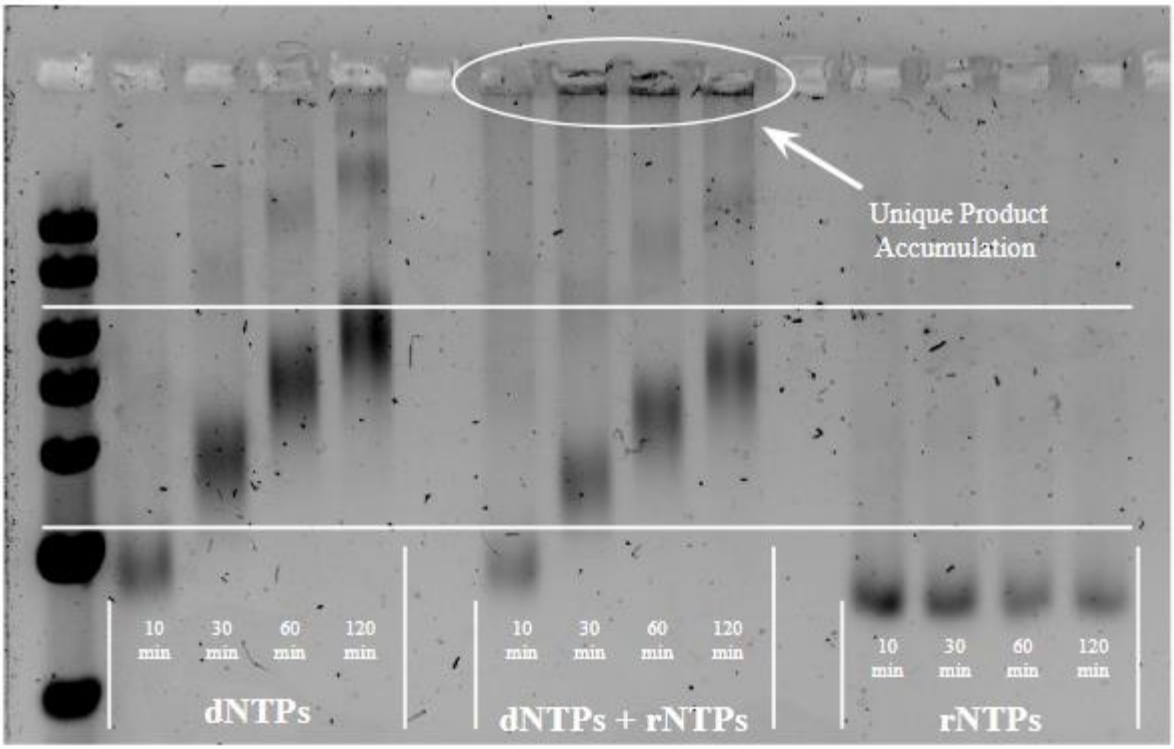

Figure 4: Agarose gel showing DNA synthesis with differing NTPs. White horizontal lines indicate where 3000 and 6400 bp dsDNA products should be. Quenched at 10-, 30-, 60-, and 120-minute intervals, we see that there is no synthesized product for just rNTPs, nor an enhancement of activity when with dNTPs.

The results in Figure 4 demonstrate that there is no synthesis for the reactions containing only rNTPs. There was also no discernible change in synthesis rates for the reactions containing both dNTPs and rNTPs compared to just dNTPs, outside of a product accumulation in the wells. The source of this product is unknown and unique to the rNTPs + dNTPs condition for this assay. The addition of rNTPs did not result in increased polymerization speed, ruling out rNTPs as a means of increasing prim-pol speed. Overall, these results show that *Larva64* exclusively incorporates dNTPs during DNA synthesis.

Protein Domain Mutations:

Prior research on *Larva64* had identified the N-terminal portion of the protein as a prim-pol domain and the C-terminal region as an ATPase/helicase domain, although the overall role of the N+C-terminal domains has yet to be fully understood (9, 13). D81A and K415A mutations were made to see how N-terminal and C-terminal domain alterations affected enzyme function. D81 is one of a number of active site residues in the prim-pol domain, and K415 is the Walker A lysine, responsible for binding to ATP (30). Alterations to the prim-pol domain should lead to a ceasing of synthesis, while alterations to the helicase domain should theoretically not affect synthesis as Folse showed that *Larva64* is not ATP dependent. Once the D81A and K415A mutants were expressed and purified, a rolling circle gel-based assay was conducted to assess kinetic activity. Linearized dsM13 and non-linearized dsM13 were in the early wells so as to show where the dsDNA product of the synthesis should be if the enzyme is capable of synthesizing a complement to the whole of the M13 plasmid, which would theoretically correspond to the 6.4 kb mark on the ladder. This was tested to see if supercoiling of DNA would affect the perceived size of enzyme synthesized products. A sample which lacked the addition of *Larva64* or its mutants served as a 0-minute time point, and the differing mutation syntheses were quenched at intervals of 10-, 30-, and 60-minutes (Figure 5).

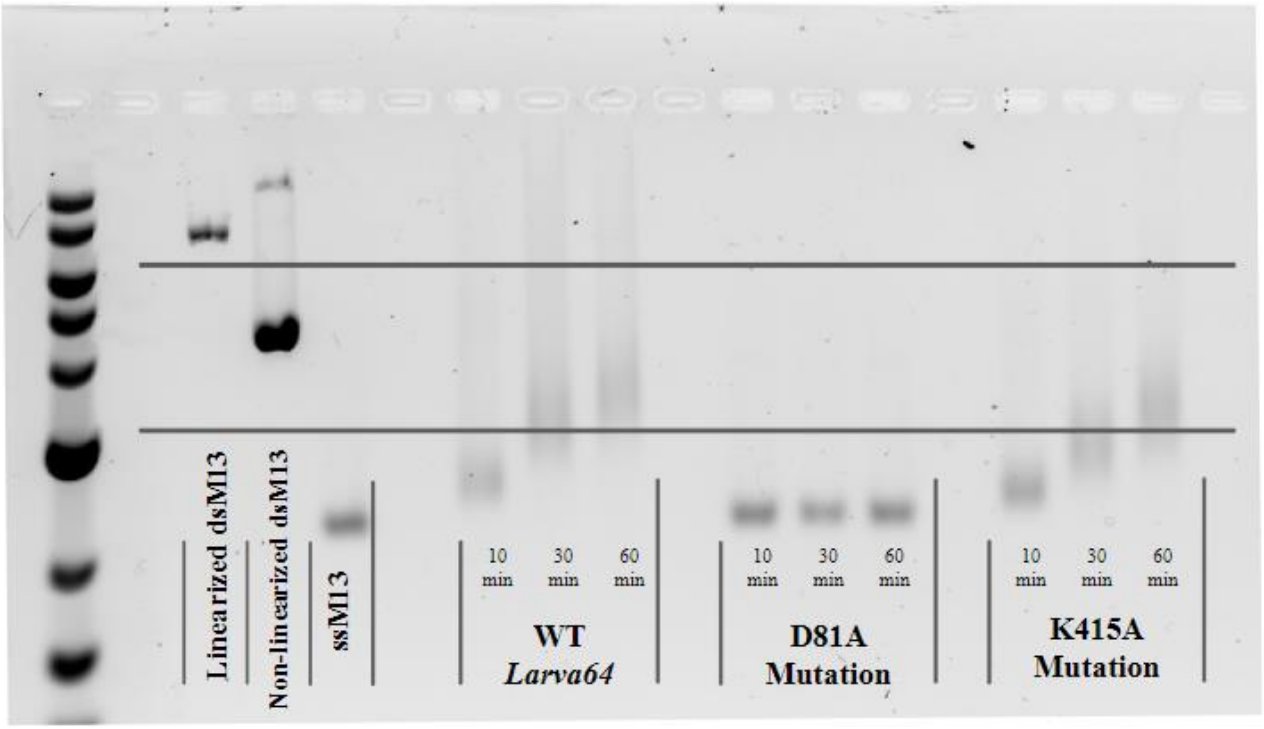

Figure 5: Agarose gel showing DNA synthesis with differing Larva64 mutations. Black horizontal lines indicate where 3000 and 6400 bp dsDNA products should be. Quenched at 10-, 30-, and 60-minute intervals, we see that there is no synthesized product for the D81A mutation.

Non-linearized dsDNA was lower on the gel than its linearized dsDNA counterpart as the overall size of the product was smaller due to supercoiling of dsDNA. This also shows that the 6.4 kb line used as a "finish point" for synthesis is fairly accurate as it falls between the linearized and non-linearized dsDNA product. The D81A mutation, which is associated with the N-terminal prim-pol domain, had no activity whatsoever, which is to be expected given its role in polymerization activity as part of the active site of the prim-pol domain. The C-terminal K415A mutation associated with the helicase/ATPase Walker A motif had activity comparable to that of the wild type (WT) prim-pol. However, there seems to be less streaking of the lanes, indicating perhaps a lack or lessening of SD synthesis. This tells us that the K415 residue, and potentially the entire C-terminal domain, is not necessary for DNA synthesis, whereas the D81 residue is, and that these two residues work independently of one another.

Optimization of Assay NaCl and pH Conditions:

All tests prior to assay condition optimizations contained 50 mM NaCl and were at a pH of 7.0. At the later stages of this research, optimizations of assay conditions were conducted to see if NaCl or pH adjustments could increase polymerization speed. These optimizations, where we systematically adjusted both the NaCl concentration and the pH, were determined using the gel-based assay. Reactions at differing buffer pH were quenched in 10-minute intervals, up to 40 minutes (Figure 6).

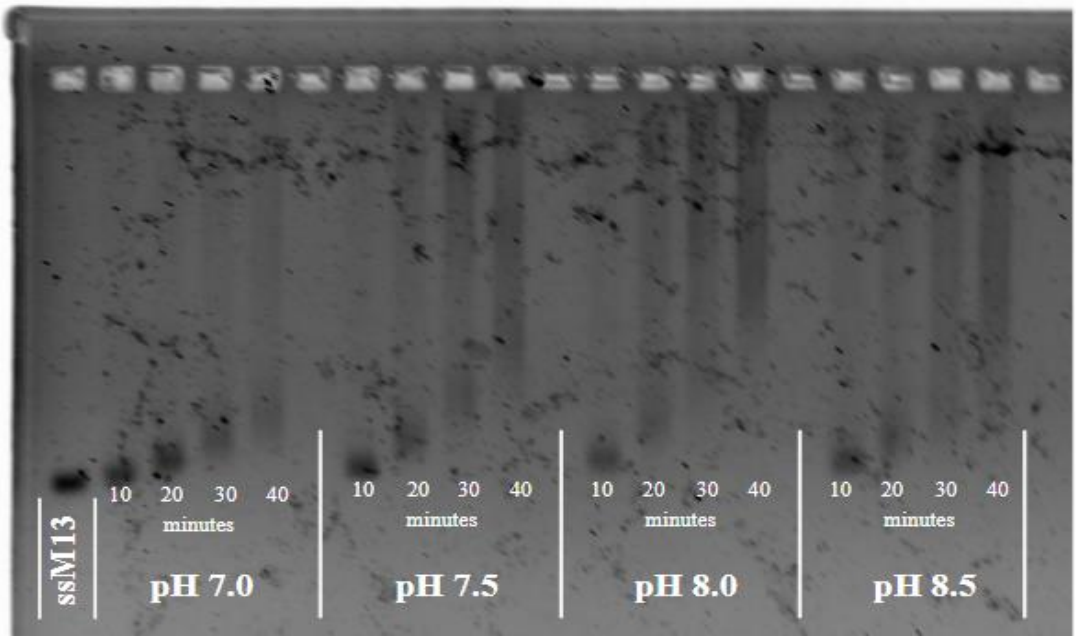

Figure 6: Agarose gel showing pH optimizations. Moving from left to right, the size of dsDNA product at identical timepoints increases with respective pH increases, with the highest being at a pH of 8.0. Spots and blurriness of the gel image are artifacts from the gel mold and the gel shifting position during imaging.

As the pH of the buffer increases, so too does the size of the dsDNA product at later time points (see lanes 7-10 and 12-15 in Figure 6). The pH 8.0 buffer appears to have the largest and darkest bands at later timepoints, indicating a larger product as well as potential SD synthesis. This

is notable as the NEB2 buffer, the buffer recommended for polymerase Klenow fragment, is at a pH of 7.9 (31). A buffer pH of 8.0 was adopted for later optimization conditions. NaCl concentration was then adjusted in an assay similar to the pH optimization, with identical quenching timepoints. NaCl concentrations were halved and doubled from the 50 mM used for previous assays, and excluded for one reaction as $MgCl_2$ would possibly be able to be used as a salt for polymerization activity in lieu of NaCl (Figure 7).

