UNDERSTANDING THE RNA-PROTEIN CONTACTS REQUIRED FOR SUCCESSFUL INFECTION OF THE SSRNA BACTERIOPHAGE Qβ

An Undergraduate Research Scholars Thesis

by

HAANIA KAKWAN

Submitted to the Undergraduate Research Scholars program at
Texas A&M University
in partial fulfillment of the requirements for the designation as an

UNDERGRADUATE RESEARCH SCHOLAR

Approved by Research Advisor: Dr. Junjie Zhang

May 2020

Major: Biomedical Sciences

TABLE OF CONTENTS

ABSTRACT

Understanding the RNA-Protein Contacts Required for Successful Infection of the ssRNA Bacteriophage Qβ

Haania Kakwan
Department of Biomedical Sciences
Texas A&M University

Research Advisor: Dr. Junjie Zhang
Department of Biochemistry and Biophysics
Texas A&M University

The single-stranded (ss) RNA bacteriophage Qβ, which infects piliated bacteria, most commonly *Escherichia coli*, requires several RNA-protein interactions for successful infection. Based on modeling the RNA secondary structure of the genome into the electron density of cryo-EM structures, interactions thought to be critical for infection were seen between RNA stem-loops at the 3' end of the genome (U1 and R1 stem loops) and the maturation protein. The purpose of this study was to validate these proposed interactions. Mutations were introduced into the R1 and U1 stem loops that change the quality of the interactions with the maturation protein, and the resulting virus-like particles' infectivity was observed.

The importance of RNA-protein interactions will be further explored by swapping the maturation protein of the closely related bacteriophage MS2 with that of Qβ. If genome-protein interactions similar to those seen in infectious Qβ are conserved, we predict that infectious virus-like particles will be produced.

The goal of this project is to understand the RNA-protein interactions required by Qβ to establish infection in order to exploit these to use Qβ as a means of RNA delivery into bacterial

cells. The ability to deliver engineered RNA molecules into bacterial cells has important therapeutic and industrial applications. RNA encoding bactericidal proteins or short interfering RNA (siRNA) fragments that interfere with critical gene expression can be used to suppress bacterial growth, which holds promise as a method of treating bacterial infections. Bacterial cell machinery can also be harnessed to produce proteins of interest by delivering RNA that encodes for them.

ACKNOWLEDGMENTS

I would like to thank my advisor Dr. Zhang and my mentors Karl Gorzelnik and Jirapat "Tee" Thongchol for their guidance and support throughout the course of this research.

I would also like to thank my friends and family for their support and encouragement throughout the project.

Lastly, I would like to thank the LAUNCH Undergraduate Research Scholars department and staff for giving me this opportunity and the department of Biochemistry and Biophysics and Texas A&M University for supporting my undergraduate research.

CHAPTER I

INTRODUCTION

Background

Qβ is a single-stranded (ss) positive-sense RNA bacteriophage that belongs to the *Leviviridae* family.[1] ssRNA phages infect Gram-negative bacteria by attaching to their F-pili, which cells subsequently retract and unwittingly pull the viral RNA in.[2] The genome of Qβ is 4,217-nucleotides (nt) long and encodes for the maturation protein (A_2), the coat protein, the read-through coat protein (A_1), and the viral subunit of the RNA-dependent RNA polymerase (replicase).[3] The near-icosahedral capsid, displays *T*=3 quasi-symmetry, where 178 copies of the coat protein have their symmetry broken by the maturation protein (Figure 1). The maturation protein protrudes through the capsid, interacting with the gRNA.

