Title
The physical and chemical parameters of in vitro packaging of heterologous RNA into virus-like particles

Author
Tanimoto, Cheylene Rachelle

Publication Date
2022

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Los Angeles

The physical and chemical parameters of *in vitro* packaging of heterologous RNA into virus-like particles

A dissertation submitted in partial satisfaction of the

requirements for the degree of Doctor of Philosophy

in Chemistry

by

Cheylene Rachelle Tanimoto

2022

ABSTRACT OF THE DISSERTATION

The physical and chemical parameters of *in vitro* packaging of heterologous RNA into virus-like particles

by

Cheylene Rachelle Tanimoto

Doctor of Philosophy in Chemistry

University of California, Los Angeles, 2022

Professor William Gelbart, Chair

The evolved strategies of viruses are wondrous and should be investigated and understood better to exploit their gene delivery "tricks" for the development of effective therapeutics. Simple RNA viruses – those consisting of a positive-sense/ready-to-translate RNA genome inside an ordered capsid made up of multiple copies of a single protein – have been used as platforms for delivery of therapeutic cargo to cells. For example, early work by several groups established that spontaneous self-assembly of virus-like particles (VLPs) can be carried out *in vitro* from purified components of RNA and the capsid protein (CP) from the plant viruses tobacco mosaic virus (TMV), cowpea chlorotic mottle virus (CCMV) and brome mosaic virus (BMV). CCMV and BMV CP can encapsulate non-viral RNAs into monodisperse, RNase-resistant spherical capsids if the length of the RNA is between 2,500 and 4,500 nucleotides (nts); longer RNAs are packaged into "multiplets" – multiple 28-nm VLPs, where RNA is shared

between two or more particles. Additionally, researchers have exploited the self-replicating nature of viruses, specifically nodamura virus (NOV), to amplify expression of genes of interest (GOIs). The GOI can vary widely from reporter genes to RNA encoding for cancer or viral antigens. This self-replicating RNA, or replicon, can be packaged into CCMV VLPs using an *in vitro* assembly method. In this study, I investigate viral strategies such as self-assembly and self-replication with the goal of delivering self-replicating functional RNAs encapsulated in spherical VLPs to mammalian cells as therapeutics.

This thesis work reports on a systematic series of experiments elucidating the physical chemical properties of self-assembled VLPs as delivery platforms for mRNA replicon therapeutics. I am interested in understanding the physical chemical parameters governing successful self-assembly – solution ionic strength, pH, protein concentration, charge and size of cargo being packaged – to develop robust therapeutics. To probe the role of these in determining RNA-protein and protein-protein interactions and ensuing assembly of nucleocapsids, interferometric scattering (iSCAT) microscopy was used to follow the assembly kinetics of individual BMV VLPs around BMV RNA1 using a range of solution ionic strengths and protein concentrations. The iSCAT experiments revealed that nucleation times become longer and more broadly distributed when RNA-protein interactions are reduced (high ionic strength) and/or protein concentrations are decreased, corresponding to an increase in nucleation barriers. To further explore the role of RNA-protein and protein-protein interactions, assemblies were carried out in a variety of assembly buffers spanning ranges of pH from 5 to 7 and of ionic strengths from 0.1 to 0.5 M, and assembly products were imaged using negative-stain electron microscopy (EM). Additionally, the salt- and pH-dependence of protein-protein interactions were characterized by differential scanning fluorimetry. To quantify the dependence of GOI protein

expression on the numbers of GOI replicons (derived from NOV RNA1) transfected into mammalian cells, cell lines were transfected with increasing masses of replicon RNA mixed with an inert carrier RNA, with the total mass of the transfected RNA held constant while increasing the fraction of replicon. Fluorescence microscopy, flow cytometry, and trypan-blue exclusion assays show a non-monotonic dose-dependence of protein expression – i.e., "less is better" – in cells transfected with replicon RNA. To package long replicon molecules – those with GOI > 1,000 nts – into RNase-resistant VLPs, the RNA needs to be compacted to fit into a BMV VLP. To probe the role of charge and size of cargo on self-assembly, long RNA molecules were pre-compacted with increasing concentrations of polyvalent cations and assembled into BMV VLPs. VLPs were analyzed by negative-stain EM and gel electrophoresis. Finally, the same replicon molecule was packaged into spherical and rod-like VLPs using the CP of CCMV and TMV, respectively, with the goal of comparing how the shape of these particles impacts their efficacy as a therapeutic delivery vehicle.

The dissertation of Cheylene Rachelle Tanimoto is approved.

Jose A. Rodriguez

Giovanni Zocchi

Shimon Weiss

William M. Gelbart, Committee Chair

University of California, Los Angeles

2022

Dedications

To my friends and family.

Table of Contents

List of Figures and Tables

Acknowledgments

I would like to acknowledge Professor Emeritus Charles M. Knobler who has served as my co-advisor and mentor but is not formally listed as a member of my committee.

Chapter 2 is a version of Garmann, R. F., Goldfain, A. M., Tanimoto, C. R., Beren, C. E., Vasquez, F. F., Knobler, C. M., Gelbart, W. M., Manoharan, V. N. (2022). Single-particle studies of the effects of RNA-protein interactions on the self-assembly of RNA virus particles. *PNAS. Under review. Preprint.* doi: 10.1101/2022.05.09.588235.

Chapter 3 is a version of Tanimoto, C. R., Beren, C. E., Garmann, R. F., Knobler, C. M., Gelbart, W. M. One-step *in vitro* self-assembly of virus-like particles as a function of pH and ionic strength. *In preparation.*

Chapter 4 is a version of Tanimoto, C. R., Thurm, A. R., Brandt, D. S., Knobler, C. M., Gelbart, W. M. (2022). The non-monotonic dose dependence of protein expression in cells transfected with self-amplifying RNA. *J. Virol.* 96 (7). doi: 10.1128/jvi.01858-21.

Chapter 5 references dynamic light scattering data (data not shown) and virus particle assemblies that will be featured in a manuscript *in preparation* by Duran Meza, A. L., Tanimoto, C. T., Chapman, A. G., Sportsman, R., Knobler, C. M., Gelbart, W. M. titled "ssRNA undergoes continuous single-molecule compaction in the presence of polyvalent cations."

Chapter 6 is a version of Tanimoto, C. T., Duran Meza, A. L., Chapman, A. G., Karan, S., Chan, S. K., Knobler, C. M., Gelbart, W. M., Steinmetz, N. F. Comparison of spherical and rod-like virus-like particles for the delivery of self-replicating mRNA therapeutics. *In preparation.*

Biographical Sketch

EDUCATION

Institution and Location	Degree	Completion Date	Field
Stanford University Stanford, CA	B.S.	06/2016	Chemistry Conc: Biochemistry
University of Los Angeles, CA Los Angeles, CA	M.S.	03/2018	Chemistry Conc: Biophysics
University of Los Angeles, CA Los Angeles, CA	Ph.D.	06/2022 (expected graduation)	Chemistry Conc: Biophysics

POSITIONS

Undergraduate Research Assistant for Professor Sharon Pitteri, 2014-2016

Canary Center at Stanford, Stanford University, CA

Graduate Student Researcher for Professor William Gelbart, 2016-2022

UCLA Department of Chemistry & Biochemistry, Los Angeles, CA

Teaching Assistant for UCLA Department of Chemistry & Biochemistry, 2016-2021

Chemical Energetics and Change (20B), Physical Chemistry: Chemical Thermodynamics (110A), Physical Chemistry: Introduction to Statistical Mechanics and Kinetics (110B), Biochemical Methods I (153L)

Applications of Nanoscience Summer Institute Fellow, 2018-2021

California Nanosystems Institute (CNSI) funded two-week summer course (CHEM 8) for high school students to learn about being a scientist and conduct nanoscience experiments

Joint Research Safety Initiative Vice President, 2019-2022

Organization of graduate students and post-doctoral fellows in the Chemistry department promoting a positive safety culture

HONORS

Michael E. Jung Excellence in Teaching Award, 05/2020

UCLA Chemistry & Biochemistry Excellence in Research Fellowship, 09/2020

UCLA Distinguished Teaching Assistant Award, 04/2021

Collegium of University Teaching Fellowship, 04/2021

Ralph and Charlene Baur Award, 05/2022

Dafni Amirsakis Memorial Dissertation Award, 05/2022

PUBLICATIONS

Garmann, R. F., Goldfain, A. M., Tanimoto, C. R., Beren, C. E., Vasquez, F. F., Knobler, C. M., Gelbart, W. M., Manoharan, V. N. (2022). Single-particle studies of the effects of ionic strength and protein concentration on the self-assembly of RNA virus particles. PNAS. Under review.

Tanimoto, C. R., Thurm, A. R., Brandt, D. S., Knobler, C. M., Gelbart, W. M. (2022). The non-monotonic dose dependence of protein expression in cells transfected with self-amplifying RNA. *J. Virol.* 96 (7). doi: 10.1128/jvi.01858-21.

CHAPTER 1

Introduction to self-replicating mRNA molecules and *in vitro* packaging of RNA by viral capsid protein

Viruses are nanoscale intracellular obligate parasites composed of a genome (DNA or RNA) surrounded by a protective protein shell, called a capsid. In many (especially, mammalian viruses) cases, a lipid bilayer membrane provides the genome with protection in addition to – or sometimes instead of – a protein capsid. A virus enters a cell, and the virus disassembles to allow the viral genome to be replicated. New virus particles assemble in the complex cellular milieu from freshly replicated genome and freshly translated capsid proteins. These new particles, of which millions of copies are made from a single starting particle, go on to infect more cells. Every aspect of this complex process has evolved over time to result in successful entry into cells, replication in cells, and assembly of protective particles specifically carrying the viral genome. Investigation of these strategies could enable the design and development of effective therapeutics. This work investigates the viral strategies of self-assembly and replication with the goal of delivering self-replicating therapeutic messenger RNA (mRNA) encapsulated in *in vitro* assembled virus-like particles (VLPs) to mammalian cells.

1.1 *In vitro* self-assembly of virus-like particles (VLPs)

Spherical VLPs

Significant to this investigation is brome mosaic virus (BMV), a plant virus from the genus bromoviridae. BMV has a multipartite genome made up of single-stranded RNA (ssRNA) molecules: RNA1 (3,234 nucleotides [nts]), RNA2 (2,865 nts), RNA3 (2,117 nts), and sub-genomic RNA4 (876 nts)[1,2]. These RNA molecules are packaged into three separate virions with

about 3,000 nts of RNA in a capsid shell composed of 180 copies of a single capsid protein (CP) with RNA1 and RNA2 in their own particles and RNA3 and RNA4 co-packaged into a single particle. Most notably, this virus – and a closely-related virus, cowpea chlorotic mottle virus (CCMV) – can spontaneously self-assemble *in vitro* from purified components[3–5]. Not only can BMV and CCMV CP encapsulate their own genomes *in vitro*, but they have also been shown to encapsulate non-viral RNA and other negatively charged cargo into virus-like particles[6–9]. These resulting particles are virus-like because they look like infectious virus particles, but they do not contain a genome that would make the particles infectious.

In a well-studied *in vitro* self-assembly protocol for BMV and CCMV VLPs, RNA and CP are mixed in a neutral pH and high ionic strength solution[10–12]. At this high ionic strength, the electrostatic interactions between the negatively charged phosphate backbone of the RNA and the positively charged arginine-rich motif (ARM) in the N-terminal tail of the CP are screened. As the mixture is dialyzed against a lower ionic strength solution, these interactions are turned on and the RNA is saturated by CP binding resulting in the formation of nucleoprotein complexes (NPCs)[10]. Then the mixture is dialyzed against a lower pH solution (pH 4.75), close to the isoelectric point of the CPs, to allow for lateral attractions between CPs leading to the formation of monodisperse, perfectly symmetric, RNase-resistant VLPs[6]. While this "two-step" assembly method (i. reducing ionic strength; ii. reducing pH) is used to make robust VLPs, BMV CP has also demonstrated assembly competence in several "one-step" solution conditions[5,13].

In the present study BMV VLP self-assembly is examined in a variety of solution conditions (ionic strength, pH, and CP concentration) to examine the role of RNA-CP and CP-CP interaction strength in successful VLP formation. In Chapter 2 (Single-particle studies of the effects of RNA-protein interactions on the self-assembly of RNA virus particles), interferometric

scattering (iSCAT) is used to follow the kinetics of individual BMV VLPs assembling in a pH 6 solution and a range of ionic strengths and protein concentrations. Chapter 3 discusses experiments using electron microscopy (EM) to visualize VLPs formed in one-step-assembly buffers of varying pH and ionic strength.

Rod-like VLPs

The other important *in vitro* assembling virus for this investigation is tobacco mosaic virus (TMV). Like BMV and CCMV, TMV is a plant virus with a ssRNA genome that can be assembled in a test tube from purified components. These three – BMV, CCMV, and TMV – are the only viruses that have been shown to assemble *in vitro* from purified components. TMV virions, however, take the shape of hollow, helical rods in which the RNA is embedded in the protein shell as opposed to being in the hollow interior[14–16]. The TMV virions are helical rods with a length of 300 nm, an inner radius of 2 nm, and an outer radius of 9 nm. The RNA inside also demonstrates helical symmetry and is embedded 2 nm into the protein shell[15,16]. Each turn of the helix is composed of about 16.3 CP, where 1 CP is interacting with 3 nts of RNA. The RNA needs to contain a specific "packaging signal" – a sequence with high affinity for CP necessary for packaging. For TMV, this sequence is termed the origin of assembly site (OAS) – a stem-loop structure located about 900 nts from the 3'-end of the wild-type genome – and has been identified as the site of nucleation for assembling virus particles[17,18]. In Chapter 6, a self-replicating mRNA molecule with the TMV OAS is packaged into both spherical and rod-like VLPs, using BMV and TMV CP, respectively, and the particles are visualized using electron microscopy.

1.2 Packaging self-replicating mRNA therapeutics into VLPs

Replicons are self-replicating RNA molecules – an RNA that encodes a protein that will replicate this RNA molecule. We derive our replicons from the genome of Nodamura virus (NOV). NOV is a ssRNA virus that infects insects. Its genome is composed of two molecules – RNA1 encodes an RNA-dependent RNA polymerase (RdRp) and RNA2 encodes a capsid protein[19]. RNA1 is a replicon because it encodes the RdRp that will replicate both RNA1 and RNA2. A gene of interest (GOI) can be added to the end of the open reading frame encoding the RdRp producing a replicon that will generate many copies of the RNA and therefore the GOI. The GOI can be reporter genes like fluorescent proteins (enhanced yellow fluorescent protein, EYFP) or luciferase, or therapeutic molecules like cancer antigens or microRNA.

The present study focuses on replicons encoding reporter genes and testing their replication in mammalian cells and packaging efficiency in spherical and rod-like VLPs. In Chapter 4, NOV-EYFP replicons are transfected in mammalian cells, with accompanying demonstration that delivery of too many replicon molecules can result in strong amplification of RNA and protein that results in cell death. The following chapters discuss packaging of RNA into spherical and rod-like VLPs. In Chapter 5, long RNA molecules are pre-compacted with polyvalent cation in an attempt to package the RNA in single 28-nm spherical particles, while in Chapter 6 NOV replicons are cloned to contain a TMV OAS such that the RNA can be packaged in rod-like TMV VLPs. Chapter 7 discusses future directions for the testing of these particles in mammalian cell culture and mice with our collaborators at the University of California, San Diego.

CHAPTER 2

Single-particle studies of the effects of ionic strength and protein concentration on the self-assembly of RNA virus particles

2.1 Abstract

Understanding the pathways by which simple RNA viruses self-assemble from their coat proteins and RNA is of practical and fundamental interest. Although RNA-protein interactions are thought to play a critical role in the assembly, our understanding of their effects is limited because the assembly process is difficult to observe directly. We address this problem by using interferometric scattering microscopy, a sensitive optical technique with high dynamic range, to follow the *in vitro* assembly kinetics of each of over 500 individual particles of brome mosaic virus (BMV) – for which RNA-protein interactions can be controlled by varying the ionic strength of the buffer. We find that when RNA-protein interactions are weak, BMV assembles by a nucleation-and-growth pathway in which a small cluster of RNA-bound proteins must exceed a critical size before additional proteins can bind. As the strength of the RNA-protein interactions increases, the nucleation time becomes shorter and more narrowly distributed until the assembly kinetics become indistinguishable from diffusion-limited adsorption. In contrast, the time to grow a capsid after nucleation varies weakly with both salt and protein concentration. These results show that the nucleation rate is controlled by RNA-protein interactions, while the growth process is driven less by RNA-protein interactions and more by protein-protein interactions and intra-protein forces. The nucleated pathway observed with the plant virus BMV is strikingly similar to that previously observed with bacteriophage MS2, a phylogenetically distinct virus

with a different host kingdom. These results raise the possibility that nucleated assembly pathways may be common to other RNA viruses.

2.2 Introduction

Since the 1950s, the question of how RNA viruses self-assemble has inspired theoretical and experimental work in many fields of basic and applied science. Simple RNA viruses, which consist of a single-stranded RNA genome inside an ordered capsid made up of multiple copies of a single protein (**Figure 1A**), have served as model systems for studying the physical principles of structural virology involving virus particles of all shapes and sizes[20,21]. But the mechanisms and pathways by which these viruses assemble into the correct structure, while avoiding the many possible malformed structures, are not yet understood.

Many different RNA viruses self-assemble nucleocapsids (most are non-infectious particles, with the exception of a rare few), though the most well-studied come from two families: *Bromoviridae*, a family of plant-infecting viruses that includes brome mosaic virus (BMV) and cowpea chlorotic mottle virus (CCMV); and *Fiersviridae* (previously *Leviviridae*), a family of bacteria-infecting viruses that includes MS2 and Qβ. These families are as distinct phylogenetically as any two RNA virus families can be, having a last common ancestor that is thought to predate the emergence of eukaryotic cells[22]. Accordingly, there are many well-established physical and biological differences among viruses in these families. Yet the four most studied members—BMV, CCMV, MS2, and Qβ—do have some structural commonalities: they are icosahedral viruses with a triangulation number (*T*) – the number of structural units per face of the icosahedron – of 3, they have no lipid envelope, and each capsid surrounds about 3000-4000 nucleotides of single-stranded RNA.

The assembly of such structures is a nontrivial process. Identical coat proteins must adopt non-equivalent positions to make a *T*=3 capsid, with some arranging in pentagonal configurations and others in hexagonal configurations[21,23,24]. Furthermore, these configurations must form in the correct proportions and positions for the capsid to close. Despite these challenges, assembly of virus-like particles of CCMV[5,25,26], BMV[5,26], and MS2[27] occurs in high yield even *in vitro* and in the absence of host-cell factors. The challenge of viruses to avoid the many possible metastable states en route to complete assembly has been likened to the Levinthal paradox of protein folding[28,29].

But unlike proteins, RNA viruses have a template for assembly: their own RNA. Current theoretical models of RNA virus self-assembly posit markedly different roles for the RNA, depending on the relative strengths of RNA-protein and protein-protein interactions, sequence-dependent RNA-protein interactions, RNA-mediated protein-protein interactions, and several other factors[30]. Although specific interactions between RNA substructures and coat proteins have been hypothesized to help the virus avoid malformed configurations[29], viruses from different families differ greatly in their RNA structures and RNA-protein interactions. It is therefore unclear whether there are common features of the assembly process for different *T*=3 viruses or if there are distinct assembly pathways that depend on RNA-protein interactions.

Recent measurements of assembly kinetics suggest the latter: that assembly of viruses from different families follows different pathways. Fluorescence correlation spectroscopy (FCS) experiments[31,32] of the kinetics of binding of MS2 coat protein and RNA indicate that assembly starts with a small cluster of RNA-bound proteins that trigger a change in the hydrodynamic radius of the RNA. In contrast, cryo-electron microscopy[10] and small-angle X- ray scattering (SAXS)[33] experiments of the assembly of CCMV coat protein and RNA show that disordered

RNA-protein complexes formed at neutral pH anneal over several thousand seconds into well-formed capsids when the pH drops below 6.

But because these experiments involve different assembly conditions and different measurement techniques, their outcomes might not reflect fundamental differences in the assembly pathways of these viruses, but rather technical differences in the methods and protocols used to study them.

Furthermore, most of the techniques that have been used do not measure the assembly process directly at the scale of individual particles because—one way or the other—they involve averaging over many particles. Such averaging can obscure the mechanisms and pathways that underpin stochastic assembly processes like viral assembly, in which each individual particle can follow its own unique sequence of intermediate states. Thus, it remains an open question whether a common assembly pathway might exist between these viruses.

Garmann, et al. recently demonstrated that interferometric scattering (iSCAT) microscopy[34] can resolve the assembly kinetics of individual virus-like particles[35], providing a method to directly measure and compare the assembly pathways of different viruses. To perform the iSCAT experiment, viral RNA molecules are tethered to the surface of a functionalized glass coverslip under the desired buffer conditions[36] (**Figure 1B**). Next, iSCAT images of the RNA-decorated coverslip are collected as viral coat proteins are injected at the desired concentration and in the appropriate buffer. As the proteins bind to the surface-tethered RNA, dark spots appear in the iSCAT images (**Figure 1C**). Subtracting the intensity associated with the RNA then yields images in which the intensity of each dark spot is proportional to the number of proteins that have accrued onto each individual RNA. Accordingly, plotting the trace of the intensity of a spot as a function of time reveals the assembly kinetics for that particle, and

plotting the collection of traces reveals the assembly kinetics for the ensemble of particles (**Figure 1D**).

Previous work[35] examined the assembly of bacteriophage MS2 and found that the assembly kinetics were consistent with a nucleation-and-growth pathway in which a small cluster of RNA-bound proteins must exceed a critical size before the binding of additional proteins becomes favorable. Despite an apparently small critical nucleus size (a few coat-protein dimers), we found that MS2 capsids grow monotonically to full or nearly full size with high yield.

Although this previous study highlighted the importance of the RNA in the assembly process, the strong and specific RNA-protein interactions in MS2 make it difficult to systematically address the central question of how the RNA affects the pathway. By contrast, the RNA in BMV interacts with the coat proteins through non-specific electrostatic interactions[37]. As a result, the strength of RNA-protein interactions can be tuned by changing the ionic strength of the buffer solution[10–12]. BMV therefore offers not only an interesting comparison to MS2—it is phylogenetically distinct but structurally similar—but also the means to understand the role of RNA-protein interactions.

In this study, we infer the assembly pathways of BMV from iSCAT measurements under different RNA-protein interaction strengths, allowing us to critically assess competing models of the assembly process. We follow the assembly trajectories of more than 500 individual virus particles under different assembly conditions, and we correlate the results with the absence and presence of ordered nucleocapsids as detected with electron microscopy. We find that BMV can assemble by a nucleation-and-growth process that is qualitatively similar to that of MS2. We show that the strength of RNA-protein interactions strongly affects the nucleation time, but only weakly affects the growth time, suggesting that RNA plays a central role in nucleating the viral

capsid, but a relatively minor role in its growth. We discuss these observations in the context of recent models and hypotheses of RNA virus self-assembly.

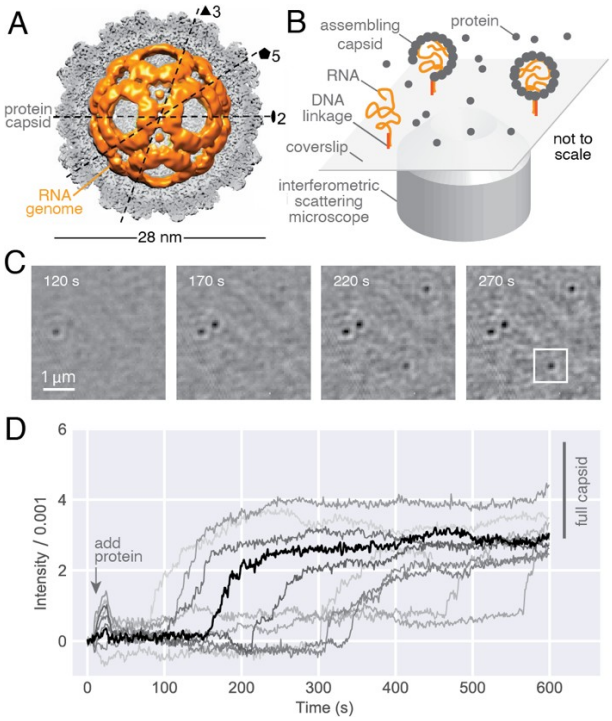

Figure 1. Overview of the system and the measurement. (A) A 3D model of BMV reconstructed from cryo-electron microscopy data[38] shows the protein capsid (gray) surrounding the RNA (gold). The model reveals most of the icosahedral capsid but only a small portion of the RNA, the rest of which adopts a disordered arrangement within the capsid. (B) A cartoon of the experiment shows viral coat proteins assembling around RNA strands that are tethered by DNA linkages to the surface of a functionalized glass coverslip. (C) The assembling proteins are imaged at 1000 Hz for 600 s using iSCAT microscopy. Each dark spot that appears in the images corresponds to proteins bound to an individual RNA strand. The darkness, or intensity, of each spot is proportional to the number of proteins bound to that RNA. The displayed images are the average of 1000 consecutive frames. (D) Traces of the intensity as a function of time (1000 frame moving average) reveal the assembly kinetics for each particle. Experimental conditions are 0.135 µmol/L protein and 250 mmol/L NaCl. The initial spike in intensity present in many of the traces is associated with vibrations introduced into the system as coat protein is injected. The thick, black trace corresponds to the boxed particle in (C). We compare the final intensities of the traces to the estimated intensity range of full capsids, which is shown as a vertical bar to the right of the traces.

2.3 Materials and Methods

2.3.1 Buffers used

Disassembly buffer: 50 mmol/L Tris-HCl, pH 7.5; 500 mmol/L $CaCl_2$, 1 mmol/L ethylenediamine tetraacetic acid (EDTA), 1 mmol/L dithiothreitol (DTT), 0.5 mmol/L phenylmethylsulfonyl fluoride (PMSF). **Protein storage buffer:** 20 mmol/L Tris-HCl, pH 7.2; 1

mol/L NaCl; 1 mmol/L EDTA; 1 mmol/L DTT; 1 mmol/L PMSF. **TAE buffer with NaCl:** 40 mmol/L Tris-HCl pH 8.3, 20 mmol/L acetic acid, 1 mmol/L EDTA; and 84 mmol/L, 167 mmol/L, or 250 mmol/L NaCl. **Assembly buffer:** 42 mmol/L 2-(N-morpholino)ethanesulfonic acid (MES), pH 6; 84 mmol/L, 167 mmol/L, or 250 mmol/L NaCl; 8.4 mmol/L $MgCl_2$; and 3 mmol/L acetic acid. For DSF measurements, we also prepared assembly buffer at pH 7 by replacing MES with sodium phosphate.

2.3.2 Synthesis of BMV RNA1

BMV RNA1 was made by in vitro transcription of the DNA plasmid pT7B1, linearized with BamHI (New England Biolabs, USA)*, with a T7 polymerase transcription system (Thermo Fisher, USA) and purified with an RNEasy Mini Kit (Qiagen, DEU), both following the manufacturers' specifications.

2.3.3 BMV coat protein purification

BMV was purified from infected barley leaves (Hordeum vulgare)[3], and coat protein was purified as described previously[39]. Nucleocapsids were disassembled by dialyzing against disassembly buffer at 4 °C overnight. The RNA was pelleted and the coat protein isolated by ultracentrifugation at 90,000 rotations per minute for 100 min at 4 °C in a Beckman TLA110 rotor. Coat protein was extracted from the supernatant and immediately dialyzed against protein storage buffer. Protein concentration and purity were assessed by UV-Vis spectrophotometry; only protein solutions with 260/280 ratios less than 0.6 were used for assembly. Protein was frozen in liquid nitrogen and stored at −80 °C until ready to use, at which point it was defrosted on ice and stored at 4 °C for up to two weeks.

2.3.4 Measuring the binding yield between fluorescently labeled 20U RNA and BMV CP using a nitrocellulose-binding assay

Fluorescently labeled 20U RNA at a concentration of 4 nmol/L was mixed with 0.043 μmol/L BMV CP_2 in TAE buffer with either 84 mmol/L, 167 mmol/L, or 250 mmol/L NaCl, and left for 30 min at room temperature. The RNA was labeled at its 5'-end with an AlexaFluor647 dye (Integrated DNA Technologies, USA). 250 μL aliquots of each RNA-protein mixture were passed through a nitrocellulose membrane (0.45 μm pore size; Thermo Fisher, USA) that was presoaked in TAE buffer with 250 mmol/L NaCl using a 96-well dot-blot apparatus (Biorad, USA) under weak vacuum. Then the membrane was washed by passing through an additional 500 μL of the corresponding assembly buffer. The amount of membrane-bound RNA was quantified at each salt concentration by measuring the fluorescence emission intensity of the AlexaFluor647 dye using a fluorescence scanner. The amount of protein bound to the membrane was constant for all salt concentrations, as determined by staining with Ponceau S solution (MilliporeSigma, USA). Because protein is retained by the membrane, but free RNA is not, the intensity of membrane-bound RNA is taken to be proportional to the yield of RNA-protein binding.