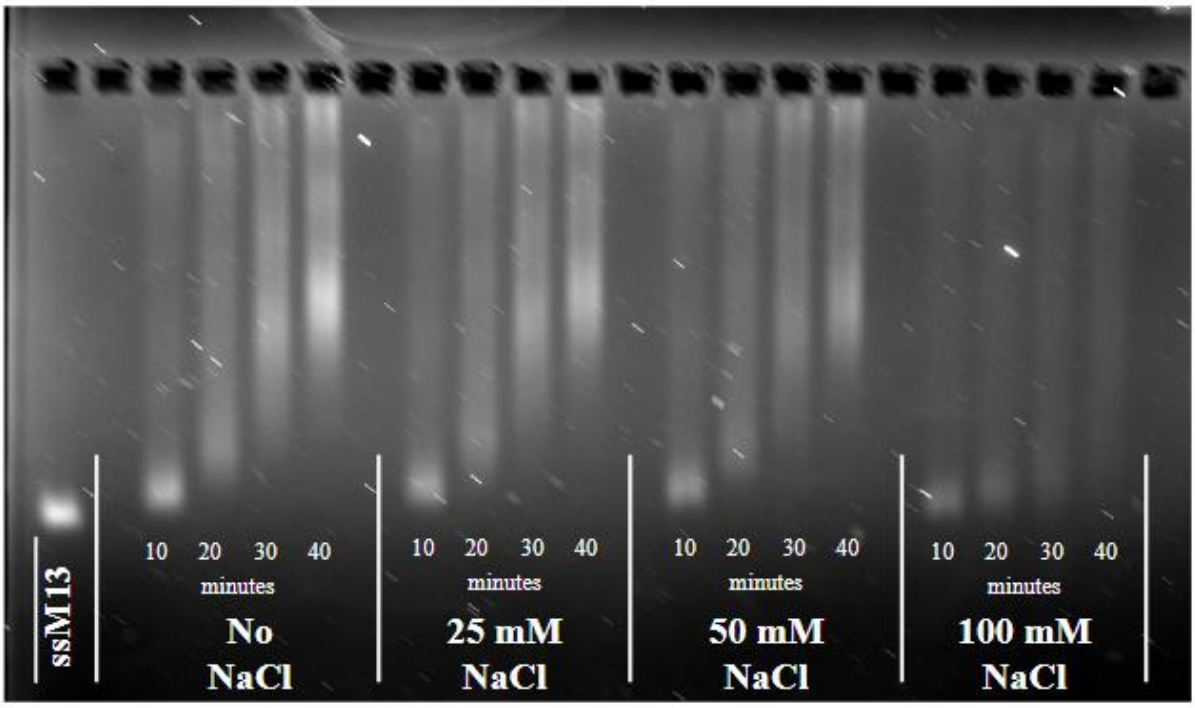

Figure 7: Agarose gel showing NaCl optimizations. Moving from left to right, the size of dsDNA products at identical timepoints increases with respective NaCl concentration increases. Spots and blurriness of the gel image are artifacts from the gel mold (debris) and the gel shifting position during imaging.

An increase in NaCl concentration to 100 mM inhibits activity, while a decrease in NaCl to 25 mM and even the outright removal of NaCl increase activity. The higher amount of NaCl likely caused an overabundance of competition for the DNA binding site of the enzyme, thus restricting activity. A buffer NaCl concentration of 25 mM was adopted for later assay conditions, due to it having a

higher activity than the 50 mM NaCl condition. These optimizations were the only conditions altered that seemed to increase prim-pol activity.

Confirmation of SD Synthesis by Larva64:

The streaking bands present in our gel-based assays indicate a distributive polymerase that only synthesizes short stretches of DNA before dissociation, the polymerase "falling off" of the DNA strand. Both T7 polymerase (a polymerase from phage T7) and *Escherichia coli* Klenow fragment are incapable of *de novo* DNA synthesis but overall are more processive polymerases (32, 33). T7 is incapable of SD synthesis, meaning that the double-stranded product produced by T7 will not be able to go past the 6.4 kb of the M13 template DNA. The demarcation line of the T7 polymerization product, shown in Figure 8 with a dashed line, shows where the final product of dsM13 is, and any product past the line is the result of SD synthesis. Klenow fragment, which is capable of SD will inevitably go past the demarcation line to serve as a positive control for what an enzyme capable of SD synthesis would look like. *Larva64* was able to prime and hand off synthesis to faster polymerases, suggesting it synthesizes a short DNA product to serve as the primer for both T7 and Klenow when solutions of the respective proteins are added to the *Larva64* assay reaction. Reactions at 2, 4, 6, 10, 30, 60, and 90 minutes show the progression of synthesis of *Larva64*, T7, and Klenow over time on a gel (Figure 8).

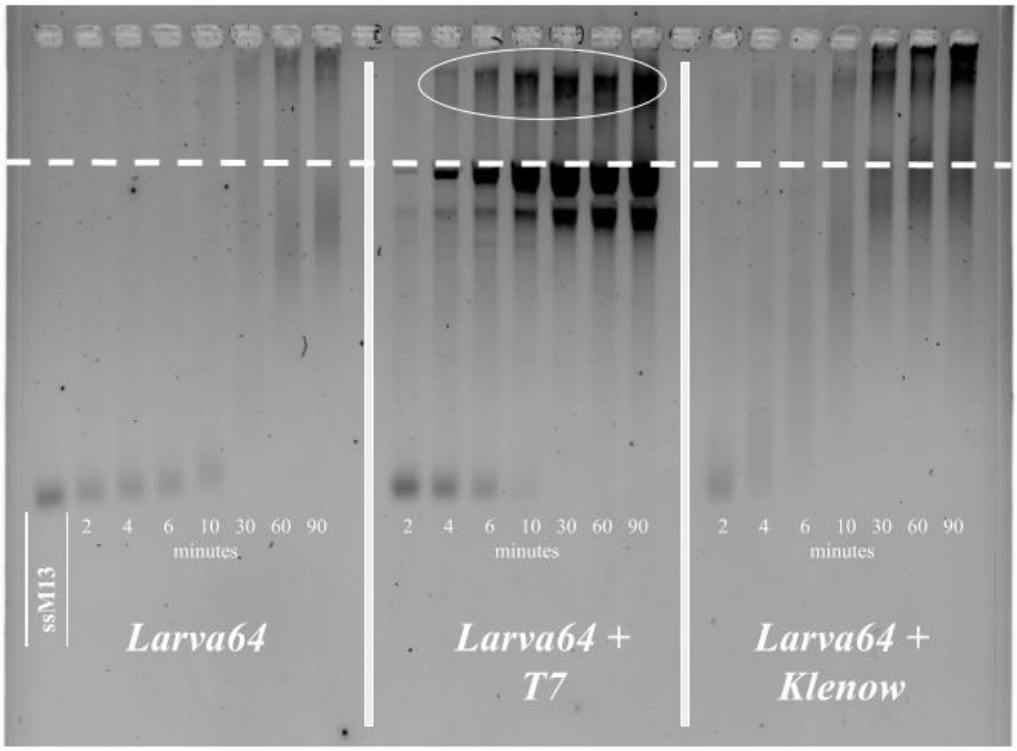

Figure 8: Agarose gel showing synthesis of Larva64, T7, and Klenow. Larva64's DNA product at longer time points goes past 6.4 kbs, showing that Larva64 is capable of strand displacement.

T7 polymerase stalls only minutes after being primed, showing that it is a much faster polymerase than either Klenow or *Larva64.* However, there is an accumulation product after the stalled T7 product shown at the dashed demarcation line. T7 is incapable of SD activity, so any product past the demarcation line, circled in white, is from SD synthesis by *Larva64* on the completed M13 dsDNA product synthesized by T7 polymerase. This is further evidence of *Larva64's* SD activity. Klenow fragment and T7 polymerase have polymerization rates of 13.5 and 130 bases/second respectively (34, 35). Klenow synthesizes dsDNA faster than *Larva64* and at later time points, the dsDNA product is larger than that of the stalled T7 synthesis dsDNA product, demonstrating SD synthesis. While *Larva64* is slow, given enough time, there are dsDNA products past the 6.4 kb demarcation line, showing that it is capable of SD synthesis. This confirmation of SD by *Larva64* is

novel to research on *Larva64* and expands upon observations of "streaking" at long time points for our gel-based assays.

Discussion

Larva64 has continued to show that it is a robust DNA primase-polymerase capable of *de novo* DNA synthesis. Attempts to increase the synthesis rate of *Larva64* were relatively unsuccessful. Neither the addition of rNTPs nor the substitution of Mg^{2+} increased the apparent rate of DNA synthesis via gel-based assays. Interestingly, later optimizations of buffer conditions, namely pH and NaCl concentration, did yield apparent increases in prim-pol activity via the gel-based assay. Interestingly, the observed increase in *Larva64* prim-pol activity via the gel-based assay appears to be an artifact or attributable to another factor. The rate of DNA synthesis, as measured by fluorescence, did not increase under our experimental conditions, which may indicate that the polymerization activity of *Larva64* may not have been optimized via the change in condition.

What is not currently apparent is whether *Larva64* synthesizes the entirety of the complement M13 strand during a single binding event, or if it does it piecewise beginning at differing starting locations after dissociating and synthesizing until the entirety of the circular ssDNA is dsDNA. Given that *Larva64* is able to hand off primers to the faster polymerases, T7 and Klenow, we know that it dissociates from the DNA strand at some point. As the prim-pol comes back to synthesis after dissociation, it is unclear where its new starting site would be, and whether it could begin SD synthesis on a small piece of dsDNA it had already synthesized without the completion of the entire plasmid. These qualitative assays have confirmed that *Larva64* continues to be a very slow polymerase. Requiring nearly an hour to synthesize 6400 bases is exceptionally slow, especially when compared to the 62991 bp genome size of phage Larva. Based on this data, Larva either requires nearly 10 hours for its prim-pol to replicate its genome once, or there are other

proteins, factors, or interactions that are occurring that are aiding in the synthesis of the Larva genome.

An interesting pattern that became apparent from the “Long Run" gel and other assays, which was greater than 60-minutes, was that *Larva64* appears to produce the double-stranded DNA product that was at or larger than 6.4 kb. This contradicts the observations of Folse (2020) who reported an apparent stalling at 3 kb, even at a time point of 60 minutes. Although the reasons for the different findings are unclear, they may be the result of better purified and folded *Larva64* enzyme or due to this research using purchased ssM13 rather than producing it in house. It appears that with enough time, *Larva64* is capable of synthesizing the entirety of the complement M13 strand. Moreover, by going beyond the 6.4 kb marker, we are able to determine that our enzyme is able to perform SD synthesis. The SD activity of *Larva64* was confirmed by comparing the dsDNA products of *Larva64* with T7 and Klenow polymerase.

The handoff of activity from *Larva64* to Klenow is of note, as Klenow is a bacterial polymerase. Many bacteriophages utilize their host polymerase proteins for their genome synthesis. This lends credence to the idea that perhaps our prim-pol, *Larva64*, serves simply as a primase and the polymerase from the host bacterium, *M. smegmatis,* synthesizes the viral genome. Future research could focus on measuring the specific handoff from *Larva64* to the host polymerase.

CHAPTER THREE: RESULTS OF ENZYME ACTIVITY QUANTIFICATION

In Vitro Quantitative Characterization of *Larva64*

The rate of dsDNA synthesis was determined for *Larva64* using PG and a dsM13 calibration curve. A handoff of synthesis from *Larva64* to a bacterial polymerase, Klenow fragment, was also characterized quantitatively using the PG assay. That handoff could be identified by a two-phase synthesis observable by an abrupt linear fluorescence increase after 5 minutes. The rate of synthesis for the Klenow fragment in our assays was then compared to the theoretical values from NEB so as to validate our assays. Lastly, a rate of dsDNA synthesis was determined for optimized assay conditions adopted in qualitative characterization. Strangely enough, the apparent increase in *Larva64* prim-pol activity via the gel-based assay seems to be somewhat of an illusion or something else entirely. The rate of DNA synthesis determined via fluorescence did not increase under our optimal conditions, which may indicate that the polymerization activity of *Larva64* may not have been optimized via the change in condition. All qualitative PG assays were performed in triplicate, and error bars shown in figures were based on a 95% confidence interval.