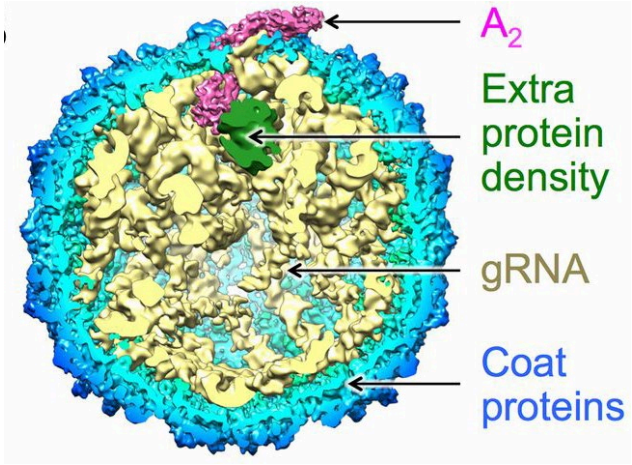

Figure 1. Cryo-EM modeled structure of Qβ. Colored pink is the MP. The β-region protruding out of the capsid is involved in adsorbing to the F-pilus during host-cell recognition and initiating cell lysis by inhibiting MurA. The α-region inside the capsid contacts the gRNA (yellow) and ICP (green). The gRNA adopts characteristic secondary and tertiary structures. Interactions between gRNA operator-like hairpins and coat proteins (blue) are distributed asymmetrically, occurring via positively charged residues at the tips of CP EF loops. These interactions have a preference for A/B dimers over C/C. Figure reproduced with permission from *PNAS* (Cui *et al.*, 2017).

However, the virion is not perfectly symmetrical, which has important implications for the mechanism of genome delivery into its host. It has been shown that the asymmetry results from interactions of the coat proteins with A_2 and the genomic RNA (gRNA).[4] Since Qβ must assemble its capsid around its gRNA, coat protein-gRNA interactions likely play a role in capsid assembly.[5] The characteristic secondary and tertiary structures the gRNA exhibits inside the capsid lends credence to this theory. In addition, the maturation protein has the dual functions of binding to the F-pilus as well as initiating cell lysis.[6]

The coat protein can adopt three conformations, A, B, and C, which form two types of dimers: A/B and C/C.[4] The gene encoding for coat protein has a "leaky" stop codon, which results in a "read-through" protein A_1 or the minor coat protein.[4] A_1 replaces about 10 copies of the major coat protein monomer in virions, consisting of an N-terminal coat protein domain and a read through domain connected by a flexible linker.[4]

The MP, A_2 takes the place of one C/C dimer in the capsid and also carries out the lysis function for Qβ by inhibiting MurA, the enzyme that catalyzes the first committed step of cell wall synthesis.[7] A_2 contains an α-region inside the capsid that contacts gRNA and a β-region above the capsid, with each domain containing high levels of α-helices and β-strands, respectively.[1] Moreover, the latest cryo-EM images also reveal an extra density at the bottom of A_2 inside the virion that fits a coat protein dimer and interacts with gRNA in a similar fashion.[7]

Expanding the case for functionalities conferred by gRNA-capsid interactions, we hypothesized that the mechanism of infection and assembly rely on interactions between specific stem loops of the gRNA and A_2 and the internal coat protein dimer (ICP) as well as those asymmetrically distributed between the gRNA and CPs. Based on our modeling of the genome into the electron density of our cryo-EM structures, interactions thought to be critical for

infection were seen between RNA stem-loops U1 and A_2 and R1 and the ICP (Figure 2). These stem loops are at the 3' end of the genome (Figure 3).

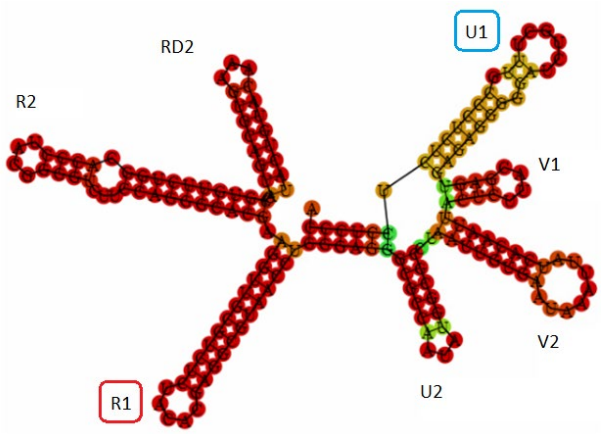

Figure 2. Predicted RNA secondary structure of 3' end. Several stem loops or "hairpin" secondary structures exist at the 3' end of the genome, the names of which are shown. U1 (circled in blue as indicated in Figure 3) interacts with the B-region of the MP. R1 (circled in red) interacts CP.