2.3.5 Measuring the apparent melting curve of BMV by differential scanning fluorimetry (DSF)

WT BMV was dialyzed against assembly buffer with pH 6 or pH 7 overnight at 4 °C. Aliquots of WT BMV at a final concentration of 0.2 mg/mL, 2.5x SYPRO orange fluorescent dye (Molecular Probes, USA), and a final salt concentration of 85 mmol/L, 167 mmol/L, or 250 mmol/L NaCl were prepared. DSF was performed in triplicate in a 96-well plate CFX Connect quantitative PCR machine (Bio-Rad, USA). All samples were heated from 25 °C to 95 °C, in 1

°C increments with a 1-min stabilization period at each temperature before the sample was measured. Excitation/emission wavelengths of 470/550 nm were used to detect the fluorescence emission of SYPRO orange binding to hydrophobic regions of coat protein exposed upon capsid disassembly. The apparent melting temperature is defined as the temperature at which the derivative of the fluorescence emission signal ($-dI/dT$) is maximal.

2.3.6 Negative-stain electron microscopy

Negative-stain electron microscopy was used to image the protein structures that assemble around RNA in solution. The assembly reaction was carried out in assembly buffer by mixing 860 nmol/L BMV CP_2 with 7.5 nmol/L BMV RNA1 and incubating at room temperature for 10 min. 6 μL of assembly reaction was deposited on glow-discharged carbon-coated copper (200-mesh) PELCO Pinpointer grids (Ted Pella, USA). After 1 min, the grids were blotted with Whatman filter paper, and then stained with 6 μL of 2 % uranyl acetate for 1 min followed by complete stain removal and storage in a desiccator overnight. Micrographs were acquired using a Tecnai G2 TF20 High-Resolution electron microscope (FEI, USA) with an accelerating voltage of 200 kV. Images were collected at 3 μm to 4 μm underfocus with a TIETZ F415MP 16-megapixel CCD camera (4000 by 4000 pixels, pixel size 15 μm).

2.3.7 iSCAT experiments

iSCAT measurements were performed as described in detail in Reference 18. In brief, a spatially filtered, 450 nm diode laser was coupled into an oil-immersion objective to illuminate a small region of the coverslip with a collimated beam. Three-dimensional active stabilization was used to extend the measurement duration. We estimate the intensity range of full BMV capsids by scaling the measured intensity range for MS2 capsids[11] by the relative mass of BMV and MS2 capsids. Coverslips were functionalized with a layer of polyethylene-glycol (PEG) molecules,

about 1 % of which were functionalized with a 20-base DNA strand (5'-GGTTGGTTGGTTGGTTGGTT-3'), to which BMV RNA1 strands were tethered using a 60-base DNA linker strand (5'-CCGTGGTCGACAAGGGATTGAACCTCGTTCCGTGGTCTACAACCAACCAACCAACCAACC-3'). Assembly kinetics experiments were performed at room temperature in assembly buffer with 84 mmol/L, 167 mmol/L, and 250 mmol/L NaCl and protein concentrations of 0.043 μmol/L, 0.135 μmol/L, and 0.427 μmol/L BMV CP_2.

2.4 Results

2.4.1 RNA-protein interactions can be controlled independently of protein-protein interactions

We first demonstrate that we can control the strength of RNA-protein interactions. By measuring the yield of RNA-protein binding between 20 nucleotide-long poly-U RNA and BMV coat-protein dimers (CP_2) using a nitrocellulose binding assay[40], we find that increasing the concentration of salt decreases the amount of RNA-protein binding (**Figure 2A**). This result is consistent with binding being driven by electrostatic interactions between the negatively charged phosphate backbone of the RNA and the positively charged N-terminus of the coat protein[37].

To test whether varying ionic strength also affects protein-protein interactions, we perform Differential Scanning Fluorimetry (DSF) on wild-type BMV particles to study their thermal stability, a measure of lateral interactions between proteins in the capsid. We find that the thermal stability is approximately constant over the range of NaCl concentrations tested (**Figure 2B**). In contrast, the stability of BMV drops sharply when the pH is increased from 6 to 7, consistent with previous results[41,42] showing that increasing pH weakens protein-protein interactions in BMV (**Figure 2B**). The absence of a change in stability with increasing salt is

important because it suggests that ionic strength only weakly affects protein-protein interactions, and therefore that ionic strength can be used to specifically modify RNA-protein interactions.

Next, we use negative-stain electron microscopy to test for capsid formation as we vary ionic strength. In assembly reactions involving BMV CP_2 and BMV RNA, we find that 84 mmol/L NaCl leads to heterogeneous assembly products, with some well-formed particles and many malformed particles, whereas 167 mmol/L and 250 mmol/L NaCl gives rise to more homogeneous spherical particles with roughly the same size and curvature as wild-type BMV (**Figure 2C**).

Taken together, these measurements show that we can control the strength of RNA-protein interactions, and that changing these interactions changes the assembly products. These results set the stage for iSCAT single-particle measurements that probe how the assembly pathways might change as RNA-protein interactions are varied.

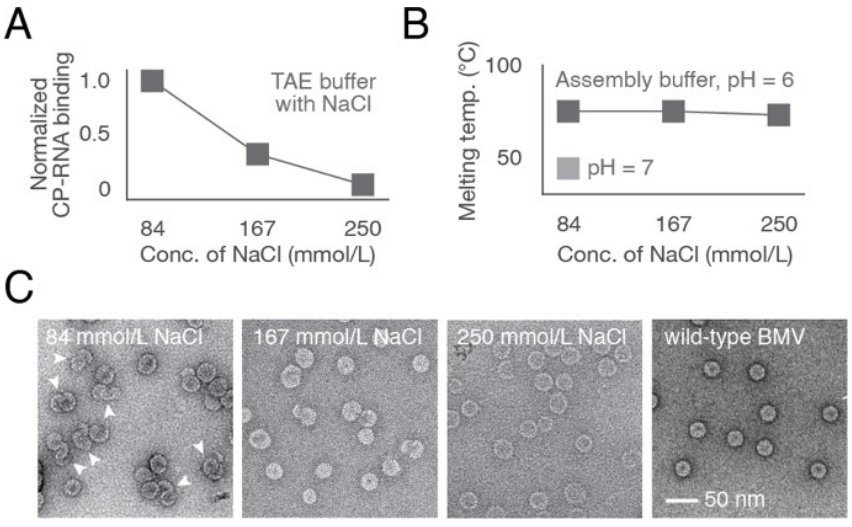

Figure 2. Ionic strength affects RNA-protein binding and the assembly process but does not affect the thermal stability of assembled particles. (A) Nitrocellulose binding measurements show that the yield of RNA-protein binding depends strongly on ionic strength. The binding yield measurements are normalized with respect to the amount of binding at 84 mmol/L NaCl. (B) Differential Scanning Fluorimetry (DSF) measurements of the melting temperature of wild-type BMV in assembly buffer at pH 6 with varying NaCl, and at pH 7 with fixed NaCl, show that the capsid stability depends weakly on ionic strength, relative to pH, across the conditions tested. (C) Uranyl acetate negative-stain transmission electron microscope images of BMV particles assembled in assembly

buffer at varying ionic strengths. The 84 mmol/L NaCl sample is notably heterogeneous, presenting many more malformed particles compared with the higher-salt assemblies, which are more monodisperse. Wild-type BMV particles are shown as a reference.

2.4.2 iSCAT measurements show that BMV assembles by a nucleation-and-growth pathway under certain conditions

We use iSCAT to measure the self-assembly kinetics of individual BMV particles as a function of the RNA-protein interaction strength, as controlled by the salt concentration (84 mmol/L, 167 mmol/L, and 250 mmol/L NaCl). We also vary the protein concentration (0.043 μmol/L, 0.135 μmol/L, and 0.427 μmol/L CP_2) for a total of 9 possible experimental conditions. However, we did not test one of these combinations (250 mmol/L NaCl and 0.043 μmol/L CP_2) because, given the results for the other 8 conditions, we expect the assembly timescales to be longer than our 600 s measurement time. In each of the 8 experimental conditions tested, we performed measurements in duplicate. In total, we measured and analyzed the assembly kinetics of 511 particles, 72 of which are shown in **Figure 3**.

We see qualitative differences among these single-particle traces for different RNA-protein interactions, most notably in the time at which each trace begins to increase rapidly (the "start time"). Consider the middle column in **Figure 3,** which corresponds to 0.135 μmol/L CP_2. When RNA-protein interactions are strongest (corresponding to 84 mmol/L NaCl), each trace begins increasing immediately after the protein is introduced. In contrast, for weaker RNA-protein interactions (corresponding to 167 mmol/L NaCl), each trace remains at a low intensity for a variable amount of time before increasing, with some traces having start times greater than 100 s at the lowest protein concentration. When RNA-protein interactions are weaker still (corresponding to 250 mmol/L NaCl), the distribution of start times extends to longer values, with some traces having start times as large as 500 s at the intermediate protein concentration.

Remarkably, one of the two experiments for 250 mmol/L NaCl did not result in any traces increasing above their initial value over the 600 s experiment.

These experiments allow us to rule out diffusion-limited accretion of proteins on the RNA as an assembly pathway for at least some of these conditions. If assembly were purely diffusion-limited[43], we would expect the onset of assembly to vary only with the concentration of protein. Instead, our results show that the start times vary with ionic strength at the same protein concentration.

The variation in start times points to a free-energy barrier to protein accretion on the RNA—in other words, a nucleation barrier[35]. In classical nucleation theory, the barrier is associated with an initially unstable cluster—here, of proteins—becoming large enough that subsequent proteins bind favorably. Because nucleation is a stochastic process, the time required to form a sufficiently large cluster varies from particle to particle. These timescales are reflected in our start-time measurements. The observation that the distribution of start times broadens and shifts to larger values as RNA-protein interactions are weakened (but protein-protein interactions are not significantly changed) reveals that nucleation in BMV is a heterogeneous process, driven at least in part by attractive interactions between the RNA and the assembling proteins.

For the strongest RNA-protein interactions there is no obvious evidence of a nucleation barrier: At 84 mmol/L NaCl, the traces have nearly identical start times, as shown in the top row of **Figure 3**. At these conditions, we see that the final intensities of the traces, a measure of the number of proteins attached to the RNA, increase with increasing protein concentration. Moreover, the final intensities tend to decrease with decreasing RNA-protein interaction strength. This result is qualitatively consistent with our electron microscopy measurements, which show heterogeneous, malformed, larger-than-wild-type particles for the strongest RNA-

protein interactions (see structures highlighted by arrowheads in **Figure 2C**), and increasingly homogeneous, spherical, roughly wild-type-size particles as the RNA-protein interaction strength decreases. Understanding why the assembly products vary with RNA-protein interaction strength requires a more quantitative analysis of the assembly kinetics, and in particular how nucleation rates compare to growth rates.

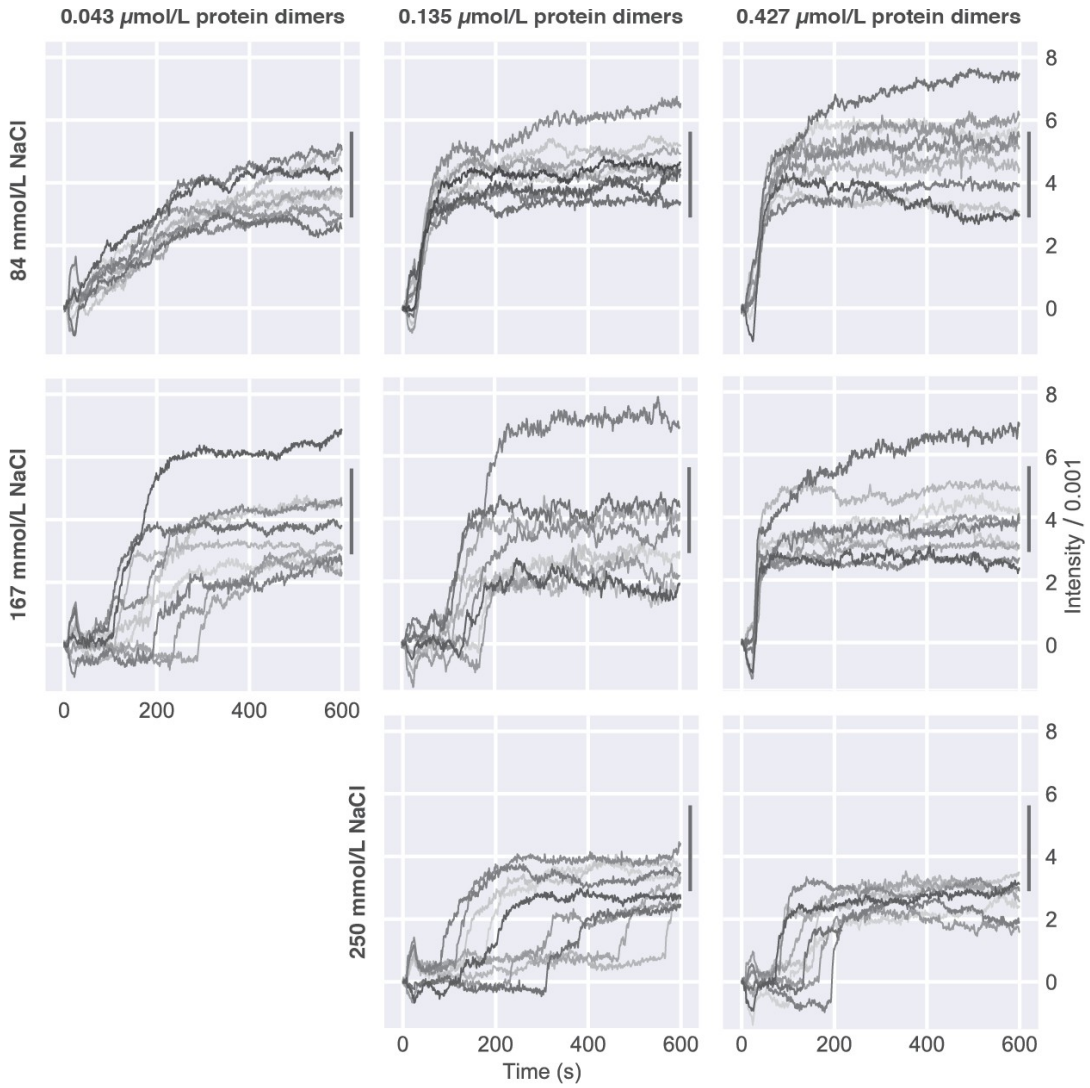

Figure 3. Single-particle measurements reveal qualitative differences in the assembly kinetics with varying ionic strength and protein concentration. Plots of traces showing intensity as a function of time from iSCAT experiments. Each trace represents the assembly of a single particle consisting of BMV coat protein and BMV RNA. We show 9 randomly selected traces at each condition. The bars to the right of the traces show the range of intensities that are consistent with a full capsid.

2.4.3 The shapes of the assembly traces reveal differences between nucleation and growth phases

To extract quantitative information from each recorded trace, we analyze the traces as shown in **Figure 4** to determine three kinetic parameters: the final intensity (I_f), the start time (t_s), and the growth time (τ_g). The values of these parameters for experiments with 0.135 μmol/L CP and varying ionic strength are shown as histograms in **Figures 5A–C**.

Consistent with our qualitative analysis of the traces, the start-time histograms show that decreasing the strength of RNA-protein interactions increases the median start times and broadens their spread, as quantified by the median absolute deviation or MAD (**Figure 5A**). We find that the median start times increase by more than a factor of 6 and the MAD of start times increases by more than a factor of 14 when we weaken the RNA-protein interactions by increasing the salt concentration from 84 mmol/L to 250 mmol/L (**Figure 5A**). By contrast, the growth times and final intensities are more narrowly distributed and less affected by changes in RNA-protein interactions, showing an increase by a factor of roughly 1.5 in the median growth times and no increase in the MAD of the growth times (**Figure 5B**). The median final intensities show a corresponding decrease in particle size by factor of over 1.5 (**Figure 5C**).

These results show that RNA-protein interactions primarily affect the nucleation phase of the assembly pathway and only weakly affect the growth phase, if at all. Specifically, stronger RNA-protein interactions reduce the characteristic nucleation time until, at sufficiently large interaction strengths, it is no longer possible to tell if the assembly pathway is nucleated because all particles begin to assemble in apparent synchrony. We discuss this point in more detail in the Discussion.

To gain further insights into the nucleation kinetics, we examine the effect of protein concentration on the start-time distribution. We plot the MAD of start times, t_{MAD}, for each protein concentration and each RNA-protein interaction strength (corresponding to values of salt concentration) in **Figure 6A**. In principle, the slope, α, of the fitted line represents the exponent in a power-law scaling $\tau_{MAD} \propto c^{-\alpha}$, where c is the protein concentration. We find that for intermediate RNA-protein interaction strength (corresponding to 167 mmol/L NaCl), the measured start times become much more narrowly distributed as the protein concentrations increase, scaling with an exponent $\alpha = 1.5 \pm 0.2$. These results are once again consistent with assembly being a nucleated process—at higher protein concentrations the law of mass action drives more proteins onto the RNA, favoring larger clusters and shortening the time needed to form a nucleus.

In contrast, for strong RNA-protein interactions (corresponding to 84 mmol/L NaCl), the start times remain narrowly distributed for all protein concentrations tested. Because the experimental uncertainties in our start-time measurements are on the order of seconds, small absolute differences in MAD values are not statistically significant. Thus, the line between nucleation and diffusion-limited aggregation becomes blurred. While it is possible that nucleation is occurring on a time scale too fast for us to measure, it is also possible that strengthening RNA-protein interactions qualitatively changes the assembly pathway such that it is no longer nucleated. We discuss these results further in the Discussion.

The growth kinetics are much less affected by varying either protein or salt concentration, as shown in **Figure 6B**. For all but one of the conditions tested, the median growth times differ by a factor of 2.5 at most. Only for 0.043 μmol/L CP_2 and 84 mmol/L NaCl does the median

growth time extend to 100 s. And while the median growth times tend to decrease with increasing protein, the amount by which they decrease is constant for all RNA-protein interaction strengths, as seen by the indistinguishable slopes of the best-fit lines in **Figure 6B** (slopes are 0.7 ± 0.3 at 84 mmol/L NaCl and 0.5 ± 0.1 at 167 mmol/L NaCl). Thus, while protein concentration appears to have a large effect on the start times, it has a much weaker effect on the growth times. Likewise, while RNA-protein interaction strength has a large effect on the start times, it does not significantly affect the growth times.

These results may explain why larger particles are favored at higher protein concentrations for all RNA-protein interactions tested, and at stronger RNA-protein interactions for all protein concentrations tested. These trends can be seen in the heat map shown in **Figure 6C**. In both cases—increasing protein concentration and increasing RNA-protein interaction strength—the nucleation times decrease relative to the growth times. Multiple nucleation events can therefore happen on the same RNA before any given nucleus has time to grow into a full capsid. Consequently, malformed structures consisting of multiple partially assembled capsids can occur. Indeed, the electron microscope images in **Figure 2C** show aggregates consisting of multiple capsid-like fragments that have the same curvature as the wild-type virus. The number and size of these aggregates decreases with decreasing RNA-protein interaction strength.

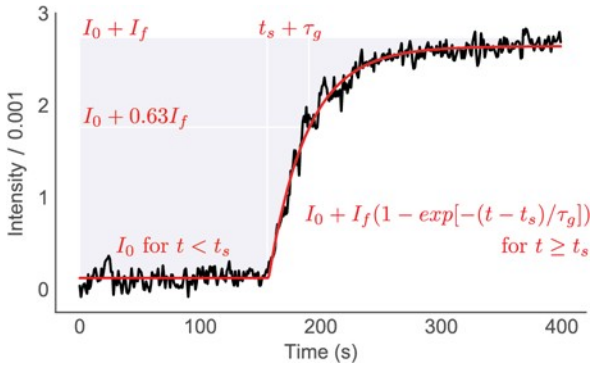

Figure 4. Extracting quantitative information from the assembly traces: I_f, t_s, and τ_g. An intensity trace of the assembly of a single particle (black curve, corresponding to the boxed particle in Figure 1C) can be fit (red curve) to a piecewise function $I(t) = \begin{cases} I_0, & t < t_s \\ I_0 + I_f\left(1 - \exp\left[-(t - t_s)/\tau_g\right]\right), & t \geq t_s \end{cases}$ where *I(t)* is the intensity as a function of time, I_0 is the initial intensity (which can be offset from zero because of microscope drift), I_f is the final intensity, t_s is the start time—the time at which the trace begins to rise from its initial value, and τ_g is the growth time (the time for the trace to reach (1-1/*e*) = 0.63 of its final value once it has started increasing). The exponential function is not intended to imply a particular assembly mechanism (such as Langmuir adsorption) but rather is chosen because it is a simple function that fits the data well for most traces, allowing us to extract a characteristic growth timescale. In this way, we determine a set of three kinetic parameters, I_f, t_s, and τ_g, for each of the traces measured in each of the experimental conditions tested.

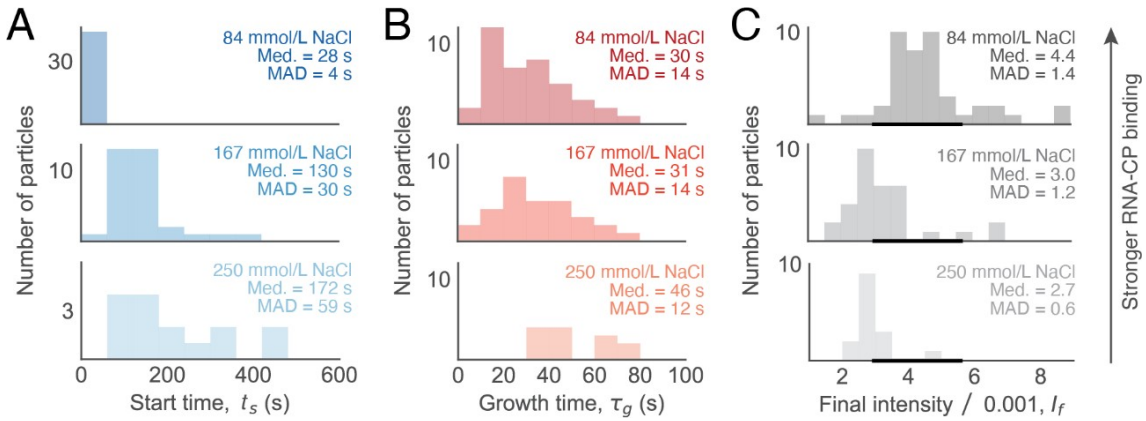

Figure 5. RNA-protein interactions strongly affect nucleation but only weakly affect growth. Histograms of start time (t_s), growth time (τ_g), and final intensity (I_f) are plotted for three experiments with 0.135 μmol/L CP_2 and either 84 mmol/L, 167 mmol/L, or 250 mmol/L NaCl: (A) Start times t_s increase and broaden with decreasing RNA-protein interaction strength; (B) growth times τ_g are less affected by RNA-protein interactions; and (C) the final intensity I_f decreases with decreasing RNA-protein interaction strength. The bold portion of the *x*-axis in (C) shows the size range for full capsids. In panels A–C, the median (Med) and median absolute deviation (MAD) of each fit parameter are listed. Both the median and MAD are robust to outliers that arise from the few traces that are fit poorly by the piecewise function.

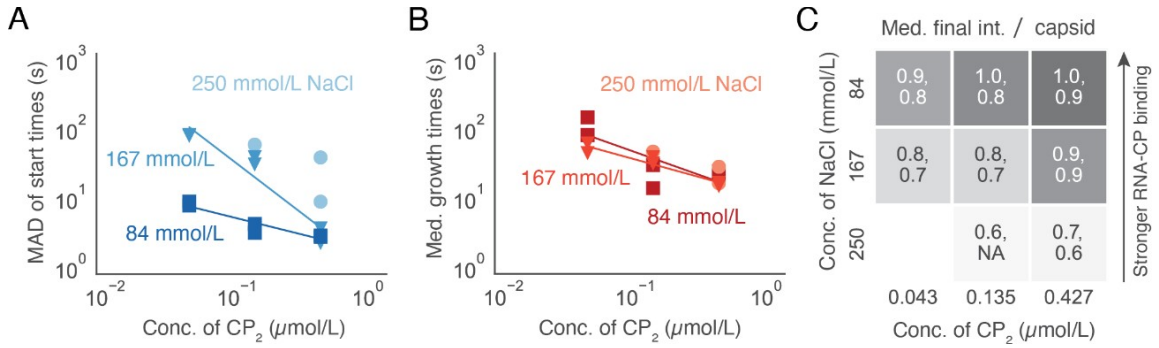

Figure 6. Experiments at varying protein concentration suggest that nucleation occurs by a collective process involving multiple proteins, while growth occurs by a lower-order process. (A) A log-log plot of the spread of start times as a function of protein concentration for each ionic strength tested. Experiments with 84 mmol/L salt are plotted as squares, 167 mmol/L salt as triangles, and 250 mmol/L salt as circles. Duplicate measurements are plotted for each set of conditions except for 250 mmol/L salt and 0.135 μmol/L CP_2, in which one of the duplicates did not yield assembly during the 600 s measurement time, and 250 mmol/L salt and 0.043 μmol/L CP_2, which was not tested. Lines have been fit to the data from the 84 mmol/L and 167 mmol/L salt experiments to show how the dependence on protein concentration becomes steeper at higher salt. (B) Log-log plot of the median growth times shows a concentration dependence that is roughly the same for all ionic strengths. (C) Heat map of the median final intensity, normalized by the intensity of a full capsid, of each duplicate measurement. The darkness of the gray color represents the magnitude of the intensity, which is proportional to the size of the assembled particle. While the intensities increase with increasing protein and decreasing salt, many of the assembled particles do not reach the size of a full capsid, indicating that they are missing proteins. These smaller particles might have partial capsids that fail to completely close around the tethered RNA strand. Duplicate measurements are separated by commas.

2.5 Discussion

We discuss our results for the self-assembly kinetics of individual BMV particles in comparison with previous results using the same technique to study bacteriophage MS2[35], focusing first on their similarities. By highlighting similarities in the assembly kinetics, we aim to identify common features of the assembly process that might be found in other RNA viruses.

The assembly traces for BMV are strikingly similar to those previously reported for MS2 (**Figure 7)** when RNA-protein interactions in BMV have been sufficiently weakened by sufficiently high salt concentrations (167 mmol/L or 250 mmol/L NaCl). For both viruses, we observe (i) broad distributions of start times that narrow with increasing protein concentration, consistent with a nucleation step; (ii) growth times that decrease with increasing protein concentration, but less rapidly than do the start times, consistent with growth involving a lower-order process; and (iii) increases in the fraction of overgrown particles with increasing protein

concentration, with overgrown particles consisting of aggregates of partially formed capsids. We note that the concentration of coat protein used in the assembly of MS2 is roughly a factor of 10 higher than that used with BMV, related to BMV coat protein having a stronger overall binding affinity for RNA.

Another interesting similarity between BMV and MS2 assembly is that the apparent critical nucleus size is small—a few proteins—in both cases. We estimate this nucleus size from the amplitude of the fluctuations prior to the start time. In nearly all traces under conditions of moderate to weak RNA-protein interactions (167 mmol/L or 250 mmol/L NaCl), these fluctuations are comparable to the noise level, which corresponds to the scattered intensity from a few CP_2. We use the term "apparent critical nucleus size" because averaging reduces our ability to observe large sub-critical fluctuations. Nonetheless, this small apparent critical nucleus is comparable to that observed for MS2, and for both viruses we find that growth proceeds monotonically for nearly all traces (**Figure 7A and B**).

These results raise the question of how either BMV or MS2 is able to grow so rapidly, in a way that ensures the capsid has the correct curvature, from such an apparently small nucleus. Nucleation by itself cannot solve the "Levinthal paradox" of virus assembly because the nucleus is too small to dictate the formation of a T=3 structure during growth. And yet, following nucleation, both BMV and MS2 display essentially monotonic growth to the size of a full capsid, with little if any disassembly or backtracking along the way. Furthermore, the curvature of the assembly products observed by electron microscopy is similar to that of wild-type capsids even when the products are malformed aggregates of partially assembled capsids. How, then, do the coat proteins add to a growing nucleus, forming both the pentamers and hexamers needed for the proper curvature, without frequent stalling events or needing to detach from the particle?

