

# UC Riverside

## UC Riverside Electronic Theses and Dissertations

### Title

Expression, Purification, and Functional Reconstitution of SARS-CoV-2 Structural Proteins: Implications for Membrane Remodeling, RNA Packaging, and p38 MAPK Signaling Dynamics

### Permalink

<https://escholarship.org/uc/item/0gt1f7jz>

### ISBN

9798190929812

### Author

Anbir, Sara

### Publication Date

2026-05-18

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA  
RIVERSIDE

Expression, Purification, and Functional Reconstitution of SARS-CoV-2 Structural  
Proteins: Implications for Membrane Remodeling, RNA Packaging, and p38 MAPK  
Signaling Dynamics

A Dissertation submitted in partial satisfaction  
of the requirements for the degree of

Doctor of Philosophy

in

Biophysics

by

Sara Anbir

June 2026

Dissertation Committee:

Dr. Thomas E. Kuhlman, Chairperson  
Dr. Bahram Mobasher  
Dr. Marcus Kaul



The Dissertation of Sara Anbir is approved:

---

---

---

Committee Chairperson

University of California, Riverside



## Acknowledgments

I am deeply grateful for today and for the incredible people who have supported and guided me throughout this journey. I began my PhD in 2020, during the height of the COVID-19 pandemic and while my mother was battling cancer. The first two years of this path were filled with unimaginable challenges. I lost my mom, and life became very difficult—but I could not have made it through without the support of an extraordinary PI.

Dr. Kuhlman is not only a brilliant scientist but also a compassionate human being, and I will be forever thankful. Even if I wasn't always the perfect student, he remained an exceptional mentor. I also want to sincerely thank Dr. Sattar Taheri Araghi, without his encouragement, this mission could not be done.

To my family: my father, who always called me "Khanoom Doctor" and believed in me unconditionally; and my beloved mother, whose strength and wisdom remain with me every day. Though she is not here, she continues to inspire me with the values she taught me—how to be kind, resilient, and brave in the face of hardship.

To my dearest sister, Anda, who stood by me like a rock during the hardest moments of my life—your endless love and support have been my anchor, and I couldn't have done this without you.

To my Liam, who came into this world and reminded me that there is still so much beauty left to discover.

And to my husband, Mir Moein Nourinia—my heart and my safe place—thank you for walking into my life and changing it forever. Your love gave me light in my darkest

days. Your faith in me never wavered, even when I had doubts myself. You gave me peace, joy, and strength—and reminded me to live in the moment and embrace the beauty of every step. This journey was mine, but you walked it with me in every way that matters.

Lastly, to my closest friends—Roudabe, Shadi, Mahnaz, Neda, Peyman, Farzam, Mehrdad, Navid, Nazi, Pedram, Aria, Shayan, Yasi, Saeed, Rubi, Arash, Shabnam and Ayeen—you are more than friends; you are family. Thank you for standing by me.

## ABSTRACT OF THE DISSERTATION

Expression, Purification, and Functional Reconstitution of SARS-CoV-2 Structural Proteins: Implications for Membrane Remodeling, RNA Packaging, and p38 MAPK Signaling Dynamics

by

Sara Anbir

Doctor of Philosophy, Graduate Program in Biophysics  
University of California, Riverside, June 2026  
Dr. Thomas E.Kuhlman, Chairperson

The SARS-CoV-2 structural proteins — membrane (M), envelope (E), nucleocapsid (N), and spike (S) — orchestrate all stages of viral assembly, budding, and release, yet the molecular mechanisms by which they remodel host membranes and drive particle formation remain incompletely understood. This dissertation presents a series of biophysical and biochemical investigations into the structure, assembly, and membrane interactions of SARS-CoV-2 structural proteins, with the goal of elucidating the physical principles underlying coronavirus particle formation.

In the first study, a SUMO-tag-based *Escherichia coli* expression system was developed to produce large quantities of full-length, native M protein. Recombinant M protein dimers were inserted into lipid bilayers and characterized by atomic force microscopy (AFM), cryo-electron microscopy (cryo-EM), and molecular dynamics (MD) simulations. AFM and cryo-EM imaging revealed M protein dimers exclusively in the compact “short” conformation, while MD simulations confirmed localized membrane thinning and curvature

induction in the vicinity of the dimers. Coarse-grained simulations further demonstrated that endodomain interactions among M protein dimers, combined with line tension from membrane thinning, can drive viral assembly and budding.

The expression platform developed for M protein was extended to the SARS-CoV-2 envelope (E) protein, yielding recombinant E protein in quantities sufficient for membrane reconstitution. Collaborative experiments using this purified E protein demonstrated that E forms multiple coexisting oligomeric species in ERGIC-mimetic lipid bilayers, induces membrane thinning and curvature, and drives concentration-dependent morphological transitions in giant vesicles. Real-time confocal microscopy further demonstrated that introduction of N protein–RNA ribonucleoprotein complexes to giant vesicles containing both M and E triggers rapid inward vesicle budding, establishing a minimal reconstituted system for SARS-CoV-2 particle assembly *in vitro*.

The second study examined the production and characterization of SARS-CoV-2 virus-like particles (VLPs) in mammalian cells. Expression of all four structural proteins at ratios reflecting viral mRNA abundance in infected cells yielded VLPs validated by Western blotting, AFM, and dynamic light scattering. RNA extracted from purified VLP fractions showed a size distribution dominated by small RNAs, consistent with selective packaging rather than nonspecific cellular RNA incorporation.

Finally, a study of p38 MAPK signal transduction demonstrated that the central human stress response kinase encodes information about experienced stresses as frequency oscillations of its activation state, and that downstream specificity is achieved through bioresonance between p38 and its target substrates.

# Contents

|                                                                                                                   |            |
|-------------------------------------------------------------------------------------------------------------------|------------|
| <b>List of Figures</b>                                                                                            | <b>xii</b> |
| <b>1 Introduction</b>                                                                                             | <b>1</b>   |
| 1.1 SARS-CoV-2 Structure and Viral Assembly                                                                       | 2          |
| 1.1.1 The Membrane (M) Protein                                                                                    | 3          |
| 1.2 Biophysical Approaches to Studying Viral Membrane Proteins                                                    | 4          |
| 1.2.1 The Envelope (E) Protein                                                                                    | 5          |
| 1.2.2 The Nucleocapsid (N) Protein and Genome Packaging                                                           | 6          |
| 1.2.3 Virus-Like Particles as a Research Tool                                                                     | 7          |
| 1.3 Cellular Stress Response and the MAPK Cascade                                                                 | 8          |
| 1.4 Dissertation Outline                                                                                          | 9          |
| <b>2 Paper I: Synthesis, Insertion, and Characterization of SARS-CoV-2 Membrane Protein Within Lipid Bilayers</b> | <b>11</b>  |
| 2.1 Introduction                                                                                                  | 12         |
| 2.2 Results                                                                                                       | 15         |
| 2.2.1 Expression and high-yield purification of SARS-CoV-2 M protein                                              | 15         |
| 2.2.2 Cryo-EM of lipid bilayer membrane reconstituted with M protein in LUVs                                      | 18         |
| 2.2.3 AFM topography of SBL reconstituted with embedded M protein                                                 | 20         |
| 2.2.4 Formation of M protein aggregate patches in the SBL                                                         | 24         |
| 2.2.5 Morphological statistics of the isolated M protein particle                                                 | 24         |
| 2.2.6 Confirmation of M protein insertion in the bilayer membrane                                                 | 27         |
| 2.2.7 Orientation of reconstituted M protein with C terminus facing the vesicle interior                          | 28         |
| 2.2.8 Reconstituted M protein dimensions and membrane thickness consistent with the short form                    | 29         |
| 2.2.9 Reduction of membrane thickness around M protein                                                            | 30         |
| 2.2.10 Membrane compression, M protein aggregation, and spontaneous membrane curvature                            | 32         |

