7. Background of Research

Many people are familiar with the double-helical structure of double-stranded nucleic acids originally proposed by Watson and Crick in the 1950s, such as that of the DNA genomes possessed by all cellular life on the planet. However, there are also many instances of nucleic acids which naturally occur as single-stranded molecules. The RNA molecules in our cells, as well as many of the RNA genomes possessed by many viruses occur as single-stranded molecules. Without a complementary strand to pair with, these RNA molecules fold in on themselves to create a range of unique structures comprised of Watson-Crick base pairs in addition to many non-Watson-Crick interactions. This structural diversity of single-stranded molecules also allows them to have a much greater functional diversity in contrast to their double-stranded counterparts. There are examples of RNA structural domains which can function as ribozymes, riboswitches, aptamers, etc.

Given the importance of RNA structure in regulating function, there is great interest amongst RNA researchers in determining the structure of these complex molecules. However, with currently available technology, the only RNA molecule larger than 1 kilobase with an existing 3D structure is that of ribosomal RNA in complex with its constituent ribosomal proteins, leading to the 2009 Nobel prize in chemistry. Of RNA structures available in the protein data bank (pdb.org), 94% of them are less than 0.25 kilobase. This demonstrates that techniques such as X-ray crystallography, NMR, or cryo-EM are inadequate for the determination of the entirety of many RNA molecules which occur in nature. Because of this, determination of the structure of long RNA molecules is generally relegated to predictive methods using a combination of chemical mapping and computational methods to arrive at the most probable structure. To overcome the limitations of current technology, I have endeavored to develop Atomic Force Microscopy (AFM) imaging as a method to reveal details about the structure of long RNA molecules.

Identification of structural domains in viral RNA molecules is of great interest to virologist as further investigation of these structures allows for additional details of viral lifecycles to be uncovered, and can then lead to the development of antiviral strategies targeting these structures or the pathways they are involved in.

8. Research methodology

I developed procedures for screening for structural domains in RNA molecules using AFM. These procedures included a method of sample preparation which allowed for images of RNA molecules containing secondary structure to be obtained in addition to the development of automated image analysis algorithms using MATLAB in order to analyze the images and extract structural information about the RNA molecules. This information includes an approximation of the region of the RNA molecules where the structural domain is located, an approximation of the size of the structural domains, and a general idea of the branching patterns/conformation of the domains. With this information, structural models of the domains were then generated by inputting the predicted range of nucleotides contained in each domain into the RNAstructure web server and screening for structures that best corresponded to the branching patterns observed in the AFM images. To test the structural models, minimal constructs of the domains were constructed and then imaged with AFM to confirm that they matched the predicted conformation for the predicted nucleotide range. Finally, to determine the functional role of the structural domains, mutations were introduced into the RNA. The ability of the mutations to disrupt the structure was assessed using AFM imaging, and then virus particles containing the mutations were reversed engineered. The effect of the mutations on viral growth kinetics and on the ability to generate spliced mRNA transcripts (RT-qPCR) is ongoing.

9. Results/impacts

Note: As much as possible, describe the contents and results of your research in a manner that is easily understandable to a non-specialist in your field. Provide a concrete description if (1) papers related to your work have been published in major academic journals, (2) particularly outstanding research results were achieved, or (3) patent applications have been made or other tangible outcomes achieved through the research.

My work has focused on development of a new method for structural characterization of RNA molecules using Atomic Force Microscopy imaging combined with automated data analysis algorithms developed in MATLAB. Since ribosomal RNA (rRNA) are the only RNA molecules larger than a kilobase with existing 3D structures (see 2009 Nobel Prize in Chemistry), proof-of-concept for this method was obtained by comparing AFM analysis of the human 28S rRNA to the existing structural models (Gilmore et al., *Nuc Acids Res*, 2017) (Fig. 1). With this analysis, we were able to show that the naked RNA molecules in the AFM images retained much of the secondary structure which is present in cry-EM images of intact ribosomes. In addition, many conformational variations were observed in these molecules that may warrant further investigation in the future.