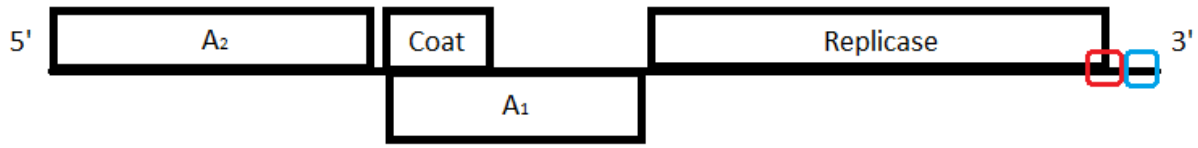

Figure 3. Schematic genome of Qβ. The genome is 4,217 nucleotides long. Indicated in blue is the site of the MP-binding loop U1. Indicated in red is the site of the ICP-binding loop R1.

Goals

The goal of this study is to further pinpoint the structural components of the Qβ genome required for producing infectious phage and to exploit these components to produce a mutant phage that retains infectivity. These goals will lead to a better understanding of the mechanisms Qβ uses to infect and the structural components that are necessary for it to do so successfully so that these processes may be taken advantage of to deliver engineered RNA molecules.

CHAPTER II

MATERIALS AND METHODS

Five different mutants of Qβ (Figures 4 and 5) were created with site-directed mutagenic PCR to determine the role stem loops U1 and R1 play in infection. Three knockout mutants (Figure 6) were also created. A ninth mutant in which all copies of the coat protein would be read and synthesized as read-through coat proteins was also created by removing the leaky stop codon at the end of the coat protein gene (Figure 6). The DNA primers used to synthesize the latter mutants are listed in Table 1. pBRT7QB plasmids (Figure 7) were used as cloning vectors.