|        |                                                                                         |    |
|--------|-----------------------------------------------------------------------------------------|----|
| 2.2.11 | Coarse-grained MD simulation of M protein–induced budding                               | 33 |
| 2.3    | DISCUSSION                                                                              | 35 |
| 2.4    | MATERIALS AND METHODS                                                                   | 38 |
| 2.4.1  | Strains and media                                                                       | 38 |
| 2.4.2  | Plasmid construction                                                                    | 39 |
| 2.4.3  | Cell growth and protein expression and purification                                     | 39 |
| 2.4.4  | Quantification of protein expression and microscopy                                     | 40 |
| 2.4.5  | Protein purification                                                                    | 40 |
| 2.4.6  | Protein refolding and dialysis                                                          | 42 |
| 2.4.7  | Cleavage of SUMO tags                                                                   | 42 |
| 2.4.8  | Quantification of yields and endotoxin levels                                           | 43 |
| 2.4.9  | SDS-PAGE and Western blots                                                              | 44 |
| 2.4.10 | Protein reconstitution into vesicles                                                    | 45 |
| 2.4.11 | Preparation of SBL samples                                                              | 46 |
| 2.4.12 | AFM imaging, force profile collection, and analysis                                     | 46 |
| 2.4.13 | Cryo-EM data collection and analysis                                                    | 47 |
| 2.4.14 | All-atom MD simulation                                                                  | 48 |
| 2.4.15 | Coarse-grained simulation                                                               | 50 |
| 2.5    | Acknowledgement                                                                         | 50 |
| 2.5.1  | funding                                                                                 | 50 |
| 2.5.2  | Author contributions                                                                    | 51 |
| 3      | SARS-CoV-2 Structural Proteins - Supplemental Information                               | 52 |
| 3.1    | Determination of Detergent Conditions for M Protein Insertion                           | 52 |
| 3.2    | Determination of Elastic Moduli from AFM Measurements                                   | 55 |
| 3.3    | Atomistic Molecular Dynamics (MD) Simulations of M-protein Stability                    | 57 |
| 3.4    | M Protein’s Impact on Membrane Thickness Through MD Simulations                         | 64 |
| 3.5    | Protein Rotation in MD Simulations                                                      | 66 |
| 3.6    | Impact of M Protein on Membrane Curvature in MD Simulations                             | 69 |
| 3.7    | Coarse-grained Molecular Dynamics Simulation of M Protein Effects on Membrane Curvature | 70 |
| 3.8    | Estimation of Line Tension from Membrane Thinning                                       | 71 |
| 4      | Paper II:VLP Assembly                                                                   | 73 |
| 4.1    | Introduction                                                                            | 75 |
| 4.2    | Materials and Methods                                                                   | 77 |
| 4.2.1  | Cells and culture conditions                                                            | 77 |
| 4.2.2  | DNA plasmid and cloning                                                                 | 79 |
| 4.2.3  | DNA transfection                                                                        | 80 |
| 4.2.4  | Collected VLPs and purification                                                         | 82 |
| 4.2.5  | DLS Analysis                                                                            | 82 |
| 4.2.6  | Atomic Force Microscopy                                                                 | 84 |
| 4.2.7  | Western Blot                                                                            | 88 |

|          |                                                                                                     |            |
|----------|-----------------------------------------------------------------------------------------------------|------------|
| 4.2.8    | RNA Incorporation in VLP                                                                            | 91         |
| 4.3      | Results and Discussion                                                                              | 93         |
| 4.3.1    | VLP Characterization                                                                                | 93         |
| 4.3.2    | RNA Packaging Analysis                                                                              | 95         |
| 4.4      | Conclusion                                                                                          | 98         |
| <b>5</b> | <b>Paper III:Human Stress Response Specificity through Bioresonance Selectivity</b>                 | <b>100</b> |
| 5.1      | introduction                                                                                        | 101        |
| 5.2      | Result                                                                                              | 103        |
| 5.2.1    | VPC directly reports p38 activation state by FRET                                                   | 103        |
| 5.2.2    | p38 and the TF-based p38 substrate PerKy-38 exhibit activation state oscillations                   | 106        |
| 5.2.3    | p38 activation state oscillates with stimulus-specific frequencies                                  | 109        |
| 5.2.4    | p38 activates substrates through bioresonance                                                       | 111        |
| 5.2.5    | Activation of specific genetic responses through dynamic bioresonant stimulation                    | 113        |
| 5.2.6    | A theoretical framework describing bioresonance                                                     | 117        |
| 5.3      | Discussion                                                                                          | 119        |
| 5.4      | Acknowledgement                                                                                     | 122        |
| 5.4.1    | funding                                                                                             | 122        |
| 5.4.2    | Author contributions                                                                                | 123        |
| <b>6</b> | <b>Human Stress Response Specificity through Bioresonance Selectivity-Supplementary Information</b> | <b>125</b> |
| 6.1      | Materials and Methods                                                                               | 125        |
| 6.1.1    | Plasmids, Constructs, and Proteins                                                                  | 125        |
| 6.1.2    | Cell lines and culturing                                                                            | 126        |
| 6.1.3    | Protein Purification and In Vitro FRET Assay                                                        | 127        |
| 6.1.4    | Western Blot Analysis                                                                               | 129        |
| 6.1.5    | Fluorescence Microscopy                                                                             | 131        |
| 6.1.6    | Pump System and Dynamic Stimulation                                                                 | 132        |
| 6.1.7    | Fourier Analysis                                                                                    | 132        |
| 6.1.8    | RNAseq                                                                                              | 133        |
| 6.2      | Supplementary Text                                                                                  | 135        |
| 6.2.1    | I. Fourier Analysis                                                                                 | 135        |
| 6.2.2    | II. Statistical Significance of Fourier Peaks                                                       | 136        |
| 6.2.3    | III. Analysis of Harmonics                                                                          | 137        |
| 6.2.4    | IV. Short Time Fourier Transform (STFT)                                                             | 143        |
| 6.2.5    | V. Forced Oscillations, Resonance, and Resonance Selectivity                                        | 145        |
| 6.2.6    | VI. Bioresonance                                                                                    | 148        |
| 6.2.7    | VII. Linear Response Theory and the p38/PerKy-38 Transfer Function                                  | 152        |

|          |                                                                                             |            |
|----------|---------------------------------------------------------------------------------------------|------------|
| 6.2.8    | <b>VIII. Bioresonant Transcriptional Regulation</b>                                         | 155        |
| 6.2.9    | <b>IX. Simulations of p38 bioresonance selectivity</b>                                      | 161        |
| 6.2.10   | <b>X. Resonance and Impedance in Alternating Current (AC) Circuits</b>                      | 164        |
| 6.2.11   | <b>The MAPK Cascade as Oscillating Phosphate Current Amplifier and Distribution Network</b> | 165        |
| <b>7</b> | <b>Conclusions</b>                                                                          | <b>169</b> |
| 7.1      | Summary of Principal Findings                                                               | 170        |
| 7.1.1    | High-Yield Expression and Membrane Reconstitution of the M Protein                          | 170        |
| 7.1.2    | Extension of the Purification Platform to E Protein and Collaborative Assembly Studies      | 171        |
| 7.1.3    | Optimized VLP Production and RNA Packaging                                                  | 173        |
| 7.1.4    | Bioresonance as a Mechanism of Signaling Specificity                                        | 174        |
| 7.2      | Cross-Cutting Themes                                                                        | 175        |
| 7.3      | Future Directions                                                                           | 176        |
| 7.4      | Closing Remarks                                                                             | 178        |



# List of Figures

|                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |    |
|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| <b>2.1 Membrane protein expression in <i>E. coli</i>.</b>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |    |
| (A to C) Fluorescence microscopy of M-mCerulean3 expression in terrific broth (TB) medium, with phase contrast image shown above and mCerulean3 fluorescence channel shown below. (A) M-mCerulean3 expressed in BL21(DE3). (B) SUMO-M-mCerulean3 expressed in BL21(DE3). (C) SUMO-M-mCerulean3 expressed in C41(DE3). All strains were imaged with identical microscope and camera settings, and images are displayed with identical lookup tables. Scale bars, 10.0 $\mu\text{m}$ . (D) Quantitative comparison of M-mCerulean3 expression in various <i>E. coli</i> strains and media [white, lysogeny broth (LB); light blue, TB; SUMO-M-mCerulean3 in LB, medium blue; SUMO-M-mCerulean3 in TB, dark blue]. Expression levels are normalized to expression of M-mCerulean3 in BL21(DE3). . . . . | 16 |
| <b>2.2 Verification of purified M protein.</b>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |    |
| (A) SDS–polyacrylamide gel electrophoresis (PAGE) of purified M sample. Lane 1, ladder; lane 2, after digestion with 0.5 mg SUMO protease per milligram of M; lane 3, digestion with 2 $\times$ amount SUMO protease. (B) Ponceau all-protein stain of protein transferred to blot from PAGE gel. Lanes same as (A). (C) Western blot of (A). (D) MALDI mass spectrum of purified M sample. . . . .                                                                                                                                                                                                                                                                                                                                                                                                  | 17 |

|                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |    |
|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| <b>2.3 Typical cryo-EM micrographs of LUVs reconstituted with M proteins.</b>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |    |
| (A) Red arrows indicate M proteins inserted in the vesicle membrane. (B and C) Schematics of LUVs with and without M reconstitution, respectively. (D) Representative two-dimensional (2D) class average image of areas without M protein in reconstituted LUVs. (E and F) Two most representative cryo-EM 2D class average images from LUV bilayer membranes reconstituted with M protein mass ratio, $R_{m,M/L} = 0.015$ . Arrows, location of M protein as bright spots. (G to I) 2D class average cryo-EM images of blank LUV bilayer membrane. Scale bar, 5.0 nm. (J) Atom number density of the short conformation M protein embedded in a bilayer membrane obtained from all-atom MD.                                                                                                                                                                                                                                                                                                                                                                                                                                                              | 19 |
| <b>2.4 AFM height image of SBL reconstituted with M protein at increasing protein-to-lipid mass ratios, <math>R_{m,M/L}</math>, imaged on mica substrates.</b>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |    |
| (A to D) Height images and representative single line profiles (green dashed lines in images) collected from supported bilayer (SBL) fabricated from LUVs that were reconstituted with M protein at $R_{m,M/L} = 0, 0.01, 0.015$ , and $0.02$ . (E) Area occupation percentage of M protein in SBL as a function of $R_{m,M/L}$ calculated from (A) to (D). (F) Representative force profiles collected from pristine mica surface (black), SBL (green), and M protein (orange) patches in lipid bilayer; the force profiles were collected during tip engagement. Scale bars: 200 nm (A–D)                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               | 22 |
| <b>2.5 Size and morphology statistics of M protein quantified from AFM images of SBL.</b>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |    |
| (A) Representative AFM height image of individual M protein reconstituted in SBL at $R_{m,M/L} = 0.01$ . (B) Zoomed-in image of (A), highlighting the protein with red color and lipid bilayer with gray color. The protruded portion of a single-M protein particle above lipid bilayer (left) is detected by setting the height of top surface of SBL as a threshold and taking the above portion. (C) Top view of the short conformation M protein (C terminus in purple) embedded in a multicomponent membrane after $1 \mu s$ into an all-atom MD simulation. Dimensional parameters—length (L), width (W), and height (h)—were measured for detected protein particles. (D to G) Histograms of L, W, L/W, and h of protein protruding above lipid membrane. Solid lines are Gaussian fits. Vertical dashed lines mark corresponding values from MD simulation for short-form M protein reconstituted with C terminus exposed above SBL. (H to J) 3D histograms each showing combined statistical survey of two out of the three dimensions. Green and blue arrows indicate dimer and higher order of oligomers, respectively. Scale bar: 50 nm (A). | 25 |