PicoGreen™ Quench Assay

PG was originally chosen for this study as a potential candidate for a fluorescent dye to be used in a real time assay (36); however, our research shows that at any detectable concentrations, PG stops dsDNA polymerization by *Larva64*. While unsuitable for real-time assays, the dsDNA concentration determination ability of PG at various quenched timepoints made it useful for determining the rate of synthesis of *Larva64*. In the literature, PG has been used to quantitatively

determine a DNA polymerization rate of T7 polymerase and Klenow fragment utilizing a quenched assay (19).

dsDNA Calibration Curve:

A PG:dsM13 calibration curve was produced to show the relationship between dsDNA concentration and PG fluorescence for our particular assay. The concentration of PG used in the development of the calibration curve was identical to that used with enzymatic rate samples. Data analysis for the calibration curve and for the samples mandated that the fluorescence of the sample with no dsDNA would be subtracted from the other samples, so as to standardize. The PG:dsM13 calibration curve has the X-axis is in units of ng of dsDNA and the Y-axis is in units of RFU (Figure 9).

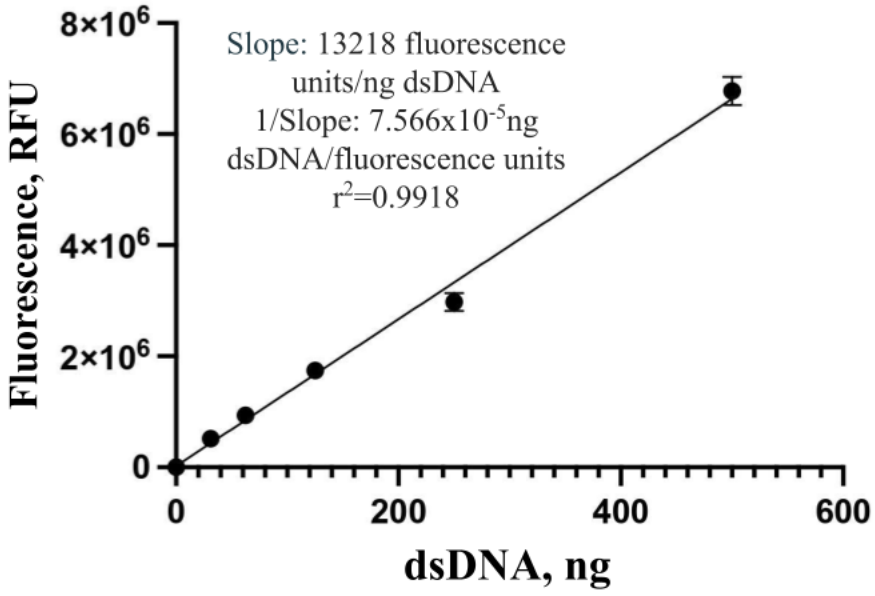

Figure 9: Calibration curve of dsM13 DNA and PG fluorescence. As dsDNA concentration goes up, so too does fluorescence. The relationship between dsDNA concentration and fluorescence via PG is linear. Slope, inverse slope, and r^2 value can be seen. Data was identical for smaller dsDNA concentrations, resulting in error bars being too small to be seen.

The resulting inverse slope of the calibration curve can be multiplied by the slope of kinetic data (fluorescence/minute) from a quenched assay to determine a rate of synthesis with units of ng dsDNA/minute.

Larva64 Rate and Polymerase Handoff:

A graph of fluorescence over time (Figure 10), demonstrates *Larva64* activity (black trend line). The slope of 144611 fluorescence units per minute was multiplied by the inverse slope of the dsDNA calibration curve, giving a rate of synthesis for *Larva64* of 10.94 ng dsDNA/minute. Results in Chapter 2, qualitative characterization, demonstrated that *Larva64* is able to hand off synthesis to the faster Klenow fragment. This bacterial polymerase was chosen as it is well characterized in the literature and has a speed slow enough, albeit faster than *Larva64*, to be accurately measured using our reagents and equipment on hand. Samples were prepared that contained Klenow and *Larva64* and quenched at rapid timepoints (Figure 10).

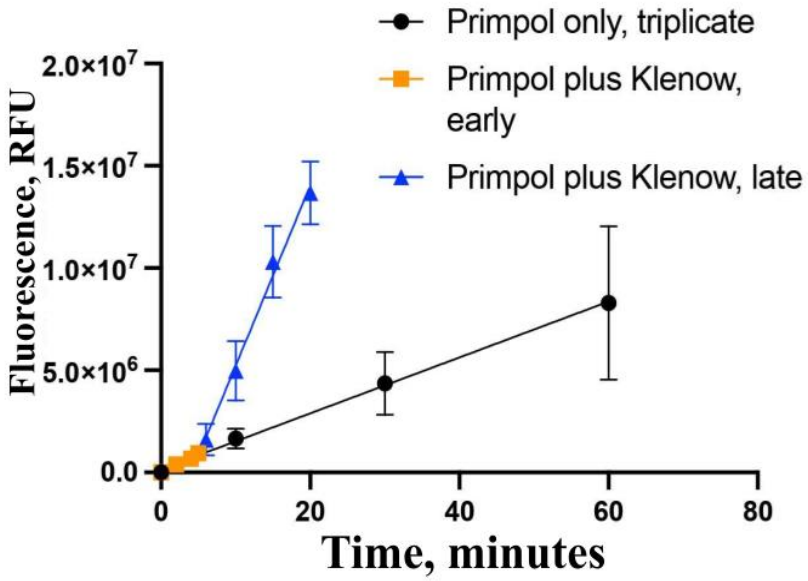

Figure 10: Rate for Larva64 (Primpol only) and Larva64+Klenow (Primpol plus Klenow). Polymerization slopes were linear. The early slope of Larva64+Klenow was nearly identical to that of Larva64 alone until 5 minutes, when there is a drastic increase in slope. The increase in slope is due to a handoff to Klenow, a faster bacterial polymerase. Error bars are based on a 95% CI.

The orange squares indicate timepoints from 0-5 minutes for *Larva64+Klenow* and blue triangles indicate timepoints from 6-20 minutes for *Larva64+Klenow.* The rate of synthesis for *Larva64+Klenow* is nearly identical to that of just *Larva64* for the first 5 minutes of synthesis. At the 6-minute mark, there is an abrupt increase in fluorescence, going to a slope of 882190 fluorescence units per minute, which translates to a rate of 66.74 ng dsDNA/min. The apparent two-phase synthesis indicates a handoff from the slower *Larva64* to that of the faster Klenow fragment. The slope of the blue line can be denoted as the synthesis rate of Klenow, a well-studied polymerase. When comparing the slopes of *Larva64* to that of the blue line, Klenow synthesis, *Larva64* is 6.5 times slower than Klenow. Prior studies have demonstrated the polymerization rate of Klenow fragment is 13.5 bases/second (34), which would make *Larva64*'s rate approximately 2.1 bases/second. A validation of the rate of Klenow in our study compared to NEB's specifications (Chapter Six: Materials and Methods) found that there is a 30.76% difference in rate between the theoretical and experimental rate, which can be classified as our experimental error. The pH condition used in the quantitative PG assays was 7.0 due to the buffer in use at the time and Klenow fragment kinetics are typically done at a pH of 7.9, so a discrepancy in speed for our experimental value compared to the theoretical is likely due to the lower pH. Overall, the rates are close enough to one another to validate that our enzyme kinetics experiments with PG are valid, as is our rate approximation for *Larva64.*

Optimized Assay Condition Rate:

During qualitative characterization of the enzyme, later optimizations of the assay buffer conditions were made, namely the change of pH to 8.0 and decreasing NaCl concentration to 25 mM. The results of the gel-based assays would suggest that the optimized conditions would have a

faster rate of synthesis due to apparent increased activity and larger dsDNA products in the gels. The rate of synthesis of *Larva64* in the new optimized assay conditions was determined to be 122405 fluorescence units per minute, which converts to 9.26 ng dsDNA/minute. Figure 11 shows the fluorescence results of the original (pH 7.0) vs. optimized assay (pH 8.0), along with results from the two-phase Klenow/*Larva64* synthesis.

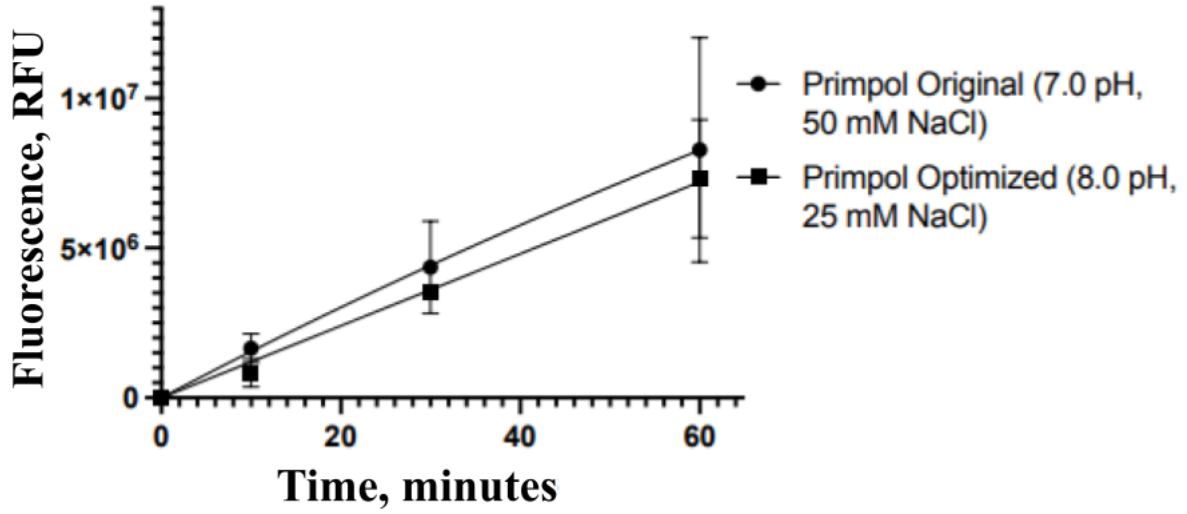

Figure 11: Rate for Larva64 optimized assay conditions (8.0 pH and 25 mM NaCl) vs. original assay conditions (7.0 pH and 50 mM NaCl). There is no statistical difference between optimized and original assay condition polymerization rate based on a 95% confidence interval.

The error bars and data from optimized and original assay conditions overlap, so while the optimized condition may be slower based sheerly off the slope, the original condition is not statistically faster based on a 95% confidence interval. The slowness of the optimized condition is exceptionally odd given the results from the qualitative characterization showing such a drastic increase in activity at a higher pH and marginally more activity at a lower NaCl concentration. However, we observed the fluorescence values for the optimized condition were far more consistent than our original assay conditions, resulting in smaller error bars.

Discussion

One of the chief goals of this research was to determine a quantifiable rate of synthesis for *Larva64*. Considering that the group where this research was conducted dissuaded use of radioactive isotopes, measuring dsDNA synthesis using radioactivity was out of the question. PG in retrospect is not necessarily the most ideal dsDNA fluorescent dye to use for the study due to its limitations in real time assays, as opposed to *SYBR^{TM}Green* which has been used to measure real time limited polymerase activity (37), although we are not certain that it could be used for our prim-pol. With that in mind, the usage of PG to measure DNA synthesis is both grounded in the literature and is relatively novel for prim-pols, with this research being the first study to use PG to measure dsDNA synthesis of a viral prim-pol. We did not investigate the use of alternative dsDNA binding fluorescent dyes in this study, but the utilization of a non-inhibitory dsDNA fluorescent binding dye such as *SYBR^{TM}Green* in future research would likely give cleaner results and show more clearly the hand-off between *Larva64* and Klenow.

The observation of a two-phased DNA synthesis in the fluorescence assay with *Larva64* and Klenow shows clearly that a handoff between the two polymerases takes place. This shows that *Larva64* would be able to hand off synthesis to a bacterial polymerase, potentially in the *M. smegmatis* host. Moreover, seeing that the rate synthesis from 0 to 5 minutes had a slope nearly identical to that of just *Larva64* shows that the prim-pol itself does not dissociate immediately after initiating synthesis. Rather it likely stays on for a short period of time before falling off and handing off synthesis to Klenow, serving as a primase for Klenow polymerase. If *Larva64* is acting solely as a primase for a bacterial polymerase, then it is slow by primase standards. Requiring 5 minutes to synthesize one primer is especially slow compared to DnaG primase from *E. coli*, which synthesizes RNA primers at a rate of one primer per second (38). Klenow requires at least a 7 base

primer to initiate synthesis (39), implying that *Larva64* synthesizes a 7-base primer in 5 minutes, or 1.4 bases/minute, to which Klenow is able to take over synthesis. This would make *Larva64* have a rate of ~0.023 base/second, substantially slower than the experimentally determined 2.1 base/second in this study. This discrepancy casts doubt on the idea that it requires 5 minutes for an adequate primer length to be synthesized, rather the "handing off" of synthesis is due to *Larva64* dissociating. Why *Larva64* would consistently dissociate after 5 minutes is still a mystery.