Fig. 1 – Proof of concept with rRNA. The cryo-EM structure of the rRNA taken from the ribovision website is shown on the left and images of human 28S ribosomal RNA molecules are shown on the left. Colored arrows point to structures which correspond to the colors of the highlighted domains in the cryo-EM structure.

With proof-of-concept for this method established, we then turned to investigating and modeling the structure of other RNA molecules. Viral RNA genomes, such as those of the Hepatitis C Virus or the Influenza Virus are an excellent application for this method given their clinical significance (Fig. 2).

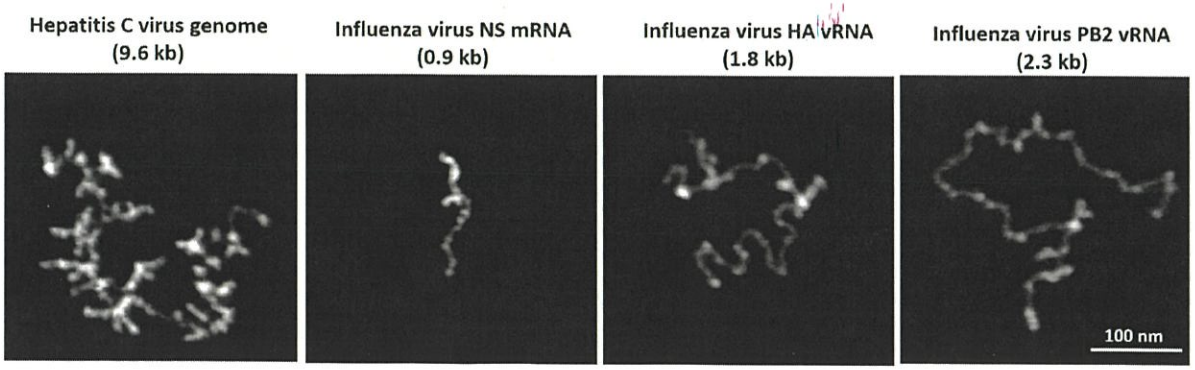

Fig. 2 – AFM of viral RNA molecules.