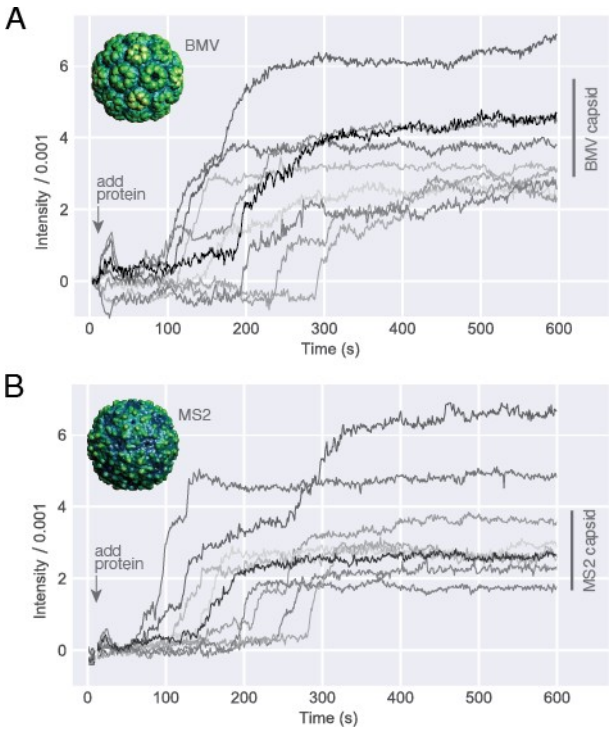

Figure 7. Comparison of the assembly traces for BMV and MS2 reveals similar nucleation-and-growth kinetics. (A) BMV assembly traces for 0.043 μmol/L protein and 167 mmol/L NaCl. (B) For comparison, we show traces recorded for bacteriophage MS2 assembly under conditions of 2 μmol/L protein and 84 mmol/L NaCl (data from (18)). Both sets of traces reveal a broad distribution of start times followed by rapid and monotonic increases in intensity. In both sets of traces, most traces plateau to a final value that is roughly consistent with a full capsid, with several traces plateauing to larger values. The bars to the right of the traces show the range of intensities that are consistent with a full capsid. These ranges are different for MS2 and BMV, reflecting the different molecular masses of their coat proteins.

A recent hypothesis is that specific subsequences on the RNA mediate the growth process to produce the correct curvature and structure. These subsequences are sometimes called "packaging signals"—a term that has come to refer to any elements of the RNA secondary structure that have high local affinity for coat protein. A model for MS2 assembly has been proposed in which upwards of 60 strategically positioned packaging signals[29,44] guide assembly of the coat proteins. While not providing direct evidence in support of this model, high-resolution cryo-electron microscopy studies of MS2 particles show a dozen or so specific interactions between the RNA genome and its surrounding protein capsid[45], suggesting that packaging signals play a role in ordering the genome. But BMV is different. The absence of such

order in the structure of the packaged BMV genome[38], and the lack of evidence of strategically positioned packaging signals in BMV RNA, argue against there being a general paradigm that depends on a special distribution of specific RNA-protein interactions.

It is therefore interesting that we observe similar assembly kinetics for BMV and MS2. Our measurements suggests that whatever role packaging signals play in the self-assembly of these two viruses, their effect appears to be at most quantitative, with the qualitative features of the assembly process determined by general non-specific interactions between the assembling proteins and the RNA. Determining the quantitative effects of packaging signals will ultimately require kinetic experiments in which the RNA sequence and secondary structure is controlled[46].

Our experiments with BMV show that one should look to protein-protein interactions and intra- protein forces, rather than RNA-protein interactions, to understand how the capsid grows after nucleation. The median growth time does not vary with salt concentration (**Figure 5B**), indicating that RNA-protein interactions are not involved in the rate limiting step of the growth phase of assembly. Protein-protein interactions or intra-protein forces must therefore direct the rate-limiting step. We have also shown that the median growth time varies less strongly with protein concentration than does the start time (**Figure 6A-B**). A simple scenario consistent with these observations is that the RNA does not play a significant role in the assembly process after nucleation. Instead, during the growth phase, new proteins bind directly to other proteins in the assembling particle. These incoming proteins bind individually, not as clusters.

Thus far we have discussed the similarities between BMV and MS2 assembly, but there are differences that arise when RNA-protein interactions are strong. Under conditions of 84 mmol/L NaCl and 0.043 μmol CP_2, experiments on BMV assembly show a small spread in start times and relatively large growth times (see upper left plot in **Figure 3**). With MS2, we found

that when the spread in start times is smaller than the average growth time, many oversized structures form. Our interpretation was that multiple nucleation events occur on the same RNA strand before the first nucleus can grow and sequester the RNA. With BMV, we observe similar overgrown structures but only at higher protein concentration (**Figure 2C**). We do not observe a significant number of overgrown structures at low protein concentration. One possible explanation is that assembly is diffusion-limited when RNA-protein interactions are strong, occurring with no nucleation barrier, and that the small observed variation between traces is due to measurement noise. Alternatively, it is possible that the assembly is nucleated, but the nucleation time is too small to resolve. In data not shown, we perform simulations to determine if our measurements are consistent with any of the above scenarios—a diffusion-limited pathway subject to measurement noise or a nucleated pathway with a small nucleation time. These simulations suggest that our measurements are consistent with either of these pathways. Thus, additional experiments that can resolve smaller spreads in start times are needed to determine the assembly pathway when RNA-protein interactions are strong.

There also remains the question of if and when an ordered capsid arises from the RNA-protein complex. Interferometric scattering measurements do not address this question because the intensity depends primarily on the number of RNA-bound proteins and weakly, if at all, on their structure. There are at least two possibilities for the pathway underlying those traces that do not show a clear separation between nucleation and growth. One is a barrier-less formation and growth of partially to completely ordered capsids, and the other is a barrier-less accretion of disordered proteins, followed by the onset of order among the bound proteins. These pathways cannot be distinguished in iSCAT because the measurement is blind to the onset of order. The same is true even when there is a clear separation between nucleation and growth processes.

Here we observe a barrier to protein accretion, but it is possible that the proteins are not ordered. An analogous situation occurs in a bulk phase transition when gas-to-solid condensation takes place through a liquid-state intermediate. Thus, while our measurements can definitively resolve the presence of a barrier to protein accretion, future structural measurements are needed to distinguish whether the barrier involves ordered or disordered proteins.

2.6 Conclusions and Future Directions

The iSCAT experiments, with their high temporal resolution and ability to resolve the kinetics of assembly of individual viral capsids, offer the most detailed view to date of virus self-assembly pathways. Although iSCAT cannot reveal at what stage the capsid becomes ordered, much information can be gleaned by combining the technique with electron microscopy and bulk assays, as we have shown.

These experiments have shown that when RNA-protein binding is weak, the assembly kinetics of BMV are similar to those of MS2, involving a nucleation phase in which small numbers of proteins bind to the RNA, followed by a monotonic growth phase in which a capsid-worth of proteins steadily accrues. When RNA-protein binding is strong, we observe no clear barrier to protein accretion and hence no separation between nucleation and growth. Because our method may obscure nucleation events that occur on fast timescales, future studies are needed to determine whether for strong RNA-protein interactions there is a qualitatively different assembly pathway involving a saturated—"en masse"[10,47]—adsorption of proteins on the RNA.

When the assembly is nucleated, we have shown that the time required to form a nucleus depends strongly on the strength of RNA-protein interactions, whereas the time needed to accrue a capsid-worth of proteins does not, suggesting that the RNA plays a more central role in the nucleation phase than in the growth phase. These results are consistent with protein accretion

being a heterogeneous process in which subcritical protein clusters form on—and are stabilized by—the RNA. This role of the RNA degree of freedom—involving both RNA-protein binding and the conformation of the RNA itself[48]—compromises the notion of the nucleus being a fixed arrangement of coat proteins, such as a hexamer[49] or a pentamer[50] of dimers, although the uncertainties in our measurements cannot rule out this possibility. Future structural experiments, or kinetic experiments that address a broader range of timescales, may clarify whether the RNA enables different nucleus structures in different conditions.

Our experiments also reveal important features of the growth process. The shapes of the iSCAT traces show that growth can take place rapidly and without significant errors starting from an apparently small critical nucleus of only a few proteins. Furthermore, the weak dependence of growth times on protein concentration suggests that growth occurs by a lower-order process, such as the addition of individual protein subunits from solution, rather than higher-order collective processes involving clusters of proteins.

The model of capsid assembly that emerges from our results is as follows: the strength of RNA-protein interactions primarily controls the kinetics of nucleation but not the local structure or curvature of the capsid, while protein-protein and intra-protein interactions control the growth phase and the emergence of the T=3 structure. This model is compatible with simulations and theory showing that the growth of a T=3 structure is driven by minimization of the elastic energy of the capsid[51,52], which in general is related to the stretching and bending of coat-protein dimers as well as of the bonds between them. The elastic-energy hypothesis could explain why BMV and MS2 show similar assembly pathways.

To test this hypothesis, future experiments and analysis might focus on the growth process rather than nucleation. Recently, BMV, MS2, and Qβ capsids of different sizes and symmetries have been observed[53–58]. Experiments that determine how such malformed or overgrown capsids form would shed light on the interactions that control growth. Another useful next step is to develop coarse-grained computer simulations[47] that include the diffusion of proteins to the surface that is inherent to the iSCAT experiments. With such models it would be possible to directly compare ensembles of simulated kinetic traces to the ensembles of traces measured with iSCAT. If agreement between these simulations and our data can be obtained, it would point the way toward a detailed mechanism of the nucleocapsid formation process.

The commonalities in the *in vitro* assembly pathways of BMV and MS2 are remarkable in light of the vast phylogenetic distance between these viruses, their different RNA structures, and the differences in the specificity and strength of RNA-protein interactions. We do not know if the same pathways are operative *in vivo*, and there is reason to suspect that, at least for MS2, there may be differences between the *in vivo* and *in vitro* pathways. Whereas wild-type BMV particles consist of only two components, the RNA genome and (180 copies of) the capsid protein, wild- type MS2 particles contain a single copy of a different gene product (the "maturation" protein) that replaces a coat protein dimer and breaks icosahedral symmetry. The *in vitro* assemblies of MS2 are reconstituted from viral RNA and capsid protein alone, without the maturation protein and with T=3 structure.

Nonetheless, the existence of two different icosahedral, T=3 capsids that assemble *in vitro* in similar ways is intriguing from both a physical and evolutionary perspective. From a physical perspective, this result suggests that the assembly of T=3 viruses might be understood through a general physical theory. From an evolutionary perspective, it highlights the question of

how icosahedral viruses with quasi-equivalent ($T > 1$) capsid subunits evolved. In MS2, for example, a single point mutation in the coat protein changes the structure of the capsid from T=3 to T=1 [58]. Because capsids are self-assembled, such mutations can in principle also change the assembly pathway. Future studies might use interferometric scattering to explore how mutations affect the assembly of virus-like particles from different virus families, with the aim of discovering whether there are conserved interactions that promote robust assembly of the T=3 structure.

2.7 Acknowledgments

We thank Ben Rogers and Abigail Chapman for many helpful discussions. Research reported in this publication was supported by the National Institute of General Medical Sciences of the National Institutes of Health (Award Number R00GM127751 to RFG) and the National Science Foundation Molecular and Cellular Biosciences Division (Awards MCB 1716925 and 2103700 to WMG). This research was partially supported by NSF through the Harvard University Materials Research Science and Engineering Center under NSF grants DMR-1420570 and DMR-2011754. AMG was partially supported by an NSF Graduate Research Fellowship under Grant DGE-1144152. We acknowledge the California Metabolic Research Foundation for its support of biochemical research at San Diego State University, and the Bauer Core Facility at Harvard University for shared experimental facilities used in this study.

CHAPTER 3

One-step *in vitro* self-assembly of virus-like particles as a function of pH and ionic strength

3.1 Abstract

Although all viruses assemble in cells, only a few have been shown to assemble in a test tube from purified components. Assembly in a test tube allows for control of the system, which will enable us to tune parameters like protein-protein and protein-RNA interactions through solution conditions and interrogate what interactions are important for successful assembly of virus-like particles (VLPs). This work investigates the role of pH and ionic strength on wild-type brome mosaic virus (BMV) particle stability by measurements of apparent melting temperature in various solution conditions. VLPs are *in vitro* self-assembled in a range of pH and ionic strength conditions and assembly products are analyzed by negative-stain electron microscopy. Wild-type BMV particle stability decreases with increasing pH but is less affected by changes in ionic strength. The particles show the highest stability at pH 6 and ionic strengths below 167 mM. The one-step assemblies at pH 6 and 167 mM NaCl also produced the highest yield of well-formed VLPs out of the many solution conditions tested. While additional conditions should still be tested (specifically lower pH solutions), these results suggest that one-step assemblies can yield well-formed VLPs within 10 minutes.

3.2 Introduction

In the 1950s and 1960s, the first and only viruses – tobacco mosaic virus (TMV), cowpea chlorotic mottle virus (CCMV), and brome mosaic virus (BMV) – demonstrated the ability to assemble *in vitro* from purified components[6,25,26,59–61]. These simple RNA viruses, consisting of a

single-stranded RNA (ssRNA) genome inside of an ordered capsid made up of multiple copies of a single capsid protein (CP), have inspired theoretical and experimental work in understanding the mechanism and pathways involved in the formation of these structurally ordered and protective particles.

There are two prevailing scenarios for spontaneous viral self-assembly of spherical capsids around a ssRNA genome according to whether protein-protein or protein-RNA interactions are dominant: (i) growth by recruitment of unbound proteins; and (ii) growth by rearrangement of bound proteins. In the first case, a critical nucleus of capsid forms from the small number of proteins that have bound to the RNA, with ensuing growth and completion of the capsid facilitated by recruitment from the solution of additional proteins to this nucleus, driven by protein-protein interactions. In the second case, the RNA is saturated by protein (because of strong protein-RNA interactions) before a critical nucleus emerges and grows by the spread of order amongst the already bound proteins. Molecular dynamics simulations predict that the assembly mechanism can be tuned by solution conditions that alter protein-protein and protein-genome interactions[47]. The recruitment mechanism is favored when protein-protein interactions are strong relative to protein-genome interactions, while the rearrangement mechanism is favored when the opposite situation prevails.

We choose to focus on the example of BMV self-assembly as it is a well-characterized spherical virus that can be *in vitro* reconstituted – and yet its assembly mechanism is not yet understood. Early work by several groups established that spontaneous self-assembly of virus-like particles (VLPs) can be carried out *in vitro* from purified components[6,26,59]. BMV (and the closely related CCMV) CP can also encapsulate non-viral RNA *in vitro* into robust, monodisperse, RNase-resistant capsids, so long as the RNA length is between about 2,500 and

4,500 nts[6]. The *in vitro* self-assembly protocol our lab has traditionally employed involves a two-step process in which RNA and CP are initially mixed in a pH 7.2, I = 1.0 M solution (ionic strength) and (i) dialyzed against a pH 7.2, I = 0.1 M solution, then (ii) dialyzed against a pH 4.75, I = 0.1 M solution. The first dialysis into a lower ionic strength solution allows for non-specific interactions between the negatively charged phosphate backbone of the ssRNA and the positively charged arginine-rich motif at the N-terminal tail of the CP, resulting in saturated adsorption of the RNA by the CP. The second dialysis into a lower-pH solution, closer to the isoelectric point of the proteins, allows the lateral interactions between CPs (that are already adsorbed onto the RNA) to become more strongly attractive (dominated by a combination of hydrophobic, hydrogen-bonding, and divalent-ion-mediated forces). This two-step protocol necessarily enforces the rearrangement assembly mechanism by controlling which interactions are favored by solution conditions (ionic strength and pH).

While our method consistently results in robust, monodisperse, RNase-resistant capsids, there are several other assembly conditions under which BMV will assemble *in vitro*. For example, Heibert et al. in 1968 assembled BMV VLPs *in vitro* by dialyzing against a pH 7.4, I = 0.01 M solution[5]. Additionally, Cuillel et al.[13] assembled BMV by dialyzing a solution initially at pH 7.8, I = 1.6 M into various lower pH and ionic strength solutions. They observed successful assembly products by neutron scattering when assembling in buffers below pH 6.5 and ionic strengths below 0.8 M[13]. BMV's ability to self-assemble under various conditions leads us to believe it is the ideal system to study how assembly mechanisms are tuned by solution conditions.

The present work investigates the role of pH and ionic strength on the stability of wild-type (WT) BMV particles and the assembly of well-formed BMV VLPs. Differential scanning

fluorimetry (DSF) is used to assess the strength of the CP-CP interactions for WT BMV in range of pH and ionic strength buffer conditions. We use these same buffers to conduct one-step (in a single buffer) *in vitro* self-assembly experiments and assess the resulting BMV VLPs using negative-stain electron microscopy. We find that protein-protein interactions are affected strongly by pH and weakly by ionic strength and that well-formed VLPs can be made in a variety of one-step solution conditions.

3.3 Materials and Methods

3.3.1 Buffers used

Disassembly buffer: 50 mmol/L Tris-HCl, pH 7.5; 500 mmol/L $CaCl_2$, 1 mmol/L ethylenediamine tetraacetic acid (EDTA), 1 mmol/L dithiothreitol (DTT), 0.5 mmol/L phenylmethylsulfonyl fluoride (PMSF). **Protein storage buffer:** 20 mmol/L Tris-HCl, pH 7.2; 1 mol/L NaCl; 1 mmol/L EDTA; 1 mmol/L DTT; 1 mmol/L PMSF. **Assembly buffers:** 42 mmol/L 2-(N-morpholino)ethanesulfonic acid (MES), pH 6; 0-1 mol/L NaCl; 8.4 mmol/L $MgCl_2$; and 3 mmol/L acetic acid. We also prepared assembly buffer at pH 6.5 and at pH 7 and 7.5 by replacing MES with sodium phosphate.

3.3.2 Synthesis of BMV RNA1

BMV RNA1 was made by in vitro transcription of the DNA plasmid pT7B1, linearized with BamHI (New England Biolabs, USA), with a T7 polymerase transcription system (Thermo Fisher, USA) and purified with an RNEasy Mini Kit (Qiagen, DEU), both following the manufacturers' specifications.

3.3.3 WT BMV and BMV capsid protein purification

WT BMV was purified from infected barley leaves (Hordeum vulgare)[3], and coat protein was purified as described previously[39]. Nucleocapsids were disassembled by dialyzing against

disassembly buffer at 4 °C overnight. The RNA was pelleted and the coat protein isolated by ultracentrifugation at 90,000 rotations per minute for 100 min at 4 °C in a Beckman TLA110 rotor. Coat protein was extracted from the supernatant and immediately dialyzed against protein storage buffer. Protein concentration and purity were assessed by UV-Vis spectrophotometry; only protein solutions with 260/280 ratios less than 0.6 were used for assembly. Protein was frozen in liquid nitrogen and stored at −80 °C until ready to use, at which point it was defrosted on ice and stored at 4 °C for up to two weeks.

3.3.4 Measuring the apparent melting curve of WT BMV by differential scanning fluorimetry (DSF)

WT BMV was dialyzed against assembly buffer at pH 6, 6.5, 7, or 7.5 overnight at 4 °C. Aliquots of WT BMV at a final concentration of 0.2 mg/mL, 2.5x SYPRO Orange fluorescent dye (Molecular Probes, USA), and a final salt concentration of 0-1 mol/L NaCl were prepared. DSF was performed in triplicate in a 96-well plate CFX Connect quantitative PCR machine (Bio-Rad, USA). All samples were heated from 25 °C to 95 °C, in 1 °C increments with a 1-min stabilization period at each temperature before the sample was measured. Excitation/emission wavelengths of 470/550 nm were used to detect the fluorescence emission of SYPRO orange binding to hydrophobic regions of coat protein exposed upon capsid disassembly. The apparent melting temperature is defined as the temperature at which the derivative of the fluorescence emission signal ($-dI/dT$) is maximal.

3.3.6 Negative-stain electron microscopy

Negative-stain electron microscopy was used to image the protein structures that assemble around RNA in solution. The assembly reaction was carried out in assembly buffer by mixing 860 nmol/L BMV CP with 7.5 nmol/L BMV RNA1 and incubating at room temperature for 10

min. 6 µL of assembly reaction was deposited on glow-discharged carbon-coated copper (200-mesh) PELCO Pinpointer grids (Ted Pella, USA). After 1 min, the grids were blotted with Whatman filter paper, and then stained with 6 µL of 2% methylamine tungstate (pH 6.4-6.8) for 1 min followed by complete stain removal and storage in a desiccator overnight. Micrographs were acquired using a Tecnai G2 TF20 High-Resolution electron microscope (FEI, USA) with an accelerating voltage of 200 kV. Images were collected at 3 µm to 4 µm underfocus with a TIETZ F415MP 16-megapixel CCD camera (4000 by 4000 pixels, pixel size 15 µm). Particles were measured manually from negative-stain electron micrographs using ImageJ[62].

3.4 Results

3.4.1 WT BMV particle stability affected by pH but less so by ionic strength

The apparent melting temperature was determined from differential scanning fluromietry measurements (**Figure 8**). The apparent melting temperatures for WT BMV particles in solutions that range in pH from 6.0 to 7.5 and NaCl concentration from 0 to 1.0 M are shown in **Figure 9**. In pH 6.0 solutions, WT BMV particles are stable to around 75°C for ionic strengths 0 to 0.17 M, then stability drops to around 70°C in ionic strengths above 0.5 M, ending with a stability around 65°C in 1.0 M. For WT BMV virions in pH 6.5 solutions, the thermal stability of the particles varied more widely with ionic strength starting below 70°C and decreasing to below 45°C at higher ionic strengths. WT BMV particles in pH 7.0 and 7.5 solutions demonstrated similar behavior across ionic strengths with thermal stability starting around 48°C and dropping to 42°C in higher ionic strength solutions.

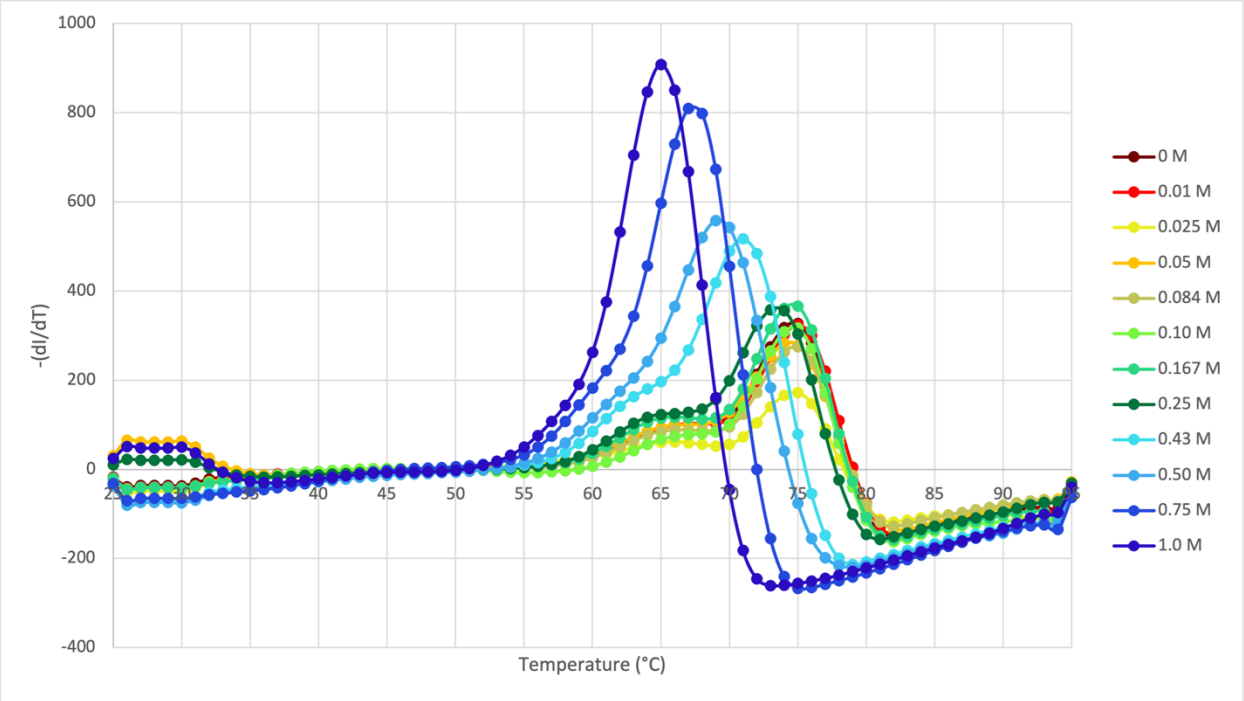

Figure 8 Plot of differential scanning fluorimetry data (-dI/dT) versus temperature. Differential scanning fluorimetry data for WT BMV at pH 6 and a range of ionic strengths from 0 to 1.0 M. The apparent melting temperature is defined as the temperature at which the derivative of the fluorescence emission signal (-dI/dT) is maximal.

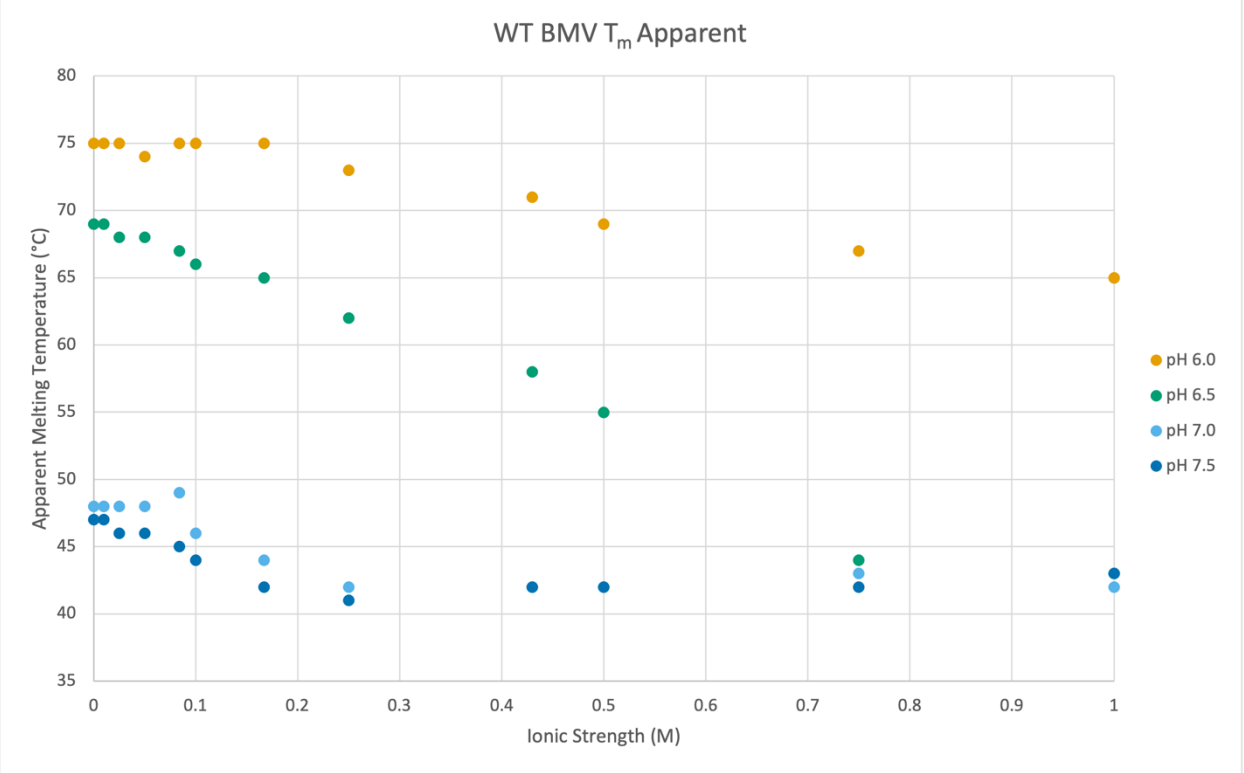

Figure 9. Apparent melting temperature of WT BMV as a function of pH and ionic strength. The apparent melting temperature for WT BMV in solutions of pH 6.0 (yellow), 6.5 (green), 7.0 (light blue), and 7.5 (dark blue) and 0, 0.01, 0.03, 0.05, 0.08, 0.1, 0.17, 0.25, 0.43, 0.50, 0.75, and 1.0 M NaCl concentrations. At pH 6.0, the particles show the highest stability (highest apparent melting temperature) over the range of ionic strengths ranging from 65-75°C. The particles have a lower apparent melting temperature at pH 6.5 and still lower at pH 7.0 and 7.5.

3.4.2 BMV virus-like particles assembled in one-step

BMV RNA1 and BMV CP were combined in an assembly solution at a specific pH and ionic strength, incubated at room temperature for 10 minutes, and then imaged using negative-stain electron microscopy. Representative electron micrographs for assemblies performed at a pH of 5.0, 6.0, or 7.5 and ionic strengths of 84, 167, 250, or 430 mM NaCl are shown in **Figure 10**. Many aggregated particles or partially assembled particles are visible, especially in the pH 5.0 assemblies. pH 6 assemblies in ionic strengths below 250 mM NaCl and the two pH 7.5 assemblies show several well-formed VLPs.