The lack of a significant increase in rate for the optimized conditions is highly intriguing. All evidence from the qualitative results indicates that there should be a drastic increase in polymerization speed using the fluorescent assay. What we observed is a non-significant change, bordering on inhibition. While not statistically slower than the original assay condition, it is exceptionally odd. One potential avenue that could explain why its polymerization rate is not faster despite the qualitative results is that higher pH and lower salt may allow for better, more consistent priming and incorporation of nucleotides. This increased consistency of priming and incorporation would lead to a cleaner dsDNA product, but the overall polymerization speed may not be affected. The quantitative rate of the optimized condition, which had smaller error bars and more consistent data, aligns with this idea. Future research with *Larva64* would likely still need to use the optimized conditions, as empirically the enzyme itself seemed to have done better at a pH of 8.0 in terms of solubility and enzymatic activity. Overall, the quantification of *Larva64*'s synthesis rate confirms and builds upon what we already know: *Larva64* is slow. Now, however, we have a number and a method to quantitatively reflect that lack of haste.

CHAPTER FOUR: BACTERIAL TWO-HYBRID ASSAY AND BIOINFORMATIC ANALYSIS

Two-Hybrid Assay of Larva Endogenous Proteins

To determine the interactions of select endogenous proteins and *Larva64*, we employed a two-hybrid assay with *Larva64* being the "bait" and other Larva proteins being the "prey". If PPi occurs, a phenotypic change will occur which will allow for growth of bacterial colonies to be observed in the presence of carbenicillin. The only PPi observed was between *Larva64* and *Larva48*, a sliding clamp protein. This interaction was evidenced by growth on a 750 µg/mL carbenicillin dosed growth plate. This interaction of prim-pol and sliding clamp is novel to the literature. Bioinformatic analysis using Phamerator's database revealed that *Larva64*-like prim-pols and *Larva48*-like sliding clamps only co-exist in cluster K5, however many genetically similar phages to Larva do not possess a sliding clamp protein. Multiple models of *Larva64* in varying stoichiometric orientations were generated, revealing a possible hexameric and monomeric conformation, as well as avenues of *Larva48* assistance in both conformations. Lastly, HHpred residue alignment showed that the different domains of *Larva64*, prim-pol and helicase domains, appear to have differing preference in stoichiometric orientation, monomer and hexamer respectively.

Rationale for Prey Protein Selection:

Four endogenous proteins were chosen for B2H, *Larva48*, *62*, *63*, and 65, all of which correspond to predicted DNA metabolism proteins in the viral proteome (Figure 12). *Larva48* is a predicted sliding clamp protein, which are capable of increasing the incorporation of nucleotides by polymerases upwards of 100-1000 fold by tethering polymerases to the template DNA strand, thus preventing dissociation and increasing processivity (40). As it stands in the literature, no interaction

between a viral prim-pol and a sliding clamp protein have been observed. *Larva62* and *63* are both annotated as HNH-endonucleases which have a striking genetic similarity to Cas (CRISPR associated) proteins. Cas proteins are endonucleases, DNA cleaving enzymes, that are part of bacterial immune defense against viruses (41). There is a category of prim-pols that are preceded by CRISPR-related proteins known as CRISPR Adjacent Primase Polymerases, CAPPs (42); these CAPPs interact with their respective CRISPR proteins to act in various DNA repair and replication roles. No CAPPs have been characterized in bacteriophage, but an interaction between a prim-pol and preceding Cas proteins in a viral proteome could reveal CAPPs as a functional category in bacteriophage. Lastly, *Larva65* is annotated as a helicase and not necessary for viral survival as determined by Folse, 2020. The identification of interactions between endogenous proteins and *Larva64* will give insight into mechanisms that may aid the *Larva64* prim-pol in DNA synthesis or reveal different functions of *Larva64* outside of genome replication which makes it necessary for viral survival.

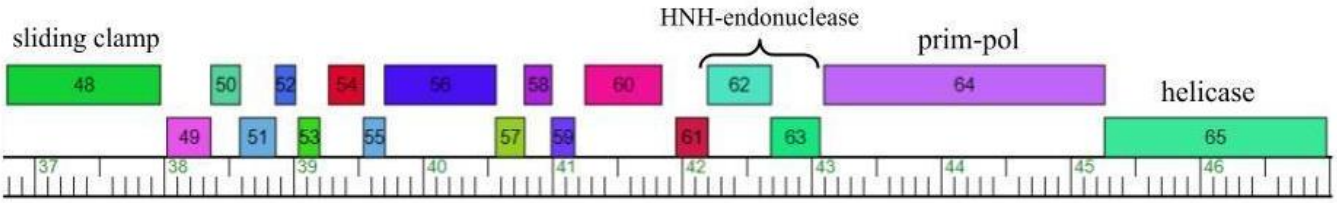

Figure 12: Larva genome map showing the locations of the proteins tested using B2H. Genes 48 (sliding clamp), 62, 63 (HNH-endonucleases), and 65 (helicase) were prey while gene 64 (prim-pol) was the bait.

Results of B2H:

Growth on a selection plate that contains carbenicillin is indicative of an interaction between *Larva64* and the other tested proteins. Growth plates which lack carbenicillin are used to observe what growth should theoretically look like if a strong PPi occurs which prompts

carbenicillin resistance. Starting on the left side of the plate, we have undiluted cells, and as we move to the right the cells are diluted serially 10-fold. The growth plate is on the left, while the selection plate (750 µg/mL carbenicillin) is on the right (Figure 13).

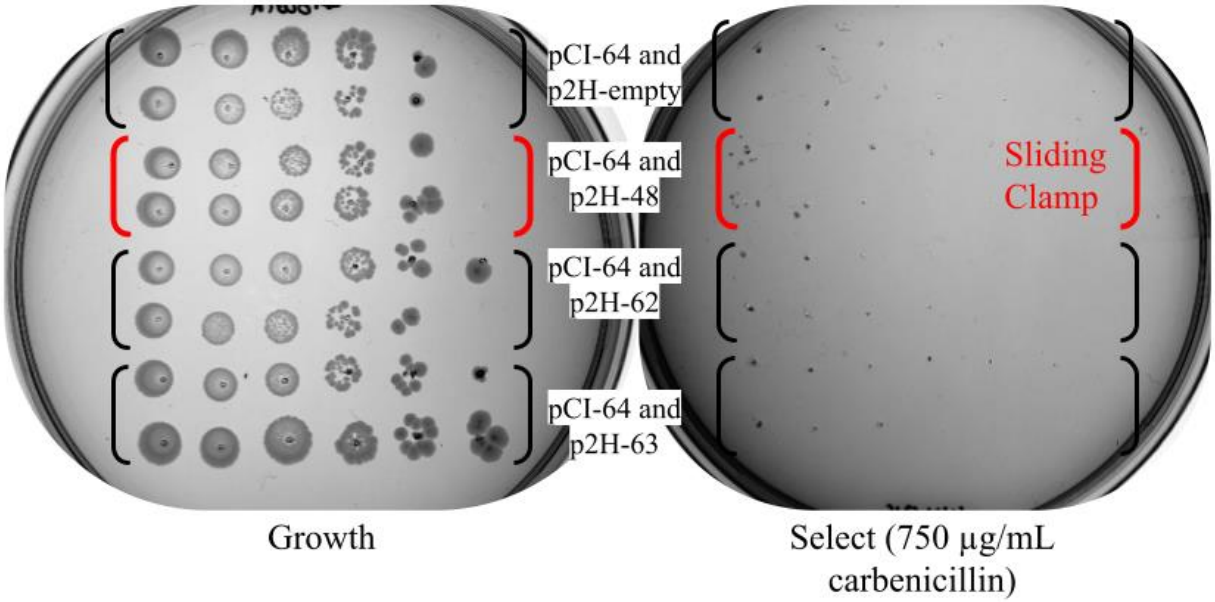

Figure 13: B2H growth plate (left) and selection plate (right) for *Larva64 used to determine PPi. Growth on the select plate for pCI-64 and p2H-48, seen in red brackets, indicates a PPi between the Larva64 prim-pol and Larva48 sliding clamp.*

While small, there are a number of colonies on the selection plate at the highest concentration of cells corresponding to a *Larva64:Larva48* PPi. There are no colonies on the Select plate for any other protein tested. The lack of interaction between the prim-pol and *Larva62, 63* seemingly rules out *Larva64* as a CAPP. The lack of interaction between the prim-pol and *Larva65*, along with Folse's CRISPRi research that showed *Larva65* was non-essential, shows that the Larva helicase may not be necessary at all for the replication of the viral genome. The small number of colonies for *Larva64:Larva48* at a high cell concentration is indicative of a weak PPi between the viral pim-pol and sliding clamp, respectively. An interaction between a sliding clamp protein and a prim-pol, even a weak interaction, is a novel development. Given that the literature has indicated that sliding

clamps can increase the speed of otherwise poor polymerases by upwards of a thousand-fold, this shows that *Larva64* may be receiving help from its sliding clamp to polymerize its viral genome.

Bioinformatic Analysis

Bioinformatic Observations of Larva48 and 64 Utilizing Phamerator Database:

Bioinformatic software *Phamerator* (Database: Actino_Draft (Version 578) was utilized to make observations about the presence of *Larva48* and *Larva64*-like proteins in other bacteriophages. All phage in cluster K5 have a conserved prim-pol identical to that of *Larva64* which appears to be unique to the K5 cluster. A handful of phages within K5 possesses a sliding clamp identical to *Larva48,* those being phage Feyre, Gengar, Guillsminger, Heftyboy, Leston, Miryou, OkiRoe, Omnicron, Paola, SoSeph, Thyatira, and Waterfoul, making up 59% of cluster K5. *Larva48*-like sliding clamp proteins can be found in a number of phages from additional clusters including AO1, AO2, AO3, AO, AR, AY, FN, K5, and 2 singletons, all with an identical pham (160319). And while many of these clusters that contain *Larva48* like sliding clamps also possess polymerases, only K5 possesses prim-pols, all being *Larva64* conserved. The clamp is completely conserved and present in all phage from clusters AO1, AO2, AO3, AO, AR, and AY. Curiously, it is only present in 66% of phage in cluster FN and 59% of phage in cluster K5, Larva being one of them. All phage in cluster K5 possess a *Larva64* like prim-pol, however, 41% of phage in cluster K5 do not possess a *Larva48*-like sliding clamp. An example of a phage in cluster K5 in which a sliding clamp is absent is phage InvictusManeo, which has a nearly identical gene map to phage Larva sans a sliding clamp, suggesting that a sliding clamp may not be necessary for comparable phage with *Larva64*-like prim-pols.

Modeling and Characterization of Larva48 and 64 Utilizing AlphaFold 3:

Utilizing the AlphaFold 3 server, multiple structure models were generated for *Larva64* with varying copy numbers, the addition of *Larva48,* and the addition of dsDNA and ssDNA. While not definitive descriptions of the structure of our proteins, the resulting structures can then be compared to closely aligned amino acid sequences of our proteins to other proteins with known structures using HHPred. By comparing predicted enzyme structures to those of comparable residue alignment, we can predict what the structure of the enzyme is like and whether it is a monomer, dimer, hexamer, etc.

The AlphaFold 3 metrics by which it measures the accuracy of its models are plDDT, pTM, and ipTM scores. plDDT is a per-atom confidence estimate on a 0-100 scale where a higher value indicates higher confidence that only considers distances to polymers. The plDDT is shown as color outputs in the image of the structure. In this scale, dark blue (plDDT > 90) indicates very high confidence of structure, light blue (90 > plDDT > 70) indicates fairly confident, yellow (70 > plDDT > 50) indicates low confidence, and orange (plDDT < 50) indicates very low confidence in structure. pTM and ipTM scores stand for predicted template modeling (pTM) score and the interface predicted template modeling (ipTM) score. Both scores are derived from a measure called the template modeling (TM) score which measures the accuracy of the entire structure (43, 44). A pTM score above 0.5 means the overall predicted fold for the complex might be similar to the true structure. Since these metric measures interactions within a complex, monomers that are modeled do not have this score. ipTM measures the accuracy of the predicted relative positions of the subunits within the complex, and is a metric shown for monomers. Values higher than 0.8 represent confident high-quality predictions, while values below 0.6 suggest a weak prediction. TM

derived scores are very strict for small structures or short chains, so for larger structures, plDDT may be more indicative of prediction quality.

