Jim L. Zhang - Graduate Research Plan Statement

Intellectual Merit – Introduction: All non-enveloped viruses contain an outer protein coat, or capsid, to encapsulate and protect the viral genome prior to infection. These capsid structures adopt a variety of forms ranging from helical filaments and icosahedrons to complex head-tail structures as seen in bacteriophages. Previously, it was thought all double-stranded RNA viruses formed icosahedral capsids. Such an arrangement efficiently organizes the multi-segmented genome and facilitates *in situ* transcription by positioning viral polymerases directly adjacent to each dsRNA molecule. The resulting single-stranded mRNA is then extruded from the viral complex for translation and viral replication while the template dsRNA molecule remains occluded – effectively evading host immunity and antiviral defenses.[1,2]

The discovery of *Colletotrichum camellia*e filamentous virus 1 (CcFV-1), a dsRNA virus with filamentous capsid morphology, presents an interesting exception to this paradigm. First identified in 2017, CcFV-1 encapsulates eight defined segments of dsRNA predicted to encode for ten proteins, which includes an RNA-dependent RNA polymerase (P1) and capsid protein (P4). Sequence analysis of P1 suggests CcFV-1 represents an evolutionary intermediate between single- and double-stranded RNA viruses. Purified virion samples imaged with electron microscopy (EM) revealed the presence of semi-rigid filaments approximately 15 nm wide with varying lengths up to 4,000 nm – the longest of any known virus.[3] The resulting particle width, however, raises several key questions concerning the structure and replication mechanism of CcFV-1. Namely, how are all eight dsRNA molecules consistently organized and encapsidated within each virion to ensure continued replication? In addition, how is the double-helix of each RNA molecule sterically accommodated within a narrow filament without energetically unfavorable distortions? Finally, how is the viral genome accessed for transcription and viral replication without complete filament disassembly and exposure of dsRNA to the host cytosol?

Given the size and filamentous symmetry of the viral assembly, single-particle cryoEM presents a promising means to obtain a high-resolution structure of the CcFV-1 capsid. Structural studies of the capsid would not only provide insight into potential mechanisms for dsRNA organization and binding but also shed light on the broad evolutionary relationships between single- and double-stranded RNA viruses. To that end, I propose the following aims:

Intellectual Merit – Research Plan:

AIM 1: Expression and purification of filamentous CcFV-1 virus-like particles. In lieu of infectious virions, which must be generated from the *Colletotrichum camellia*e host, recombinant CcFV-1 virus-like particles (VLPs) from overexpressed capsid protein represent a more accessible target for structural analysis. In my preliminary work, I have generated two polyhistidine-tagged P4 constructs, one at each protein termini (Figure 1A), for cloning and expression in *Escherichia coli*. Initial screenings indicated both tagged constructs are soluble and behave well *in vitro.* Visualization of crude cell lysate with transmission EM (TEM) revealed the expected filamentous particles (Figure 1B). However, nickel-based affinity purification was found to only isolate capsid protein oligomers upon analysis with size-exclusion chromatography, suggesting the termini of the protein are ultimately buried and thus unresponsive to affinity-based purification in the assembled VLP (Figure 1C).

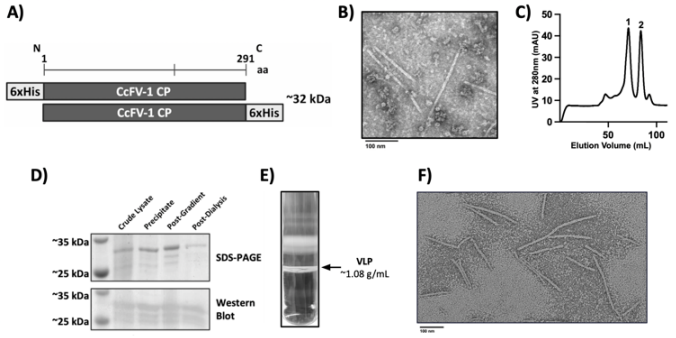

Figure 1. Design, expression, and purification of CcFV-1 P4. (**A**) Diagram outlining construct design. (**B**) TEM micrograph of crude cell lysate. (**C**) Size-exclusion profile of affinity-purified P4. (**D**) SDS-PAGE and Western blot of VLP purification fractions. (**E**) Density gradient following ultracentrifugation. (**F**) TEM micrograph of VLP fractions from (E).