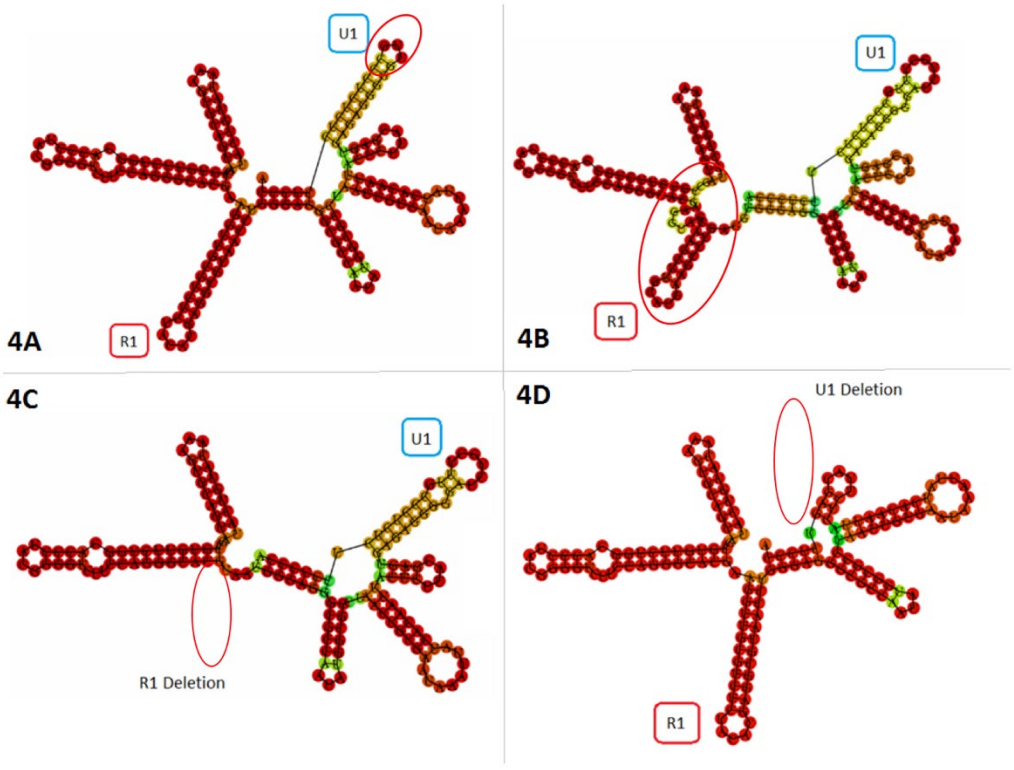

Figure 4. R1 and U1 mutants that do not plaque. Four different mutants of Qβ were created by site-directed mutagenesis. Changes to the wild-type secondary structures are circled in red. **A. U1 loop change.** Seven nucleotides were deleted from U1. **B. R1 scrambled. C. R1 deletion.** Replicase truncation. **D. U1 Deletion.**

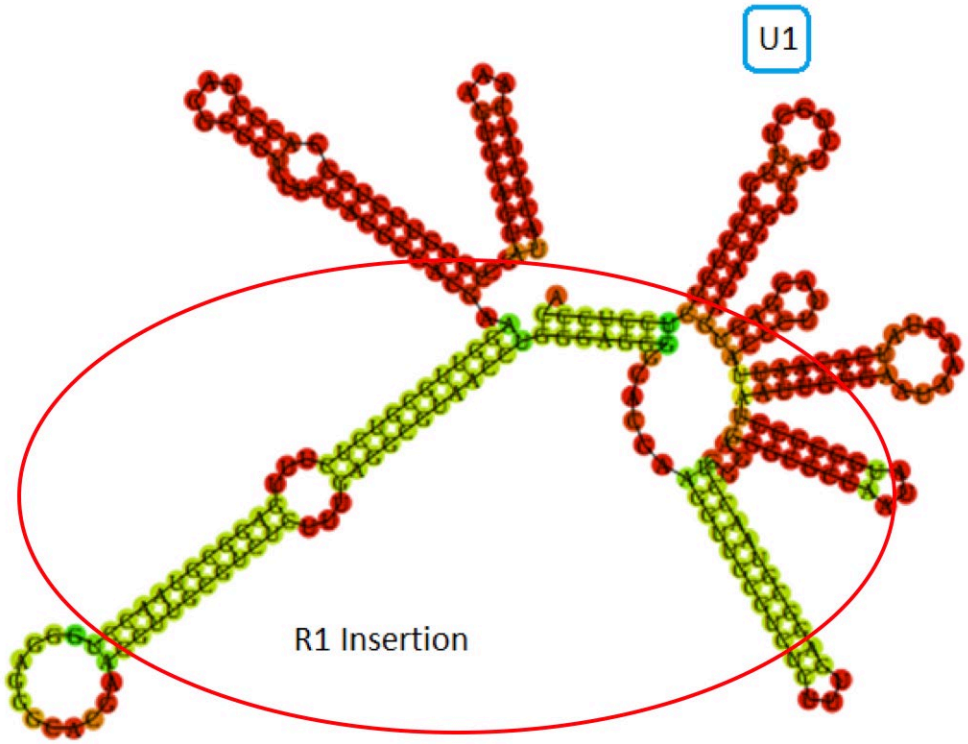

Figure 5. R1 insertion: mutant that had no effect on plaque formation. 77 nucleotides were inserted into R1. Changes to the wild-type secondary structure are circled in red.