Among the models generated with AlphaFold 3, the ones that will be discussed in this paper are *Larva64* as a monomer, a dimer, and a hexamer. *Larva48* has not had a structural determination study conducted on it specifically, however sliding clamps like it are firmly established as a dimer, and is included essentially as a positive control for our modeling (45) . AlphaFold 3 models for *Larva64* as a monomer, dimer, hexamer and *Larva48* dimer along with respective pTM and iPTM values show the confidence of the models and the potential structure of the proteins (Figure 14).

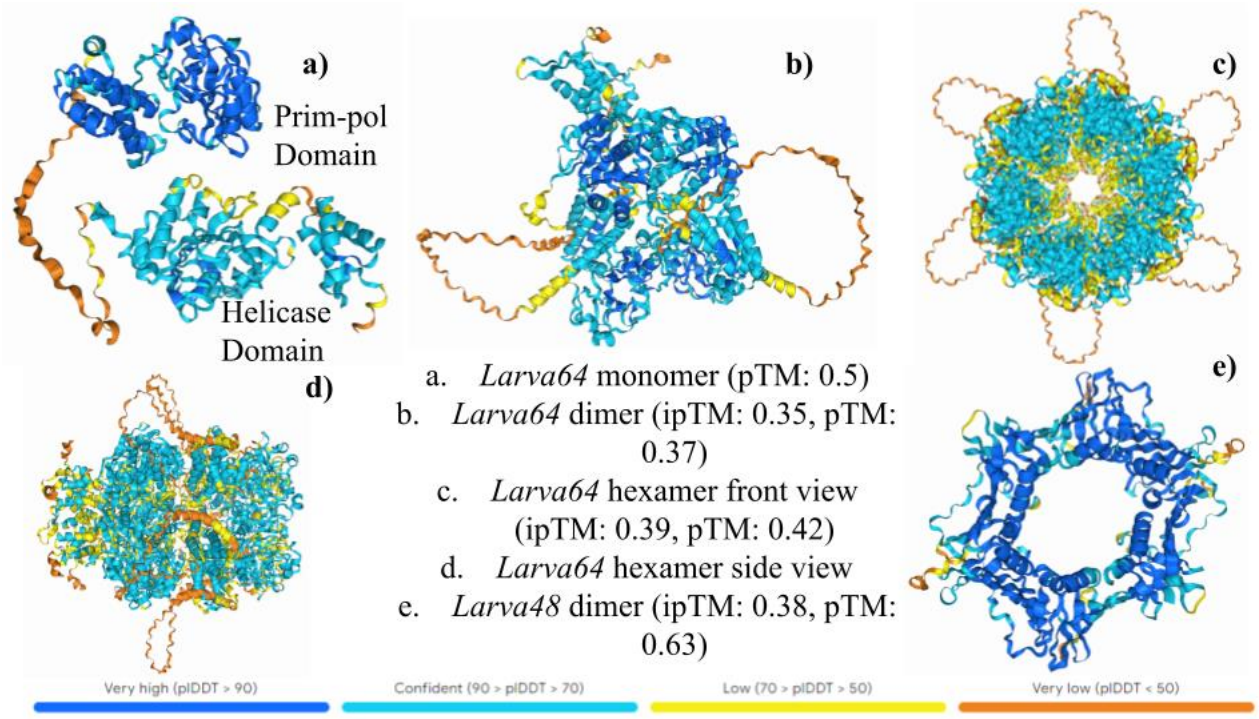

Figure 14: AlphaFold 3 models for Larva64 and Larva48. a) shows Larva64 monomer, b) shows Larva64 dimer and c) and d) show Larva64 hexamer (front and side view). e) shows Larva48 dimer. The key of colors as they correspond to plDDT values are found at the bottom of the figure.

From these models, especially in Figure 14a we can observe a fairly confident structure of the helicase and prim-pol domains, with a very low confidence strand of disordered residues

connecting the two. Of the *Larva64* models, the monomer has the highest pTM score, followed by the hexamer and dimer. *Larva48* which as a sliding clamp can be firmly described as a dimer, has a higher pTM value than all structures shown indicating high confidence in the protein structure. Folse modeled *Larva64* as a monomer, but never as a dimer or hexamer. The hexamer confirmation of *Larva64* is notable as a "hole" can be observed in the middle of the structure that could likely accommodate ssDNA and dsDNA. To determine the accommodation of DNA in the structure of *Larva64* and *Larva48*, models were produced with the addition of 60mer dt40 ssDNA and dsDNA (Figure 15).

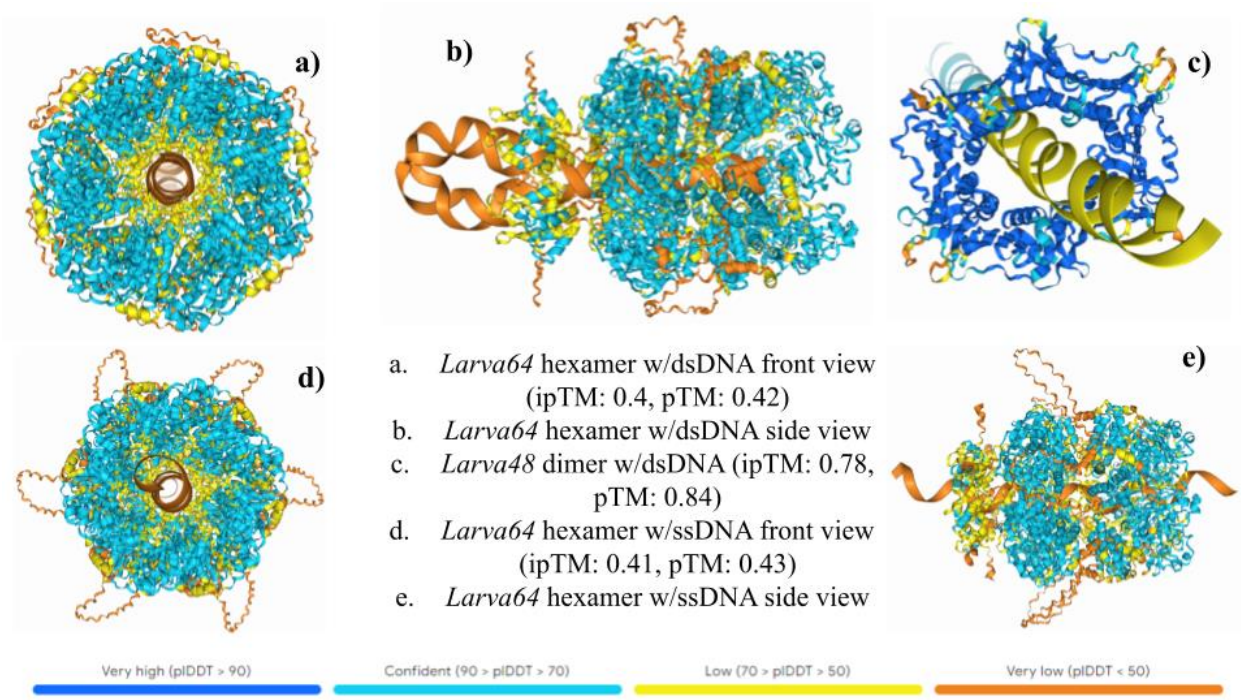

Figure 15: AlphaFold 3 models for Larva64 and Larva48 with 60mer DNA. a) and b) show Larva64 hexamer with 60mer dsDNA. c) shows Larva48 with 60mer dsDNA. d) and e) show Larva64 hexamer with 60mer ssDNA. The key of colors as they correspond to plDDT values are found at the bottom of the figure.

As would be expected, *Larva48* dimer with dsDNA running through the center (Figure 15c) has the highest ipTM and pTM scores, having a higher score than the protein structure without dsDNA. *Larva64* hexamer with dsDNA had a slightly higher ipTM score than its non-DNA added

counterpart. *Larva64* hexamer with ssDNA had a slightly higher ipTM and pTM score than its non-DNA added or dsDNA added counterparts. This shows that a hexameric model of *Larva64* can accommodate both dsDNA and ssDNA within its center. Interestingly, the hexamer model of *Larva64* with dsDNA appears to show the splitting of dsDNA into ssDNA and its subsequent recombination back to dsDNA. This may be a modeling of the helicase splitting and prim-pol recombination of DNA.

In our B2H assays, we observed a weak interaction between *Larva64* and *Larva48*, so we decided to model the two most promising *Larva64* structures (monomer and hexamer) with the addition of *Larva48* dimer. Moreover, one model of *Larva64* monomer with *Larva48* dimer also has the addition of 60mer dsDNA for *Larva48* to wrap around. The idea is that the sliding clamp could interact with the prim-pol in a way that holds the monomer prim-pol to the DNA so as to prevent dissociation and slowing synthesis of DNA. The models of *Larva64* (monomer and hexamer) with *Larva48* dimer were generated (Figure 16).

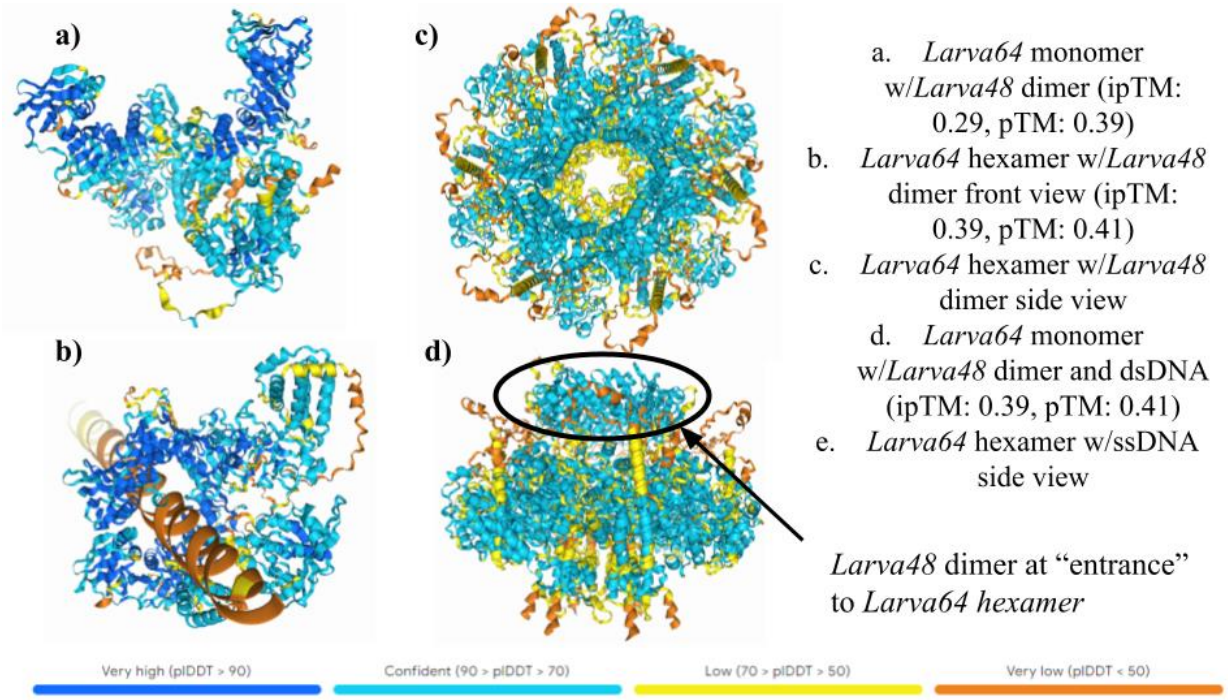

Figure 16: AlphaFold 3 models for Larva64 and Larva48 combinations. a) and b) show Larva64 monomer with Larva48 dimer. b) has the addition of 60mer dsDNA. c) and d) show Larva64 hexamer with Larva48 dimer. The key of colors as they correspond to plDDT values are found at the bottom of the figure.