**2.6 Analysis of membrane thickness around individual M protein particles at  $R_m, M/L = 0.01$  and through MD simulation.**

(A) Example of individual M protein particles reconstituted in SBL. Height profiles were taken along the green dashed lines starting from the center of the protein. (B) View from above of membrane thickness in nm from 500 ns to 1  $\mu$ s of an all-atom MD simulation for the short conformation M protein embedded in a multicomponent membrane. (C) Schematics of membrane thinning around M protein particles. (D) Side view of the MD simulation along the y axis, with the protein in purple, where average thickness is shown dependent on distance from the protein. (E) Solid green line shows the experimentally measured averaged height profile of membrane around individual M protein particles. Inset shows around 100 height profiles obtained from 25 individual protein particles. Here,  $x = 0$  nm is defined by the boundary between the M protein particle and lipid membrane. The blue points represent vertically shifted values of membrane thickness obtained through MD shown in (B). Scale bar: 25 nm (A). . . . .

31

**2.7 Coarse-grained simulations of M proteins embedded in a flat membrane.**

(A) Schematic illustration of M proteins embedded in lipid bilayers. The top figure shows the actual size of the lipid bilayer and M proteins [Protein Data Bank (PDB): 7VGR (Zhang et al. 2022)], and the bottom figure shows the coarse-grained model used in the simulation. (B) Schematic figure (left) and simulation snapshot (right) for the case that the M proteins have no endodomain interactions ( $\epsilon_{M3-M3} = 0$ ). The M proteins aggregate and form a small bump due to the line tension, but no further budding is observed in the simulation. (C) Schematic figure (left) and simulation snapshot (right) for the case that M proteins have endodomain interactions ( $\epsilon_{M3-M3} = 8kBT$ ). The interaction helps M proteins overcome the energy barrier and the membrane buds into a spherical shell. For each M protein, the domain particles have a diameter size of 1.0a (M1), 1.0a (M2), and 0.8a (M3), where a is the system length unit corresponding to 3 nm. For both simulations, the transmembrane domain M2 interacts with M2 with strength  $\epsilon_{M2-M2} = 4kBT$ . The membrane is represented as a triangular lattice where rigid particles (M0) occupy the vertices due to the incompressibility of the lipid membrane. The M0 particles interact with M with the strength  $\epsilon_{M-M} = 4kBT$ . The rest of particles have excluded volume repulsion defined through a cutoff Lennard-Jones potential. . . . .

34

|     |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |    |
|-----|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| 3.1 | <b>Dynamic light scattering (DLS) study of vesicle size as a function of detergent concentration.</b>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |    |
|     | (A) Intensity distribution vs. hydrodynamic size measured with DLS for as-extruded lipid vesicles in biological buffer. (B) Schematics of the three stages of vesicle dissolution with increasing detergent concentration. (C) and (D) are hydrodynamic size of vesicles measured as functions of TX-100 and DDM detergent concentrations, respectively. The empty and filled stars represent the detergent-saturation concentration, $C_{sat}$ and vesicle dissolution concentration, $C_{sol}$ . (E) and (F) are $C_{sat}$ and $C_{sol}$ observed at different lipid concentrations for TX-100 and DDM, respectively. . . . .                                                                                                           | 53 |
| 3.2 | <b>Elastic modulus of SBL and M protein patches obtained from nanoindentation technique.</b>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |    |
|     | (A) Histogram of Young's modulus obtained from analyzing the AFM nanoindentation force profiles using Hertzian contact model. (B) and (C) are the force profiles obtained from SBL and protein patches, respectively. The thick solid lines represent the force profile calculated using the mean elastic modulus values from (A). . . . .                                                                                                                                                                                                                                                                                                                                                                                                | 56 |
| 3.3 | <b>Initial Simulation Schematic for M Protein Embedded in a Multi-component <math>\sim 25</math> nm x 25 nm Membrane.</b>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |    |
|     | (A) A view from above looking down on the lower leaflet, with the elliptical shape of the C-terminal for the short form shown in purple.(B) A view from the side.(C) Diagram of a long form M protein dimer embedded in a membrane, where the differing protein colors refers to the monomers within the dimer (created with BioRender.com). . . . .                                                                                                                                                                                                                                                                                                                                                                                      | 58 |
| 3.4 | Atom Number Density, Excluding Hydrogen Atoms, from MD Simulations for the Short and Long Forms . . . . .                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 | 59 |
| 3.5 | <b>RMSD and <math>R_g</math> of Four Different M Protein Simulations</b>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |    |
|     | The long/short (blue/yellow) form membrane simulations and long/short (green/purple) form solvent simulations. (A) Root mean square distance (RMSD) represents the difference of each $C_\alpha$ atom at the corresponding time with its original position squared, summed over every $C_\alpha$ atom, and then divided by the number of $C_\alpha$ atoms and taken to the half power. (B) Total radius of gyration $R_g$ , defined as the radial distance to a point in which the moment of inertia would match the actual inertia if the body was condensed there, representing how compact the protein is. RMSD and $R_g$ were calculated using the GROMACS commands <code>gmx rms</code> and <code>gmx gyrate</code> respectively. 60 |    |

|                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |    |
|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| <b>3.6 Long and short conformation dimensions obtained throughout MD simulations.</b>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |    |
| These calculations were performed with MDAnalysis. (A) Protein height, determined by finding the difference between an average of the five highest and lowest $C_\alpha$ atoms throughout the simulation. (B) The difference in height between an average of the five furthest away C-terminal $C_\alpha$ atoms and the phospholipid heads within 2 nm of the protein. (C) Schematic showing the different characteristic quantities (created with BioRender.com). (D) Transmembrane domain height, determined by subtracting (B) and (F) from (A). (E) Width of the protein, the maximum distance between $C_\alpha$ atoms of like residues. (F) The height difference between the N-terminal and the membrane, calculated similarly to (B). Averages were taken over last 100 ns, where the corresponding error is standard deviation. . . . . | 62 |
| <b>3.7 Dimensions of the M Protein When Fitted to an Ellipse</b>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |    |
| (A)-(B) and (D)-(E) were obtained by calculating the inertia tensor of every protein atom above the height of the corresponding leaflet, diagonalizing it, and then assuming a uniform protein atom density in the shape of an ellipse. (A)-(B) correspond to the long axis of the C-terminal and N-terminal respectively, while (D)-(E) is the short axis of the C-terminal and N-terminal. (c) A schematic of the principal axes of the protein, with the ellipse fit for each terminal shown for the M protein (created with BioRender.com). Averages were taken over last 100 ns, where the corresponding error is standard deviation. . . . .                                                                                                                                                                                               | 63 |
| <b>3.8 Long to Short Axis Ratios from M Protein Ellipse Fit</b>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |    |
| (A) Repeating the schematic shown in Fig. 3.7. (B) Ratios of the C-terminal fit, Fig. 3.7A and 3.7D, and (C) ratio of the N-terminal fit, Fig. 3.7B and 3.7E. Averages were taken over last 100 ns, where the corresponding error is standard deviation. . . . .                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 | 64 |
| <b>3.9 The Average Membrane Thickness is Shown for the Last 500 ns of Each Simulation</b>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |    |
| (A) – (C) Visual from above in nm, with a radial plot from the center of the box or center of the protein shown to the right. Each blue dot represents the average thickness at a given bin, orange defines the 95% confidence interval for a linear line of best fit, and black is a radial average of average thickness. Every atom in the protein at every frame averaged over is shown in gold. (C) Shows the non-shifted radial plot used in Fig. 3.7e. . . . .                                                                                                                                                                                                                                                                                                                                                                             | 65 |