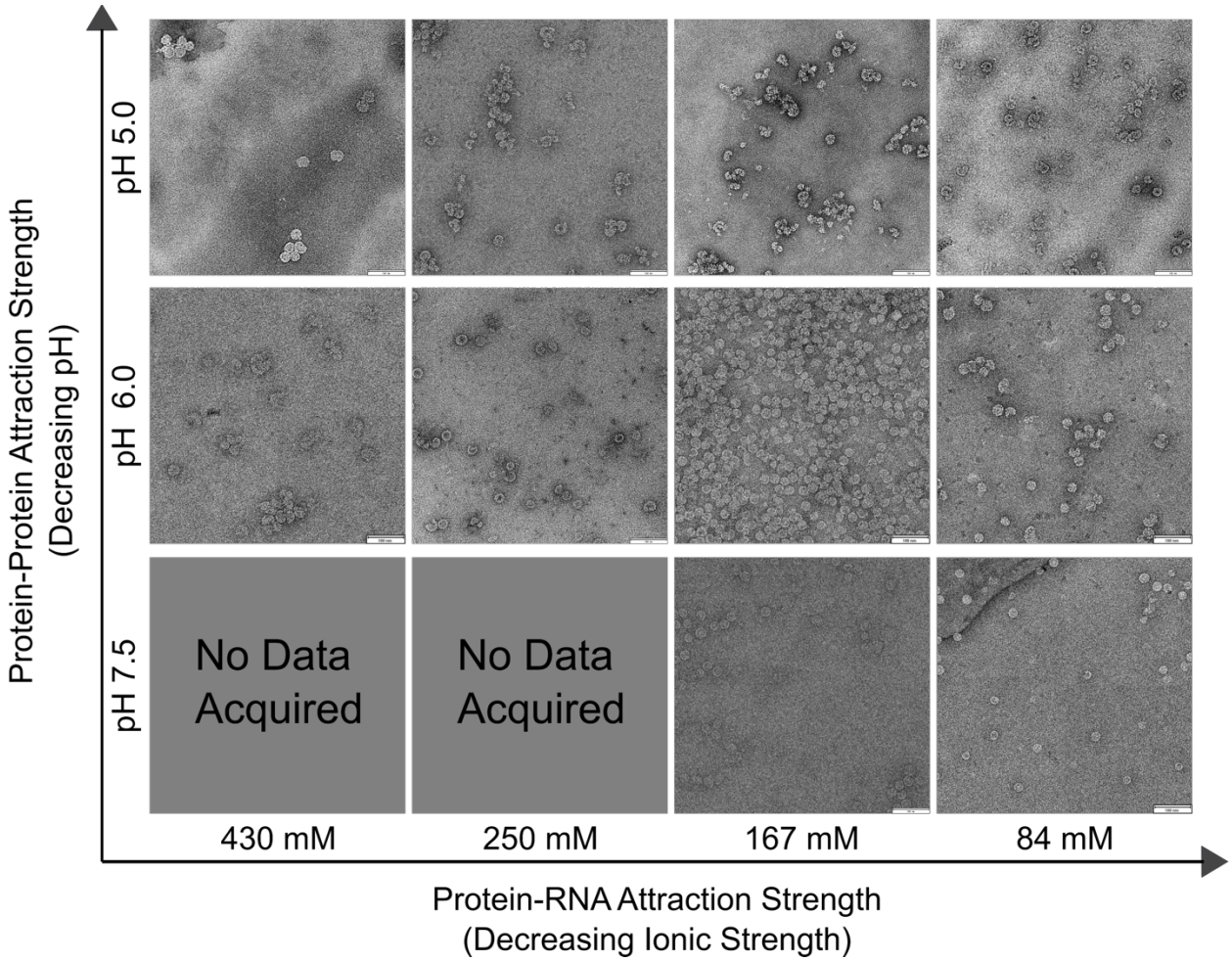

Figure 10. Negative-stain electron micrographs of virus-like particles assembled in buffers of pH 5.0, 6.0, or 7.5 and ionic strengths of 84, 167, 250, and 430 mM. Negative-stain electron micrographs are organized by increasing protein-RNA attraction strength (decreasing ionic strength) on the x-axis and increasing protein-protein attraction strength (decreasing pH) on the y-axis. Assemblies were not performed for assembly conditions pH 7.5, I = 430 mM and pH 7.5, I = 250 mM. Scale bar is 100 nm.

The diameter of particles from electron micrographs of assemblies performed in pH 6.0 solutions were measured and histograms of the particle size distributions are shown in **Figure 11**. For pH 6.0 assemblies, the average particle diameter measured at 84, 167, 250, and 430 mM NaCl is 28.66 ± 3.78 nm, 26.01 ± 2.27 nm, 26.55 ± 3.96 nm, and 29.98 ± 4.06 nm, respectively. There were fewer well-formed particles to measure in the higher ionic strength assemblies, with as few as 100 measured for the 250 mM NaCl sample.

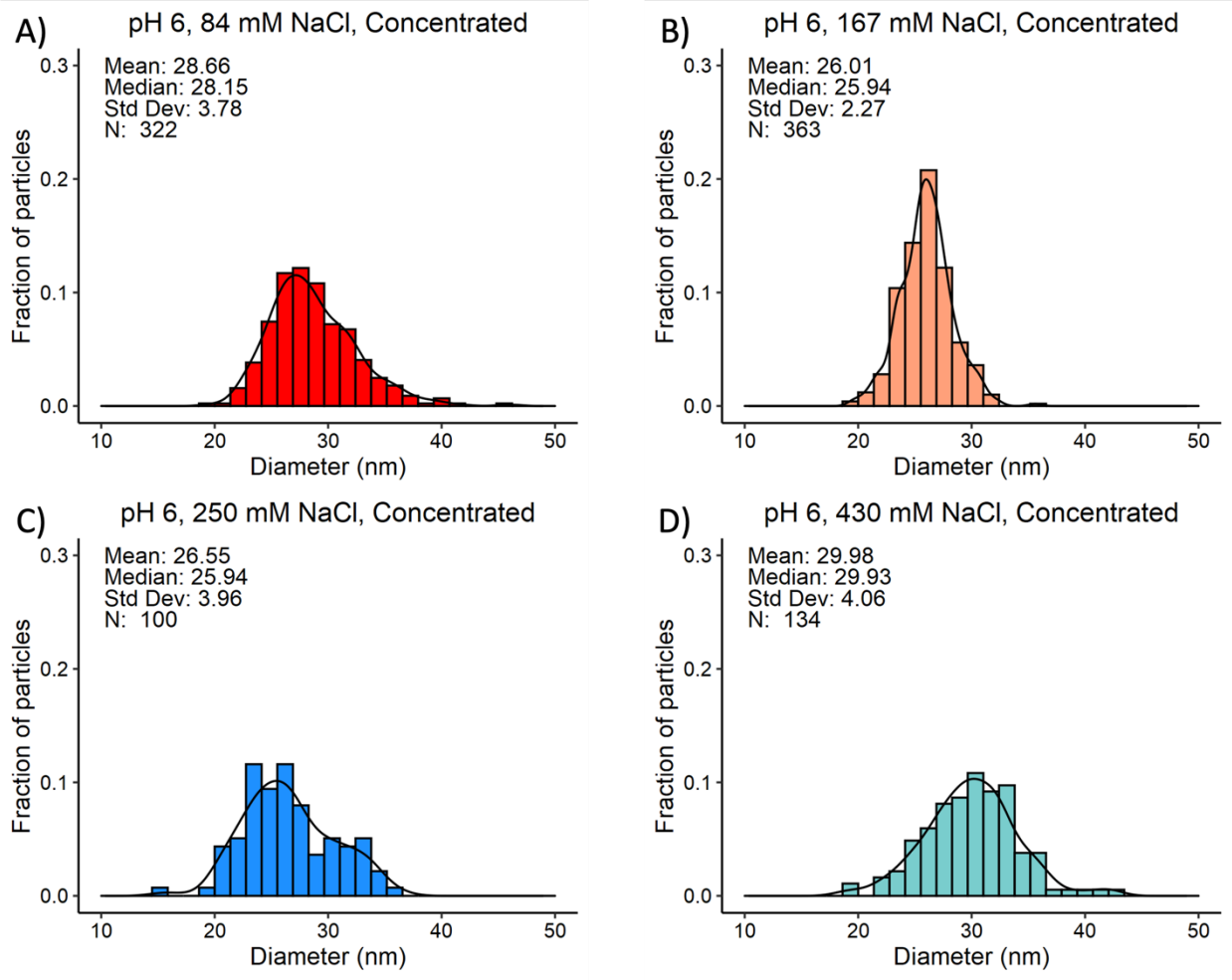

Figure 11. Histograms of the size distribution of BMV virus-like particles assembled at pH 6 and 84, 167, 250, or 430 mM NaCl. Histograms of the geometric mean of orthogonal measurements of the diameter of the particles are shown as diameter versus fraction of particles. The mean, median, standard deviation, and number of particles (N) are shown for each data set.

3.5 Discussion

BMV capsids are held together by lateral attractions between CP dimers. The lateral attractions between CP are pH-dependent due to electrostatic repulsions generated by deprotonation of adjacent acidic residues on amino acids at neutral pH[63]. The 30°C decrease in thermal stability of WT BMV at elevated pH values (75°C in pH 6.0, 45°C in pH 7.5) illustrates a decrease in effective lateral CP-CP attractions. The thermal stability of the virions varied less across ionic strengths, decreasing about 10°C from 0 M NaCl to 1.0 M NaCl for pH 6.0, 7.0, and

7.5. For pH 6.5, the thermal stability of the particles decreased significantly (25°C) across ionic strength solutions.

Well-formed VLPs were observed in assemblies performed at pH 6 and pH 7.5 and ionic strengths of 84 mM and 167 mM NaCl. Aggregation of particles was observed in assemblies performed in pH 5.0 when CP-CP attractions are strong. The most well-formed VLPs were observed in assemblies at pH 6.0, 167 mM NaCl, suggesting weaker CP-CP and CP-RNA interactions favor formation of well-formed particles.

3.6 Conclusions and Future Directions

These experiments examined the role of pH and ionic strength on the assembly of BMV VLPs and suggest that well-formed VLPs can form in under 10 minutes at pH 6.0, 167 mM NaCl. There are likely other conditions of moderate CP-CP and CP-RNA attraction that will produce well-formed, monodisperse VLPs. While we begin to see aggregation of particles at low pH when CP-CP attractions are high, there are notable well-formed particles when the ionic strength is sufficiently high (430 mM NaCl). A more systematic study and robust analysis of VLPs formed in one-step assembly processes could aid the understanding of RNA virus assembly mechanisms.

CHAPTER 4

Non-monotonic dose dependence of protein expression in cells transfected with self-amplifying RNA

4.1 Abstract

Self-amplifying(sa) RNA molecules – "replicons" – derived from the genomes of positive-sense RNA viruses are receiving increasing attention as gene and vaccine delivery vehicles. This is because mRNA forms of genes of interest can be incorporated into them and strongly amplified, thereby enhancing target protein expression. In this report, we demonstrate a non-monotonic dependence of protein expression on the mass of transfected replicon, in contrast to the usual, monotonic case of non-sa RNA transfections. We lipotransfected a variety of cell lines with increasing masses of enhanced yellow fluorescent protein (eYFP) as a reporter gene in sa form and found that there is a "sweet spot" at which protein expression and cell viability are optimum. To control the varying mass of transfected replicon RNA for a given mass of lipofectamine, the replicons were mixed with a "carrier" RNA that is neither replicated nor translated; the total mass of transfected RNA was kept constant while increasing the fraction of the replicon from zero to one. Fluorescence microscopy studies showed that the optimum protein expression and cell viability are achieved for replicon fractions as small as one tenth of the total transfected RNA, and these results were quantified by a systematic series of flow cytometry measurements.

4.2 Introduction

RNA-based therapeutics and technologies are of great interest due to their versatility and, upon delivery to mammalian cells, their ability to be directly translated in the cytosol into

proteins of interest[64,65]. However, the resulting gene expression from conventional mRNAs is often quite low, as RNAs are quickly degraded by cellular ribonucleases (RNases) and therefore exist only transiently in targeted cells[66]. To improve upon the duration and yield of protein synthesis from mRNAs, Biddlecome, et al.[67] have recently exploited a self-amplifying RNA (saRNA) molecule – a replicon – to deliver therapeutic genes of interest, taking advantage of RNA replicons constructed[68] from the Nodamura virus (NOV).

NOV is a positive-sense single-stranded RNA (+ssRNA) insect virus whose segmented genome consists of two RNA molecules: RNA1 (3.2 knt in length), which codes for the RNA-dependent RNA polymerase (RdRp), and RNA2 (1.3 knt), which codes for the viral capsid protein[19]. Replication of NOV proceeds through the initial translation of the RdRp. This polymerase then synthesizes the reverse complement, or negative-sense, genome from which more copies of the positive-sense genome are transcribed[69]. Inserting a gene of interest (GOI) at the end of the RdRp open reading frame in RNA1, and separating it from the RdRp gene by a T2A ribosome-shifting sequence, results in a therapeutic replicon – an RNA that is self-amplifying (sa) by virtue of it being translated directly into a gene product, the RdRp, that binds and replicates it[68], including the GOI. NOV and the related flock house virus have been studied extensively and are noted for their ability to replicate to high levels in mammalian cells, making the NOV replicon an attractive therapeutic candidate for a host of human diseases[70–72]. Replicons made from mammalian +ssRNA viruses, in particular the alphavirus species Sindbis virus, Venezuelan equine encephalitis virus, and Semliki forest virus (SFV), have been more widely developed for vaccine purposes[73], but we have pursued nodavirus-derived examples because they are short enough to be *in vitro* self-assembled into spherical virus-like particles using the capsid protein from bromoviruses[67].

Because the replicons made from RdRp genes of +ssRNA viruses do not encode for any structural proteins normally found in a viral genome, e.g., the capsid protein encoded by NOV RNA2, they are classified as non-infectious RNA molecules and are not true viral genomes. However, their mechanisms of replication are very similar to those of actual viruses, so upon delivery to cells they can induce a "pseudoinfection" that is less toxic to cells while still displaying many characteristics of natural viral infections[74]. Although NOV replicons are shown to have limited cytotoxicity in most cell types, their self-replicating nature nevertheless raises concerns of therapeutic overload if delivered at the traditionally high doses of RNA normally employed in transfection[75].

Cationic-lipid-based transfection is one of the most commonly used ways to deliver nucleic acids to adherent cell lines and is often the first step in testing an mRNA-based therapeutic like a replicon[76]. Lipid-based transfection agents contain a mix of cationic and neutral lipids that self-assemble into "lipoplexes" – complexes of lipid and mRNA whose positive surface charge interacts with the anionic plasma membrane of cultured cells, facilitating delivery of the RNA into the cytoplasm. There, mRNA is translated by the ribosomal machinery and, in the case of replicons, subsequently undergoes the self-amplification cycle in which one replicon molecule generates thousands of copies of itself (with concomitant strong expression of its gene product – RNA replicase – along with those of any added genes of interest).

In typical transfection experiments, each lipoplex contains several hundreds of copies of the RNA to be delivered[77]. For transient mRNAs, this delivery influx is necessary to overcome inevitable RNA degradation, which leads to moderate gene expression. However, when hundreds of saRNAs are delivered in this way, high levels of RNA replication (and ensuing translation of the saRNA) can be detrimental, even cytotoxic, for the cell. In a natural viral infection, after all,

a single copy of the viral genome in the cytoplasm is sufficient to induce cytotoxicity on a cellular level; by this logic, much lower dosages of replicon therapeutics should be required – and desired – for effective expression of the GOI in cell culture. In accord with this, efforts to prepare replicon-persistent cells have been successful to the extent that RdRp mutations of the +ssRNA-viral genome and to the corresponding replicon result in lower levels of RNA replication and hence less cytotoxicity[78,79].

In order to control the numbers of RNA molecules delivered to culture cells, it is not sufficient to simply change the mass of RNA in the transfection mix. This is because the lipoplexes change their size and structure for different absolute and/or relative amounts of RNA and/or lipid[77]. Accordingly, instead of transfecting only the RNA of interest, we mix it with a second RNA and keep fixed the total masses (hence charges) of RNA and of lipid components: only the relative masses of the "active" RNA and the "carrier" RNA are varied. More explicitly, the active RNA is an saRNA containing a reporter mRNA gene along with the replicase mRNA gene whose protein product strongly amplifies the saRNA; the carrier RNA is an mRNA that is weakly translated because it is uncapped[69] and its gene product has virtually no effect on the host cell. By keeping constant the total mass of RNA (and of transfection lipids), we ensure that the lipoplexes taken up by the cells remain constant in their structure and transfection properties, with their composition of "active" and "carrier" molecules equal to that prepared in the transfection mix. In this way, because the total mass of RNA in the lipoplexes and the total number of lipoplexes are constant, the number of replicons delivered to the cells is simply proportional to the number in the transfection mix, which we control.

The aim of this study is to determine the dependence of gene-of-interest (GOI) protein expression on numbers of GOI replicons transfected into each of different cell types, taking

advantage of the above-described way of controlling the relative numbers of transfected replicons. More explicitly, we used – as our replicon – RNA1 of Nodamura (NOV), which encodes the RdRp that replicates it in mammalian as well as insect cells. An eYFP reporter gene was added to RNA1 so that it too is replicated as soon as the RdRp gene is translated to give RNA replicase **(Figure 12).** As the carrier RNA we used an *in vitro* transcribed, uncapped, form of RNA1 of the plant virus, brome mosaic virus (BMV), which encodes an RdRp that is not functional in mammalian cells, referred to as B1. By varying the relative masses of the two RNAs, for a fixed total mass of RNA (and transfection lipid), the changing number of transfected replicons was controlled. Increasing numbers of the reporter-gene replicons were transfected in this way into different adherent cell lines, which were then monitored for effective eYFP production by fluorescence microscopy and by flow cytometry, and assayed for cell cytotoxicity and death by Trypan-Blue-exclusion staining.

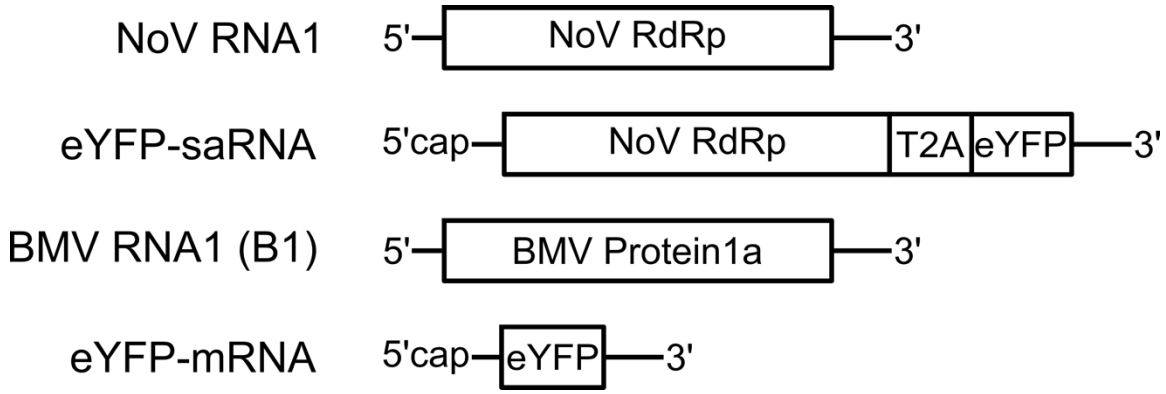

Figure 12: Schematic of RNA constructs. A replicon (eYFP-saRNA) was constructed by adding an enhanced yellow fluorescent protein (EYFP) gene to the end of the open reading frame of Nodamura virus RNA1 (NOV RNA1) encoding for an RNA-dependent RNA polymerase (RdRp) and separating the genes by a T2A ribosome-shifting sequence. 5'cap indicates the RNA was capped during *in vitro* transcription. BMV RNA1 (B1) was used as a "carrier" RNA during transfections. eYFP-mRNA encodes for eYFP.

4.3 Materials and Methods

4.3.1 Synthesis of NOV eYFP replicon RNA (eYFP-saRNA)

The plasmid encoding the NOV eYFP replicon, pNoda-B1-FPG, described in detail by Gitlin et al.[68] was constructed and generously provided by Dr. Gitlin, for which we are grateful. After plasmid linearization with enzyme XbaI (New England BioLabs), the RNA was transcribed *in vitro* using a T7 polymerase mMessage mMachine Kit (Thermo Fisher) and purified with an RNEasy Mini Kit (Qiagen), both following the manufacturer's specifications. The mMessage Kit was used to produce capped eYFP-saRNA – and eYFP mRNA, see Synthesis of eYFP mRNA below – because we want them to be translated in transfected cells.

4.3.2 Synthesis of BMV RNA1 (B1)

Uncapped BMV RNA1 (B1) was made by *in vitro* transcription of the DNA plasmid pT7B1, described in detail in Cadena-Nava, et al.[6], linearized with BamHI (New England BioLabs) with a T7 polymerase MegaScript Kit (Thermo Fisher) and purified with an RNEasy Mini Kit (Qiagen) according to the manufacturer's specifications. The higher-yield but non-capping MegaScript Kit was chosen because the B1 RNA should not be translated after transfection into cells[69].

4.3.3 Synthesis of eYFP mRNA

Capped eYFP mRNA was made by *in vitro* transcription of the DNA plasmid pSP64eYFP, linearized with EcoRI (New England BioLabs), with a SP6 polymerase mMessage mMachine Kit (Thermo Fisher) and purified with an RNEasy Mini Kit (Qiagen), both following manufacturers' specifications. All plasmids are available upon request.

4.3.4 Transfection of BHK-21, HEK-293, and CHO-K1 adherent cells

BHK-21, HEK-293, and CHO-K1 cells (all from ATCC) were each split 24 hours prior to transfections and grown to 70-90% confluence into 6-well or 24-well plates in 2 mL or 500 µL Dulbecco's Modified Eagle Medium (DMEM, Thermo Fisher) for BHK-21 and HEK-293 cells, and F12K medium (Thermo Fisher) for CHO-K1 with 10% Fetal Bovine Serum (FBS, Thermo Fisher) and 1% penicillin/streptomycin (Thermo Fisher). Cells were transfected with 2.5 µg (6-well) or 0.5 µg (24-well) of total RNA per well using, respectively, 10 or 2 µL Lipofectamine 2000 (Thermo Fisher) according to manufacturer's specification. Negative control experiments included: mock-transfected cells that were treated with the corresponding amount of lipofectamine only, B1-transfected cells that were treated with the full amount (2.5 µg or 0.5 µg) of B1 RNA in lipofectamine, and control cells were not treated with any lipofectamine or RNA (untransfected). The RNA transfection amounts are shown in **Table 1** and vary both in amounts of eYFP self-amplifying RNA (eYFP-saRNA) and non-sa/uncapped B1 RNA, with the total RNA mass staying constant. Lipofectamine reagent was diluted 1:10 in OptiMEM medium (Thermo Fisher) with the total volume determined by the number of wells being transfected. The RNA masses were mixed together as described in Table 1 and diluted to the same volume of diluted lipofectamine with OptiMEM. The diluted RNA and diluted lipofectamine were mixed in a 1:1 ratio and incubated at room temperature for 5 minutes. The cell media was aspirated, the transfection mixture was overlaid onto the confluent cells, and the cells were incubated for 8 hours at 37°C in 5% CO_2. The transfection medium was removed after 8 hours and the cells were fed with 2 mL or 500 µL of DMEM (for BHK-21 and HEK-293 cells, F12K for CHO-K1 cells) with FBS and antibiotics; this medium was left for the required time period before being

analyzed using fluorescence microscopy, flow cytometry, and trypan-blue-exclusion assays.

Each transfection was performed in biological triplicate.

Table 1. Amounts of eYFP replicon, B1, and total RNA included in each treatment of the transfection experiment using 24-well plates. All values were increased 5-fold in transfection experiments using 6-well plates.

Transfection Condition	eYFP replicon mass (ng)	B1 RNA mass (ng)	Total RNA mass (ng)	Moles of replicon ($x10^{-15}$)	Moles of B1 RNA ($x10^{-15}$)	Total moles of RNA ($x10^{-15}$)	Moles B1 per mole of replicon	Volume of Lipofectamine 2000 Reagent per well (μL)
Mock	0	0	0	0	0	0		0
Lipofectamine-only	0	0	0	0	0	0		2
B1	0	500	500	0	469	469		2
5 ng eYFP-saRNA	5	495	500	4	464	468	116	2
50 ng eYFP-saRNA	50	450	500	37	422	459	11	2
250 ng eYFP-saRNA	250	250	500	185	234	419	1.3	2
500 ng eYFP-saRNA	500	0	500	370	0	370	0	2

4.3.5 Flow cytometry analysis

The BHK-21 and CHO-K1 cells were harvested at 8, 16, 24 and 48-hours post-transfection. After aspirating the cell medium, the cells were trypsinized by incubation at 37°C for 3-5 minutes in 250 µL 0.25% Trypsin-EDTA (Thermo Fisher). After quenching the trypsin with 250 µL DMEM (+FBS +antibiotic), the cells were collected by centrifugation (5 minutes, 200 xg), washed twice with 500 µL of Dulbecco's phosphate-buffered saline (DPBS, Thermo Fisher), and then fixed by resuspending in 500 µL 1% paraformaldehyde (VWR). Cells were analyzed using the MACSQuant® Analyzer 10 Flow Cytometer (Miltenyi Biotec) to collect forward and side scatter profiles for live/dead cell analysis and fluorescence intensity distributions for eYFP detection (488-nm excitation). The resulting data were analyzed with FlowJo software and subsequent visualization in RStudio using the ggplot2 function. Cells were gated for viability based on forward and side scatter, and then a manual gate was imposed on the EYFP fluorescence using the machine's GFP laser settings to determine the percentage of fluorescent cells as compared to untransfected controls. A total of 24,000 events was recorded for each sample of transfected cells. Statistical significance between all combinations of RNA treatments was determined by a two-way ANOVA with repeated measures followed by a Tukey's Honest Significant Difference using the TukeyHSD function in R (https://www.rdocumentation.org/packages/stats/versions/3.6.2/topics/TukeyHSD) to determine the significance of differences between treatments, which was determined at a p-value of 0.001.

4.3.6 Fluorescence microscopy of transfected cells

Live cells were examined by fluorescence microscopy at 8, 16, 24, 36, 48 and 72-hours post-transfection using an Amscope IN300TC-FL microscope at 20x magnification. Brightfield images were taken using a 5MP CCD camera and processed using the ISCapture software with

an exposure setting of 200 milliseconds. Fluorescence images were taken using 450-490 nm excitation and 520 nm emission filters with a 2.2 second exposure time using the same camera and software for processing.

4.3.7 Trypan-blue-exclusion assays

Transfected cells were examined at 8, 16, 24, 48 and 72-hours post-transfection. The medium was aspirated from one well of each transfection condition, the cells were trypsinized by incubating at 37°C for 3-5 minutes in 250 µL of 0.25% Trypsin-EDTA. After incubation, the trypsin was quenched with 250 µL DMEM (+FBS +antibiotic) as described above in Section 2.5. After thorough mixing, 10 µL of the cells was added to 10 µL of 0.4% Trypan-Blue stain (Gibco), and 10 µL of the resulting mixture was added to a hemocytometer and examined under an Amscope IN300TC-FL microscope with a 20x objective. The cells were counted by eye and the percentage of live cells calculated, assuming that live cells have excluded the stain completely and dead cells have stained blue completely. Statistical significance between all combinations of RNA treatments was determined by a two-way ANOVA with repeated measures followed by a Tukey's Honest Significant Difference test using a p-value cutoff of 0.001. In fluorescence microscopy studies, HEK-293 cells showed fluorescence expression profiles similar to that of BHK-21 cells (data not shown), but due to difficulty keeping the cells attached to plates throughout the entire transfection, the HEK-293 cells were excluded from further analyses.

4.4 Results

4.4.1 eYFP expression in BHK-21 cells peaked at small fractions of transfection

BHK-21 cells were transfected with 5, 50, 250, or 500 ng of eYFP-saRNA in 24-well plates, with 500 ng representing a typical RNA amount used in standard lipofectamine transfections. In each case (see **Table 1**) the replicon RNA was mixed with an uncapped mRNA

(RNA1 from brome mosaic virus, termed B1) so as to have a total of 500 ng of RNA in the transfection mix. The cells remained in culture for 72 hours post-transfection, with both fluorescence and brightfield images taken at 6 designated time points to assay fluorescent protein production and cell health.

Similarly, transfections were carried out with mixtures of non-sa eYFP mRNA and uncapped B1 RNA, with 5, 50, 250, and 500 ng of eYFP mRNA mixed with B1 RNA to a total of 500 ng, but a quantitative dose dependence could not be deduced, because at short times and low masses the fluorescence was too weak to be observed. All further discussion will refer to transfections with the *replicon* form of the eYFP mRNA (eYFP-saRNA), which showed a non-monotonic dependence of fluorescence on the mass of transfected replicon, as discussed below.