As modeled, *Larva64* monomer with *Larva48* dimer without DNA (Figure 16a) gives a structure that bears little resemblance, or ipTM and pTM scores, to the original structures of its constituent proteins. However, when in the presence of dsDNA (Figure 16b), we observe the sliding clamp enveloping the DNA strand with *Larva64* attached to the clamp with a section of the prim-pol touching the dsDNA. The pTM score of the prim-pol/dsDNA/clamp complex is also higher than that of the monomer by itself, albeit lower than just *Larva48* dimer with dsDNA. This suggests that the structure of *Larva64* is potentially more stable as a monomer when coupled with *Larva48* sliding clamp around DNA. This supports the hypothesis that *Larva64* is aided by the *Larva48* clamp to not dissociate from the DNA strand so as to synthesize DNA at a faster rate. As shown in previous models, *Larva64* hexamer has a hole in the center which can accommodate DNA similar

to that of *Larva48* dimer. When the *Larva64* hexamer and *Larva48* dimer are modeled together, the sliding clamp sits at the entrance to the hexameric "hole". This could potentially serve as a support structure for *Larva64* hexamer, however, the lower ipTM and pTM scores than the non-combined counterpart cast doubt upon that idea. It seems that structurally, *Larva64* as a monomer by itself is preferred to that of a dimer or hexamer confirmation and is aided by the addition of *Larva48* dimer sliding clamp. *Larva64* as a hexamer is fascinating and still a possible structure as it can accommodate the envelopment of dsDNA/ssDNA. Based solely on the predicted structure of *Larva64* it seems that a monomer conformation is the most likely, although it may fluctuate between a monomeric and hexameric conformation based on activity. *Larva48*'s ring structure was accurately modeled as well and shown to interact with both monomer and hexamer confirmations of *Larva64*.

Characterization of Larva64 Structure Utilizing HHpred Alignment:

The entirety of the amino acid sequence for *Larva64* was submitted to HHpred (Job ID:1084060). HHpred is an online interactive server that predicts protein structure and function using amino acid sequences by detecting remote protein homology (46). A section of the results from HHpred were included in this study (Figure 17). There were hundreds of alignments found for the C-terminal helicase domain, with only a handful of alignments found for the prim-pol domain. For the sake of brevity, we will focus on the top 4 aligned proteins in the prim-pol domain and the top 4 aligned proteins from the helicase domain and their global stoichiometry (monomer, dimer, trimer, etc.)

Figure 17: HHpred displayed top results for amino acid sequence alignment. The top 4 aligned sequences in the N-terminal prim-pol domain are 3M1M_A, 1RO2_A, 6A9W_A, and 7OM0_C. The top 4 aligned sequences from the C-terminal helicase domain are from 1NLF_C, 1U94_A, 2DR3_A, and 8FWJ_D (Credit: HHpred Job ID:1084060).

Unsurprisingly, the top 3 proteins aligned for the prim-pol domain are from other prim-pols: 3M1M_A, 1RO2_A, and 6A9W_A. 3M1M_A and 1RO2_A from *Sulfolobus islandicus,* while 6A9W_A is from phage NrS-1. All of which exist as monomers (47, 48, 49). However, 7OM0_C comes from a primase-helicase from *Staphylococcus aureus,* which exists as a hexamer (50). The top 4 proteins aligned for the helicase domain were 1NLF_C, 1U94_A, 2DR3_A, and 8FWJ_D, all of which are DNA binding proteins (51, 52, 53, 54) that exist as hexamers. 1NLF_C and 1U94_A are from *E. coli,* 2DR3_A is from *Pyrococcus horikoshii OT3,* and 8FWJ_D is from *Cereibacter sphaeroides.*

The general observation seems to be that the helicase domain and helicase associated proteins prefer to be hexamers while the prim-pol domain prefers to be a monomer. Herein lies the glaring contradiction and conflict within *Larva64* itself based on this research: a prim-pol domain that wants to be a monomer that is fused to a helicase domain that wants to be a hexamer. The results of the HHpred analysis and alignment shed light onto the work done utilizing AlphaFold 3. *Larva64* seemingly can exist as a monomer or a hexamer; however, it is unknown which domain has precedence for the structure of the protein itself or whether the structure changes depending on the activity of the protein being done.

Discussion

The interaction between *Larva64* and *Larva48* is intriguing due to its novelty. Never before in the literature has there been an interaction reported between a prim-pol and a sliding clamp protein. It is also very exciting that sliding clamps can often increase the speed of otherwise poorly performing polymerases, of which *Larva64* appears to be. However, deeming the sliding clamp protein the key to an increase in *Larva64* polymerization activity is speculative at best. First, phage Larva does not possess an endogenous clamp loader, which is needed for the sliding clamp to load onto the DNA (55). While the sliding clamp is present in the genome, and presumably, in the viral proteome, for it to be loaded onto the DNA strand outside of co-opting a clamp loader from the host is not likely. Presumably, if proteins are already being co-opted from the host, why couldn't phage Larva borrow a polymerase from *M. smegmatis*, allow *Larva64* to prime, and then handoff synthesis to the much faster bacterial polymerase? The lack of an endogenous clamp loader lends suspicion to it being *Larva64's* needed speed boost. Secondly, multiple other bacteriophages that possess a prim-pol identical to *Larva64* lack a sliding clamp protein. If phage InvictusManeo is able to survive without a sliding clamp, it would stand to reason that sliding clamps are not necessary for the speeding up of polymerization needed for *Larva64*-like prim-pols. Nevertheless, the results of the B2H show that *Larva64* does weakly interact with the *Larva48* sliding clamp. Phage Larva likely possesses the sliding clamp for some reason that is advantageous for viral survival, rather than being the result of sheer chance. Further research with CRISPRi could silence the *Larva48* gene and show whether it is necessary for the survival of phage Larva.

Using AlphaFold we were able to model the structure of *Larva64* as a monomer, a dimer, and hexamer. Out of those options, *Larva64* as a monomer and a hexamer seem to be the most likely conformations. As a hexamer, *Larva64* has a "hole" which can accommodate ssDNA and

dsDNA. Phage primers and prim-pols in hexameric confirmation have precedence, as the T7-primase-helicase and NrS-1 prim-pol exist as hexameric rings (5, 56). As a monomer, there is a disordered section of amino acids between the two domains that theoretically could allow the prim-pol and helicase domain to wrap themselves or nestle themselves within the structure of DNA. Having modeled *Larva64* along with the structure of *Larva48* dimer, we have seen potential interactions that could assist in the polymerization activity of our prim-pol. The sliding clamp with *Larva64* monomer in the presence of double-stranded DNA, appears to hold the prim-pol in a steady location so as to conduct synthesis. This potentially could assist with the polymerization activity of *Larva64* at a higher rate. A hexameric conformation may be stabilized by a *Larva48* sliding clamp, given that the clamp in our model sits at the entrance of the hexameric "hole" which would accommodate DNA. HHpred analysis of *Larva64* reveals that for the different domains, helicase and prim-pol, there is a disparity between the two domains as to what their stoichiometric structure ought to be. The helicase domain prefers to be a hexamer, while the prim-pol domain prefers to be a monomer. The fusion of the prim-pol and helicase domain in *Larva64*, may have created an internal conflict for the structure of the protein itself.

Larva64 more than likely exists as a monomer or potentially a hexamer above other orientations, however further analysis will need to be conducted to confirm either confirmation. A hexameric orientation would allow theoretically for the prim-pol to hold on firmly to the DNA strand for synthesis, but the six prim-pol domains that want to be monomers may sterically hinder or compete with one another. There's a possibility that the inherent slow speed of *Larva64* is due to the internal conflict of structure for the protein itself. If the main activity center of DNA polymerization wants to be a monomer, while half of its structure wants to be a hexamer, then this could cause conflict that causes the protein to work slowly by "falling on and off" of the DNA

strand. A way to determine if this is true, would be to truncate *Larva64* to the 326 (not including disordered region) or 388 (including disordered region) amino acid residue, so as to isolate the prim-pol domain. These could be modeled, along with other prim-pol or sliding clamp truncations as an avenue of future research. If the conflict of structure is the reason why the prim-pol is so slow, then it should theoretically have a faster rate. While the K415 mutation did theoretically deactivate the helicase domain, the remaining structures that would want to be a hexamer still exist which means that the conflict of structure (monomer vs. hexamer) would still cause a low polymerization rate. This is an intriguing avenue for future research.

CHAPTER FIVE: CONCLUSIONS AND FUTURE WORK

Qualitative characterization utilizing gel-based assays confirmed and revealed novel characteristics of *Larva64*. The activity of *Larva64* remains slow, robust, and Mg^{2+} dependent. Efforts to speed up synthesis via the addition of rNTPs, alternate divalent metal ions, and mutations to the prim-pol and helicase domain were unsuccessful, as were efforts to synthesize a double-stranded product using solely rNTPs. Optimization of assay conditions altering pH and NaCl concentrations, 8.0 and 25 mM respectively, showed increased synthesis activity in gel-based assays. *Larva64* was also shown capable of priming synthesis for faster polymerases, both viral and bacterial, and confirmed to have SD activity. A polymerization hand-off to a faster bacterial polymerase is a possible method of phage Larva's genome replication. The dsDNA observed in this research from *Larva64* were larger than the apparent 3kb stalled products reported by Folse in 2020. This is possibly due to the usage of higher quality ssM13 that was purchased from NEB rather than being made in house.

Quantitative characterization of *Larva64* was able to give a rate of synthesis utilizing a PG-based fluorescence assay. This research is the first to utilize PG to quantify a viral prim-pol polymerization activity. *Larva64:*Klenow polymerization handoff was confirmed and quantified using the PG based fluorescence assay. The polymerization rate of the optimized conditions was not significantly faster or slower than that of its original counterpart, which is curious given results from the qualitative research. B2H testing revealed an interaction between *Larva64* and *Larva48*, a sliding clamp protein. An interaction between a prim-pol and a sliding clamp protein is novel to the literature, and it is possible that the sliding clamp protein is aiding polymerization by *Larva64*. Bioinformatics analysis from Phamerator, however, casts some doubt upon that hypothesis, as many bacteriophages within the same cluster, such as InvicusManeo, have conserved prim-pols

identical to *Larva64* that do not possess a sliding clamp. AlphaFold 3 modeling as well as HHpred analysis suggest that *Larva64* either exists as a monomer or a hexamer. In either orientation, there appears to be a structural mechanism by which the *Larva48* sliding clamp could assist the prim-pol. The conflicting structural preferences for the prim-pol and helicase domain may affect the speed of the polymerase as well. The continued slow speed of *Larva64* and interactions with endogenous and non-endogenous *M. smegmatis* proteins support the conclusion that *Larva64* is likely not the sole protein responsible for phage Larva genomic replication.

This work has opened up multiple avenues of future research. Understanding the different parts of *Lava64* itself will give a better picture of how it interacts with the proteins around it, be it endogenous or in the host. While the K415A mutation looked similar to WT in our assays, a strand displacement assay with the K415A mutation would be intriguing given its comparative lack of streaking. The K415A mutation affects the helicase domain of the prim-pol, so SD synthesis could theoretically be hampered with the alteration of the helicase domain. Truncations of the WT prim-pol would also further elaborate on the role of each of the domains in *Larva64*, namely truncating the protein at the 326 or 388 residues to see how polymerization is affected without the helicase structure present. Moreover, AlphaFold 3 modeling of the truncated domains could further support the idea of conflicting structural preferences. The addition of heparin, which binds to the DNA binding site of polymerases, after an assay had begun could reveal roughly at what time point *Larva64* dissociates from the DNA strand. Knowing the rate of synthesis, as well as a time when the prim-pol falls off DNA could reveal the size of a primer that *Larva64* is generating for handoff to a bacterial host. Moreover, determining whether *Larva64* is able to hand off synthesis to the *M. smegmatis* polymerase would be worth investigating, albeit difficult as the *M. smegmatis*

polymerase has multiple constituent parts that would each have to be individually expressed and purified (57).

While PG has been an effective dsDNA fluorescent dye for our purposes, it is by no means the most ideal dye for future research. The inability to perform real time measurement of dsDNA concentration, and therefore an enzymatic rate, in a sample introduces a wide variety of potential errors. Time points can often be too late based on the response time of the researcher quenching or depending on the number of samples being tested. While quenching in time points of minutes is acceptable for slow enzymes such as *Larva64*, it becomes impractical for enzymes that complete synthesis in a matter of seconds. Utilizing *SYBR™Green* or any other fluorescent dye that can measure polymerization in real time using fluorescence needs to be pursued in future research. An increased resolution of data from a real time assay could be greatly beneficial at identifying points where SD occurs or points where the prim-pol dissociates from the DNA strand.

B2H testing could be utilized to measure interactions between *Larva64* and other *M. smegmatis* proteins, elaborating on the research of Folse in 2020. Further B2H testing of *Larva64* and other endogenous proteins that interact with DNA would be beneficial for understanding the viral genome replication machinery of Larva in a greater sense. Additional confirmation and replication of data for the interaction between *Larva64* and *Larva48* needs to be pursued. Given that the interaction between a prim-pol and a sliding clamp is novel to the literature, this observation must be strongly supported by evidence. With the addition of a clamp loader with *Larva64* and *Larva48* present in an assay, we may observe a drastic increase in the speed needed for *Larva64* to replicate the viral genome efficiently. Further bioinformatics analysis of conserved prim-pols identical to *Larva64* and the presence/absence of *Larva48* conserved sliding clamps would go a long way in determining how sliding clamps interact with other prim-pols. This could

include comparisons between the phage Larva and phage InvictusManeo, both bioinformatically and biochemically. All future research with *Larva64*, and other viral prim-pols for that matter, should have the goal of uncovering the interactions between host and viral proteins that make viral replication possible and most effective. By understanding these interactions and systems, we can better customize and tailor bacteriophage treatments for a not-so-distant antibiotic-resistant world.