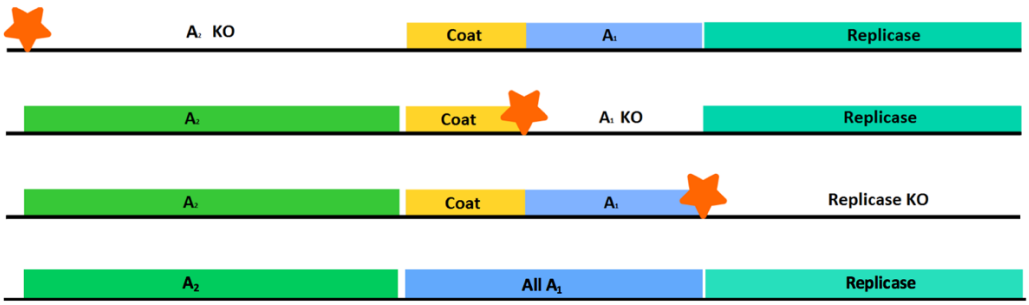

Figure 6. Knockout mutants. Gene knockouts were introduced into Qβ cDNA in pBRT7QB using site-directed mutagenesis. pBRT7QB has a T7 promoter, but the plasmid transformed into cells without a T7 RNA polymerase to reduce expression from the plasmid to basal levels. Once there is a single transcript made from host RNA polymerases, the replicase can be translated and subsequently hijack host proteins to initiate replication. In a non-expression strain there is a titer of ~10^{11} plaque forming units/ml.

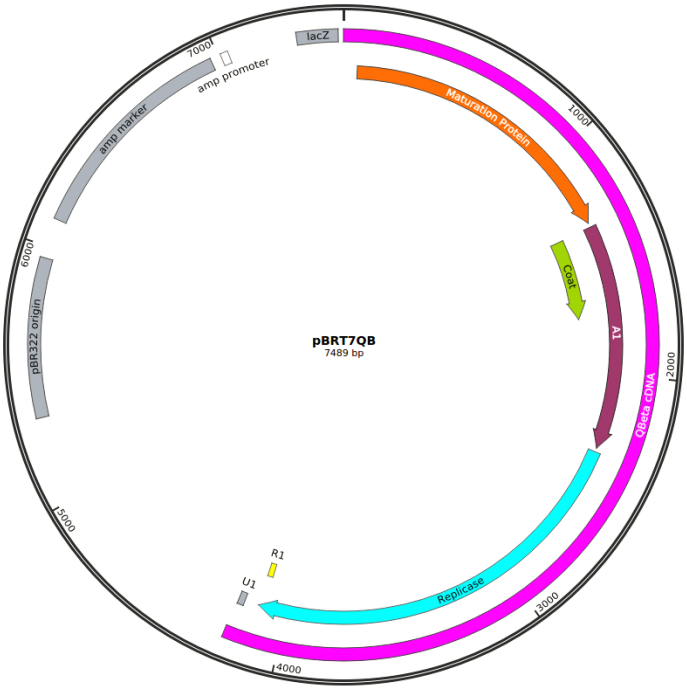

Figure 7. pBRT7QB vector. Obtained from Sean Leonard, University of Texas, Austin. Contains a T7 promoter at the start of the Qβ cDNA, an ampicillin resistance marker, and a pBR322 origin of replication.

Table 1. DNA primers used to synthesize Knockout and All A_1 mutants. These primers were ordered from Integrated DNA Technologies and used at a concentration of 10 µM in the PCRs.

Mutant	Forward Primer Sequence
A_2 KO	CTGAGTATAAGAGGACATTTTCCTAAATTACCGCGTGG
A_1 KO	GATCAGCTGAACCCAGCGTATTAAACACTGCTCATTGCCGGTG
Replicase KO	GTAACTAAGGATGAAATGCCCGTCTAAGACAGCATCTTCGC
All A_1	CCAGCGTATTGGACACTGCTCATTGC

The 50 µL PCRs included 1 µL of template, 1 µL of dNTPs, 2 µL of DMSO, 1 µL of Phusion DNA polymerase enzyme, 10 µL of 10X Phusion GC Buffer, 2 µL of the respective 10 µM Forward and Reverse primers, primer, and 31 µL of water. The PCR was set up as shown in Table 2.

Table 2. PCR parameters.

Initial Denaturation	Denaturation	Annealing	Elongation	Final Elongation
30 seconds, 98°C	30 seconds, 98°C	30 seconds, 60°C	7 minutes, 72°C	10 minutes, 72°C
	18 cycles			