|      |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |    |
|------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| 3.10 | <b>The Average Membrane Midplane Height is Shown with Respect to the Average Value Along the Boundary</b>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |    |
|      | (A) – (C) Visual from above in nm, with radial plots to the right showing the average heights of each bin (blue) in comparison to either the radial distance from the center of the membrane or the radial distance from the center of the protein. Every atom in the protein at every frame averaged over is shown in gold. . . . .                                                                                                                                                                                                                                                                                                                                                                                      | 67 |
| 3.11 | <b>Protein Rotation for Both Forms Relative to Short Axis of the Corresponding M Protein Throughout MD Simulations</b>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |    |
|      | (A) Diagram displaying rotation in the x-y plane, where $\theta$ defines the angle between the short axis of the protein and the x-axis. The yellow lines represent projections of the long and short axis of the protein (created with BioRender.com). (B) Time evolution of the long/short (red/blue) form orientation ( $\theta$ ) with respect to the x-axis. (C) Shows the frequency of angles throughout the simulation for both short and long forms. . . . .                                                                                                                                                                                                                                                      | 68 |
| 3.12 | <b>Coarse-grained Simulations of M Proteins Embedded in the ER-Golgi Intermediate Compartment (ERGIC)</b>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |    |
|      | (A) M proteins embedded in the bottom of ERGIC have no endodomain interactions, and no budding is observed in this scenario. (B) M proteins have endodomain interactions, proteins bud into ERGIC and form a spherical shell in 4000s. (C) M proteins have no endodomain interactions, a budding process happens when N proteins are added. Note that N proteins interact with M proteins endodomain and effectively introduce endodomain interactions. (D) M proteins have no endodomain interactions, a budding happens when a short RNA genome (Length = 400a) is added. The M proteins interact with the genome through screened electrostatic interactions, which effectively cause endodomain interactions. . . . . | 71 |
| 4.1  | Schematic overview of the SARS-CoV-2 virus-like particle (VLP) production workflow. HEK293T cells were transiently transfected with plasmids encoding the four structural proteins (M, N, E, and S) at optimized expression ratios. VLPs were harvested from the culture supernatant, purified by ultracentrifugation through a sucrose cushion, and validated by Western blotting, dynamic light scattering (DLS), and atomic force microscopy (AFM). . . .                                                                                                                                                                                                                                                              | 78 |
| 4.2  | <b>Expression level of Nucleocapsid protein</b>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |    |
|      | (A) Expression level of the Nucleocapsid protein in HEK293T cells transfected with a plasmid encoding Nucleocapsid fused to the mScarlet fluorescent tag, measured 12 hours post-transfection.(B) Expression level at 24 hours post-transfection.(C) Expression level at 48 hours post-transfection. .                                                                                                                                                                                                                                                                                                                                                                                                                    | 80 |

|     |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |    |
|-----|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| 4.3 | Dynamic light scattering size distribution of SARS-CoV-2 VLPs (M:N:E expression condition) after RNase H treatment. Hydrodynamic diameter measurements were performed using a Malvern Zetasizer Nano ZSP in TNE buffer at 25 °C. VLPs exhibit a hydrodynamic diameter in the range of 100–146 nm, consistent with native SARS-CoV-2 virion dimensions. RNase H treatment was applied prior to measurement to remove surface-associated RNA, ensuring that the size distribution reflects intact VLP particles. . . . .                                                                     | 83 |
| 4.4 | Comparison of VLP hydrodynamic size distributions before and after RNase H treatment. In both conditions, VLPs maintain a hydrodynamic diameter in the range of 100–146 nm, consistent with native SARS-CoV-2 virion dimensions. DLS measurements demonstrate that RNase H treatment does not significantly alter VLP size or integrity, confirming that changes in RNA content following treatment reflect removal of externally associated RNA rather than disruption of the particles themselves. . . . .                                                                               | 83 |
| 4.5 | AFM diameter analysis of a SARS-CoV-2 VLP produced under the MNE (3:12:2) plasmid ratio condition. <b>(Top left)</b> AFM topographic image showing a representative VLP (~100 nm lateral diameter, ~10 nm height) adsorbed on APTES-functionalized mica. <b>(Top right)</b> Horizontal line profile (X-direction) across the particle center with Gaussian fit; FWHM = 88 nm. <b>(Bottom left)</b> Vertical line profile (Y-direction) with Gaussian fit; FWHM = 73 nm. <b>(Bottom right)</b> Zoomed topographic view of the same particle with a 100 nm diameter scale indicator. . . . . | 86 |
| 4.6 | AFM line profile of a representative VLP from the MNE (3:12:2) condition. The blue trace shows the raw height profile; the red dashed curve is a Gaussian fit (FWHM = 83 nm). The particle height above the local baseline (−4.1 nm) is approximately 17.5 nm, reflecting partial flattening upon adsorption to the mica substrate. . . . .                                                                                                                                                                                                                                                | 87 |
| 4.7 | Three-dimensional AFM surface rendering of a SARS-CoV-2 VLP from the MNE (3:12:2) condition. <b>(Left)</b> Side-angle view showing the dome-shaped morphology; the 50 nm scale bar indicates the particle radius (~100 nm diameter), with a maximum height of approximately 8.7 nm. <b>(Right)</b> Top-down view confirming the roughly circular footprint of the adsorbed particle. . . . .                                                                                                                                                                                               | 87 |

- 4.8 **Western blot analysis of nucleocapsid (N) protein incorporation into purified VLPs under different plasmid transfection conditions.** Purified VLP fractions and controls were resolved by SDS-PAGE and immunoblotted with an anti-N antibody. **MNES (3:12:2:5):** VLPs produced by co-transfection of M, N, E, and S expression plasmids at a 3:12:2:5 molar ratio. **MNE (3:12:2):** VLPs produced with M, N, and E plasmids at a 3:12:2 molar ratio, showing the strongest N incorporation ( $\sim 50$  kDa). **MNE (3:3:3):** VLPs produced with M, N, and E plasmids at an equimolar ratio, showing reduced N incorporation. **MSE:** VLPs produced without the N plasmid; absence of the N band confirms specificity of incorporation. **MSE+N:** VLPs produced by co-transfection of a single polycistronic MSE plasmid with a separate N expression plasmid; no detectable N band in the purified fraction, indicating that this plasmid ratio is incompatible with efficient N packaging into VLPs. **Control:** concentrated conditioned media harvested from untransfected cells processed identically to VLP samples, confirming the absence of non-specific background. **Media (–FBS):** serum-free conditioned media control. Molecular weight markers (kDa) are indicated. All blots are representative of  $n = 3$  independent experiments. . . . . 89
- 4.9 **Western blot analysis of membrane (M) protein incorporation into purified VLPs under different plasmid transfection conditions.** Purified VLP fractions and controls were resolved by SDS-PAGE and immunoblotted with an anti-M antibody. **MNES (3:12:2:5):** VLPs produced by co-transfection of M, N, E, and S expression plasmids at a 3:12:2:5 molar ratio. **MNE (3:12:2):** VLPs produced with M, N, and E plasmids at a 3:12:2 molar ratio, showing a prominent M monomer band at  $\sim 25$  kDa. **MNE (3:3:3):** VLPs produced with M, N, and E plasmids at an equimolar ratio. **MSE:** VLPs produced with M, S, and E plasmids. **MSE+N (two lanes):** VLPs produced by co-transfection of a single polycistronic MSE plasmid with a separate N expression plasmid, shown as biological replicates. **Control:** concentrated conditioned media harvested from untransfected cells processed identically to VLP samples. M protein is detected as a monomer at  $\sim 25$  kDa, with additional bands corresponding to SDS-stable dimers ( $\sim 50$  kDa) and higher-order oligomers ( $\sim 100$  kDa), consistent with the known oligomeric behavior of the M protein. Molecular weight markers (kDa) are indicated. All blots are representative of  $n = 3$  independent experiments. . . . . 90
- 4.10 **RNA quantification of VLP samples following RNA extraction.** Nucleic acid concentration ( $\text{ng}/\mu\text{l}$ ) was measured across all VLP conditions after RNA extraction. Samples include MSE, MSE+N, MNES, MNE (3:12:2), and MNE (3:3:3) preparations, with RNaseH-treated replicates included as negative controls to confirm RNA identity. All samples yielded detectable RNA, with concentrations ranging from approximately 8 to 32  $\text{ng}/\mu\text{l}$ . The MSE+N condition showed the highest RNA yield ( $\sim 32$   $\text{ng}/\mu\text{l}$ ). NEW RNA labels indicate independently prepared biological replicates. . . . . 92



**4.11 RNA quantification of VLP samples following RNA extraction. RNA Size Distribution of SARS-CoV-2 Virus-Like Particle (VLP)-Associated RNA Species Analysed by Agilent Bioanalyzer Small RNA Chip.** Capillary electrophoresis gel image depicting the size distribution of RNA extracted from purified SARS-CoV-2 VLP fractions following RNase H treatment, analysed across sixteen samples (A1–H1 and A2–H2) using an Agilent Bioanalyzer Small RNA chip. Size markers range from 25 to 6,000 nucleotides (nt), with the lower internal standard indicated by the green fluorescent band at 25 nt. RNA integrity numbers (RIN<sup>e</sup>) were not calculable for any sample (denoted “–”), consistent with the absence of intact ribosomal RNA species and the heterogeneous nature of VLP-encapsidated RNA cargo. The dominant signal observed across samples corresponds to small RNA species in the approximately 20–40 nt range, with neither 18S nor 28S ribosomal RNA peaks detected, confirming the absence of bulk cytoplasmic RNA contamination carried over during VLP purification (Kim et al. 2021). The enrichment of this discrete small RNA population, rather than a broad degradation smear, is consistent with selective encapsidation of host- or viral-derived small RNA species within VLPs, potentially mediated through interactions with the N protein RNA-binding domains or passive packaging during vesicle budding (Masters 2006; Swain et al. 2022). Inter-sample variability in higher molecular weight signal (notably D1 and E2) may reflect differences in VLP yield or residual RNA species co-purified within vesicular fractions. Red triangles denote individual sample injection points. . . . .