CHAPTER SIX: MATERIALS AND METHODS

In Vitro Qualitative and Quantitative Characterization of *Larva64*

Larva64 protein used in characterization was either a -80/-20 °C frozen stock from 2020 or *E. coli* expressed and purified in-house using the method from Folse in 2020. Both ssM13 and dsM13 were purchased from New England Biolabs. DNA synthesis assays were performed in a 37°C water bath in 1.5 mL microcentrifuge tubes at a protein (*Larva64*) concentration of 2.5 µM. Reactions for pre-optimized assay conditions were carried out under the following conditions: 25 mM Tris pH 7.0, 50 mM NaCl, 1 mM DTT, 100 µg/mL BSA, 1 mM dNTPs, 0.5 micrograms of M13 ssDNA, and 10 mM $MgCl_2$. Reactions for optimized assay conditions were carried out under the following conditions: 25 mM Tris pH 8.0, 25 mM NaCl, 1 mM DTT, 100 µg/mL BSA, 1 mM dNTPs, 0.5 micrograms of M13 ssDNA, and 10 mM $MgCl_2$. Variable components included 0.1667 units of Klenow fragment in NEB2 buffer, 10 mM $MnCl_2$, $CaCl_2$, $NiCl_2$, $CoCl_2$, $ZnSO_4$, as indicated, and 1 mM rNTPs, as indicated. Metal salt solutions were prepared in house. D81A and K415A mutant proteins were tested at identical concentrations to WT prim-pol, 2.5 µM. Reactions were quenched with either a 2X buffer containing 100 mM EDTA, 0.4% SDS, and 10% glycerol (experiments corresponding to Figures 2, 3, 4, and 5) or simply 100 mM EDTA in equal measure to the sample (experiments corresponding to Figures 6, 7, 8, 10 and 11). D81A and K415A mutations were created using custom primers, PCR, and WT prim-pol DNA. These mutations were expressed in *E. coli* and purified in-house using the method from Folse in 2020. Qualitative samples were analyzed by running quenched samples of each timepoint overnight on a 0.8% agarose gel in a buffer-recirculating gel rig at 20V in TAE buffer. Ultraviolet imaging of the gel DNA products within, were conducted on a Biorad ChemiDoc Imager.

Quantitative samples, 20 µL quenched samples on a 96 well plate, were analyzed 10 minutes after the addition of 5 µL of 5X PG dye and measured on a SpectraMax iD5 plate reader. Plate reader settings were Ex, Em: 485, 535 nm (filter engaged), and PMT Gain: Low. A 60 ppb quinine solution (Ex, Em: 349, 461 nm) served as a fluorescence standard between days to account for differing instrument emission intensity. Fluorescence data was analyzed using Microsoft Excel, first standardizing all data based on the quinine standard, then subtracting the "0-minute" fluorescence from a day's respective data sets. The dsDNA calibration curve was generated by a serial dilution of 500 ng dsM13 and the addition of PG and EDTA identical to that of our assays. Graphing and statistical analysis were performed using Prism10 software, with 95% confidence intervals as a metric of statistical significance and error bars in our figures.

Klenow Fragment PG Rate Validation:

We obtained Klenow fragment from New England Biolabs (NEB), and per the manufacturer "One unit is defined as the amount of enzyme that will incorporate 10 nmol of dNTP into acid insoluble material in 30 minutes at 37°C." (58). Using NEB's units system, we calculate how many moles of dNTPs are incorporated per minute and converted to dNMPs. We then calculate the mass of dNMPs that are in the dsDNA strand, and therefore the mass of dsDNA will be produced per minute (ng dsDNA/minute), the unit of enzyme speed measurement for our study.

For our assays, we start with a purchased stock from NEB that is 5000 units/mL Klenow fragment, 5 units per µL. This is then diluted 1:30 in NEB2 buffer and water. 1 µL of diluted Klenow fragment is equal to 0.1667 units of Klenow fragment. In our testing, we used 5 µL of diluted Klenow fragment in our "polymerase handoff" PG assay, meaning 0.833 units of Klenow fragment. Upon incorporation into the dsDNA strand, a dNTP becomes a dNMP with the loss of

two phosphates. The moles of dNTPs incorporated are identical to the moles of dNMPs. Assuming the average dNMP is 327 g/mol, we understand that for one unit of Klenow fragment, $3.27x10^{-6}$ grams of dNMP are incorporated every 30 minutes, meaning that for every 1 minute, $1.09x10^{-7}$ grams of dNMP are incorporated. Converting to nanograms, 109 ng dNMP are incorporated by 1 unit of Klenow fragment. At our Klenow fragment concentration of 0.833 units, the enzyme should incorporate 91 ng dNMP per minute. The mass of dsDNA and incorporated dNMPs should be roughly conserved, meaning the mass of incorporated dNMPs should equal the mass of dsDNA. By this logic, one unit of Klenow fragment should produce 109 ng dsDNA per minute, and our concentration of Klenow fragment used in testing should produce 91 ng dsDNA per minute. This is the theoretical rate of synthesis for Klenow fragment using NEB's specifications and our solution preps. The slope of fluorescence for Klenow fragment using our PG assay is 882190 fluorescence units/minute. When multiplied by the inverse slope of the dsDNA calibration curve ($7.66x10^{-5}$ ng dsDNA/fluorescence units), it provides a rate of 66.74 ng dsDNA/minute, which is the experimental rate of synthesis for Klenow fragment for our PG assay.

Bacterial Two-Hybrid Analysis

The B2H protocol used in this research was provided by Dr. Maria Gainey based on the research of Dr. Danielle Heller at the HHMI SEA-PHAGES program (59). pCI plasmid fused to *Larva64* served as the bait, while p2Hα plasmids fused to *Larva48, 62, 63,* and *65* served as the prey. *Larva64:* pCI served as the bait and as it was the protein of interest, we wanted to observe PPi with. Alterations to this protocol were the variation of carbenicillin concentrations (500 µg/mL, 750 µg/mL, 1500 µg/mL). Only the 750 µg/mL plate is shown in this paper as the higher concentrations of carbenicillin yielded no growth.

Bioinformatic Analysis

Bioinformatic analysis using Phamerator was done through observation from the "Actino_Draft" Phamerator database (www.phamerator.org) on 06/13/2024 using the "Cluster" feature. Alpha Fold 3 (alphafoldserver.com) modeling predictions were done by pasting the amino acid sequences of *Larva64* and *Larva48* into the software and modifying the copy numbers to alter stoichiometry (monomer, dimer, etc.), using default settings. For models that required ssDNA, a DT40 60mer sequence was used, and for dsDNA, the reverse complement of that sequence was added to the program. HHpred analysis was done by pasting the amino acid sequence of *Larva64* into the tool and observing the matched/aligned results (Job ID: 1084060) from the PDB_mmCIF70_16_Aug database, the default database for HHpred.

REFERENCES

1. Iyer, L. M.; Koonin, E. V.; Leipe, D. D.; Aravind, L. Origin and Evolution of the Archaeo-Eukaryotic Primase Superfamily and Related Palm-Domain Proteins: Structural Insights and New Members. *Nucleic Acids Res*. **2005**, 33 (12), 3875–3896.
2. Boldinova, E. O.; Belousova, E. A.; Gagarinskaya, D. I.; Maltseva, E. A.; Khodyreva, S. N.; Lavrik, O. I.; Makarova, A. V. Strand Displacement Activity of Primpol. *International Journal of Molecular Sciences* **2020**, *21* (23), 9027. DOI:10.3390/ijms21239027.
3. Garafutdinov, R. R.; Sakhabutdinova, A. R.; Gilvanov, A. R.; Chemeris, A. V. Rolling Circle Amplification as a Universal Method for the Analysis of a Wide Range of Biological Targets. *Russian Journal of Bioorganic Chemistry* **2021**, *47* (6), 1172–1189. DOI:10.1134/s1068162021060078.
4. Boldinova, E. O.; Baranovskiy, A. G.; Esyunina, D.; Tahirov, T. H.; Makarova, A. V. DNA Synthesis across DNA Hairpins by Human Primpol. *DNA Repair* **2024**, *142*, 103741. DOI:10.1016/j.dnarep.2024.103741.
5. Chen, X.; Su, S.; Chen, Y.; Gao, Y.; Li, Y.; Shao, Z.; Zhang, Y.; Shao, Q.; Liu, H.; Li, J.; Ma, J.; Gan, J. Structural Studies Reveal a Ring-Shaped Architecture of Deep-Sea Vent Phage NRS-1 Polymerase. *Nucleic Acids Research* **2020**, *48* (6), 3343–3355. DOI:10.1093/nar/gkaa071.
6. Mislak, A. C.; Anderson, K. S. Insights into the Molecular Mechanism of Polymerization and Nucleoside Reverse Transcriptase Inhibitor Incorporation by Human Primpol. *Antimicrobial Agents and Chemotherapy* **2016**, *60* (1), 561–569. DOI:10.1128/aac.02270-15.

7. Roach, D. R.; Debarbieux, L. Phage Therapy: Awakening a Sleeping Giant. *Emerging Topics in Life Sciences* **2017**, *1* (1), 93–103. DOI:10.1042/etls20170002.
8. Hatfull, G. F. Mycobacteriophages. *Microbiology Spectrum* **2018**, *6* (5). DOI:10.1128/microbiolspec.gpp3-0026-2018.
9. Dana, R.; Gray, V. The Actinobacteriophage Database: Larva https://phagesdb.org/phages/Larva/ (accessed Oct 29, 2023).
10. Tanner, N. A.; Loparo, J. J.; Hamdan, S. M.; Jergic, S.; Dixon, N. E.; van Oijen, A. M. Real-Time Single-Molecule Observation of Rolling-Circle DNA Replication. *Nucleic Acids Research* **2009**, *37* (4). DOI:10.1093/nar/gkp006.
11. Cann, A. J. Genomes. *Principles of Molecular Virology (Genomes)* **2016**, 59–104. DOI:10.1016/b978-0-12-801946-7.00003-1.
12. Lopata; Jójárt; Surányi; Takács; Bezúr; Leveles; Bendes; Viskolcz; Vértessy; Tóth. Beyond Chelation: EDTA Tightly Binds Taq DNA Polymerase, Mutt and dUTPase and Directly Inhibits Dntpase Activity. *Biomolecules* **2019**, *9* (10), 621. DOI:10.3390/biom9100621.
13. Nathan Folse. BIOINFORMATIC AND IN VITRO CHARACTERIZATION OF PRIMASE-POLYMERASE ENZYMES FROM VIRUSES THAT INFECT ACTINOBACTERIAL HOSTS. Masters Thesis, Western Carolina University, Cullowhee, NC, April **2020**.
14. Singh, A.; Patel, S. S. Quantitative Methods to Study Helicase, DNA Polymerase, and Exonuclease Coupling during DNA Replication. *Methods in Enzymology* **2022**, 75–102. DOI:10.1016/bs.mie.2022.03.011.

15. Joyce, C. M. Techniques Used to Study the DNA Polymerase Reaction Pathway. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics* **2010**, *1804* (5), 1032–1040. DOI:10.1016/j.bbapap.2009.07.021.

16. Held, P. https://www.agilent.com/cs/library/applications/fluorometric-quantitation-dsDNA-using-PicoGreen-5994-2888EN-agilent.pdf (accessed 2023-11-01).

17. Dragan, A. I.; Casas-Finet, J. R.; Bishop, E. S.; Strouse, R. J.; Schenerman, M. A.; Geddes, C. D. Characterization of PicoGreen Interaction with Dsdna and the Origin of Its Fluorescence Enhancement upon Binding. *Biophysical Journal* **2010**, 99 (9), 3010–3019. DOI:10.1016/j.bpj.2010.09.012.

18. Sheppard, E. C.; Rogers, S.; Harmer, N. J.; Chahwan, R. A universal fluorescence-based toolkit for real-time quantification of DNA and RNA nuclease activity **2019**. DOI:10.1101/548628.

19. Driscoll, M. D.; Rentergent, J.; Hay, S. A Quantitative Fluorescence-based Steady-state Assay of Dna Polymerase. *The FEBS Journal* **2014**, 281 (8), 2042–2050. DOI:10.1111/febs.12760.