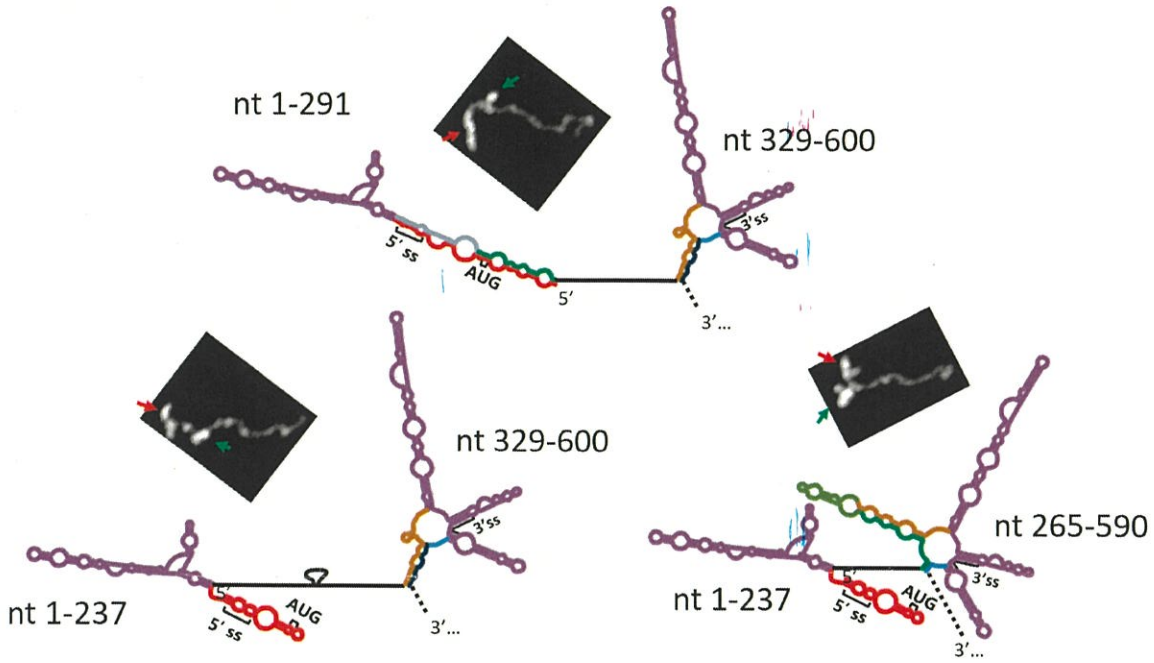

Fig. 3 – Structural models of various coexisting conformations of the Influenza Virus NS mRNA.

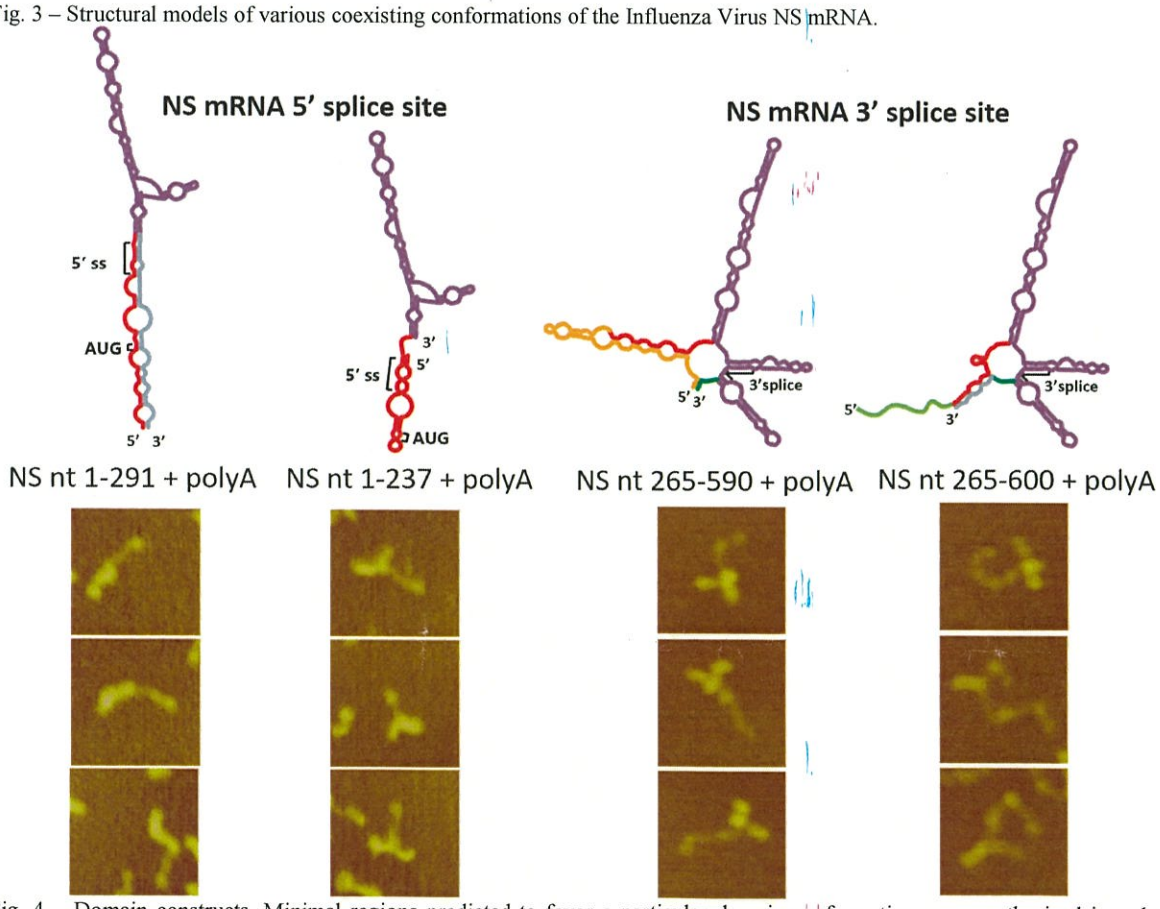

Fig. 4 – Domain constructs. Minimal regions predicted to favor a particular domain conformation were synthesized in order to validate the models displayed in Fig. 3. The molecules were polyadenylated for ease of visualization.