97

**5.1 VPC, a direct FRET-based reporter of p38 activation state.**

(A) FRET state of two resting HEK-293T cells expressing VPC (VPC293). (B) The same two cells, immediately after hyperosmotic shock induced by the addition of 120 mM sorbitol. (C) VPC (dark blue, from VPC293 cells) and p38 $\alpha$  (light blue, from untransfected HEK-293T cells) phosphorylation dynamics measured using Western blot and ensemble-averaged FRET (red,  $N \approx 1,500$  cells). Shaded region is the 95% confidence bound for FRET data, while error bars represent standard deviation of three biological Western blot replicates. See Fig. S1 for example Western blots. (D) In vitro FRET of purified VPC protein: as purified (VPC), VPC + 0.25 mM ATP-Mg (VPC + ATP), VPC + 0.25 mM ATP-Mg + 100 U/mL Antarctic Phosphatase (VPC + dePhos), VPC + 0.25 mM ATP-Mg + 200 nM MKK6 kinase (VPC + MKK6), and VPC + 0.25 mM ATP-Mg + 200 nM MKK6 kinase + 100 U/mL Antarctic Phosphatase. Values are means of 15 measurements; error bars represent SEM. (E) Dynamic ensemble-averaged FRET response of VPC293 cells upon hyperosmotic (green: 40 mM, orange: 80 mM, blue: 120 mM sorbitol) or hypoosmotic (red: 2 $\times$  dilution with ultrapure H<sub>2</sub>O) shock. Shock was induced at  $t = 0$ , indicated by the vertical gray line. Points represent the average of three biological replicates; shaded region indicates 95% confidence bounds. (F) Normalized height of the first peak in the VPC FRET signal upon induction of hyperosmotic shock via addition of sorbitol, as in (E). Blue points: mean of three biological replicates; error bars: standard deviation. Red line: best fit to the function  $1 - e^{-b[\text{Sorbitol}]}$ , with best-fit value  $b = 0.01 \pm 0.002 \text{ mM}^{-1}$ . Hyperosmotic shock with 400 mM sorbitol results in cell death. Scale bar: 20  $\mu\text{m}$  (A, B).

105

## 5.2 PerKy-38 and p38 exhibit phosphorylation state oscillations with stimulus-specific frequencies.

(A–B) Interpretation of single-cell FRET signals. (A) Activation of a p38 protein by phosphorylation discretely increases the FRET signal from that molecule. (B) Individual cells contain many p38 molecules; hence, the single-cell FRET signal is proportional to the number of active p38 molecules. (C) Oscillating p38 activation state shown for a single VPC293 cell stimulated by hyperosmotic shock with sorbitol. The first image ( $t = 0$ ) is 35 hours after addition of 120 mM sorbitol. Scale bar: 20  $\mu\text{m}$  (C). (D) FRET time traces of five fields of view (FOVs) of PerKy293 cells exposed to 300 ng/mL IL-1 $\beta$ . (E) Ensemble-averaged Fourier spectrum of 70 FOVs of PerKy293 (as in D). Shaded region: 95% confidence bounds; dashed gray line: 99% confidence interval for peak magnitudes. Asterisks indicate statistically significant peaks. (F) Ensemble-averaged Fourier power spectrum of 100 FOVs of PerKy293 cells exposed to 2 nM SARS-CoV-2 Spike S1. Note the absence of statistically significant peaks. (G) FRET time traces of five FOVs of VPC293 cells exposed to 300 ng/mL IL-1 $\beta$ . (H) Ensemble-averaged Fourier spectrum of 66 FOVs of VPC293 cell line (as in G). Asterisks indicate statistically significant peaks; gray dashed line represents the 99% confidence level. (I) Ensemble-averaged Fourier power spectrum of 70 FOVs of VPC293 cells exposed to 2 nM SARS-CoV-2 Spike S1 protein. Note the statistically significant peak at  $f = 0.45 \pm 0.01 \text{ hr}^{-1}$  ( $p = 2.3 \times 10^{-4}$ , marked with an asterisk), with no significant peak at  $f = 0.36 \text{ hr}^{-1}$ . (J) Statistics of local maxima in Fourier spectra of unstimulated VPC293 cells (see SM II). Integration over frequency yields a normal distribution (right), used to quantify  $z$  score;  $z$  = number of standard deviations from the mean. Gray: mean; blue:  $z = 2$ ; red:  $z = 3$ . (K) Oscillatory frequencies in individual cells for different stimuli. Top: aggregated spectrum. Expected frequencies of the  $n$ th harmonic [ $f_n = (n + 1)f_0$ ,  $n = 0, 1, 2, \dots$ , and  $f_0 = 0.09 \text{ hr}^{-1}$ ] are indicated. Blue:  $2 \leq z < 3$ ; red:  $z \geq 3$ . (L) Inharmonicity of identified frequencies. Points: band center of mass; error bars: band width at  $z$  score boundary. Blue:  $2 \leq z < 3$ ; red:  $z \geq 3$ ; yellow:  $z > 10$ . Gray shaded region: interval of  $\pm f_0/4$  (see SM III).

107

### 5.3 p38 drives target activation through forced resonant oscillations.

Real (A) and Imaginary (B) parts of the transfer function,  $\psi(\omega)$ , between p38 [ $y(\omega)$ ] and PerKy-38 [ $x(\omega)$ ],  $\psi(\omega) = x(\omega)/y(\omega)$ ; see SM VII and Fig. S6. Blue points: data,  $N = 1,092$  ratios of Fourier spectra from VPC293 and PerKy293 cells; blue shaded region: 95% confidence region; red lines: fits to the expected form of the transfer function of a driven simple harmonic oscillator (see SM VII, Eqs. S7.3, S7.7; cf. Fig. S6A,B). (C) Normalized modulus of the transfer function between p38 and PerKy-38, showing PerKy-38 amplitude response as a function of p38 driving frequency (see SM V, Eq. S5.3, and SM VII, Eq. S7.8; cf. Fig. S6C). Fits yield damping parameter  $\gamma = 0.00423 \pm 0.00009$  (SEM)  $\text{hr}^{-1}$  and quality factor  $Q \equiv \omega_0/\gamma = 535 \pm 11$ . (D–F) Driving the p38 MAPK system with periodic hyperosmotic shock. Time series (left) and Fourier spectra (right) of the input sorbitol flow (red, top, measured by rhodamine B fluorescence), p38 response (orange, middle, VPC FRET), and PerKy-38 output (green, bottom, PerKy-38 FRET) when cells are subjected to periodic hyperosmotic stress by oscillatory flow of sorbitol between 0 mM and 40 mM at a frequency of (D)  $0.27 \text{ hr}^{-1}$ , (E)  $0.36 \text{ hr}^{-1}$ , and (F)  $0.54 \text{ hr}^{-1}$ . The corresponding driving frequencies relative to the modulus of the transfer function are indicated by red arrows in (C).

112

### 5.4 Activation of specific genetic responses through dynamic bioresonant stimulation.

(A) Clustergram of expression levels of 3,336 statistically significant ( $\text{padj} < 0.01$ ) differentially expressed genes with  $|\log_2(\text{fold change})| > 1$ , as determined by RNA sequencing of untreated negative control samples (Control), samples osmotically stressed with continuously applied sorbitol (20 mM or 40 mM), and samples driven with sorbitol oscillating between 0–40 mM at a frequency of either  $0.54 \text{ hr}^{-1}$  or  $0.36 \text{ hr}^{-1}$ , as in Figure 5.3D–F. All RNA-seq was performed using two biological replicates for each sample. (B–C) Volcano plots of RNA-seq results from samples periodically driven at a frequency of  $0.54 \text{ hr}^{-1}$  (B) and  $0.36 \text{ hr}^{-1}$  (C). Purple points: genes differentially expressed regardless of driving frequency ( $N = 1,144$ ); blue points: genes differentially expressed only at  $0.54 \text{ hr}^{-1}$  ( $N = 184$ ); red points: genes differentially expressed only at  $0.36 \text{ hr}^{-1}$  ( $N = 195$ ). (D–E) Enrichment analysis of genes uniquely expressed at a driving frequency of  $0.36 \text{ hr}^{-1}$  [red points in (C)], showing enriched (D) regulatory pathways and (E) disease-related pathways. Data sources include KEGG, Reactome (REAC), and WikiPathways (WP). Point size increases linearly with the number of intersecting genes (range: 5–20). Terms are filtered for brevity and relevance; full lists of enriched terms are provided in Figures S8–S10.