20. Ahn, S. PicoGreen Quantitation of DNA: Effective Evaluation of Samples Pre- or Post-PCR. *Nucleic Acids Research* **1996**, *24* (13), 2623–2625. DOI:10.1093/nar/24.13.2623.

21. Zhang, S. B.; Yang, S.; Vidyasagar, S.; Zhang, M.; Casey-Sawicki, K.; Liu, C.; Yin, L.; Zhang, L.; Cao, Y.; Tian, Y.; Swarts, S.; Fenton, B. M.; Keng, P.; Zhang, L.; Okunieff, P. PicoGreen Assay of Circular DNA for Radiation Biodosimetry. *Radiation Research* **2015**, *183* (2), 188–195. DOI:10.1667/rr13556.1.

22. Holden, M. J.; Haynes, R. J.; Rabb, S. A.; Satija, N.; Yang, K.; Blasic, J. R. Factors Affecting Quantification of Total DNA by UV Spectroscopy and PicoGreen Fluorescence.

Journal of Agricultural and Food Chemistry **2009**, *57* (16), 7221–7226. DOI:10.1021/jf901165h.

23. Dove, S. L.; Hochschild, A. A Bacterial Two-Hybrid System Based on Transcription Activation. *Protein-Protein Interactions* **2004**, 231–246. DOI:10.1385/1-59259-762-9:231.
24. Singh, A.; Van Slyke, A. L.; Sirenko, M.; Song, A.; Kammermeier, P. J.; Zipfel, W. R. Stoichiometric Analysis of Protein Complexes by Cell Fusion and Single Molecule Imaging. *Scientific Reports* **2020**, *10* (1). DOI:10.1038/s41598-020-71630-6.
25. Lopez de Saro, F. Regulation of Interactions with Sliding Clamps during DNA Replication and Repair. *Current Genomics* **2009**, *10* (3), 206–215. DOI:10.2174/138920209788185234.
26. Alberts B, Johnson A, Lewis J, et al. Molecular Biology of the Cell. 4th edition. New York: Garland Science; **2002**. Analyzing Protein Structure and Function. Available from: https://www.ncbi.nlm.nih.gov/books/NBK26820/
27. Abramson, J.; Adler, J.; Dunger, J.; Evans, R.; Green, T.; Pritzel, A.; Ronneberger, O.; Willmore, L.; Ballard, A. J.; Bambrick, J.; Bodenstein, S. W.; Evans, D. A.; Hung, C.-C.; O'Neill, M.; Reiman, D.; Tunyasuvunakool, K.; Wu, Z.; Žemgulytė, A.; Arvaniti, E.; Beattie, C.; Bertolli, O.; Bridgland, A.; Cherepanov, A.; Congreve, M.; Cowen-Rivers, A. I.; Cowie, A.; Figurnov, M.; Fuchs, F. B.; Gladman, H.; Jain, R.; Khan, Y. A.; Low, C. M.; Perlin, K.; Potapenko, A.; Savy, P.; Singh, S.; Stecula, A.; Thillaisundaram, A.; Tong, C.; Yakneen, S.; Zhong, E. D.; Zielinski, M.; Žídek, A.; Bapst, V.; Kohli, P.; Jaderberg, M.; Hassabis, D.; Jumper, J. M. Accurate Structure Prediction of Biomolecular Interactions with Alphafold 3. *Nature* **2024**, 630 (8016), 493–500. DOI:10.1038/s41586-024-07487-w.

28. Vashishtha, A. K.; Wang, J.; Konigsberg, W. H. Different Divalent Cations Alter the Kinetics and Fidelity of DNA Polymerases. *Journal of Biological Chemistry* **2016**, 291 (40), 20869–20875. DOI:10.1074/jbc.r116.742494.

29. Forslund, J. M.; Pfeiffer, A.; Stojkovič, G.; Wanrooij, P. H.; Wanrooij, S. The Presence of rNTPs Decreases the Speed of Mitochondrial DNA Replication. *PLOS Genetics* **2018**, 14 (3). DOI:10.1371/journal.pgen.1007315.

30. Sharma, D.; Say, A. F.; Ledford, L. L.; Hughes, A. J.; Sehorn, H. A.; Dwyer, D. S.; Sehorn, M. G. Role of the Conserved Lysine within the Walker a Motif of Human DMC1. *DNA Repair* **2013**, *12* (1), 53–62. DOI:10.1016/j.dnarep.2012.10.005.

31. Biolabs, N. E. NebufferTM 2. https://www.neb.com/en-us/products/b7002-nebuffer-2 (accessed 2024-07-21).

32. Singh, K.; Srivastava, A.; Patel, S. S.; Modak, M. J. Participation of the Fingers Subdomain of *Escherichia coli* DNA Polymerase I in the Strand Displacement Synthesis of DNA. *Journal of Biological Chemistry* **2007**, 282 (14), 10594–10604. DOI:10.1074/jbc.m611242200.

33. Chowdhury, K.; Tabor, S.; Richardson, C. C. A Unique Loop in the DNA-Binding Crevice of Bacteriophage T7 DNA Polymerase Influences Primer Utilization. *Proceedings of the National Academy of Sciences* **2000**, 97 (23), 12469–12474. DOI:10.1073/pnas.230448397.

34. Goel, A.; Frank-Kamenetskii, M. D.; Ellenberger, T.; Herschbach, D. Tuning DNA "Strings": Modulating the Rate of DNA Replication with Mechanical Tension. *Proceedings of the National Academy of Sciences* **2001**, *98* (15), 8485–8489. DOI:10.1073/pnas.151261198.

35. Maier, B.; Bensimon, D.; Croquette, V. Replication by a Single DNA Polymerase of a Stretched Single-Stranded DNA. *Proceedings of the National Academy of Sciences* **2000**, *97* (22), 12002–12007. DOI:10.1073/pnas.97.22.12002.

36. Blotta, I.; Prestinaci, F.; Mirante, S.; Cantafora, A. Quantitative Assay of Total DsDNA with PicoGreen Reagent and Real-Time Fluorescent Detection. *Annali dell'Istituto Superiore di Sanità* **2005**, *41* (1), 119–123.

37. Ponchel, F.; Toomes, C.; Bransfield, K.; Leong, F. T.; Douglas, S. H.; Field, S. L.; Bell, S. M.; Combaret, V.; Puisieux, A.; Mighell, A. J.; Robinson, P. A.; Inglehearn, C. F.; Isaacs, J. D.; Markham, A. F. Real-Time PCR Based on SYBR-Green I Fluorescence: An Alternative to the TaqMan Assay for a Relative Quantification of Gene Rearrangements, Gene Amplifications and Micro Gene Deletions. *BMC Biotechnology* **2003**, 3 (1), 18. DOI:10.1186/1472-6750-3-18.

38. Wu, P.; Zehnder, J.; Schröder, N.; Blümmel, P. E.; Salmon, L.; Damberger, Fred. F.; Lipps, G.; Allain, F. H.-T.; Wiegand, T. Initial Primer Synthesis of a DNA Primase Monitored by Real-Time NMR Spectroscopy. *Journal of the American Chemical Society* **2024**, *146* (14), 9583–9596. DOI:10.1021/jacs.3c11836.

39. Zhao, G.; Guan, Y. Polymerization Behavior of Klenow Fragment and & Taq DNA Polymerase in Short Primer Extension Reactions. *Acta Biochimica et Biophysica Sinica* **2010**, *42* (10), 722–728. DOI:10.1093/abbs/gmq082.

40. Mulye, M.; Singh, M. I.; Jain, V. From Processivity to Genome Maintenance: The Many Roles of Sliding Clamps. *Genes* **2022**, 13 (11), 2058. DOI:10.3390/genes13112058.

41. Makarova, K. S.; Koonin, E. V. Annotation and Classification of CRISPR-Cas Systems. *Methods in Molecular Biology* **2015**, 47–75. DOI:10.1007/978-1-4939-2687-9_4.

42. Zabrady, K.; Zabrady, M.; Kolesar, P.; Li, A. W.; Doherty, A. J. CRISPR-Associated Primase-Polymerases Are Implicated in Prokaryotic CRISPR-Cas Adaptation. *Nature Communications* **2021**, 12 (1). DOI:10.1038/s41467-021-23535-9.

43. Zhang, Y.; Skolnick, J. Scoring Function for Automated Assessment of Protein Structure Template Quality. *Proteins: Structure, Function, and Bioinformatics* **2004**, *57* (4), 702–710. DOI:10.1002/prot.20264.

*44. Xu, J.; Zhang, Y. How Significant Is a Protein Structure Similarity with Tm-Score = 0.5? Bioinformatics **2010**, 26 (7), 889–895. DOI:10.1093/bioinformatics/btq066.*

45. *Crystal structure of a putative DNA polymerase III beta subunit (eubrec_0002; ERE_29750) from eubacterium rectale ATCC 33656 at 2.26 a resolution* **2011**. DOI:10.2210/pdb3t0p/pdb.

46. Soding, J.; Biegert, A.; Lupas, A. N. The HHPRED Interactive Server for Protein Homology Detection and Structure Prediction. *Nucleic Acids Research* **2005**, *33* (Web Server). DOI:10.1093/nar/gki408.

47. Vannini, A.; Beck, K.; Lipps, G.; Cramer, P. *Crystal structure of the primase-polymerase from Sulfolobus Islandicus* **2010**. DOI:10.2210/pdb3m1m/pdb.

48. Lipps, G.; Weinzierl, A. O.; von Scheven, G.; Buchen, C.; Cramer, P. *Bifunctional DNA primase/polymerase domain of orf904 from the archaeal plasmid Prn1- triple mutant F50M/L107M/L110M Manganese Soak* **2004**. DOI:10.2210/pdb1ro2/pdb.

49. Guo, H. J.; Li, M. J.; Wang, T. L.; Wu, H.; Zhou, H.; Xu, C. Y.; Liu, X. P.; Yu, F.; He, J. H. *Structure of the bifunctional DNA primase-polymerase from phage NRS-1* **2019**. DOI:10.2210/pdb6a9w/pdb.

50. Qiao, C. C.; Mir-Sanchis, I. *Structure of primase-helicase in sapi5* **2022**. DOI:10.2210/pdb7om0/pdb.

51. Xu, H.; Sträter, N.; Schröder, W.; Böttcher, C.; Ludwig, K.; Saenger, W. Structure of DNA Helicase Repa in Complex with Sulfate at 1.95 Å Resolution Implicates Structural Changes to an `open' Form. *Acta Crystallographica Section D Biological Crystallography* **2003**, *59* (5), 815–822. DOI:10.1107/s0907444903004025.

52. Xing, X.; Bell, C. E. Crystal Structures of *Escherichia coli* RecA in a Compressed Helical Filament. *Journal of Molecular Biology* **2004**, *342* (5), 1471–1485. DOI:10.1016/j.jmb.2004.07.091.

53. Bagautdinov, B.; Kunishima, N. *Crystal structure of RecA superfamily ATPase PH0284 from Pyrococcus horikoshii OT3* **2006**. DOI:10.2210/pdb2dr3/pdb.

54. Pitsawong, W.; Pádua, R. A.; Grant, T.; Hoemberger, M.; Otten, R.; Bradshaw, N.; Grigorieff, N.; Kern, D. From Primordial Clocks to Circadian Oscillators. *Nature* **2023**, *616* (7955), 183–189. DOI:10.1038/s41586-023-05836-9.

55. Hedglin, M.; Kumar, R.; Benkovic, S. J. Replication Clamps and Clamp Loaders. *Cold Spring Harbor Perspectives in Biology* **2013**, *5* (4), a010165–a010165. https://doi.org/10.1101/cshperspect.a010165.

56. Kato, M.; Frick, D. N.; Lee, J.; Tabor, S.; Richardson, C. C.; Ellenberger, T. A Complex of the Bacteriophage T7 Primase-Helicase and DNA Polymerase Directs Primer Utilization. *Journal of Biological Chemistry* **2001**, *276* (24), 21809–21820. DOI:10.1074/jbc.m101470200.

57. Ditse, Z.; Lamers, M. H.; Warner, D. F. DNA Replication in Mycobacterium Tuberculosis. Microbiology Spectrum 2017, 5 (2). DOI:10.1128/microbiolspec.tbtb2-0027-2016.

58. Biolabs, N. E. DNA Polymerase I, Large (Klenow) Fragment. https://www.neb.com/en-us/products/m0210-dna-polymerase-i-large-klenow-fragment (accessed 2024-07-21).
59. Protocol 4.1: 2-Hybrid Selection Assay. https://seagenes.helpdocsonline.com/4-1 (accessed 2024-11-04).