After running the PCR, 10 µL of each reaction were run in 1% agarose gels in TAE buffer at 130V for 20 minutes and visualized with ethidium bromide. After confirming PCR results, the remaining 40 µL PCR were digested with 1 µL Dpn1 enzyme for 1 hour at 37°C to degrade methylated DNA. The samples were then purified according to the Qiagen PCR Purification Kit, after which 2 µL of each sample were transformed into 50 µL of the DH10B strain of competent *E. coli* cells using electroporation. The cells were allowed to recover in 1 mL of LB at 37°C before being plated on LB agar with ampicillin to select for successfully transformed cells. The plated cultures were incubated at 37°C for 24 hours, after which isolated colonies (several for each mutant) were inoculated into 5 mL of liquid LB with ampicillin and grown again at 37°C for 24 hours. These cultures were mini-prepped according to the QIAprep Spin Miniprep Kit and then sent for sequencing. Those samples that contained the desired mutations, i.e., whose sequences matched the sequences of the primers used to synthesize them at the respective positions in the genome, were re-transformed into DH10B *E. coli*, recovered, plated, selected for, inoculated, and cultured as described previously. 5 mL liquid cultures were spun down using a centrifuge, cell pellets were discarded, and dilutions of the supernatant were made up to the 10^{-11} dilution. These dilutions were spot-titered on wild-type *E. coli* lawns plated on LB agar with

tetracycline. Assembly and presence of virus-like particles in the supernatants and expression of A_2 were confirmed separately with CsCl preparations and Western blot analysis, respectively.

CHAPTER III

RESULTS

Eight of the nine mutations, U1 loop change, R1 scrambled, R1 deletion, U1 deletion, A_2 KO, A_1 KO, Replicase KO, and All A_1, did not produce plaques, indicating that the mutants could not infect the host *E. coli* cells. One mutation, the R1 insertion, resulted in no difference in plaque formation when compared to wild-type Qβ, indicating that that mutant remained infectious (Figure 8).