115

## 5.5 An Integrated Conceptual Framework Describing MAPK Bioresonance and Output.

(A) When stimuli are received by the cell, receptors activate the MAPK cascade, which we treat as a “black box” signal generator (generator symbol) that sinusoidally drives p38 oscillations (orange oval) at different frequencies. p38 functions as a transmitter (antenna symbol) whose oscillatory frequencies resonantly drive oscillations of downstream targets (green, blue, and yellow ovals), which we model as two-component biological oscillators with natural frequencies  $f_i$ . Resonantly activated substrates then activate or repress genes regulated by those substrates. The combined changes in gene expression lead to the overall correct (blue) or incorrect (green) cellular stress response. (B–D) Simulated time traces of p38 (orange) resonantly interacting with two downstream targets, Substrate A and Substrate B, with natural frequencies  $f_A = 0.18 \text{ hr}^{-1}$  (green) and  $f_B = 0.36 \text{ hr}^{-1}$  (blue). Shown are examples with: (B) p38 driven at  $f = 0.09 \text{ hr}^{-1}$ , which generates harmonics on-resonance with Substrate A; (C) p38 driven at  $f = 0.36 \text{ hr}^{-1}$ , on-resonance with Substrate B; (D) p38 driven at  $f = 0.27 \text{ hr}^{-1}$ , off-resonance with both substrates. Transparent lines: individual simulations; solid lines: mean of 7 simulations simulating one FOV. Fourier spectra are averages of 14 simulated FOVs; see SM IX. (E) A quantitative model of p38 [ $y(t)$ , orange] driving oscillations in two interacting TFs [ $x_0(t)$ , green and yellow]. TFs are phosphorylated with rates  $\zeta_0$ ; TFs react with each other with rates  $\alpha_0$ ; TFs spontaneously dephosphorylate (null symbol) with rates  $\beta_0$ . See SM VI. (F) Equivalent representation of (E) in terms of AC circuit elements, oscillating phosphate current in p38 (orange) inducing chemical potentials,  $\mu_0(t)$ , in downstream TFs, with spontaneous dephosphorylation analogous to resistive heat losses. See SM VI and SM X–XI. (G) Representation of the MAPK cascade as a phosphate current amplifier and distribution network. Phosphate currents ( $I_{\text{RMS}}$ , lines) are amplified or suppressed along paths through the cascade based upon the net frequency-dependent impedance,  $Z$ , of each MAPK element (circles) along each path. . . . .

118

## 6.1 Western blot analysis of p38 $\alpha$ and VPC phosphorylation

(A) Representative immunoblots as a function of time following addition of 300 nM IL-1 $\beta$ , showing (from top to bottom): phospho-VPC, phospho-p38 $\alpha$ , total VPC, total p38 $\alpha$ , and GAPDH as a loading control. (B) SDS–PAGE gel of purified VPC–2 $\times$ Strep stained with SYPRO. (C) Western blot analysis of purified VPC–2 $\times$ Strep. . . . .

128

|            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |     |
|------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| <b>6.2</b> | <b>Extended Data Fig. 2. PerKy293+SB203580 Control Statistics</b>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |     |
|            | (A) Representative FRET time traces of PerKy293 cells treated with 300 ng mL <sup>-1</sup> IL-1 $\beta$ and 10 $\mu$ M SB203580. (B) Ensemble-averaged power spectrum of $N = 450$ PerKy293 cells under the conditions shown in panel (A). The blue dashed line indicates the 99% confidence interval determined using the statistical procedure described below. (C) Log-log plot of the PerKy293 power spectrum under IL-1 $\beta$ and SB203580 treatment. The red dashed line represents a fit to a power-law $f^{-m}$ , yielding $m = 1.74 \pm 0.34$ . The red shaded region denotes the best fit $\pm$ fitting error. (D) Histogram of deviations of the PerKy293 power spectrum, $\Phi_1$ , from the power-law fit, $\Phi_0$ . Dark blue bars indicate deviations from the best fit, while the light blue shaded region corresponds to deviations from the best fit $\pm$ the errors shown in panel (B). The solid red line shows the best fit to a log-normal distribution, and the shaded red region indicates the uncertainty in the log-normal fit. . . . . | 138 |
| <b>6.3</b> | <b>PerKy293 Response to Continuous Hyperosmotic Shock and TNF Stimulation</b>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |     |
|            | (A) Representative FRET time traces from five fields of view (FOVs) of PerKy293 cells treated with 120 mM sorbitol. (B) Ensemble-averaged power spectrum of PerKy293 cells treated with 120 mM sorbitol. (C) Representative FRET time traces from five FOVs of PerKy293 cells treated with 200 nM TNF. (D) Ensemble-averaged power spectrum of PerKy293 cells treated with 200 nM TNF. . . . .                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        | 139 |
| <b>6.4</b> | <b>VPPC293 Control Statistics</b>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |     |
|            | (A) Representative FRET time traces from five fields of view (FOVs) of VPPC293 control cells treated with 300 ng mL <sup>-1</sup> IL-1 $\beta$ . (B) Ensemble-averaged power spectrum of VPPC293 cells. The blue dashed line indicates the 99% confidence interval determined using the statistical procedure described below. (C) Log-log plot of the VPPC293 power spectrum. The red dashed line represents a fit to a power-law $f^{-m}$ , yielding $m = 1.86 \pm 0.20$ . The red shaded region denotes the best fit $\pm$ fitting error. (D) Histogram of deviations of the VPPC293 power spectrum, $\Phi_1$ , from the power-law fit, $\Phi_0$ . Dark blue bars indicate deviations from the best fit, while the light blue shaded region corresponds to deviations from the best fit $\pm$ the errors shown in panel (B). The solid red line shows the best fit to a log-normal distribution, and the shaded red region indicates the uncertainty in the log-normal fit. . . . .                                                                                | 140 |

|                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |     |
|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| <b>6.5 Harmonic Analysis of Stimulus Spectra.</b>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |     |
| (A) Distances (yellow) of identified frequency bands (red, as in Figure 5.2K) measured relative to the tines of a Dirac comb with evenly spaced harmonics (green). (B) Mean distance of frequency bands to the nearest tine of the Dirac comb as a function of the fundamental frequency $f_0$ . The best-fit fundamental frequency corresponds to the minimum of this curve (red), yielding $f_0 = 0.090 \pm 0.002 \text{ h}^{-1}$ . (C) Distribution of the mean distance of anharmonically (randomly) placed frequency bands from their nearest harmonic, expressed in units of the fundamental frequency $f_0$ . The distribution follows a Bates distribution on the interval $[0, 1/2]$ and is shown for different numbers of frequency bands. As 11 frequency bands are observed in the aggregate spectrum (Figure 5.2K, top), $n = 11$ (purple curve) was used to calculate $p$ -values, with distances measured from the center of mass of each band (red dashed line). | 141 |
| <b>6.6 Time Dependence of the Frequency Response of p38 to Various Stimuli.</b>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |     |
| Short Time Fourier Transform (STFT) spectrograms illustrating how the frequency response of p38 changes over time when subjected to various stimuli. Note that by adding resolution along the time axis, resolution along the frequency axis is reduced; hence these spectrograms are not directly comparable to the time-independent Fourier spectra presented elsewhere, which are effectively (but not mathematically identical to) averages over the time axis of these spectrograms. (A) Unstimulated VPC293 control. (B) $300 \text{ ng mL}^{-1}$ IL-1 $\beta$ . (C) $300 \text{ ng mL}^{-1}$ IL-1 $\beta$ + SB203580. (D) $200 \text{ ng mL}^{-1}$ TNF. (E) $2 \text{ nM}$ SARS-CoV-2 Spike S1. (F) Hyperosmotic shock induced by addition of $120 \text{ mM}$ sorbitol.                                                                                                                                                                                                  | 144 |
| <b>6.7 Transfer Function of a Forced, Lightly Damped Linear Oscillator.</b>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |     |
| (A) Real part, (B) imaginary part, and (C) modulus of the transfer function expected for a forced, lightly damped linear oscillator, 6.37, using the parameters extracted from fitting the p38/PerKy-38 transfer function. Plots are scaled by an arbitrary parameter to allow for direct comparison to Fig5.3. Note that, compared to Fig5.3, the real and imaginary parts are swapped and the imaginary part inverted, exactly as predicted when treating the system as coupled chemical reactions, eqs. S7.7.                                                                                                                                                                                                                                                                                                                                                                                                                                                                 | 147 |

## 6.8 Bioresonant Transcriptional Regulation

(A) Concentration of active resonant (red) and off-resonant transcription factor (blue line) over time, 6.74. The dissociation constant for binding to a cis-regulatory region,  $K$ , is indicated by a horizontal black line. The duty cycle,  $D$  6.85, is calculated as the ratio of the red to blue shaded areas over one period. (B) The regulatory function  $F_{\text{reg}}$  6.73 plotted as a function of active transcription factor concentration for various values of the Hill coefficient,  $m$ . Blue,  $m = 1$ ; orange,  $m = 3$ ; green,  $m = 10$ ; red,  $m = 100$ . The “duty cycle” approximation corresponds to large values of  $m$  resulting in a switch-like step function, e.g., the red curve. (C) The regulatory function  $F_{\text{reg}}$  plotted as a function of time (i.e., the integrand of 6.76 for various values of the Hill coefficient,  $m$ , for resonant (thick lines) and off-resonant (thin lines) transcription factors. The ratio of areas under the thick versus thin lines represents the enhancement of gene expression when a transcription factor is on-resonance versus off-resonance. Parameter values used:  $f = 100$ ,  $B_0/K = 0.73$ ,  $B_1/K = 0.001$ ,  $f_0 = 0.36 \text{ h}^{-1}$ ,  $\gamma = 0.004$ . (D) The gene expression fold change as a function of the driving frequency and parameter combination  $(1 - B_0/K)/(B_1/K)$  under the duty cycle approximation, 6.84 and 6.85. Note the narrow range of the  $x$ -axis, indicating that gene expression is enhanced only very close to the resonant frequency of the transcription factor. Parameter values used:  $f = 100$ ,  $f_0 = 0.36 \text{ h}^{-1}$ ,  $\gamma = 0.004$ . . . . .