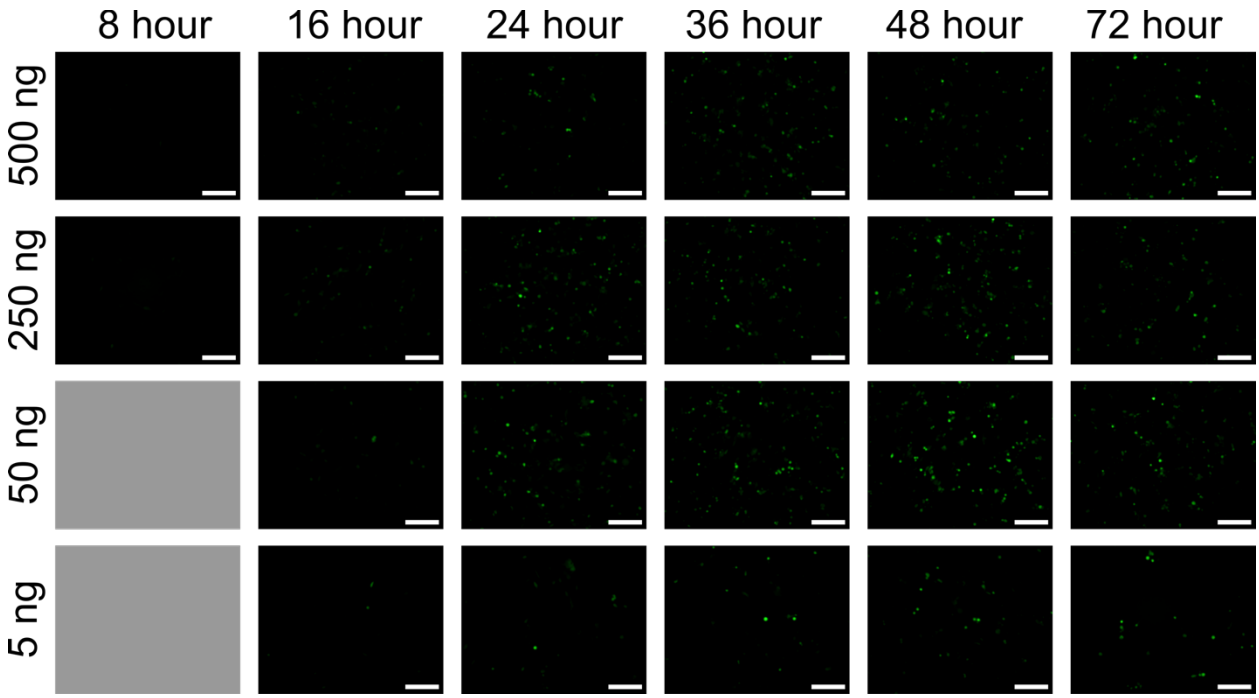

Figure 13: Representative fluorescence images collected at 8, 16, 24, 36, 48 and 72-hours post-transfection for BHK-21 cells transfected with 5, 50, 250, or 500 ng of eYFP-saRNA. Scale bar = 100 μm. Gray boxes indicate no image collected because no fluorescence was observed (specifically, at the 8-hour time points for 5 and 50 ng transfections).

The images in **Figure 13** showed that cell fluorescence peaked broadly at 36 to 48 hours post-transfection, and that the highest fluorescence was seen for transfections with 50 ng of eYFP-saRNA. Significantly, cells transfected with as much as 500 ng of eYFP-saRNA fluoresced less strongly than those treated with 250 or 50 ng, and showed fluorescence levels comparable to those in cells transfected with as little as 5 ng. Brightfield images acquired (data not shown) did not show significant changes in the number of cells per image. For more quantitative and reliable studies of protein expression levels and determination of their dose dependence, however, the flow cytometry assays are preferred over fluorescence imaging.

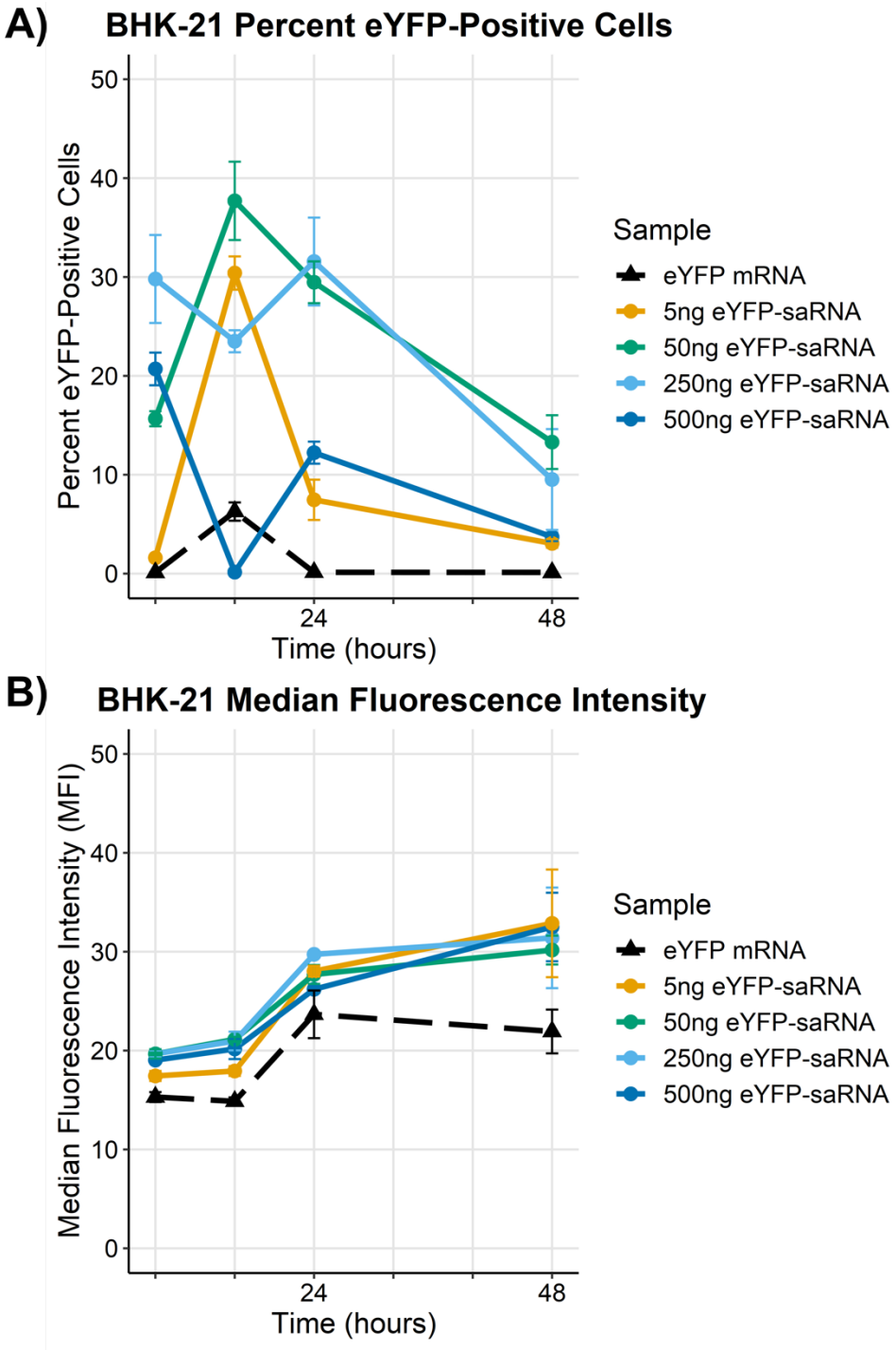

Figure 14: A. Flow cytometry data of %-eYFP-positive BHK-21 cells at 8, 16, 24 and 48-hours post-transfection with 5, 50, 250, or 500 ng of eYFP-saRNA or eYFP mRNA. **B.** Median fluorescence intensity of eYFP-positive cells at each time point for each RNA amount. All samples when compared to the negative control (eYFP mRNA) had a p-value < 0.001.

Flow cytometry analysis (**Figure 14A**) confirmed the fluorescence microscopy data, with the 50 ng saRNA transfection amount showing generally the highest percentage of eYFP-positive cells. Additionally, the percentage of eYFP-positive cells for samples transfected with 5 and 500 ng of eYFP-saRNA did not differ statistically (p-value = 0.66) but were both significantly different from the percentage of eYFP-positive cells seen in samples transfected with 50 and 250 ng of eYFP-saRNA (p-values ≤ 4.69E-13). Expression appeared to peak much earlier than in the microscopy time course (16 hours post-transfection vs. 36 hours), which is likely attributable to the higher sensitivity of detection of the flow cytometer versus that of the fluorescence microscope camera, as well as the very bright signal emitted by dying cells that only become visible later in the time course. The cells appeared to express eYFP very strongly starting early in the time course, but at levels not detectable by the fluorescence microscope, which also is in line with previous work[67] showing the peak of saRNA amplification at 16 hours post-transfection. All saRNA-transfected samples had significantly higher eYFP-positive cells than the eYFP-mRNA control (p-values ≤ 1.24E-08).

The cells transfected with 500 ng saRNA showed very low levels of fluorescence, comparable only to those transfected with 5 ng saRNA, for the majority of the time course. The drops in expression in the two highest (250 ng and 500 ng) saRNA transfection amounts that occur at 16 hours post-transfection are likely explained by the higher saRNA amounts causing more of a strain on cellular resources when the saRNA reaches its peak of amplification, limiting fluorescent protein expression at those times. Transfection efficiency appeared to average around 30% positive cells, which is expected with lipoplex-based transfections. **Figure 14B** shows the median fluorescence intensity (MFI) of each transfection. The MFI for each of the saRNA treatments were statistically different from the mRNA control (p-values ≤ 3.44E-06) but are not

statistically different from one another (p-values ≥ 4.86E-01). The saRNA transfection traces showed a gradual increase over time that reflects the brighter fluorescence of dead cells at later times. This signifies that using lower doses of saRNA does not negatively affect the amount of protein produced per cell.

4.4.2 BHK-21 cell viability correlates inversely with saRNA dosage

Trypan-blue-exclusion assays were performed on BHK-21 cells transfected with the same increasing (5, 50, 250, or 500 ng) amounts of eYFP-saRNA, again at 5 time points up to 72 hours post-transfection. Numbers of live and dead cells were counted to calculate the percentage of live cells that remained in each well (**Figure 15**) at the designated times.

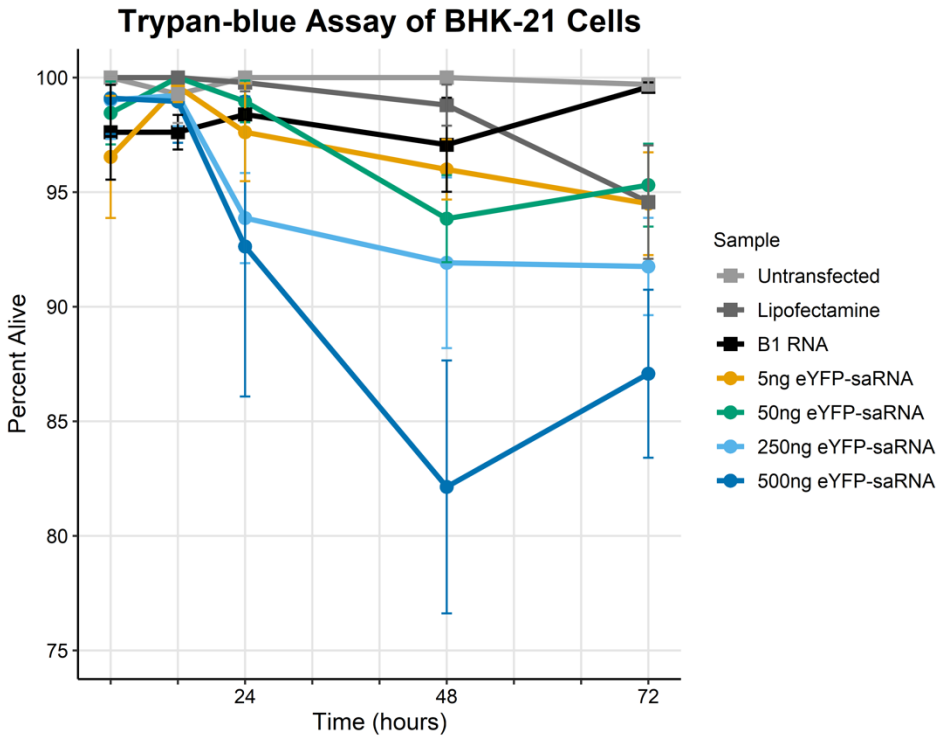

Figure 15: Trypan-blue-exclusion assays of BHK-21 cells collected at 8, 16, 24, 48 and 72-hours post-transfection with 5, 50, 250, or 500 ng of eYFP-saRNA or 500 ng B1 RNA only. Control samples include untransfected cells, cells transfected with lipofectamine reagent only (no RNA), and cells transfected with 500 ng of B1 RNA. Samples of 250 and 500 ng eYFP-saRNA when compared to the negative controls (untransfected, lipofectamine, and B1 RNA) had a p-value < 0.01.

As expected, the highest viabilities were found in cells transfected with the smallest amount of RNA: cells transfected with 5 ng eYFP-saRNA were close to 100% alive at early time points and saw only a slight drop to 95% by 72 hours post-transfection. Cells transfected with 50 and 250 ng eYFP-saRNA had very similar viability profiles, with both remaining above 90% alive even at 72 hours post-transfection, although the 250 ng-transfected cells saw a quicker drop in viability at 24 hours post-transfection. Most noticeably, cells treated with 500 ng eYFP-saRNA were significantly less viable at the eYFP expression peak than the other three treated wells (48 hours post-transfection), dropping to about 80% alive before recovering slightly to about 85% alive at 72 hours post-transfection. The control samples, untransfected, mock-transfected (treated with lipofectamine only), and B1-transfected cells remained over 95% viable for the duration of the experiment. Both the 250 and 500 ng-transfected cells showed viability profiles that differ from those of the controls (p-values $\leq$ 8.73E-04). Brightfield images of cells confirmed the trypan exclusion data, with lower-dosed cells appearing significantly healthier by the end of the 72-hour time course and those treated with 500 ng of eYFP-saRNA appearing clumped and distressed.

4.4.3 eYFP expression in CHO-K1 cells peaks at earlier time points and is observable for smaller amounts of saRNA delivery

CHO-K1 cells were transfected with the same amounts of active eYFP-saRNA and carrier B1 RNA as described above for BHK-21 cells. Cell fluorescence was overall stronger in this line, but also peaked much earlier in the time course (**Figure 16**), sharply at about 16 hours post-transfection as compared to the later, broader maximum seen in BHK-21 cells. Fluorescence was again strongest in the cells transfected with 50 ng of eYFP -saRNA, somewhat lower in those treated with 250 ng of saRNA, and still lower in those transfected with 500 ng of saRNA.

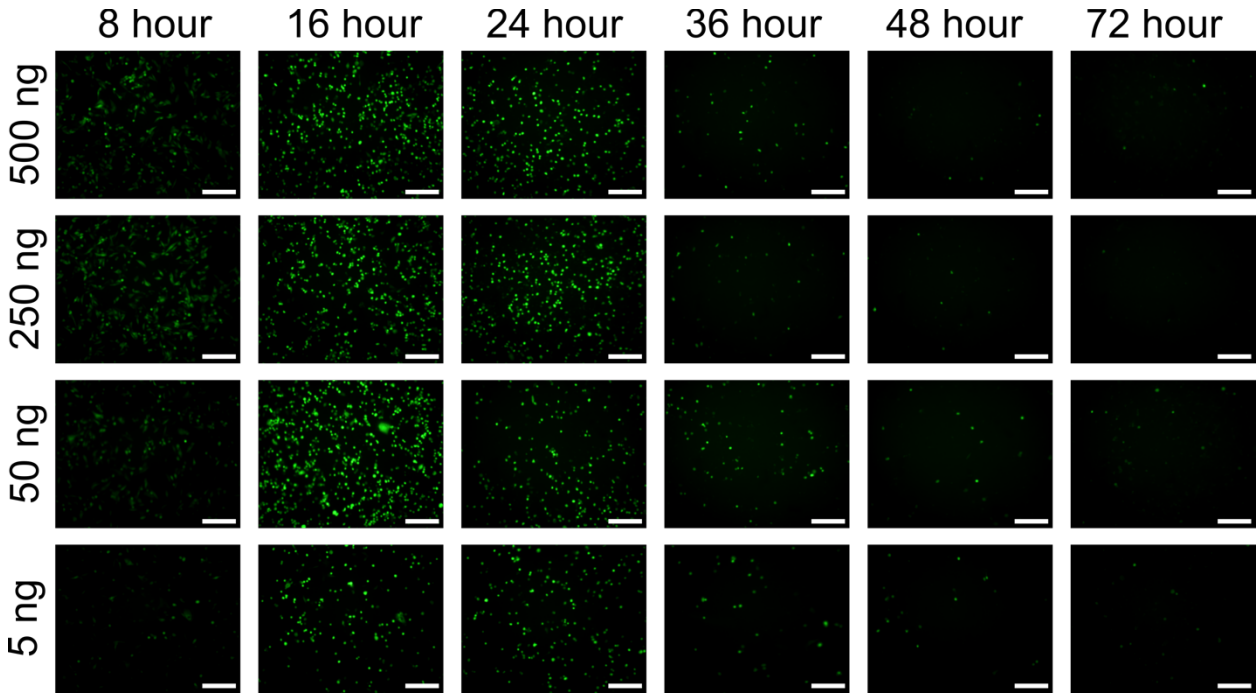

Figure 16: Representative fluorescence images collected at 8, 16, 24, 36, 48 and 72-hours post-transfection for CHO-K1 cells transfected with 5, 50, 250, or 500 ng of eYFP-saRNA. Scale bar = 100 μm.

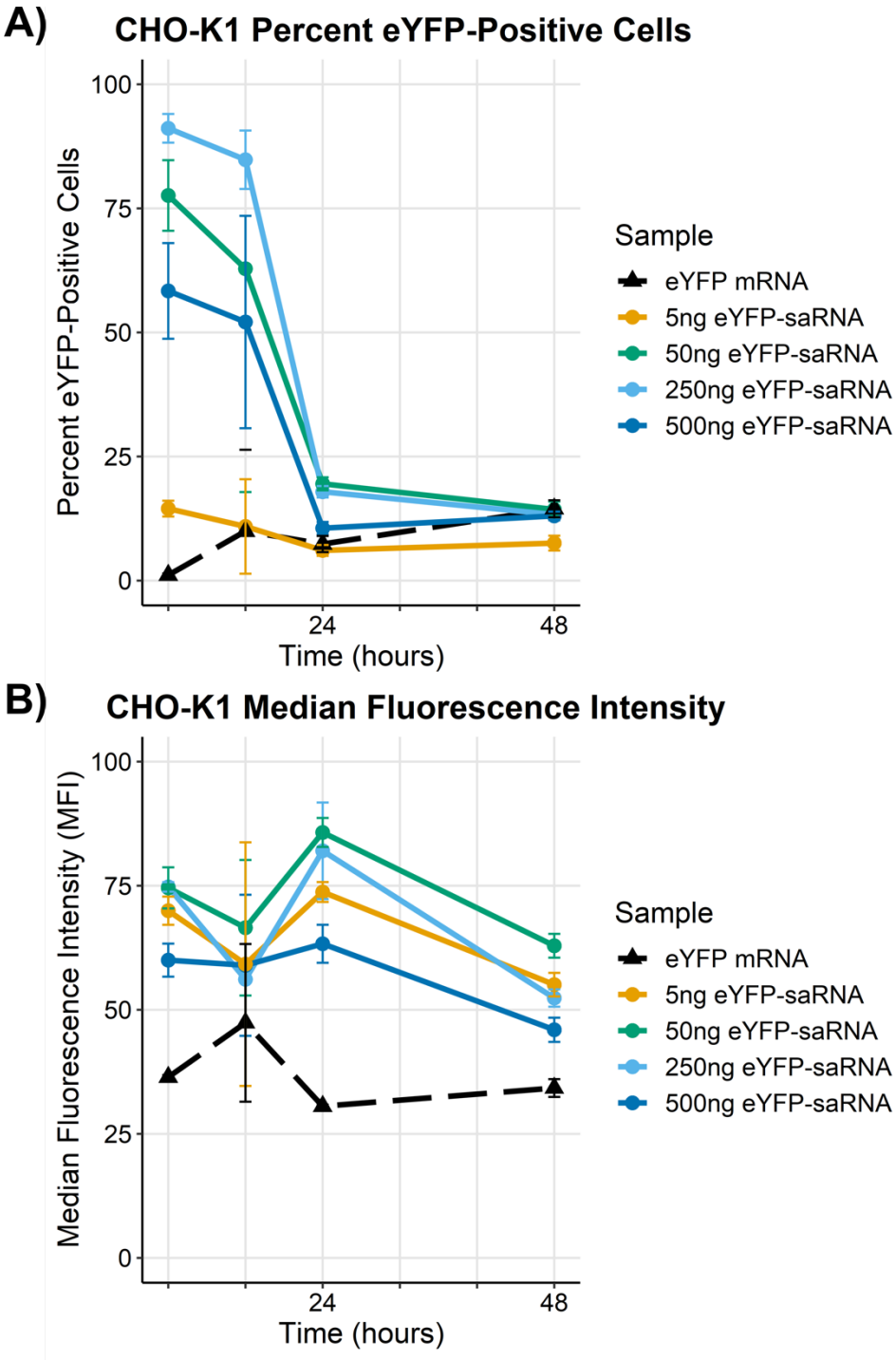

Figure 17: A. Flow cytometry data of %-eYFP-positive CHO-K1 cells at 8, 16, 24 and 48-hours post-transfection with 5, 50, 250, or 500 ng of eYFP-saRNA or EYFP mRNA. Samples of 50, 250 and 500 ng eYFP-saRNA when compared to the negative control (eYFP-mRNA) had a p-value < 0.01. **B.** Median fluorescence intensity of eYFP-positive cells at each time point for each RNA amount. All samples when compared to the negative control (eYFP-mRNA) had a p-value < 0.01

As in the BHK-21 cells, flow cytometry analysis (**Figure 17A**) confirmed the fluorescence data with 250 ng and 50 ng saRNA showing the strongest expression of eYFP. The percentage of eYFP positive cells was again seen to peak earlier in the time course (8 hours post-transfection) than they appeared to in the fluorescence microscopy (16 hours post-transfection), likely owing to the more sensitive measurement in flow cytometry. Transfection efficiency appeared higher in this cell line than in BHK-21 cells, with higher overall levels of eYFP positive cells. The cells treated with 50, 250, or 500 ng of saRNA showed significantly higher percentages of eYFP-positive cells when compared to the eYFP-mRNA control (p-values ≤ 1.54E-04), while the 5 ng saRNA-treated cells had a similar percentage of eYFP-positive cells when compared to the control.

As with the BHK-21 cells, the MFI of CHO-K1 cells transfected with saRNA were significantly higher than the eYFP-mRNA control **(Figure 17B)**. Interestingly, the MFI in CHO-K1 cells was more affected by the dose of saRNA used in transfection than for the BHK-21 cells. The 50 ng saRNA amount showed the strongest overall fluorescence signal, followed by the 250 ng saRNA dose, meaning that not only were more cells producing protein at these lower RNA transfection amounts, but each cell was individually producing protein at a stronger level. Additionally, the 50 ng saRNA treatment showed a significantly higher MFI when compared to the 500 ng saRNA treatment (p-value = 5.95E-04). The fluorescence intensity again peaked relatively late in the time course as dead, highly fluorescent cells began to appear. Additionally, the median fluorescence intensity was much higher for CHO-K1 than for BHK-21 cells, indicating that CHO-K1 cells may be more amenable to transfection (as also indicated by their higher transfection efficiency) and better able to tolerate high levels of saRNA amplification.

4.4.4 CHO-K1 cell viability drops sharply with increasing dosage of eYFP-saRNA

Trypan-blue-exclusion assays were again performed on CHO-K1 cells transfected as above, and live percentages were calculated as in BHK-21 cells (**Figure 18**). Again, an inverse trend of survival was seen with increasing saRNA doses. In these cells, viability dropped below 25% for the highest RNA amount delivered (500 ng) at the peak of expression, 16 hours post-transfection, and hovered around 25% for cells treated with 250 ng of eYFP-saRNA. Viabilities for both higher dosages rose gradually back to 35%, for 500 ng, or 70%, for 250 ng, by 72 hours post-transfection, as healthy cells continued to grow and divide. The viability of cells transfected with 50 ng, those that expressed fluorescent protein most robustly, dropped to about 50% at 16 hours post-transfection, but rose back to 80% by the end of the time course. The viabilities of the 50, 250, and 500 ng eYFP-saRNA treated samples were all significantly different from those of control samples (untransfected, lipofectamine-transfected, and B1 RNA transfected) with p-values $\leq$ 2.22E-16. Additionally, the 50 ng saRNA-treated cell viability was statistically different from that of the 250 and 500 ng saRNA-treated samples (p-values $\leq$ 2.22E-16).

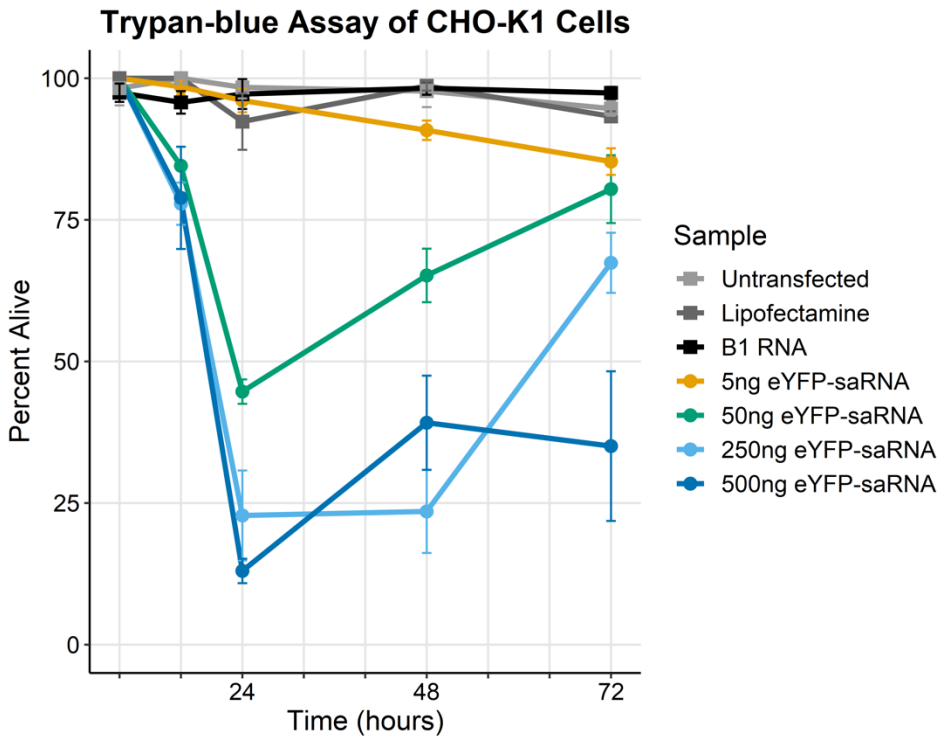

Figure 18: Trypan-blue-exclusion assays of CHO-K1 cells collected at 8, 16, 24, 48 and 72-hours post-transfection with 5, 50, 250, or 500 ng of eYFP-saRNA or 500 ng B1 RNA only. Control samples include untransfected cells, cells transfected with lipofectamine reagent only (no RNA), and cells transfected with 500 ng of B1 RNA. Samples of 50, 250 and 500 ng eYFP-saRNA when compared to the negative controls (untransfected, lipofectamine, and B1 RNA) had a p-value < 0.01.

In contrast, cells transfected with 5 ng eYFP-saRNA showed a steady decrease in viability over time, with viability still at 95% at the peak of protein expression but dropping to 85% by 72 hours post-transfection. As before, controls (untransfected, mock-transfected, and B1-transfected cells) remained highly viable for the duration of the experiment and the 5 ng eYFP-saRNA-treated samples were not statistically different from the controls (p-values ≥ 0.23). Brightfield images again validated the trypan data, with cells treated with higher RNA doses becoming much less adherent at earlier time points and appearing overall less healthy throughout the course of the experiment.