Nonetheless, viral capsid assemblies are of significantly greater size and density compared to most cellular protein complexes. I hypothesized this physical feature can be exploited to isolate VLPs without conventional affinity-mediated protein purification. Thus, protein precipitation and density gradient ultracentrifugation presented a promising alternative for VLP purification. Following serial treatments of

salt and pelleting, it was determined that CcFV-1 VLPs precipitate at a concentration of 10% ammonium sulfate (w/v) from crude cell lysate (Figure 1D). The resulting sample was subjected to gradient ultracentrifugation with the VLP banding at ~1.08 g/mL, a lower density than the infectious virion due to the absence of dsRNA (Figure 1E). Subsequent TEM imaging revealed the expected filamentous particles (Figure 1F). However, both sample purity and VLP stability remain a concern for downstream structural analysis. Future work will improve sample purity via additional self-generated gradients and attempt to stabilize the isolated VLP with co-expression of other CcFV-1 proteins. Self-generated gradients enable finer separation by density, while bioinformatic analyses indicates CcFV-1 P5 contains a viral tegument-like domain, suggesting it may play a non-structural accessory role in filamentous particle assembly.

AIM 2: Single-particle cryoEM analysis and structural determination of CcFV-1 VLPs. Single-particle analysis presents the most promising means of obtaining a high-resolution structure of the CcFV-1 VLP. Here, purified samples of protein are plated on holey carbon-based grids and flash-frozen in a layer of vitreous ice. The resulting grid is then imaged under cryogenic conditions, obtaining micrographs depicting 2D CcFV-1 particles. These particles are then classified, averaged, and subject to reconstruction via computational processing, resulting in a 3D electron density map. Due to CcFV-1's heterogeneous particle length, accommodations must be made during 2D particle classification where the helical repeats within each filament will be analyzed instead of the particle's entirety. Thus, a single VLP can be expected to contribute multiple particles for classification and averaging.[4]

Obtaining a high-resolution structure of <3 Å will require specialized 300 kV cryo-TEM equipment for imaging and data collection, in addition to powerful computing services. These services are available across the graduate programs I am applying to, in addition to nearby institutions such as the University of Texas Health Science Center. Thus, I will establish a collaboration between my current lab and cryoEM facilities to support the structural analysis of CcFV-1.

AIM 3: Biochemical characterization of the CcFV-1 polymerase-capsid (P1-P4) complex. All icosahedral dsRNA viruses package copies of viral RNA-dependent RNA polymerase (RdRP) within the capsid structure to facilitate *in situ* transcription. Determining how CcFV-1 RdRP (P1) is incorporated into the P4 helical capsid complex, if at all, is crucial to characterizing mechanisms for viral transcription and replication. To that end, I will first conduct a pull-down assay to assess for protein-protein interactions between CcFV-1 P1 and P4. Previous work has already led to the cloning and expression of Strep-tagged P1 within *E. coli*. Repurposing the size-exclusion purified P4 oligomers from Aim 1, I will incubate the two proteins and immobilize P1 on Strep-binding resin. If P4 co-eludes with P1, that would suggest the CcFV-1 polymerase and capsid engage in the expected interactions for a dsRNA virus. If I do not observe any binding between the two lone proteins, I hypothesize that motifs within the viral dsRNA may function as a mediator for interaction.[3] Thus, I will repeat the pull-down assay in the presence of CcFV-1 dsRNA segments, obtained either from collaborators or commercial *de novo* synthesis. Following confirmation of protein-protein interactions, I will co-express P1 and P4 to assess for P4 incorporation in fully assembled, filamentous CcFV-1 VLPs. These complexes can be similarly purified and analyzed as described in Aims 1 and 2. A detailed structure of the P1-P4 VLP complex would provide key clues towards mechanisms of dsRNA occlusion, access, and transcription within CcFV-1.

If the two proteins do not interact, P1 remains a key target for analysis. RNA binding assays can still be used to identify viral genomic sequences required for initiation and packaging, while preliminary work suggests lone P1 remains a complex of suitable size for single-particle cryoEM analysis.

Broader Impacts and Significance: Research on CcFV-1 has the potential to reveal new paradigms for dsRNA virus assembly and replication, expanding the field of virology. In addition, many dsRNA viruses such as the *Reoviridae* family represent significant pathogens to human health. Given the ubiquitous distribution of viruses, it is highly unlikely CcFV-1 represents the only case of a filamentous, dsRNA virus. It is thus critical to obtain the fundamental knowledge associated with their maintenance and replication as novel members and their potential pathogenicity are discovered.

References: (1) Borodavka, A., Desselberger, U. & Patton, J. T. *Current Opinion in Virology* **33**, 106–112 (2018). (2) Vabret, N. & Blander, J. M. *Front. Immunol.* **4**, (2013). (3) Jia, H. *et al. Nat Commun* **8**, 168 (2017). (4) Cheng, Y., Grigorieff, N., Penczek, P. A. & Walz, T. *Cell* **161**, 438–449 (2015).