156



# Chapter 1

## Introduction

The emergence of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) in late 2019 precipitated a global pandemic of extraordinary scale, resulting in hundreds of millions of infections and millions of deaths worldwide ([World Health Organization 2023](#); [Dong et al. 2020](#)). Beyond its acute clinical presentation as coronavirus disease 2019 (COVID-19), SARS-CoV-2 has revealed the ongoing vulnerability of global health systems to novel zoonotic coronaviruses and exposed critical gaps in our mechanistic understanding of this class of pathogen. Despite the remarkable success of mRNA-based vaccines and the rapid deployment of antiviral therapeutics ([Polack et al. 2020](#); [Baden et al. 2021](#)), the emergence of new variants and the persistence of long COVID as a clinical syndrome underscore the need for a deep, molecular-level understanding of how SARS-CoV-2 assembles, replicates, and interacts with its host. This dissertation addresses that need through a series of biophysical and biochemical investigations into the structural proteins of SARS-CoV-2, with a focus on how they remodel host membranes and cooperate to form new viral particles.

## 1.1 SARS-CoV-2 Structure and Viral Assembly

SARS-CoV-2 is an enveloped, positive-sense single-stranded RNA virus of approximately 30 kb in length, belonging to the *Betacoronavirus* genus of the family *Coronaviridae* (Wu et al. 2020; Zhou et al. 2020). The virion is organized around four primary structural proteins: the membrane (M) protein, the envelope (E) protein, the nucleocapsid (N) protein, and the spike (S) protein. Together, these proteins coordinate every stage of the viral life cycle, from genome packaging and particle formation to host cell entry (McBride et al. 2014; Masters 2006).

The S protein, a large trimeric glycoprotein embedded in the viral envelope, mediates receptor recognition and membrane fusion (Walls et al. 2020; Wrapp et al. 2020). It binds with high affinity to the host angiotensin-converting enzyme 2 (ACE2), triggering the conformational changes necessary for viral entry (Shang et al. 2020a; Wang et al. 2020; Walls et al. 2020). Because S protein is the principal target of neutralizing antibodies and the basis of all current vaccine platforms, it has received by far the most research attention of the four structural proteins (Krammer 2020; Dai & Gao 2020).

In contrast, the M, E, and N proteins, which are responsible for the physical assembly of the virion, remain substantially less characterized. Coronavirus assembly does not occur at the plasma membrane, as is the case for many other enveloped viruses, but instead proceeds at the endoplasmic reticulum–Golgi intermediate compartment (ERGIC) — a specialized intracellular compartment with a lipid composition distinct from either the ER or the Golgi (Klumperman et al. 1994; Stertz et al. 2007; Siu et al. 2008). This spatial specificity is enforced by the intrinsic Golgi-targeting signals encoded in M and E,

which concentrate these proteins at the ERGIC membrane where budding occurs (Nal et al. 2005; Cohen et al. 2011). The resulting budded particles then traffic through the secretory pathway and are released by exocytosis (Ujike & Taguchi 2015).

### 1.1.1 The Membrane (M) Protein

The M protein is the most abundant structural component of the SARS-CoV-2 virion and serves as the central organizer of viral assembly (Neuman et al. 2011a; Schoeman & Fielding 2019; McBride et al. 2014). It spans the membrane three times and exposes a small glycosylated N-terminal ectodomain on the outer surface of the virion, while a large C-terminal endodomain faces the viral interior (Rottier et al. 1984; Armstrong et al. 1984). Through its endodomain, M protein interacts with all other structural proteins: it binds N protein to anchor the ribonucleoprotein complex to the viral envelope, it cooperates with E protein to drive membrane curvature and particle budding, and it incorporates S protein into nascent virions (Narayanan et al. 2000; He et al. 2004; Tseng et al. 2010; Neuman et al. 2011a). These interactions make M protein the organizing scaffold around which the coronavirus particle is assembled.

Despite its central importance, M protein has been among the most difficult viral proteins to study *in vitro*. Its extreme hydrophobicity, arising from three transmembrane helices, renders it prone to aggregation and largely insoluble in standard aqueous buffers, severely limiting the yield and purity achievable by conventional expression methods (Kuo & Masters 2002; Thomas 2020). As a result, structural and functional data on M protein lagged far behind that of other viral proteins, and the molecular mechanisms by which it remodels membranes and drives budding remained poorly understood.

## 1.2 Biophysical Approaches to Studying Viral Membrane Proteins

A recurring challenge in the structural biology of viral membrane proteins is the need to study them in an environment that recapitulates the physical properties of the membrane in which they function (White et al. 2001; Cross et al. 2011). Detergent-solubilized proteins may adopt conformations that differ from their native membrane-embedded states, and cell-based assays average over large and heterogeneous populations. The experiments described in this dissertation address these challenges by combining high-yield recombinant protein expression with direct *in situ* imaging of proteins reconstituted into biologically relevant supported lipid bilayers and giant unilamellar vesicles.

Atomic force microscopy (AFM) provides nanometer-scale topographic imaging of membrane proteins in liquid at physiological temperature, enabling direct visualization of individual protein complexes in their lipid bilayer environment without the need for staining, fixation, or cryogenic preservation (Müller & Dufrêne 2008; Casuso et al. 2012). Cryo-electron microscopy (cryo-EM) complements AFM by providing three-dimensional structural information at near-atomic resolution, albeit on flash-frozen rather than native-state samples (Cheng 2015; Nogales 2016). Molecular dynamics (MD) simulations, both all-atom and coarse-grained, extend experimental observations by providing atomic-resolution models of protein–lipid and protein–protein interactions, including membrane deformations and the thermodynamic driving forces behind protein oligomerization and budding (Marrink & Tieleman 2013; Lindahl et al. 2001). Together, these complementary techniques provide a multiscale picture of protein behavior that neither approach could achieve alone.

Dynamic light scattering (DLS) and Western blotting provide quantitative characterization of VLP size distributions and structural protein content, respectively, enabling validation of assembly efficiency across different expression conditions (Stetefeld et al. 2016; Mahmood & Yang 2012). RNA analysis by microfluidic capillary electrophoresis (Agilent Bioanalyzer) and spectrophotometry establishes the concentration, purity, and size distribution of nucleic acid cargo packaged within VLPs, laying the groundwork for future sequencing-based studies of selective genome packaging.

### 1.2.1 The Envelope (E) Protein

The E protein is the smallest structural protein of SARS-CoV-2, comprising only 75 amino acids organized into a short N-terminal domain (NTD), a hydrophobic transmembrane domain (TMD), and a longer C-terminal domain (CTD) (Schoeman & Fielding 2019; DeDiego et al. 2007). Although E protein is present at low copy number in the virion compared to M, genetic deletion studies have demonstrated that it is indispensable for efficient viral release and normal virion morphology (DeDiego et al. 2014; Alvarez et al. 2010; Schoeman & Fielding 2019). Its CTD harbors a PDZ-binding motif that disrupts normal cellular functions by sequestering host PDZ-domain proteins (Jiménez-Guardenõ et al. 2014), and both M and E must be co-expressed to produce virus-like particles in cell-based assembly assays (Vennema et al. 1996; Corse & Machamer 2000). Like M protein, E protein has been difficult to study *in vitro* because its hydrophobic transmembrane domain renders it prone to aggregation in standard aqueous buffers, and its low copy number in the virion makes biochemical characterization from native sources impractical. Access to recombinant E pro-

tein in quantities sufficient for membrane reconstitution has therefore been a prerequisite for understanding its structural role.

The SUMO-tag-based bacterial expression system developed in this dissertation for M protein was directly extended to produce recombinant E protein at high yield. This purified E protein was provided to collaborators, who carried out membrane reconstitution experiments in ERGIC-mimetic supported lipid bilayers and giant unilamellar vesicles. These collaborative experiments revealed that E protein forms multiple coexisting oligomeric species — monomers, dimers, pentamers, and higher-order assemblies — in a native-like membrane environment, providing direct structural evidence that addresses the long-standing controversy over E protein’s oligomeric state ([Pervushin et al. 2009](#); [Surya et al. 2018](#); [Cao et al. 2021](#)). E protein was further shown to induce membrane thinning and curvature, and to drive concentration-dependent morphological transitions in giant vesicles. Most significantly, real-time confocal microscopy by collaborators demonstrated that introduction of pre-assembled N protein–RNA ribonucleoprotein complexes to giant vesicles containing both M and E proteins triggers rapid inward vesicle budding, establishing a minimal reconstituted system for SARS-CoV-2 particle assembly *in vitro*. These findings, enabled by the protein expression platform developed in this dissertation, are being prepared for separate publication.

### **1.2.2 The Nucleocapsid (N) Protein and Genome Packaging**

The N protein binds and condenses the approximately 30 kb positive-sense RNA genome through its N-terminal RNA-binding domain and undergoes liquid–liquid phase separation to form the ribonucleoprotein (RNP) complex that occupies the viral interior

(Savastano et al. 2020; Perdikari et al. 2020; Iserman et al. 2020). N protein also interacts directly with M protein through its C-terminal domain, coupling genome packaging to the membrane-budding machinery (Narayanan et al. 2000; He et al. 2004). While the structural organization of the N protein–RNA complex and the determinants of selective genomic RNA packaging are active areas of investigation (Khoo et al. 2021; Yao et al. 2020), the relationship between N protein assembly, RNP morphology, and the physical process of membrane deformation during budding is not yet understood at a mechanistic level.