The cells used in these experiments are known to be easily transfectable and were chosen because all three lines are adherent and amenable to lipoplex-based transfection protocols. HEK-

293 cells were also exposed to the same transfection scheme as described for BHK-21 and CHO-K1 cells. They showed fluorescence expression profiles very similar to those of BHK-21 cells but were excluded from further data analysis due to their similarity to BHK-21s and due to difficulty in keeping the cells attached to plates throughout the entire transfection. Despite the differences in the peak fluorescence expression times between BHK-21 and CHO-K1 cells, both cell types fluoresced most strongly, and were relatively viable, when transfected with 50 ng of eYFP replicon RNA diluted with 450 ng of B1 RNA (**Figure 19**). Brightfield images corresponding to the fluorescence microscopy images are included to demonstrate that fluorescence intensity changes are due to changes in cell expression and not due to a change in the number of cells per image. This is a significantly lower mass amount than recommended in most mRNA transfection protocols and underscores the relevance of the self-amplifying nature of a replicon when considering its protein expression *in vitro* and its dosage efficacy as a therapeutic.

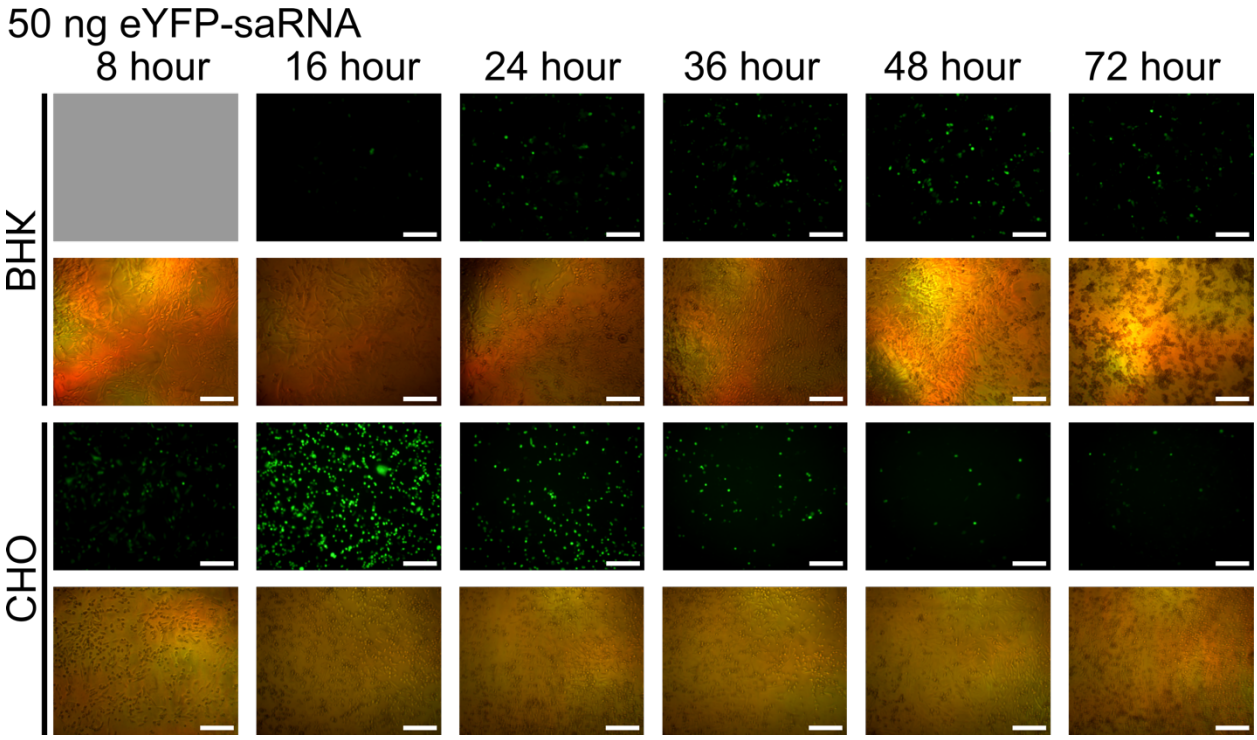

Figure 19: Representative fluorescence and brightfield images collected at 8, 16, 24, 36, 48 and 72-hours post-transfection for BHK-21 and CHO-K1 cells transfected with 50 ng of eYFP-saRNA supplemented with 450 ng of carrier RNA. Brightfield images show cell health and the presence of cells despite fluorescence not being observed. Scale bar = 100 μm.

4.5 Discussion

Lipoplex-forming transfection agents have previously been shown to be particularly useful in the case of transfection of viral-derived RNAs. Because the replication mechanisms employed by saRNAs are similar to those of the viral genomes from which they are derived, these RNAs are more likely to cause cell toxicity, weakening and death upon transfection[80]. More physically disruptive methods of transfection such as electroporation, which causes permeabilization of the cell membranes through sudden application of an intense electric field, can cause high amounts of cell death even in mock DNA transfections and may be particularly ill-suited for adherent cell lines if the shocked cells fail to reattach. These alternate transfection methods could cause cell death to the extent of invalidating results due to aggressive genetic drift spurred by the survival of only a few, genetically aberrant, cells[80]. As evidenced by our results, it is important even with gentler lipoplex transfections to optimize protocols for replicon and genomic viral RNAs and adjust for cell viability. More work could be done to titrate not only the RNA amount necessary for optimal translation, but also the corresponding lipoplex-forming reagent amount; if the reagent amount were appropriately scaled down, this would eliminate the use of dilution of active RNA with the carrier RNA (here, B1) and would increase cell viability to escape the somewhat toxic effects of excess lipid[81].

The non-monotonic dose response seen in our results favorably demonstrates the potential of this replicon system as a vaccine, immune system stimulatory agent, or other therapeutic vector, depending on the GOI chosen[82,83]. Since low amounts of replicon RNA are shown to be both unusually effective in GOI expression and relatively nontoxic in adherent cells

– significantly more so than larger amounts of RNA – this alleviates much of the targeting and delivery concern normally involved in mRNA therapeutics. Although it will still be necessary to protect our replicon therapeutics in a delivery vehicle, as we have shown previously with plant-virus-derived virus-like-particles, it is clear that very low numbers of delivery particles will be required for uptake per cell in order for effective RNA replication and translation to occur[67].

4.6 Conclusions

This work emphasizes the importance of an understanding of viral replication pathways and replicon behavior in cells when considering *in vitro* testing systems for self-amplifying molecules. Replicons have long been considered an interesting avenue of RNA-based technologies, and many viral genomes have been adopted for use in studying replication and expression systems[84]. However, little systematic optimization has been done in order to maximize their efficacy in cell cultures and to understand how they will best be suited for use in clinical applications, although there have been investigations of the effects of increased viral MOI (rather than increased replicon numbers) on protein expression level using recombinant SFV[6,85]. The examples considered in the present work involve cell lines commonly used both for studying viral replication and for testing gene expression in general. The clear existence of a "sweet spot" – a low-to-intermediate dose of saRNA that yields maximum viability and GOI expression, across all lines tested – strongly underscores the fact that the amplification of replicon RNA must be taken into account in transfection protocols. The ability to achieve such high levels of expression from very low mass amounts of saRNA provides exciting opportunities for the design of effective therapeutics that are not reliant for function on huge delivery payloads for function.

4.7 Acknowledgements

We thank Prof. Otto Yang for allowing access to the MACSQuant® Analyzer 10 that was generously provided by the James B. Pendleton Charitable Trust, Dr. Balamurugan Arumugam for assistance with flow cytometry data processing, and Adi Mukund for assistance with data processing and visualization. Funding: This work was supported by NSF Grant MCB17-16925 to WMG from the National Science Foundation Molecular and Cellular Biosciences Division.

CHAPTER 5

Packaging overlong RNAs into virus-like particles using polyvalent cations

5.1 Abstract

In vitro assembled virus-like particles (VLPs) are of interest as delivery vehicles for therapeutic molecules. This work investigates the packaging of long RNA molecules – those above 3,000 nucleotides – into spherical, 28-nm VLPs by pre-compacting the RNA with polyvalent cation. RNA was mixed with the quadrivalent cation, spermine, and analyzed by gel electrophoresis to determine charge ratios at which RNA becomes compacted by spermine. Pre-compacted RNA (at a 1 to 0.25 RNA to spermine charge ratio) was packaged into VLPs using brome mosaic virus capsid protein. VLPs assembled with normal RNA and pre-compacted RNA were visualized using negative-stain electron microscopy, and the number of singlets, doublets, triplets, and higher-order multiplets were counted for each sample. Pre-compaction of RNA with spermine results in a marginal increase in the number of singlet VLPs and a reduction in the number of higher-order multiplets. Packaging RNA with spermine has the potential to compact RNA such that it can fit into an RNase-resistant singlet VLP.

5.2 Introduction

Virus-like particles (VLPs) – non-infectious, self-assembled, homogeneous nanoparticles derived from the coat proteins (CP) of viral capsids – are being explored as effective delivery vehicles for the targeted delivery of therapeutic cargo[86,87]. Cowpea chlorotic mottle virus (CCMV) and the closely related brome mosaic virus (BMV) are two single-stranded RNA (ssRNA) plant viruses whose capsid proteins can be used to *in vitro* self-assemble around non-

viral RNA into spherical, 28-nm diameter VLPs[5,6,25,26,59]. The non-viral ssRNA can be between lengths of 100-12,000 nts and the resulting VLPs will be around 28-nm in diameter. For ssRNAs shorter than 2,500 nucleotides (nts), multiple RNAs are packaged into one 28-nm particle. For ssRNAs between 2,500 and 4,500 nts, one RNA is packaged into one 28-nm particle. ssRNAs longer than 4,500 nts are packaged into "multiplets" – multiple 28-nm VLPs, where RNA is shared between the particles[6]. These multiplets are not RNase-resistant because RNA is strung between the particles and susceptible to degradation by RNases. The inability to protect longer RNAs (>4,500 nts) in spherical VLPs is a limitation of this therapeutic delivery system. It is especially limiting for the delivery of self-replicating RNA therapeutics[67,88], where the self-replicating machinery accounts for about 3,000 nts, limiting genes of interest (GOIs) lengths to less than 1,000 nts.

Polyvalent cations can be used to compact nucleic acid into smaller structures, and it is hypothesized that ssRNA can be pre-compacted before *in vitro* assembly such that longer ssRNA molecules can be packaged into spherical, 28-nm VLPs. Spermine is a polyamine that is a polycation at physiological pH with four positive charges. While spermine is found in all eukaryotic cells as it is involved in cellular metabolism, a large concentration of spermine is found in sperm cells where a large DNA genome needs to be compacted to fit into a small cell[89–91]. Spermine is also found in turnip yellow mosaic virus (TYMV), another ssRNA plant virus. Spermine compacts the TYMV genome such that a 6,400 nt RNA molecule can fit inside a 28-nm spherical capsid[92] (note: TYMV has not been shown to assemble *in vitro*). Notably, CCMV CP assembles around 6,400 nt ssRNA, but the resulting VLPs are doublets that are not RNase-resistant[6].

Taking inspiration from nature, this work aims to use spermine, a polyvalent cation, to pre-compact ssRNA molecules longer than 4,500 nts before *in vitro* assembling BMV VLPs. Six RNAs were used ranging in length from 3,000 to 12,000 nts. RNA was mixed with increasing amounts of spermine and analyzed by gel electrophoresis to assess the compaction of RNA as a function of charge ratio. The 1 to 0.25 RNA to spermine charge ratio was used for assembly experiments because it compacted the RNA, but also allowed for enough negative charges on the RNA for the *in vitro* assembly to occur successfully. These experiments demonstrate that BMV VLPs can form around RNA pre-compacted with polyvalent cation and that pre-compaction marginally increases the number of singlets (and reduces the number of higher order multiplets).

5.3 Materials and Methods

5.3.1 Buffers used

Disassembly buffer: 50 mmol/L Tris-HCl, pH 7.5; 500 mmol/L $CaCl_2$, 1 mmol/L ethylenediamine tetraacetic acid (EDTA), 1 mmol/L dithiothreitol (DTT), 0.5 mmol/L phenylmethylsulfonyl fluoride (PMSF). **Protein storage buffer:** 20 mmol/L Tris-HCl, pH 7.2; 1 mol/L NaCl; 1 mmol/L EDTA; 1 mmol/L DTT; 1 mmol/L PMSF. **RNA assembly buffer (RAB):** 50 mmol/L Tris-HCl, pH 7.2, 50 mmol/L NaCl, 10 mmol/L KCl, 5 mmol/L $MgCl_2$. **Virus storage buffer (VSB):** 50 mmol/L sodium acetate, pH 4.5, 8 mmol/L magnesium acetate. **Tris-Acetic Acid-EDTA (TAE):** 40 mmol/L Tris-HCl, pH 8, 20 mmol/L acetic acid, 1 mmol/L EDTA.

5.3.2 Synthesis of RNAs

The RNAs used in these experiments are BMV RNA1 (3,324 nts), NOV-R.Luc (4,413 nts), NOV-STING* (4,638 nts), TMV (6,395 nts), SIN-19 (8,985 nts), and FL-SIN (11,703 nts). BMV RNA1 was made by *in vitro* transcription of the DNA plasmid pT7B1, linearized with

BamHI (New England Biolabs, USA), with a T7 polymerase transcription system (Thermo Fisher, USA). NOV-R.Luc and NOV-STING* were made by *in vitro* transcription of the DNA plasmid pNod1-R.Luc and pNod1-STING*-T2A, linearized with XbaI (New England Biolabs, USA), with a T7 polymerase transcription system. TMV was made by *in vitro* transcription of the DNA plasmid p30B, linearized with XhoI (New England Biolabs, USA), with a T7 polymerase transcription system. SIN-19 and FL-SIN were made by *in vitro* transcription of the DNA plasmid pSindbis-19 and pTE12, linearized with XhoI (New England Biolabs, USA), with a SP6 polymerase transcription system. All RNA was purified with an RNEasy Mini Kit (Qiagen, DEU) following the manufacturers' specifications.

5.3.3 Pre-compaction of RNA with spermine

Spermine was stored in degassed water overlayed with argon gas and frozen at -80°C until used for experiments. Frozen RNA was thawed on ice and thermally denatured by heating the sample to 90°C at a rate of 1°C per second, holding at 90°C for 1 second, then cooling to 4°C at a rate of -1°C per second. Thermally denaturing and slowly cooling the RNA allows for disruption of any duplexes formed between RNA molecules in the freezing process and for refolding of the RNA. The RNA was then mixed with spermine at a specified charge ratio and incubated at 4°C for 30 minutes. 1 RNA molecule contributes 1 negative charge per nucleotide (from the phosphate backbone of RNA), so the number of negative charges per molecule will depend on the RNA length:spermine contributes 4 positive charges per molecule.

5.3.4 Gel electrophoresis of RNA

750-1000 ng of RNA was mixed with 2X RNA loading dye and fully denatured by heating the sample to 65°C at a rate of 1°C per second, holding at 65°C for 10 minutes, then cooling to 4°C at a rate of -1°C per second. If the sample contained RNA mixed with spermine, the spermine

was added at the specific charge ratio and incubated at 4°C for 30 minutes. The RNA was then mixed with 2.5 μL of 200X GelRed and loaded into the 0.8% agarose gel in TAE. The gel ran at 100 V for 1.5 hours and was visualized for the nucleic acid stain with an excitation wavelength of 302 nm.

5.3.5 BMV capsid protein purification

WT BMV was purified from infected barley leaves (Hordeum vulgare)[3], and coat protein was purified as described previously[39]. Nucleocapsids were disassembled by dialyzing against disassembly buffer at 4 °C overnight. The RNA was pelleted, and the coat protein isolated by ultracentrifugation at 90,000 rotations per minute for 100 min at 4 °C in a Beckman TLA110 rotor. Coat protein was extracted from the supernatant and immediately dialyzed against protein storage buffer. Protein concentration and purity were assessed by UV-Vis spectrophotometry; only protein solutions with 260/280 ratios less than 0.6 were used for assembly. Protein was frozen in liquid nitrogen and stored at −80 °C until ready to use, at which point it was defrosted on ice and stored at 4 °C for up to two weeks.

5.3.6 *In vitro* self-assembly of BMV virus-like particles

RNA and CP were mixed in protein storage buffer at a 1:4.8 RNA to CP mass ratio to a final RNA concentration of 30 ng/μL and then dialyzed overnight (12-16 hours) against RAB at 4°C. The assembly was dialyzed against VSB at 4°C for at least 6 hours after which VLPs were collected for analysis. For samples that were assessed for RNA protection, VLPs were mixed with RNaseA and incubated at 4°C for 1 hour. Digestion of RNA was stopped by the addition of RNase Inhibitor and the sample was washed through a 100 kDa MW-cutoff Amicon filter purify the remaining RNase-resistant VLPs.

5.3.7 Negative-stain electron microscopy

6 μL of VLP sample was deposited on glow-discharged carbon-coated copper (200-mesh) PELCO Pinpointer grids (Ted Pella, USA). After 1 min, the grids were blotted with Whatman filter paper, and then stained with 6 μL of 2 % uranyl acetate for 1 min followed by complete stain removal and storage in a desiccator overnight. Micrographs were acquired using a Tecnai G2 TF20 High-Resolution electron microscope (FEI, USA) with an accelerating voltage of 200 kV. Images were collected at 3 μm to 4 μm underfocus with a TIETZ F415MP 16-megapixel CCD camera (4000 by 4000 pixels, pixel size 15 μm). Electron micrographs were manually analyzed to count the number of singlets, doublets, triplets, and higher-order multiplets in each of the images acquired.

5.4 Results

5.4.1 RNA is compacted with spermine

In the interest of examining the packaging of overlong (>4,500 nt) RNAs, these experiments involve BMV RNA1 (3,324 nts), NOV-R.Luc (4,413 nts), NOV-STING* (4,638 nts), TMV (6,395 nts), SIN-19 (8,985 nts), and FL-SIN (11,703 nts). The *in vitro* transcribed RNA was analyzed by gel electrophoresis with a ssRNA ladder to confirm the lengths of the RNA (**Figure 20**).

Each RNA was mixed with increasing concentrations of spermine corresponding to specific RNA-to-spermine charge ratios and subjected to gel electrophoresis to determine the effect of spermine on RNA charge and size (**Figure 21**). Similar trends are observed across all RNA lengths in that the RNA bands first appear to run slower, closer to the top of the gel (negative electrode) than the control samples (lanes 1 and 2), but then around lanes 5 or 6, corresponding to charge ratios of 1 to 0.05 and 0.075, respectively, the bands appear to run faster

or closer to the bottom of the gel (positive electrode). When at the highest charge ratios 1-to-1 and 1-to-2 (lanes 11 and 12), some RNAs begin to demonstrate aggregation where the sample does not leave the well at the top of the gel. An RNA to spermine charge ratio of 1 to 0.25 (lane 8) was used for *in vitro* assembly experiments.

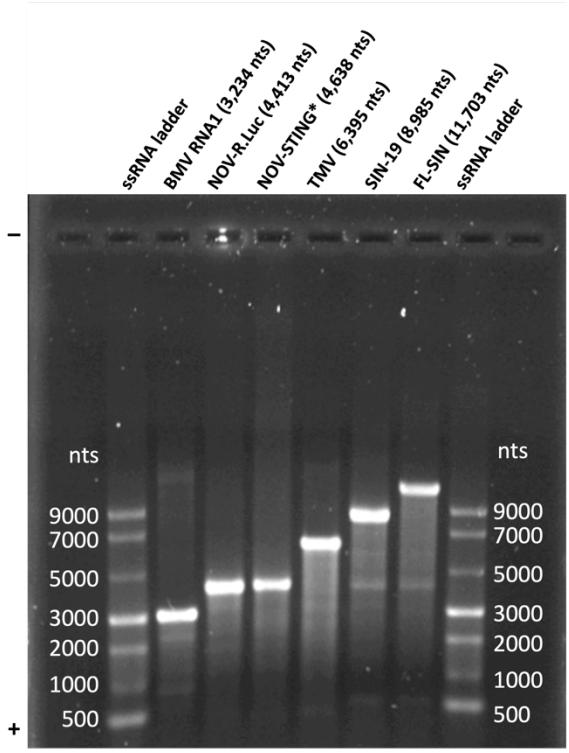

Figure 20: Gel electrophoresis of *in vitro* transcribed RNA. The RNAs used in these experiments are BMV RNA1 (3,324 nts), NOV-R.Luc (4,413 nts), NOV-STING* (4,638 nts), TMV (6,395 nts), SIN-19 (8,985 nts), and FL-SIN (11,703 nts) and are shown in the gel above with flanking single-stranded RNA ladders with standard RNA bands of 500, 1,000, 2,000, 3,000, 5,000, 7,000, and 9,000 nts. Samples were stained with GelRed for visualization of nucleic acids.

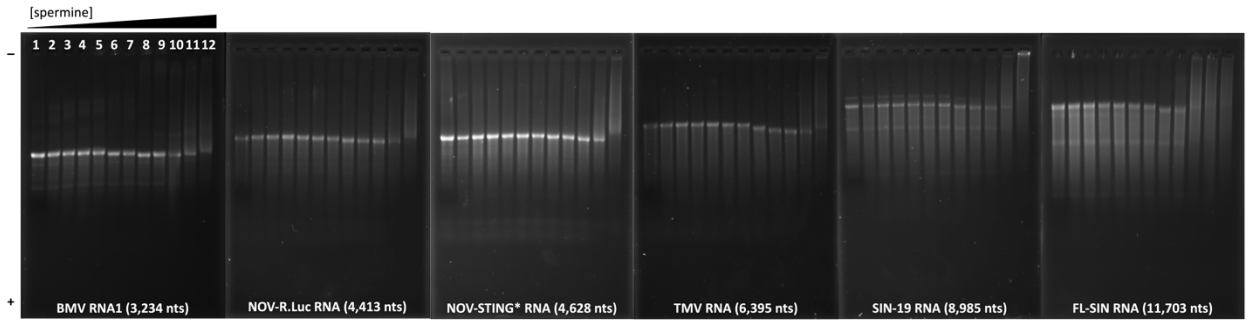

Figure 21: Gel electrophoresis of RNAs incubated with increasing concentrations of spermine. Each of the six gels contains samples of one RNA – BMV RNA1 (3,324 nts), NOV-R.Luc (4,413 nts), NOV-STING* (4,638 nts), TMV (6,395 nts), SIN-19 (8,985 nts), or FL-SIN (11,703 nts) – mixed with spermine concentrations. Lane 1 is RNA with no spermine, fully denatured with formaldehyde, Lane 2 contains RNA with no spermine, Lanes 3-12 contain RNA mixed with spermine at the following RNA to spermine charge ratios, 1 to: 0.01, 0.025, 0.05, 0.075, 0.10, 0.25, 0.50, 0.75, 1, and 2.

5.4.2 RNA compacted with spermine can be packaged into BMV VLPs

BMV VLPs were made with all six RNAs that were mixed with spermine in a 1 to 0 or 1 to 0.25 RNA to spermine charge ratios. Representative negative-stain electron micrographs of assembly products are presented in **Figure 22** with assemblies without spermine on top and assemblies with spermine on the bottom. Particles observed in electron micrographs were manually counted, with at least 200 particles counted per sample, and classified as singlets, doublets, triplets, or higher order (4+) multiplets. A stacked bar plot showing these distributions for all six RNAs in the two spermine conditions are shown in **Figure 23**. The same data is plotted as a line plot of RNA length versus fraction of particles in **Figure 24**.

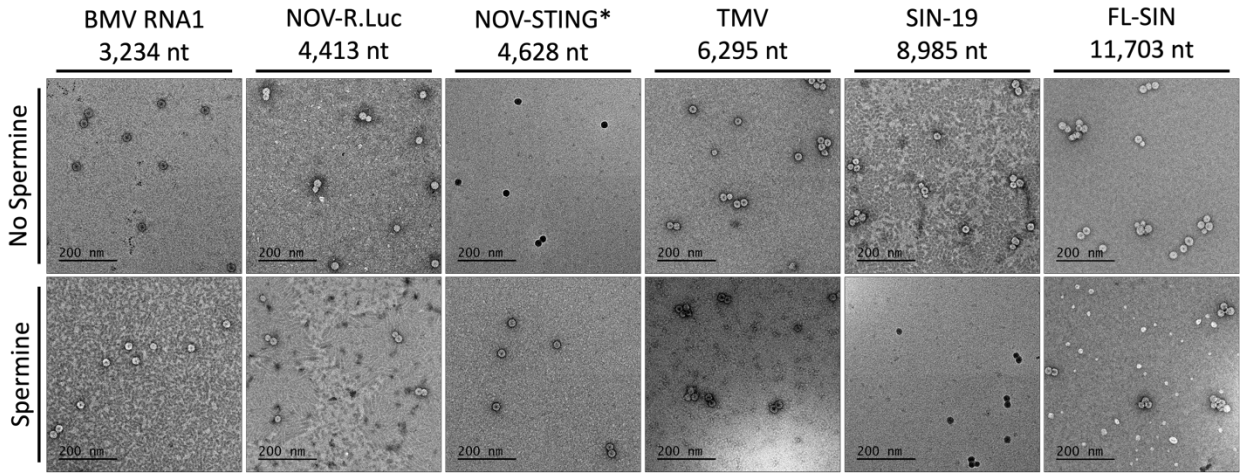

Figure 22: Electron micrographs of RNA assembled with BMV capsid protein in the absence or presence of spermine. RNAs – BMV RNA1 (3,324 nts), NOV-R.Luc (4,413 nts), NOV-STING* (4,638 nts), TMV (6,395 nts), SIN-19 (8,985 nts), or FL-SIN (11,703 nts) – were assembled into virus-like particles (VLPs) using BMV capsid protein (top) or pre-compacted with spermine (at a 1 to 0.25 RNA to spermine charge ratio) and then packaged into VLPs (bottom).

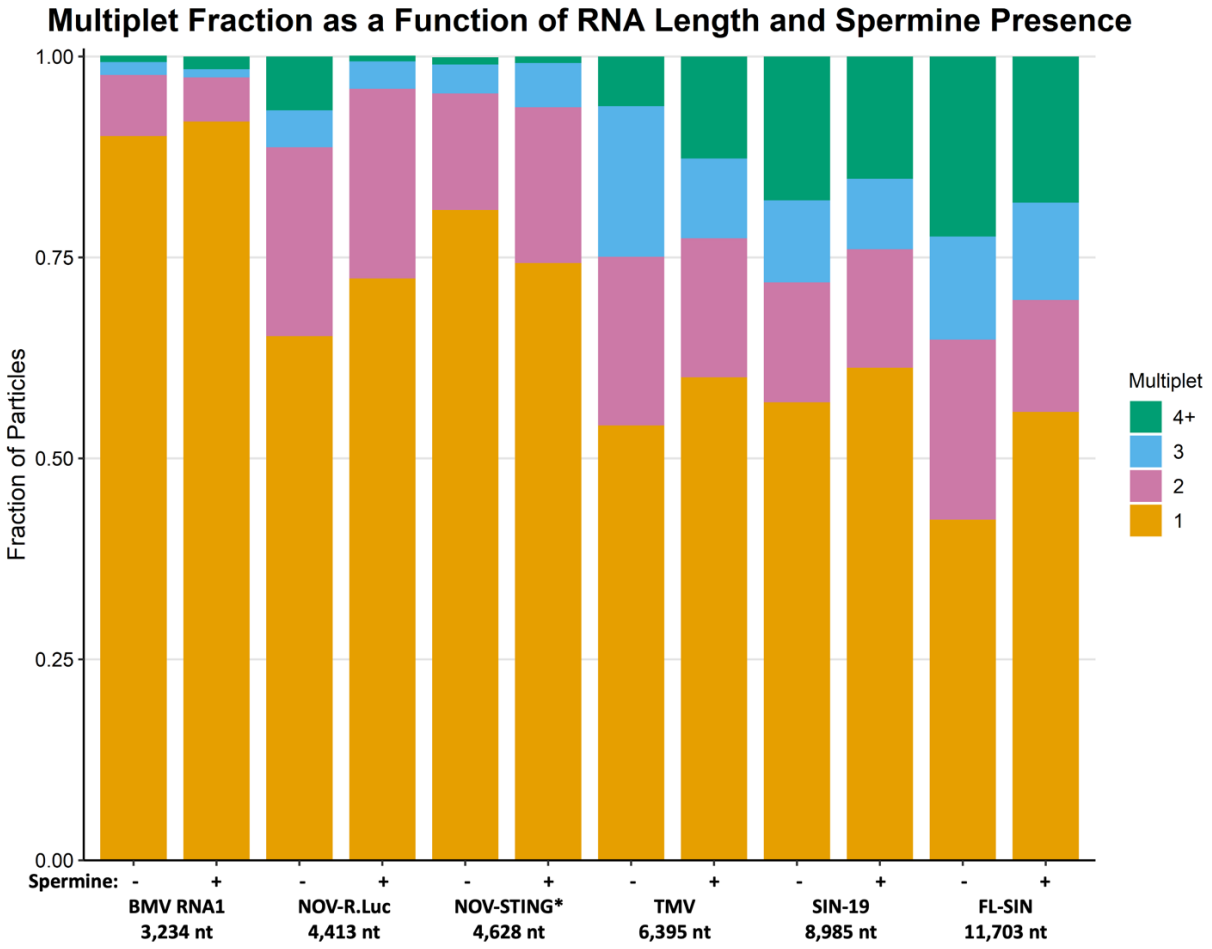

Figure 23: Stacked bar plot of multiplet virus-like particle distributions for assemblies of RNA with and without spermine. Electron micrographs of virus-like particles (VLPs) assembled from RNAs – BMV RNA1 (3,324 nts), NOV-R.Luc (4,413 nts), NOV-STING* (4,638 nts), TMV (6,395 nts), SIN-19 (8,985 nts), or FL-SIN (11,703 nts) – either pre-compacted with spermine (1 to 0.25 RNA to spermine charge ratio) or not were analyzed. The number of singlet (orange), doublet (pink), triplet (light blue), or higher order multiplets (4+, green) were counted for each assembly and plotted as a fraction of the total particles for that sample.