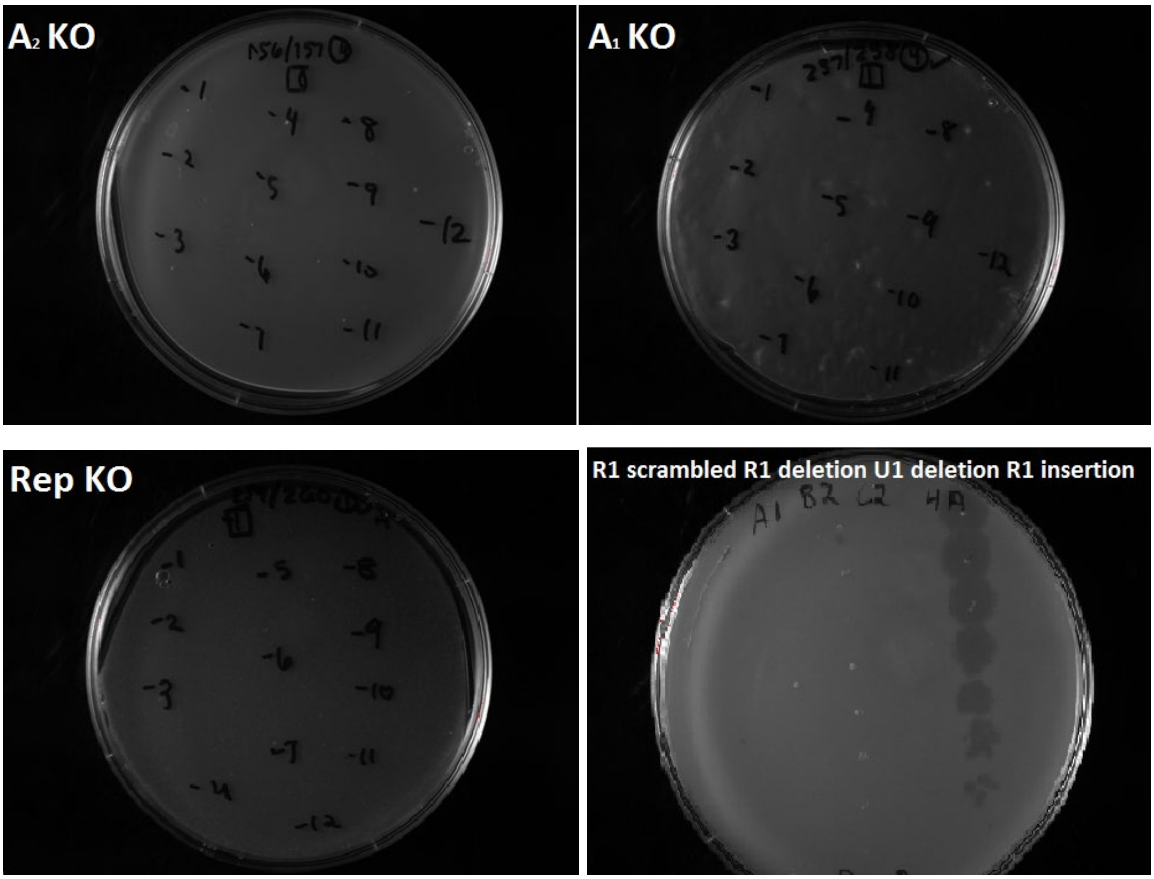

Figure 8. Spot titers of A_2 KO, A_1 KO, Replicase KO, R1 scrambled, R1 deletion, U1 deletion, and R1 insertion. The first six mutants shown did not produce any plaques at even the highest concentrations. The last mutant produced as many plaques as the wild-type virus. (U1 loop change and All A_1 not shown).

CHAPTER IV

CONCLUSION

Conclusions

Titer results indicate that both U1 and R1 stem loops are gRNA secondary structures necessary for successful assembly of an infectious particle. The interaction with the MP and ICP may mediate the transfer of the genome into the host cell or proper assembly of the virion. The knockouts of various genes within Qβ validate previous studies showing that all genes are essential for infection. Why A_1, which is only present in 3-10 copies per virion, is important for assembly/infection is unknown and will be further investigated. On the other extreme, it was also shown that infectious particles cannot be assembled if all copies of the coat protein are read-throughs.

These results lend support to the theory that viral genomic structure and genome-capsid interactions are functionally significant. Qβ, an exceptionally small virus with only four proteins, must capitalize on these genomic functionalities to be able to complete the various functions in its life cycle.

Future Directions

Understanding the complex mechanisms of viral replication present in Qβ could provide a model for gene delivery, with applications ranging from gene therapy to antibiotic design. I will use the gene knockouts I made to determine if the replicase can replicate foreign RNAs placed inside the genome. I will also determine which regions of A_1 are necessary for infection through a truncation analysis. I believe the presence of copies of A_1 in the capsid have important effects on its symmetry, but the region of the genome that encodes for it may have certain, more

defined areas that are critical due to interactions with the capsid or their effects on the genome's secondary structure as a whole.

REFERENCES

1. Morais MC. Breaking the symmetry of a viral capsid. *Proceedings of the National Academy of Sciences of the United States of America* 113: 11390–11392, 2016.

2. Rumnieks J. Protein-RNA Interactions in the Single-Stranded RNA Bacteriophages. *Subcell Biochem* 88:281-303, 2018.

3. Kashiwagi A, Yomo T. Ongoing phenotypic and genomic changes in experimental coevolution of RNA bacteriophage Qβ and Escherichia coli. PLoS Genet. 2011;7(8):e1002188.

4. Gorzelnik KV, Cui Z, Reed CA, Janaka J, Young R, Zhang J. Asymmetric cryo-EM structure of the canonical Allolevivirus Qβ reveals a single maturation protein and the genomic ssRNA in situ. *Proceedings of the National Academy of Sciences of the United States of America* 113: 11519–11524, 2016.

5. Klovins, J., Berzins, V., & Duin, J. V. A long-range interaction in Qβ RNA that bridges the thousand nucleotides between the M-site and the 3′ end is required for replication. *RNA,4*(8): 948-957, 1998.

6. Rumnieks J, Tars K. Bacteriophage Qbeta maturation protein. J Mol Biol. November 2017. doi:10.2210/pdb5mnt/pdb.

7. Cui Z, Gorzelnik KV, Chang J-Y, Langlais C, Janaka J, Young R, Zhang J. Structures of Qβ virions, virus-like particles, and the Qβ–MurA complex reveal internal coat proteins and the mechanism of host lysis. *Proceedings of the National Academy of Sciences of the United States of America* 114: 11697–11702, 2017.