Of the RNA molecules which were observed with AFM, one of the particularly striking ones was the 0.9 kb influenza NS mRNA, which reproducibly contained two prominent structural domains localized around the 5' and 3' splice sites. Both structural domains had both branched and unbranched conformations. Application of the MATLAB-based automated algorithms to these molecules allowed for the prediction of nucleotides contained in each domain conformation, and an approximation of the general morphology and length of the branches within each domain conformation. We used this data to select models predicted by the RNAstructure web server which best matched the parameters defined by AFM analysis (Fig. 3).

To confirm the models, minimal constructs of each structural domain were synthesized by T7 transcription, and AFM was used to confirm that they matched the expected morphology for each conformation (Fig. 4). Finally, we wanted to investigate the function of he structural domains. Based on their localization in the mRNA molecule, we hypothesized that the structural domains are involved in regulation of alternative splicing of this mRNA molecule. To investigate this, mutations were introduced into the region of the structural domains to disrupt the stability of the various domain conformations. AFM imaging of the mutant RNA molecules suggested that these mutations affected the equilibrium between the conformations localized at their splice sites. To assess how these mutations affect the viral lifecycle, the mutations were then reverse engineered into the virus particles. The mutant virus was then used to infect MDCK cells, RNA was then isolated at various time points over the course of 10 following infection. Then, both semi-quantitative and quantitative RT-PCR was used to assess if the mutations have an effect on the amount of splicing. By the end of the fellowship, I was able to lay the groundwork for these experiments, but I have not been able to repeat the experiments enough times to get conclusive results. Interestingly, in my attempt to generate PCR standards, I clearly identified four bands corresponding to NS mRNA suggested that there may actually be at least three splice isoforms of the influenza NS segment. Although the ability of the NS1 segment to be alternatively spliced to form the NS2 isoform has been well established, there has only been one prior report of the NS3 segment, and no reports of a possible 4th isoform. Although additional experiments are needed to confirm that this is actually another NS segment and not an artifact of the PCR reaction, it is possible that there may be a fourth previously unidentified protein encoded by the influenza virus. I hope that the Noda lab will continue my project to understand the process of splicing in the influenza virus and the role of RNA structure in this process.

10. Research Presentations during the period of the fellowship (Name of the conference, title, place, date)

- 6th Negative Stand Virus Japan meeting, Analysis of structural elements in influenza virus mRNA using Atomic Force Microscopy, Okinawa, 2017
- Bermuda Principles Impact on Splicing Conference, Nanoimaging of secondary structural domains at the splice sites of influenza virus mRNA, Bermuda, 2017
- 62nd Annual Biophysical Society meeting, A nanoimaging approach for identification of the secondary structural domains in long ssRNA molecules, San Francisco, CA, 2018
- 23rd Annual RNA society meeting, Nanoimaging of structural domains in long single-stranded RNA molecules, University of California Berkeley, 2018

11. A list of paper published during or after the period of the fellowship, and the names of the journals in which they appeared (Please fill in the format below). Attach a copy of each article if available.

Author(s)	Title	Name of Journal	Volume	Page	Date	Note
Jamie L. Gilmore, Katashi Deguchi, Kunio Takeyasu	Nanoimaging of RNA molecules with Atomic Force Microscopy	Microscopy and imaging science: practical approaches to applied research and education	7	300-306	Feb 2017	
Jamie L. Gilmore, Aiko Yoshida, James A. Hejna, Kunio Takeyasu	Visualization of conformational variability in the domains of long single-stranded RNA molecules	Nucleic Acids Research	45	8493-8507	Aug 21, 2017	

12. Awards during the period of the fellowship (Name of the award, Institution, date etc.)
RNA society best poster award 2018