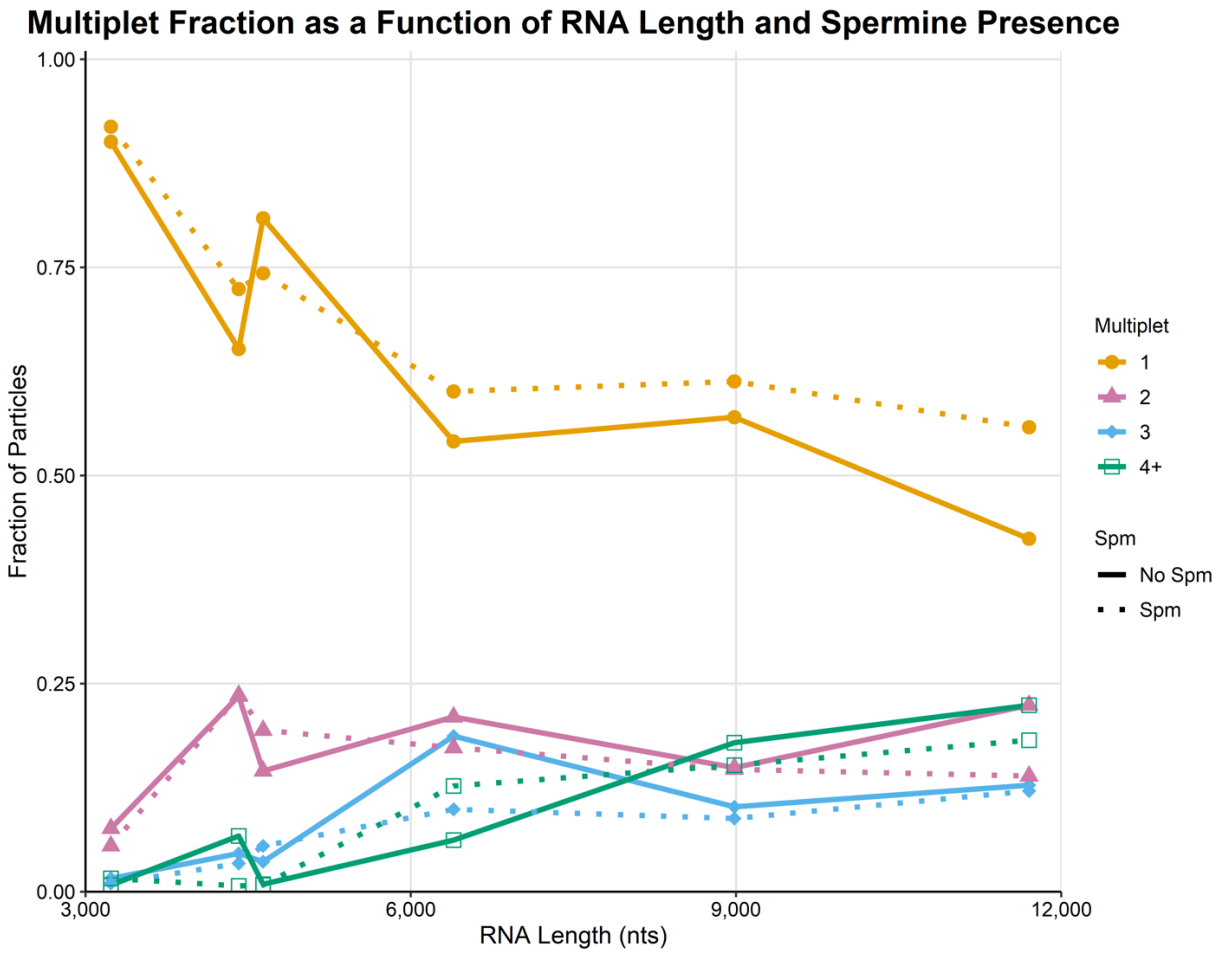

Figure 24: Line plot of multiplet virus-like particle distributions for assemblies of RNA with and without spermine. Electron micrographs of virus-like particles (VLPs) assembled from RNAs – BMV RNA1 (3,324 nts), NOV-R.Luc (4,413 nts), NOV-STING* (4,638 nts), TMV (6,395 nts), SIN-19 (8,985 nts), or FL-SIN (11,703 nts) – either pre-compacted with spermine (dotted lines, 1 to 0.25 RNA to spermine charge ratio) or not (solid lines) were analyzed. The number of singlet (orange circles), doublet (pink triangles), triplet (light blue diamonds), or higher order multiplets (4+, green outlined squares) were counted for each assembly and plotted as a fraction of the total particles for that sample as a function of the RNA length.

For the VLPs assembled, the number of singlets increases and the number of higher-order multiplets decreases in the presence of spermine. However, the increase in the number of singlet VLPs observed is very marginal and is not observed in the NOV-STING* (4,638 nts) case. For BMV RNA1, the singlet fraction increases from 0.90 to 0.92 in the presence of spermine, for NOV-R.Luc the fraction increases by 0.07 (0.65 to 0.72), for TMV the increase is 0.06. (0.54 to 0.60), for SIN-19 the increase is 0.04 (0.57 to 0.61), and for FL-SIN the fractional increase in

singlets is the highest at 0.14 (0.42 to 0.56). The fraction of singlets for NOV-STING* decreases in the presence of spermine from 0.81 to 0.74. The number of higher-order multiplets (doublets, triplets, and four-plus multiplets) decreases for all samples in the presence of spermine, except the NOV-STING* case. Again, the changes are marginal with a decrease in the fraction of particles of 0.02 (BMV RNA1), 0.07 (NOV-R.Luc), 0.06 (TMV), 0.04 (SIN-19), and 0.13 (FL-SIN), with an increase of 0.07 for NOV-STING* RNA.

5.5 Discussion

Spermine compaction of RNA occurs because the positive charges on the spermine interact with the negative charges from the phosphate backbone of the RNA which screens the negative charges on the RNA molecule and therefore the repulsions from those negative charges. The screening of repulsions allows for the RNA to become more compact. Gel electrophoresis analysis of RNAs mixed with spermine illustrate the compaction of RNA with increasing spermine compaction as observed by bands that run faster in the gel (closer to the bottom). The porous gel network separates molecules by their size with smaller particles travelling through the network faster than larger particles. The electric field applied allows for separation of particles by charge. If particles are the same size, but different charge, the more negatively charged molecule will travel farther because it will move faster through the gel. If particles are the same charge, but different sizes, then the smaller molecule will travel farther. When there are competing effects, as in the present case, molecules that have a smaller negative charge can travel farther than those with more negative charge if they are sufficiently small. As the spermine concentration increases, the overall negative charge of the RNA is decreased, so if the RNA was not getting compacted (i.e., the same size) but had different charge, we would expect that samples with more spermine would not travel as far as those without spermine. However, what

we see is that with increasing spermine concentrations, the RNA runs farther in the gel, indicating a smaller effective size.

These results are further confirmed with dynamic light scattering (DLS) measurements, in which the effective hydrodynamic diameter of the particles is measured (data not shown). DLS results confirm that RNA is continuously compacted with increasing spermine concentration. For example, for NOV-STING* RNA molecule around 60 nm in diameter with no spermine present is compacted to a molecule as small as 30 nm in diameter at charge ratios as high as 1 to 2 RNA to spermine. At a charge ratio of 1 to 0.25, those used for the *in vitro* assembly experiments, NOV-STING* RNA has an effective diameter of 42 nm.

While RNA compaction was observed with increasing spermine concentrations, *in vitro* assembly experiments only resulted in marginal increases in the number of singlet BMV VLPs observed. In order for spherical VLPs to be effective therapeutic delivery vehicles, the RNA it is packaging must be protected from RNA degradation. Even with RNA as short as NOV-R.Luc (4,413 nts), the fraction of particles that are not RNase-resistant even when the RNA is pre-compacted with spermine is upwards of 0.25. RNAs even longer than 4,500 nts are less likely to fit into singlet VLPs with the spermine charge ratio tested.

5.6 Conclusions and Future Directions

It is possible that higher concentrations of spermine could result in enough compaction to package long RNAs into singlet VLPs. Additional experiments can be performed with higher spermine concentrations, but enough negative charges need to remain on the RNA such that the *in vitro* self-assembly process can occur. It is understood that the negative charges on the phosphate backbone of the RNA interact with the positively charged N-terminal tails of the CP which enables packaging of RNA into capsids[10,11]. Additionally, after successful packaging of

pre-compacted RNA into VLPs, the RNA needs to be assessed for functionality in cells to ensure the spermine does not interfere with successful translation or replication of therapeutic mRNAs.

CHAPTER 6

Packaging the same self-replicating RNA inside spherical and cylindrical virus-like particles

6.1 Abstract

The first viruses discovered to assemble *in vitro* are tobacco mosaic virus (TMV) and cowpea chlorotic mottle virus (CCMV). These viruses have inspired countless experiments to understand how they assemble *in vitro* to form virus-like particles (VLPs) and how we might use these VLPs as effective, targetable delivery vehicles for therapeutics. In these experiments, the same RNA molecule, NOV-R.Luc-OAS – a self-replicating mRNA derived from nodamura virus RNA1 that contains a reporter gene (Renilla luciferase) and a sequence for the TMV origin of assembly site (OAS) – is packaged by CCMV capsid protein (CP) and TMV CP to produce spherical and rod-like VLPs, respectively. The resulting VLPs were imaged using negative-stain electron microscopy. These results confirm the necessity of the TMV OAS for successful packaging of RNA into rod-like particles and that NOV-R.Luc-OAS can be packaged by both CCMV and TMV CP.

6.2 Introduction

Assembly of infectious virus particles in a test tube is unique to a few single-stranded RNA (ssRNA) plant viruses, cowpea chlorotic mottle virus (CCMV), brome mosaic virus (BMV), and tobacco mosaic virus (TMV). These viruses have been shown to self-assemble *in vitro*[5,14,25,26,59–61] and are of interest as delivery vehicles for therapeutics[86,87].

CCMV assembles *in vitro* around non-viral RNA in a two-step protocol resulting in spherical, monodisperse, 28-nm virus-like particles (VLPs)[6,26]. CCMV capsid protein (CP) has

been shown to assemble around non-viral RNA of a variety of lengths, always resulting in 28-nm VLPs. When the RNA length is between 2,500 and 4,500 nucleotides (nts), one RNA molecule is packaged into one VLP. However, when the RNA is longer than 4,500 nts, the RNA is packaged into multiplets, or multiple 28-nm VLPs where the RNA is shared between the particles[6].

TMV, on the other hand, assembles *in vitro* to form hollow, helical rod-like VLPs[60,61]. The rods have an inner diameter of 4 nm and an outer diameter of 18 nm with the length of the rod depending on the length of the RNA being packaged[14–16]. The packaged RNA forms a helix that is embedded about 2 nm into the protein shell. TMV CP can encapsulate non-viral RNA of any length if the RNA contains a specific recognition sequence, or packaging signal, called the origin of assembly site (OAS), at which VLP nucleation begins[17,18,93]. The VLP grows by addition of CP in which the 5' end of the RNA is pulled into the growing rod[17,60].

This work aims to package the same RNA in a spherical VLP made with CCMV CP and a rod-like VLP made with TMV CP with the ultimate goal of comparing the efficiency of these particles as delivery vehicles for self-replicating RNA molecules, or replicons. RNA1 from nodamura virus has been edited to include a gene of interest (GOI) in the open reading frame encoding the RNA-dependent RNA polymerase. The nodamura replicon containing a Renilla luciferase (R.Luc) reporter gene is referred to as NOV-R.Luc. This construct was then manipulated to include a TMV OAS (NOV-R.Luc-OAS), so that TMV CP can package the RNA. Both the NOV-R.Luc and NOV-R.Luc-OAS RNAs were mixed with TMV CP and resulting VLP rods were assessed demonstrating successful packaging of RNA that contains the OAS sequence. NOV-R.Luc-OAS RNA was mixed with CCMV CP and TMV CP to produce both spherical and rod-like VLPs.

6.3 Materials and Methods

6.3.1 Buffers used

Disassembly buffer: 50 mmol/L Tris-HCl, pH 7.5; 500 mmol/L $CaCl_2$, 1 mmol/L ethylenediamine tetraacetic acid (EDTA), 1 mmol/L dithiothreitol (DTT), 0.5 mmol/L phenylmethylsulfonyl fluoride (PMSF). **Protein storage buffer:** 20 mmol/L Tris-HCl, pH 7.2; 1 mol/L NaCl; 1 mmol/L EDTA; 1 mmol/L DTT; 1 mmol/L PMSF. **RNA assembly buffer (RAB):** 50 mmol/L Tris-HCl, pH 7.2, 50 mmol/L NaCl, 10 mmol/L KCl, 5 mmol/L $MgCl_2$. **Virus storage buffer (VSB):** 50 mmol/L sodium acetate, pH 4.5, 8 mmol/L magnesium acetate. **Tris-Acetic Acid-EDTA (TAE):** 40 mmol/L Tris-HCl, pH 8, 20 mmol/L acetic acid, 1 mmol/L EDTA. **TMV Assembly Buffer (TAB):** 75 mM sodium phosphate buffer, pH 7.2.

6.3.2 Synthesis of RNAs

NOV-R.Luc and NOV-R.Luc-OAS were made by *in vitro* transcription of the DNA plasmid pNod1-R.Luc and pNod1-R.Luc-OAS, linearized with XbaI (New England Biolabs, USA), with a T7 polymerase transcription system and purified with an RNEasy Mini Kit (Qiagen, DEU) both following the manufacturers' specifications.

6.3.3 CCMV capsid protein purification

WT CCMV was purified from infected cowpea plants (Vigna unguiculata black-eyed pea)[3], and coat protein was purified as described previously[39]. Nucleocapsids were disassembled by dialyzing against disassembly buffer at 4 °C overnight. The RNA was pelleted and the coat protein isolated by ultracentrifugation at 90,000 rotations per minute for 100 min at 4 °C in a Beckman TLA110 rotor. Coat protein was extracted from the supernatant and immediately dialyzed against protein storage buffer. Protein concentration and purity were assessed by UV-Vis spectrophotometry; only protein solutions with 260/280 ratios less than 0.6 were used for

assembly. Protein was frozen in liquid nitrogen and stored at −80 °C until ready to use, at which point it was defrosted on ice and stored at 4 °C for up to two weeks.

6.3.4 TMV capsid protein purification

TMV CP was stored in TAB and was kindly provided by our collaborators in the lab of Prof. Nicole Steinmetz at University of California, San Diego.

6.3.5 *In vitro* self-assembly of CCMV virus-like particles

RNA and CP were mixed in protein storage buffer at a 1:4.2 RNA to CP mass ratio to a final RNA concentration of 30 ng/μL and then dialyzed overnight (12-16 hours) against RAB at 4°C. The assembly is dialyzed against VSB at 4°C for at least 6 hours after which VLPs were collected for analysis. For samples that were assessed for RNA protection, VLPs were mixed with RNaseA and incubated at 4°C for 1 hour. Digestion of RNA was stopped by the addition of RNase Inhibitor and the sample was washed through a 100 kDa MW-cutoff Amicon filter purify remaining RNase-resistant VLPs.

6.3.6 *In vitro* self-assembly of TMV virus-like particles

RNA and CP were mixed in a TAB at a 1:20 RNA to CP mass ratio to a final RNA concentration of 20 ng/μL and incubated overnight (16+ hours) at 30°C. For samples that were assessed for RNA protection, VLPs were mixed with RNaseA and incubated at 4°C for 1 hour. Digestion of RNA was stopped by the addition of RNase Inhibitor and the sample was washed through a 100 kDa MW-cutoff Amicon filter purify remaining RNase-resistant VLPs.

6.3.7 Negative-stain electron microscopy

6 μL of VLP sample was deposited on glow-discharged carbon-coated copper (200-mesh) PELCO Pinpointer grids (Ted Pella, USA). After 1 min, the grids were blotted with Whatman filter paper, and then stained with 6 μL of 2 % uranyl acetate for 1 min followed by complete

stain removal and storage in a desiccator overnight. Micrographs were acquired using a Tecnai G2 TF20 High-Resolution electron microscope (FEI, USA) with an accelerating voltage of 200 kV. Images were collected at 3 μm to 4 μm underfocus with a TIETZ F415MP 16-megapixel CCD camera (4000 by 4000 pixels, pixel size 15 μm).

6.4 Results

6.4.1 TMV origin of assembly site (OAS) necessary for packaging nodamura replicon

The TMV OAS is a stem-loop structure located about 900 nts from the 3'-end of the wild-type (WT) TMV genome. Virion assembly initiates at the OAS because TMV CP recognizes this sequence with very high selectivity. The NOV-R.Luc RNA construct has been edited to make NOV-R.Luc-OAS, which contains about 230 nts of the WT TMV genome (in the region of the OAS) after the Renilla luciferase gene. Both NOV-R.Luc and NOV-R.Luc-OAS were mixed with TMV CP in assembly buffer overnight at 30°C, and then treated with RNaseA. The remaining RNase-resistant VLPs were purified and visualized using negative-stain electron microscopy. Representative electron micrographs (EMs) of VLPs containing NOV-R.Luc-OAS (top) and NOV-R.Luc (bottom) are shown in **Figure 25**. VLPs containing NOV-R.Luc-OAS RNA show polydispersity in rod length, with most rods around 200 nm in length. EMs of VLPs made with RNA lacking an OAS, however, showed short rods often less than 30 nm in length or TMV CP disks (18-nm diameter spheres), which is the predominant form of the CP at neutral pH[60].

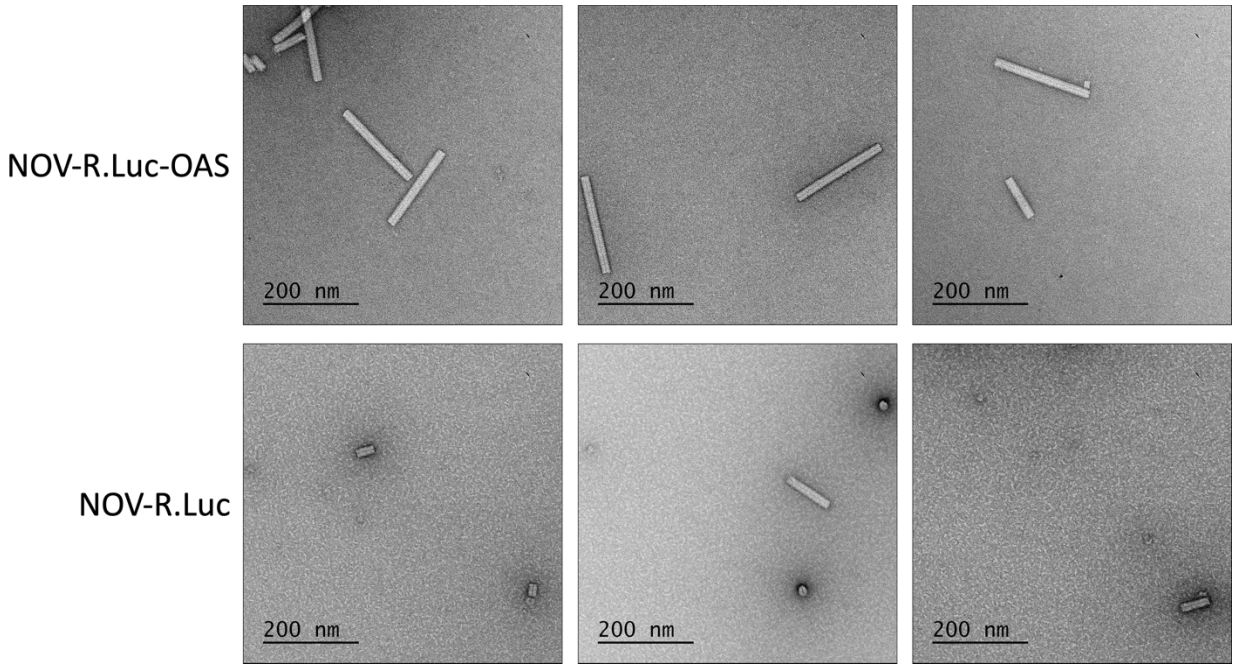

Figure 25: Electron micrographs of TMV virus-like particles assembled around RNA with and without an origin of assembly site (OAS). NOV-R.Luc-OAS RNA and NOV-R.Luc RNA were mixed with TMV capsid protein to assemble rod-like virus-like particles and imaged using negative-stain electron microscopy.

6.4.2 NOV-R.Luc-OAS can be packaged into spherical and rod-like virus-like particles

NOV-R.Luc-OAS RNA was packaged with CCMV CP and TMV CP to produce VLPs in the shape of spheres and rods, respectively. Representative EMs of VLPs formed around NOV-R.Luc-OAS are shown in **Figure 26**. Both singlet and doublet spherical VLPs are observed in assemblies of NOV-R.Luc-OAS with CCMV CP. Rod-like TMV VLPs observed are polydisperse in length, but are on average around 200 nm in length.

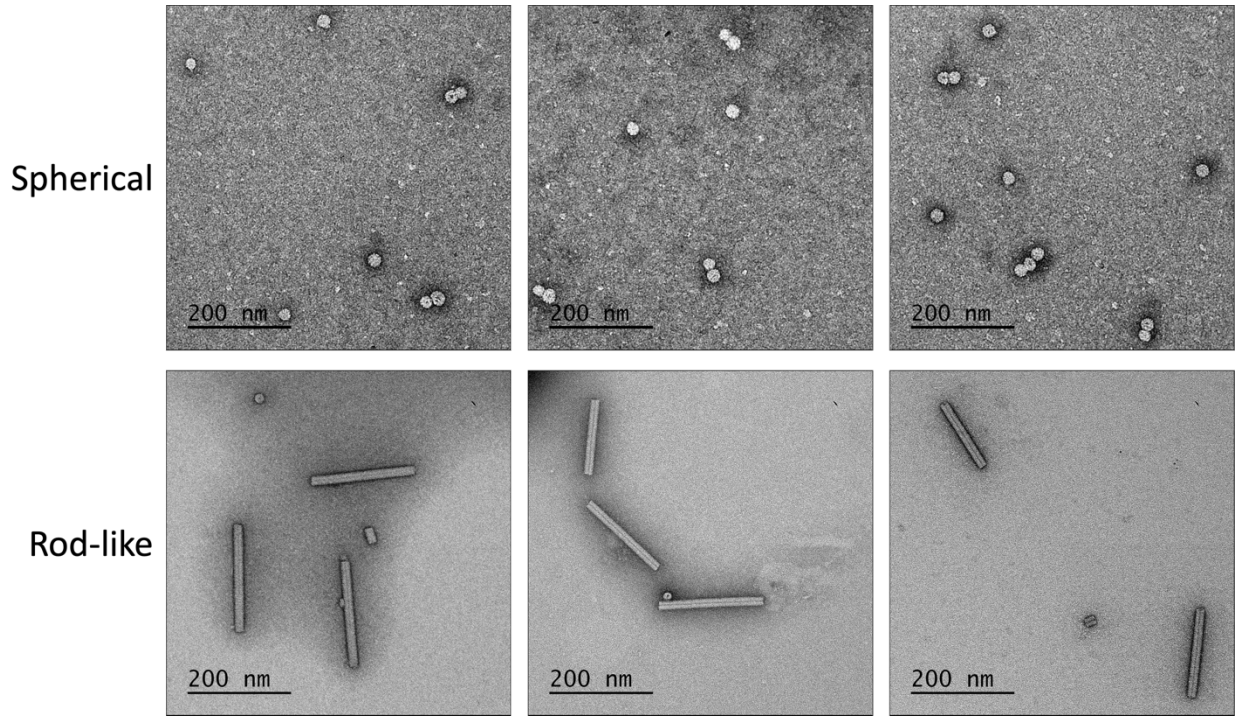

Figure 26: Electron micrographs of spherical and rod-like virus-like particles assembled around the same RNA. NOV-R.Luc-OAS RNA was packaged with CCMV capsid protein (CP) and TMV CP to produce spherical (top) and rod-like (bottom) virus-like particles.

6.5 Discussion

These results demonstrate the successful packaging of the same RNA molecule, NOV-R.Luc-OAS, into both spherical and rod-like VLPs. The OAS fragment inserted into the NOV-R.Luc replicon was sufficient for initiation of assembly of rod-like VLPs by TMV CP. The polydispersity observed in TMV VLPs could be the result of incomplete packaging of RNA or breakage of VLPs. Meanwhile, the singlet and doublet VLPs observed when NOV-R.Luc-OAS is packaged with CCMV CP is expected as NOV-R.Luc-OAS RNA is around 4,500 nts. It is important to note that CCMV VLPs were not RNase-treated as it is expected that the doublets formed would not be RNase-resistant.

6.6 Conclusions and Future Directions

While we have successfully packaged a self-replicating RNA molecule into spherical and rod-like VLPs, more work remains to further characterize and test these particles. For example, CCMV and TMV VLPs can be compared in their packaging efficiency, thermal stability, and RNA protection. Additionally, these particles should be tested in mammalian cell lines and animal models to assess their ability to release their genetic material in cells. While both particles can be used for delivery of therapeutic genes, it would be interesting to study how their shape (sphere or rod) influences the efficacy of the particles.

CHAPTER 7

Summary and perspective on future work

This work explores *in vitro* self-assembly of brome mosaic virus (BMV), cowpea chlorotic mottle virus (CCMV) and tobacco mosaic virus (TMV) in their ability to package and protect self-replicating RNAs. In Chapters 2 and 3, one-step *in vitro* self-assembly of BMV virus-like particles (VLPs) is explored to investigate the roles of RNA-capsid protein (CP) and CP-CP attraction strength, controlled by ionic strength and pH solution conditions, on the assembly of well-formed, spherical VLPs. Additional interferometric scattering (iSCAT) experiments (presented in Chapter 2) could be performed at some of the pH and ionic strength conditions explored in Chapter 3, furthering our understanding of the interplay between RNA-CP and CP-CP interactions in successful virus particle assembly.

Attempts to better understand the *in vitro* self-assembly process have as a goal the formation of more robust, monodisperse VLPs that we can use to package therapeutic, self-replicating mRNAs. The self-replicating mRNA, or replicon, introduced in Chapter 4 contains a reporter gene as the gene of interest (GOI); however, any GOI may be inserted into the construct. For example, replicons have been made to contain other reporter genes (EYFP, mCherry, Renilla luciferase), an ovalbumin epitope (a model antigen for testing vaccines)[67], and microRNA. Self-replication of these genes of interest is strong when transfected into mammalian cells, but more research needs to be conducted on the therapeutic ability of these replicons when packaged in VLP form.

Experiments to pre-compact long RNAs so they fit into RNase-resistant, single 28-nm diameter VLPs made of BMV CP, presented in Chapter 5, resulted in only a marginal increase in

the number of singlets observed. High concentrations of spermine should be tested to see if enough charges on the RNA can be screened such that the molecule is compact enough to fit into a 28-nm capsid. However, spermine concentrations that are too high may not result in successful assembly because negative charges are necessary on the RNA in order for BMV CP to bind. Nevertheless, assemblies with higher spermine concentrations should be tested as they could increase the fraction of singlet particles observed. Additionally, other polyvalent cations can be tested like spermidine (+3 charge) or cobalt hexamine (+3 charge). Furthermore, pre-compacted RNAs packaged in VLPs need to be transfected into mammalian cells to assess whether the genetic material is released from VLPs and if the genetic material is able to be translated despite being bound by spermine.

If packaging overlong RNAs into spherical VLPs by pre-compacting with spermine proves to be unsuccessful, TMV rod-like VLPs may serve as an alternative delivery vehicle to CCMV or BMV VLPs. Chapter 6 introduces the idea of packaging the same replicon into spherical and rod-like VLPs using CCMV CP and TMV CP, respectively. While we have demonstrated the ability to package the same RNA molecule in these two differently shaped particles, more work needs to be done to compare and test these particles in their ability to package RNA efficiently, protect the packaged genetic material, and delivery genetic material to cells. Future experiments with our collaborators in the lab of Professor Nicole Steinmetz at University of California, San Diego aim to test these replicon-containing spherical and rod-like VLPs in mammalian cell culture and mice.

Bibliography

(1) Ahlquist, P.; Janda, M. CDNA Cloning and in Vitro Transcription of the Complete Brome Mosaic Virus Genome. *Molecular and cellular biology* **1984**.

(2) Allison, R. F.; Janda, M.; Ahlquist, P. Infectious in Vitro Transcripts from Cowpea Chlorotic Mottle Virus CDNA Clones and Exchange of Individual RNA Components with Brome Mosaic Virus. *J Virol* **1988**, *62* (10), 3581–3588. https://doi.org/10.1128/JVI.62.10.3581-3588.1988.

(3) Bancroft, J. B. The Self-Assembly of Spherical Plant Viruses. *Advances in Virus Research* **1970**, *16* (C), 99–134. https://doi.org/10.1016/S0065-3527(08)60022-6.

(4) Bancroft, J. B.; Bracker, C. E.; Wagner, G. W. Structures Derived from Cowpea Chlorotic Mottle and Brome Mosaic Virus Protein. *Virology* **1969**, *38* (2), 324–335. https://doi.org/10.1016/0042-6822(69)90374-2.

(5) Hiebert, E.; Bancroft, J. B.; Bracker, C. E. The Assembly in Vitro of Some Small Spherical Viruses, Hybrid Viruses, and Other Nucleoproteins. *Virology* **1968**, *34* (3), 492–508. https://doi.org/10.1016/0042-6822(68)90069-X.

(6) Cadena-Nava, R. D.; Comas-Garcia, M.; Garmann, R. F.; Rao, A. L. N.; Knobler, C. M.; Gelbart, W. M. Self-Assembly of Viral Capsid Protein and RNA Molecules of Different Sizes: Requirement for a Specific High Protein/RNA Mass Ratio. *Journal of Virology* **2012**, *86* (6), 3318–3326. https://doi.org/10.1128/JVI.06566-11.

(7) Cadena-Nava, R. D.; Hu, Y.; Garmann, R. F.; Ng, B.; Zelikin, A. N.; Knobler, C. M.; Gelbart, W. M. Exploiting Fluorescent Polymers to Probe the Self-Assembly of Virus-like Particles. *J Phys Chem B* **2011**, *115* (10), 2386–2391. https://doi.org/10.1021/jp1094118.

(8) Rao, A. L. N.; Cadena-Nava, R. D.; Gelbart, W. M.; Knobler, C. M.; Comas-Garcia, M. In Vitro Quantification of the Relative Packaging Efficiencies of Single-Stranded RNA Molecules by Viral Capsid Protein. *Journal of Virology* **2012**. https://doi.org/10.1128/jvi.01695-12.

(9) Hu, Y.; Zandi, R.; Anavitarte, A.; Knobler, C. M.; Gelbart, W. M. Packaging of a Polymer by a Viral Capsid: The Interplay between Polymer Length and Capsid Size. *Biophysical Journal* **2008**, *94* (4), 1428–1436. https://doi.org/10.1529/biophysj.107.117473.

(10) Garmann, R. F.; Comas-Garcia, M.; Gopal, A.; Knobler, C. M.; Gelbart, W. M. The Assembly Pathway of an Icosahedral Single-Stranded RNA Virus Depends on the Strength of Inter-Subunit Attractions. *Journal of Molecular Biology* **2014**, *426* (5), 1050–1060. https://doi.org/10.1016/J.JMB.2013.10.017.

(11) Garmann, R. F.; Comas-Garcia, M.; Koay, M. S. T.; Cornelissen, J. J. L. M.; Knobler, C. M.; Gelbart, W. M. Role of Electrostatics in the Assembly Pathway of a Single-Stranded RNA Virus. *Journal of Virology* **2014**, *88* (18), 10472. https://doi.org/10.1128/JVI.01044-14.

(12) Garmann, R. F.; Comas-Garcia, M.; Knobler, C. M.; Gelbart, W. M. Physical Principles in the Self-Assembly of a Simple Spherical Virus. *Accounts of Chemical Research* **2016**, *49* (1), 48–55. https://doi.org/10.1021/acs.accounts.5b00350.

(13) Cuillel, M.; Berthet-Colominas, C.; Timmins, P. A.; Zulauf, M. Reassembly of Brome Mosaic Virus from Dissociated Virus - A Neutron Scattering Study. *European Biophysics Journal* **1987**. https://doi.org/10.1007/BF00263681.

(14) Franklin, R. E.; Holmes, K. C. Tobacco Mosaic Virus: Application of the Method of Isomorphous Replacement to the Determination of the Helical Parameters and Radial

Density Distribution. *Acta Crystallographica* **1958**, *11* (3), 213–220. https://doi.org/10.1107/S0365110X58000529.
(15) Namba, K.; Pattanayek, R.; Stubbs, G. Visualization of Protein-Nucleic Acid Interactions in a Virus. *Journal of Molecular Biology* **1989**, *208* (2), 307–325. https://doi.org/10.1016/0022-2836(89)90391-4.
(16) Ge, P.; Zhou, Z. H. Hydrogen-Bonding Networks and RNA Bases Revealed by Cryo Electron Microscopy Suggest a Triggering Mechanism for Calcium Switches. *Proceedings of the National Academy of Sciences of the United States of America* **2011**, *108* (23), 9637–9642. https://doi.org/10.1073/pnas.1018104108.
(17) Butler, P. J. G. Self-Assembly of Tobacco Mosaic Virus: The Role of an Intermediate Aggregate in Generating Both Specificity and Speed. *Philosophical Transactions of the Royal Society B: Biological Sciences* **1999**, *354* (1383), 537–550. https://doi.org/10.1098/rstb.1999.0405.
(18) Wege, C.; Koch, C. From Stars to Stripes: RNA-Directed Shaping of Plant Viral Protein Templates—Structural Synthetic Virology for Smart Biohybrid Nanostructures. *WIREs Nanomedicine and Nanobiotechnology* **2020**, *12* (2), e1591. https://doi.org/10.1002/wnan.1591.
(19) Venter, P. A.; Schneemann, A. Nodaviruses. In *Encyclopedia of Virology*; Mahy, B. W. J., van Regenmortel, M. H. V., Eds.; Academic Press, 2008; pp 430–438.
(20) Crick, F. H. C.; Watson, J. D. Structure of Small Viruses. *Nature* **1956**, *177* (4506), 473–475. https://doi.org/10.1038/177473a0.
(21) Caspar, D. L.; Klug, A. Physical Principles in the Construction of Regular Viruses. *Cold Spring Harbor symposia on quantitative biology* **1962**, *27*, 1–24. https://doi.org/10.1101/SQB.1962.027.001.005.
(22) Wolf, Y. I.; Kazlauskas, D.; Iranzo, J.; Lucía-Sanz, A.; Kuhn, J. H.; Krupovic, M.; Dolja, V. V.; Koonin, E. V. Origins and Evolution of the Global RNA Virome. *mBio* **2018**, *9* (6). https://doi.org/10.1128/MBIO.02329-18/ASSET/F6B6021E-A7FB-4E6F-A45B-A03C3274CB66/ASSETS/GRAPHIC/MBO0061842030010.JPEG.
(23) Johnson, J. E.; Speir, J. A. Quasi-Equivalent Viruses: A Paradigm for Protein Assemblies. *J Mol Biol* **1997**, *269* (5), 665–675. https://doi.org/10.1006/jmbi.1997.1068.
(24) Twarock, R.; Luque, A. Structural Puzzles in Virology Solved with an Overarching Icosahedral Design Principle. *Nat Commun* **2019**, *10* (1), 4414. https://doi.org/10.1038/s41467-019-12367-3.
(25) Bancroft, J. B.; Hiebert, E. Formation of an Infectious Nucleoprotein from Protein and Nucleic Acid Isolated from a Small Spherical Virus. *Virology* **1967**, *32* (2), 354–356. https://doi.org/10.1016/0042-6822(67)90284-x.
(26) Bancroft, J. B.; Hills, G. J.; Markham, R. A Study of the Self-Assembly Process in a Small Spherical Virus Formation of Organized Structures from Protein Subunits in Vitro. *Virology* **1967**, *31* (2), 354–379. https://doi.org/10.1016/0042-6822(67)90180-8.
(27) Sugiyama, T.; Hebert, R. R.; Hartman, K. A. Ribonucleoprotein Complexes Formed between Bacteriophage MS2 RNA and MS2 Protein in Vitro. *Journal of Molecular Biology* **1967**, *25* (3), 455–463. https://doi.org/10.1016/0022-2836(67)90198-2.
(28) Dykeman, E. C.; Stockley, P. G.; Twarock, R. Solving a Levinthal's Paradox for Virus Assembly Identifies a Unique Antiviral Strategy. *Proceedings of the National Academy of Sciences* **2014**, *111* (14), 5361–5366. https://doi.org/10.1073/pnas.1319479111.

(29) Twarock, R.; Bingham, R. J.; Dykeman, E. C.; Stockley, P. G. A Modelling Paradigm for RNA Virus Assembly. *Current Opinion in Virology* **2018**, *31*, 74–81. https://doi.org/10.1016/J.COVIRO.2018.07.003.
(30) Dragnea, B. Watching a Virus Grow. *Proceedings of the National Academy of Sciences of the United States of America*, 2019, *116*, 22420–22422. https://doi.org/10.1073/pnas.1915986116.
(31) Borodavka, A.; Tuma, R.; Stockley, P. G. Evidence That Viral RNAs Have Evolved for Efficient, Two-Stage Packaging. *Proceedings of the National Academy of Sciences of the United States of America* **2012**, *109* (39), 15769–15774. https://doi.org/10.1073/pnas.1204357109.
(32) Borodavka, A.; Tuma, R.; Stockley, P. G. A Two-Stage Mechanism of Viral RNA Compaction Revealed by Single Molecule Fluorescence. *RNA Biology* **2013**, *10* (4), 481–489. https://doi.org/10.4161/rna.23838.
(33) Chevreuil, M.; Law-Hine, D.; Chen, J.; Bressanelli, S.; Combet, S.; Constantin, D.; Degrouard, J.; Möller, J.; Zeghal, M.; Tresset, G. Nonequilibrium Self-Assembly Dynamics of Icosahedral Viral Capsids Packaging Genome or Polyelectrolyte. *Nature Communications* **2018**, *9* (1), 1–9. https://doi.org/10.1038/s41467-018-05426-8.
(34) Ortega-Arroyo, J.; Kukura, P. Interferometric Scattering Microscopy (ISCAT): New Frontiers in Ultrafast and Ultrasensitive Optical Microscopy. *Phys Chem Chem Phys* **2012**, *14* (45), 15625–15636. https://doi.org/10.1039/c2cp41013c.
(35) Garmann, R. F.; Goldfain, A. M.; Manoharan, V. N. Measurements of the Self-Assembly Kinetics of Individual Viral Capsids around Their RNA Genome. *Proceedings of the National Academy of Sciences of the United States of America* **2019**, *116* (45), 22485–22490. https://doi.org/10.1073/pnas.1909223116.
(36) F. Garmann, R.; Sportsman, R.; Beren, C.; N. Manoharan, V.; M. Knobler, C.; M. Gelbart, W. A Simple RNA-DNA Scaffold Templates the Assembly of Monofunctional Virus-Like Particles. *Journal of the American Chemical Society* **2015**, *137* (24), 7584–7587. https://doi.org/10.1021/jacs.5b03770.
(37) Ni, P.; Wang, Z.; Ma, X.; Das, N. C.; Sokol, P.; Chiu, W.; Dragnea, B.; Hagan, M.; Cheng Kao, C. An Examination of the Electrostatic Interactions between the N-Terminal Tail of the Brome Mosaic Virus Coat Protein and Encapsidated RNAs. *Journal of Molecular Biology* **2012**. https://doi.org/10.1016/j.jmb.2012.03.023.
(38) Beren, C.; Cui, Y.; Chakravarty, A.; Yang, X.; Rao, A. L. N.; Knobler, C. M.; Hong Zhou, Z.; Gelbart, W. M. Genome Organization and Interaction with Capsid Protein in a Multipartite RNA Virus. *Proceedings of the National Academy of Sciences of the United States of America* **2020**, *117* (20), 10673–10680. https://doi.org/10.1073/pnas.1915078117.
(39) Annamalai, P.; Rao, A. L. N. Dispensability of 3' TRNA-like Sequence for Packaging Cowpea Chlorotic Mottle Virus Genomic RNAs. *Virology* **2005**, *332* (2), 650–658. https://doi.org/10.1016/J.VIROL.2004.12.009.
(40) Stockley, P. G. Filter-Binding Assays. *Methods in molecular biology (Clifton, N.J.)* **1994**, *30*, 251–262. https://doi.org/10.1385/0-89603-256-6:251.
(41) Incardona, N. L.; Kaesberg, P. A PH-INDUCED STRUCTURAL CHANGE IN BROMEGRASS MOSAIC VIRUS. *Biophysical journal* **1964**, *4* (1 Pt 1), 11–21. https://doi.org/10.1016/S0006-3495(64)86766-7.

(42) Hema, M.; Murali, A.; Ni, P.; Vaughan, R. C.; Fujisaki, K.; Tsvetkova, I.; Dragnea, B.; Cheng Kao, C. Effects of Amino-Acid Substitutions in the Brome Mosaic Virus Capsid Protein on RNA Encapsidation. */ 1433 MPMI* **2010**, *23* (11), 1433–1447. https://doi.org/10.1094/MPMI.
(43) Hagan, M. F. A Theory for Viral Capsid Assembly around Electrostatic Cores. **2009**. https://doi.org/10.1063/1.3086041.
(44) Stockley, P. G.; Twarock, R.; Bakker, S. E.; Barker, A. M.; Borodavka, A.; Dykeman, E.; Ford, R. J.; Pearson, A. R.; Phillips, S. E. V.; Ranson, N. A.; Tuma, R. Packaging Signals in Single-Stranded RNA Viruses: Nature's Alternative to a Purely Electrostatic Assembly Mechanism. *Journal of Biological Physics* **2013**, *39* (2), 277. https://doi.org/10.1007/S10867-013-9313-0.
(45) Dai, X.; Li, Z.; Lai, M.; Shu, S.; Du, Y.; Zhou, Z. H.; Sun, R. In Situ Structures of the Genome and Genome-Delivery Apparatus in a Single-Stranded RNA Virus. *Nature* **2017**, *541* (7635), 112–116. https://doi.org/10.1038/nature20589.
(46) Song, Y.; Gorbatsevych, O.; Liu, Y.; Mugavero, J.; Shen, S. H.; Ward, C. B.; Asare, E.; Jiang, P.; Paul, A. V.; Mueller, S.; Wimmer, E. Limits of Variation, Specific Infectivity, and Genome Packaging of Massively Recoded Poliovirus Genomes. *Proceedings of the National Academy of Sciences* **2017**, *114* (41), E8731–E8740. https://doi.org/10.1073/pnas.1714385114.
(47) Perlmutter, J. D.; Perkett, M. R.; Hagan, M. F. Pathways for Virus Assembly around Nucleic Acids. *Journal of Molecular Biology* **2014**, *426* (18), 3148–3165. https://doi.org/10.1016/j.jmb.2014.07.004.
(48) Singaram, S. W.; Garmann, R. F.; Knobler, C. M.; Gelbart, W. M.; Ben-Shaul, A. Role of RNA Branchedness in the Competition for Viral Capsid Proteins. *Journal of Physical Chemistry B* **2015**, *119* (44). https://doi.org/10.1021/acs.jpcb.5b06445.
(49) Speir, J. A.; Munshi, S.; Wang, G.; Baker, T. S.; Johnson, J. E. Structures of the Native and Swollen Forms of Cowpea Chlorotic Mottle Virus Determined by X-Ray Crystallography and Cryo-Electron Microscopy. *Structure* **1995**, *3* (1), 63–78. https://doi.org/10.1016/S0969-2126(01)00135-6.
(50) Zlotnick, A.; Aldrich, R.; Johnson, J. M.; Ceres, P.; Young, M. J. Mechanism of Capsid Assembly for an Icosahedral Plant Virus. *Virology* **2000**, *277* (2), 450–456. https://doi.org/10.1006/VIRO.2000.0619.
(51) Panahandeh, S.; Li, S.; Marichal, L.; Leite Rubim, R.; Tresset, G.; Zandi, R. How a Virus Circumvents Energy Barriers to Form Symmetric Shells. *ACS Nano* **2020**, *14* (3), 3170–3180. https://doi.org/10.1021/acsnano.9b08354.
(52) Panahandeh, S.; Li, S.; Dragnea, B.; Zandi, R. Virus Assembly Pathways Inside a Host Cell. *ACS Nano* **2022**, *16* (1), 317–327. https://doi.org/10.1021/acsnano.1c06335.
(53) Bond, K.; Tsvetkova, I. B.; Wang, J. C. Y.; Jarrold, M. F.; Dragnea, B. Virus Assembly Pathways: Straying Away but Not Too Far. *Small* **2020**, *16* (51), 1–7. https://doi.org/10.1002/smll.202004475.
(54) de Martín Garrido, N.; Crone, M. A.; Ramlaul, K.; Simpson, P. A.; Freemont, P. S.; Aylett, C. H. S. Bacteriophage MS2 Displays Unreported Capsid Variability Assembling T = 4 and Mixed Capsids. *Mol Microbiol* **2020**, *113* (1), 143–152. https://doi.org/10.1111/mmi.14406.

(55) Biela, A. P.; Naskalska, A.; Fatehi, F.; Twarock, R.; Heddle, J. G. Programmable Polymorphism of a Virus-like Particle. *Commun Mater* **2022**, *3*, 7. https://doi.org/10.1038/s43246-022-00229-3.
(56) Chang, J.-Y.; Gorzelnik, K. V.; Thongchol, J.; Zhang, J. Structural Assembly of Qβ Virion and Its Diverse Forms of Virus-like Particles. *Viruses* **2022**, *14* (2), 225. https://doi.org/10.3390/v14020225.
(57) Ruszkowski, M.; Strugala, A.; Indyka, P.; Tresset, G.; Figlerowicz, M.; Urbanowicz, A. Cryo-EM Reconstructions of BMV-Derived Virus-like Particles Reveal Assembly Defects in the Icosahedral Lattice Structure. *Nanoscale* **2022**, *14* (8), 3224–3233. https://doi.org/10.1039/D1NR05650F.
(58) Asensio, M. A.; Morella, N. M.; Jakobson, C. M.; Hartman, E. C.; Glasgow, J. E.; Sankaran, B.; Zwart, P. H.; Tullman-Ercek, D. A Selection for Assembly Reveals That a Single Amino Acid Mutant of the Bacteriophage MS2 Coat Protein Forms a Smaller Virus-like Particle. *Nano Lett.* **2016**, *16* (9), 5944–5950. https://doi.org/10.1021/acs.nanolett.6b02948.
(59) Berthet-Colominas, C.; Cuillel, M.; Koch, M. H. J.; Vachette, P.; Jacrot, B. Kinetic Study of the Self-Assembly of Brome Mosaic Virus Capsid. *European Biophysics Journal* **1987**. https://doi.org/10.1007/BF00263680.
(60) Klug, A. The Tobacco Mosaic Virus Particle: Structure and Assembly. *Philosophical Transactions of the Royal Society B: Biological Sciences* **1999**, *354* (1383), 531–535. https://doi.org/10.1098/rstb.1999.0404.
(61) Culver, J. N. TOBACCO MOSAIC VIRUS ASSEMBLY AND DISASSEMBLY: Determinants in Pathogenicity and Resistance. *Annual Review of Phytopathology* **2002**, *40* (1), 287–308. https://doi.org/10.1146/annurev.phyto.40.120301.102400.
(62) Schneider, C. A.; Rasband, W. S.; Eliceiri, K. W. NIH Image to ImageJ: 25 Years of Image Analysis. *Nat Methods* **2012**, *9* (7), 671–675. https://doi.org/10.1038/nmeth.2089.
(63) Running, W. E.; Ni, P.; Kao, C. C.; Reilly, J. P. Chemical Reactivity of Brome Mosaic Virus Capsid Protein. *Journal of Molecular Biology* **2012**, *423* (1), 79–95. https://doi.org/10.1016/j.jmb.2012.06.031.
(64) Tews, B. A.; Meyers, G. Chapter 2 Self-Replicating RNA. *Methods in Molecular Biology 1499*. https://doi.org/10.1007/978-1-4939-6481-9_2.
(65) Damase, T. R.; Sukhovershin, R.; Boada, C.; Taraballi, F.; Pettigrew, R. I.; Cooke, J. P. The Limitless Future of RNA Therapeutics. *Frontiers in Bioengineering and Biotechnology* **2021**, *9*, 161. https://doi.org/10.3389/FBIOE.2021.628137/BIBTEX.
(66) Schott, J.; Stoecklin, G. Networks Controlling MRNA Decay in the Immune System. *Wiley interdisciplinary reviews. RNA* **2010**, *1* (3), 432–456. https://doi.org/10.1002/WRNA.13.
(67) Biddlecome, A.; Habte, H. H.; McGrath, K. M.; Sambanthamoorthy, S.; Wurm, M.; Sykora, M. M.; Knobler, C. M.; Lorenz, I. C.; Lasaro, M.; Elbers, K.; Gelbart, W. M. Delivery of Self-Amplifying RNA Vaccines in in Vitro Reconstituted Virus-like Particles. *PLoS ONE* **2019**. https://doi.org/10.1371/journal.pone.0215031.
(68) Gitlin, L.; Hagai, T.; LaBarbera, A.; Solovey, M.; Andino, R. Rapid Evolution of Virus Sequences in Intrinsically Disordered Protein Regions. *PLOS Pathogens* **2014**, *10* (12), e1004529. https://doi.org/10.1371/JOURNAL.PPAT.1004529.
(69) Ball, L. A.; Amann, J. M.; Garrett, B. K. *Replication of Nodamura Virus after Transfection of Viral RNA into Mammalian Cells in Culture*; 1992; pp 2326–2334.

(70) Miller, D. J.; Ahlquist, P. Flock House Virus RNA Polymerase Is a Transmembrane Protein with Amino-Terminal Sequences Sufficient for Mitochondrial Localization and Membrane Insertion. *Journal of Virology* **2002**, *76* (19), 9856–9867. https://doi.org/10.1128/JVI.76.19.9856-9867.2002/ASSET/FEF6EBCF-253B-41DD-9F79-0A2E2B2A4E68/ASSETS/GRAPHIC/JV1920403009.JPEG.
(71) Ball, L. A. *Requirements for the Self-Directed Replication of Flock House Virus RNA 1*; 2; 1995; Vol. 69, pp 720–727.
(72) Venter, P. A.; Schneemann, A. Review Recent Insights into the Biology and Biomedical Applications of Flock House Virus. https://doi.org/10.1007/s00018-008-8037-y.
(73) Schlesinger, S. Alphavirus Vectors for Gene Expression and Vaccines. *Current Opinion in Biotechnology* **1999**, *10* (5), 434–439. https://doi.org/10.1016/S0958-1669(99)00006-3.
(74) Lindenbach, B. D.; Sgro, J.-Y.; Ahlquist, P. Long-Distance Base Pairing in Flock House Virus RNA1 Regulates Subgenomic RNA3 Synthesis and RNA2 Replication. *Journal of Virology* **2002**, *76* (8), 3905–3919. https://doi.org/10.1128/JVI.76.8.3905-3919.2002/ASSET/A1CC6114-BE57-4352-9F97-A26EDB74C3CC/ASSETS/GRAPHIC/JV0822141008.JPEG.
(75) Johnson, K. L.; Ball, L. A. Induction and Maintenance of Autonomous Flock House Virus RNA1 Replication. *Journal of Virology* **1999**, *73* (10), 7933. https://doi.org/10.1128/jvi.73.10.7933-7942.1999.
(76) Malone, R. W.; Felgner, P. L.; Verma, I. M. Cationic Liposome-Mediated RNA Transfection. *Proceedings of the National Academy of Sciences of the United States of America* **1989**, *86* (16), 6077. https://doi.org/10.1073/PNAS.86.16.6077.
(77) Leonhardt, C.; Schwake, G.; Stögbauer, T. R.; Rappl, S.; Kuhr, J. T.; Ligon, T. S.; Rädler, J. O. Single-Cell MRNA Transfection Studies: Delivery, Kinetics and Statistics by Numbers. *Nanomedicine: Nanotechnology, Biology, and Medicine* **2014**, *10* (4), 679–688. https://doi.org/10.1016/j.nano.2013.11.008.
(78) Agapov, E. V.; Frolov, I.; Lindenbach, B. D.; Pra´gaipra´gai, B. M.; Schlesinger, S.; Rice, C. M. Noncytopathic Sindbis Virus RNA Vectors for Heterologous Gene Expression. *Cell Biology* **1998**, *95*, 12989–12994.
(79) Frolov, I.; Agapov, E.; Hoffman, T. A.; Prágai, B. M.; Lippa, M.; Schlesinger, S.; Rice, C. M. Selection of RNA Replicons Capable of Persistent Noncytopathic Replication in Mammalian Cells. *Journal of virology* **1999**, *73* (5), 3854–3865. https://doi.org/10.1128/JVI.73.5.3854-3865.1999.
(80) Barreau, C.; Dutertre, S. P.; Paillard, L.; Osborne, H. B. Liposome-Mediated RNA Transfection Should Be Used with Caution. https://doi.org/10.1261/rna.191706.
(81) Brito, L. A.; Kommareddy, S.; Maione, D.; Uematsu, Y.; Giovani, C.; Berlanda Scorza, F.; Otten, G. R.; Yu, D.; Mandl, C. W.; Mason, P. W.; Dormitzer, P. R.; Ulmer, J. B.; Geall, A. J. Self-Amplifying MRNA Vaccines. *Advances in genetics* **2015**, *89*, 179–233. https://doi.org/10.1016/BS.ADGEN.2014.10.005.
(82) Lundstrom, K. Replicon RNA Viral Vectors as Vaccines. *Vaccines 2016, Vol. 4, Page 39* **2016**, *4* (4), 39. https://doi.org/10.3390/VACCINES4040039.
(83) Khromykh, A. A. Replicon-Based Vectors of Positive Strand RNA Viruses. *Current Opinion in Molecular Therapeutics* **2000**, *2*, 555–569.
(84) Blasey, H. D.; Lundström, K.; Tate, S.; Bernard, A. R. *Recombinant Protein Production Using the Semliki Forest Virus Expression System*; Kluwer Academic Publishers, 1997; Vol. 24, pp 65–72.

(85) Schlaeger, E. J.; Lundstrom, K. Effect of Temperature on Recombinant Protein Expression in Semliki Forest Virus Infected Mammalian Cell Lines Growing in Serum- Free Suspension Cultures. *Cytotechnology* **1998**, *28* (1–3), 205. https://doi.org/10.1023/A:1008006504379.
(86) Rohovie, M. J.; Nagasawa, M.; Swartz, J. R. Virus-like Particles: Next-generation Nanoparticles for Targeted Therapeutic Delivery. *Bioeng Transl Med* **2017**, *2* (1), 43–57. https://doi.org/10.1002/btm2.10049.
(87) Chung, Y. H.; Cai, H.; Steinmetz, N. F. Viral Nanoparticles for Drug Delivery, Imaging, Immunotherapy, and Theranostic Applications. *Adv Drug Deliv Rev* **2020**, *156*, 214–235. https://doi.org/10.1016/j.addr.2020.06.024.
(88) Tanimoto, C. R.; Thurm, A. R.; Brandt, D. S.; Knobler, C. M.; Gelbart, W. M. The Nonmonotonic Dose Dependence of Protein Expression in Cells Transfected with Self-Amplifying RNA. *Journal of Virology* **2022**, *96* (7), e01858-21. https://doi.org/10.1128/jvi.01858-21.
(89) Hud, N. V.; Downing, K. H. Cryoelectron Microscopy of λ Phage DNA Condensates in Vitreous Ice: The Fine Structure of DNA Toroids. *Proceedings of the National Academy of Sciences* **2001**, *98* (26), 14925–14930. https://doi.org/10.1073/pnas.261560398.
(90) Pegg, A. E. Functions of Polyamines in Mammals. *J Biol Chem* **2016**, *291* (29), 14904–14912. https://doi.org/10.1074/jbc.R116.731661.
(91) Mandal, S.; Mandal, A.; Johansson, H. E.; Orjalo, A. V.; Park, M. H. Depletion of Cellular Polyamines, Spermidine and Spermine, Causes a Total Arrest in Translation and Growth in Mammalian Cells. *Proceedings of the National Academy of Sciences* **2013**, *110* (6), 2169–2174. https://doi.org/10.1073/pnas.1219002110.
(92) Mitra, S.; Kaesberg, P. Interaction of Polyamines with Turnip Yellow Mosaic Virus RNA. *Biochemical and Biophysical Research Communications* **1963**, *11* (2), 146–151. https://doi.org/10.1016/0006-291X(63)90082-2.
(93) Butler, P. J. G. Y. 1984. The Current Picture of the Structure and Assembly of Tobacco Mosaic Virus. *Journal of General Virology 65* (2), 253–279. https://doi.org/10.1099/0022-1317-65-2-253.