### 1.2.3 Virus-Like Particles as a Research Tool

Because live SARS-CoV-2 requires biosafety level 3 containment and because the intact virus resists many of the *in vitro* manipulations needed for quantitative biophysical characterization, virus-like particles (VLPs) have emerged as a powerful surrogate system (Noad & Roy 2003; Ludwig & Wagner 2007). VLPs are non-infectious, self-assembled structures composed of viral structural proteins that recapitulate the morphology and biochemical composition of native virions without containing the viral genome (Lua et al. 2014; Zeltins 2013). They can be produced under biosafety level 2 conditions in mammalian cell expression systems and subjected to the full range of biochemical and biophysical analyses applicable to purified macromolecular assemblies.

A critical variable in VLP production is the relative stoichiometry of the structural proteins provided during assembly. Recent work has demonstrated that expressing M, N, E, and S proteins at ratios that reflect viral mRNA abundance in productively infected cells — rather than at equal amounts — substantially improves VLP yield, spike incorporation, and structural fidelity (Swann et al. 2020; Boson et al. 2021). This finding points to quantitative

control of protein expression as a key determinant of assembly efficiency, and it establishes biologically calibrated VLPs as the appropriate model system for studies of SARS-CoV-2 particle formation and RNA packaging.

### 1.3 Cellular Stress Response and the MAPK Cascade

A fourth study included in this dissertation departs from the structural virology focus of the preceding chapters to address a fundamental question in cell signaling: how can a single kinase, p38 mitogen-activated protein kinase (MAPK), selectively activate the correct subset of downstream targets from among hundreds of possible substrates in response to diverse environmental stresses? Aberrant p38 signaling is implicated in a broad range of diseases, including Alzheimer’s disease, amyotrophic lateral sclerosis, cancer, and neuroHIV, making the mechanistic basis of its specificity an important open question with direct therapeutic relevance ([Cuenda & Rousseau 2007](#); [Zarubin & Han 2005](#); [Ashwell 2006](#)).

This study is included in the dissertation for two reasons. First, it represents independent collaborative research conducted during the same period as the SARS-CoV-2 work and constitutes a co-authored contribution to the peer-reviewed literature. Second, and more substantively, it illustrates a general principle that is relevant to the coronavirus assembly work: that biological specificity can be achieved not only through structural complementarity between interacting molecules, but also through the dynamics and temporal organization of molecular interactions. The bioresonance mechanism described in [Chapter 5](#) — in which p38 encodes information about experienced stresses as frequency oscillations of its activation state, and downstream specificity is achieved when p38 oscillations resonate



with the natural frequencies of target substrates — represents a new conceptual framework for understanding how cells process and transmit information through signaling cascades (Levine et al. 2013; Purvis & Lahav 2013).

## 1.4 Dissertation Outline

This dissertation is organized as three research chapters, each corresponding to a manuscript that has been published or is in preparation, preceded by this general introduction and followed by a conclusions chapter.

**Chapter 2** presents the expression, purification, and biophysical characterization of SARS-CoV-2 M protein. Using a SUMO-tag-based *E. coli* expression system, full-length native M protein was produced at yields orders of magnitude higher than previously reported methods. Purified M protein dimers were reconstituted into supported lipid bilayers and characterized by AFM, cryo-EM, and all-atom and coarse-grained MD simulations. The chapter establishes that M protein dimers adopt exclusively the compact “short” conformation in lipid bilayers, induce localized membrane thinning and curvature, and that endodomain interactions among dimers, combined with line tension from membrane thinning, can drive viral assembly and budding. The expression platform developed in this chapter was subsequently extended to E protein purification, enabling the collaborative assembly studies described in Section 1.2 above.

**Chapter 3** describes the production and characterization of SARS-CoV-2 VLPs in mammalian cells. Expression of all four structural proteins — M, N, E, and S — at ratios reflecting viral mRNA abundance in productively infected cells yields VLPs that

are validated by Western blotting, AFM, and dynamic light scattering. AFM imaging of individual VLP particles reveals dome-shaped particles with lateral diameters of 80–100 nm and morphology consistent with intact, membrane-enveloped vesicles. RNA isolated from purified VLP fractions exhibits a size distribution dominated by small RNAs, consistent with selective packaging rather than nonspecific incorporation of cellular RNA, and establishes a biochemical foundation for future sequencing-based characterization of VLP RNA cargo.

**Chapter 4** presents a study of p38 MAPK signal transduction, demonstrating that the central human stress response kinase encodes information about experienced stresses as frequency oscillations of its activation state. A novel FRET-based direct reporter of p38 activation state (VPC) is introduced and validated, and a mathematical framework based on the analogy between bioresonant signaling cascades and alternating current electronic circuits is developed to describe how frequency-selective amplification confers downstream target specificity. Periodic stimulation of cells at different frequencies is shown to activate distinct genetic programs, providing direct experimental evidence that bioresonance is a functionally significant mechanism of signaling specificity in living cells.

## Chapter 2

# Paper I: Synthesis, Insertion, and Characterization of SARS-CoV-2 Membrane Protein Within Lipid Bilayers

**Author Contributions:** This chapter contains a paper published in *Science Advances*. Sara Anbir and Yuanzhong Zhang are co-first authors. Sara Anbir developed a novel method for achieving native full-length M protein expression and purification, led experimental design, and contributed to data analysis. Yuanzhong Zhang performed atomic force microscopy (AFM) experiments, large unilamellar vesicle (LUV) experiments, and protein characterization. Molecular dynamics simulations and associated data analysis were performed by the group of Roya Zandi, including contributions from Michael E. Colvin and

Ajay Gopinathan. Additional molecular dynamics simulations were performed by Joseph McTiernan. Siyu Li, Michael Worcester, and Pratyasha Mishra contributed to experimental work. Umar Mohideen and Thomas E. Kuhlman provided research supervision.

**Abstract** Throughout history, coronaviruses have posed challenges to both public health and the global economy; nevertheless, methods to combat them remain rudimentary, primarily due to the absence of experiments to understand the function of various viral components. Among these, membrane (M) proteins are one of the most elusive because of their small size and challenges with expression. Here, we report the development of an expression system to produce tens to hundreds of milligrams of M protein per liter of *Escherichia coli* culture. These large yields render many previously inaccessible structural and biophysical experiments feasible. Using cryo-electron microscopy and atomic force microscopy, we image and characterize individual membrane-incorporated M protein dimers and discover membrane thinning in the vicinity, which we validated with molecular dynamics simulations. Our results suggest that the resulting line tension, along with predicted induction of local membrane curvature, could ultimately drive viral assembly and budding.

## 2.1 Introduction

Coronaviruses have caused severe health crises and economic challenges over the centuries. Although the likelihood of another coronavirus type emerging after the eradication of one has always been high, research on them has often diminished due to limited success in understanding fundamental elements of their functioning and life cycle. This is in sharp contrast to progress on other viruses like HIV and hepatitis B virus. One of the most

notable factors contributing to our limited knowledge of coronaviruses is the difficulty in performing in vitro experiments on viral proteins due to challenges posed by the small size of the proteins and low yield expression methods for them. As an example, while the membrane (M) protein of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) is the most abundant protein with a crucial role in the formation of the virus, it is much less well characterized than most other SARS-CoV-2 structural proteins. Here, we take important steps toward developing a deeper understanding of M protein structure and function in the physiologically relevant membrane environment by developing an expression technique with orders of magnitude higher yield, inserting the protein into lipid membranes and characterizing the protein-membrane configurations using a combination of atomic force microscopy (AFM), cryo-electron microscopy (cryo-EM), and molecular dynamics (MD) simulations. Specifically, using a small ubiquitin-related modifier (SUMO) tag-based expression system, we produced and purified substantial quantities of full-length, native M protein in *Escherichia coli*. As far as we are aware, this is the only reported method for producing substantial quantities of full-length, native M protein, and our method is of low cost and produces large quantities of pure native, full-length protein with low endotoxin levels. We examined the purity of our M protein product using Western blot and matrix-assisted laser desorption/ionization (MALDI) mass spectrometry and showed that it has low endotoxin activity. Our method produces tens to hundreds of milligrams of protein per liter of culture, with these high yields enabling many previously inaccessible experiments, including small-angle x-ray scattering and cryo-EM.

In addition, we developed a reliable method for inserting purified M protein dimers into a

suspended lipid membrane in a homogeneous orientation, and we used cryo-EM and AFM to image the M protein dimers in the membrane. We compared our experimental results to all-atom MD simulations of the M protein dimers in a lipid membrane to provide structural information about the protein and the surrounding membrane to aid in interpreting the imaging results. While some prior experimental studies([Zhang et al. 2022](#))revealed two distinct M protein structures, a “short” and a “long” form, other studies only showed the existence of a single-M protein structure([Dolan et al. 2022](#)), similar to the short form. Our AFM results, together with our atomistic MD simulations, indicate that M protein dimers within our supported lipid bilayer are entirely in the short form identified in earlier structural studies.

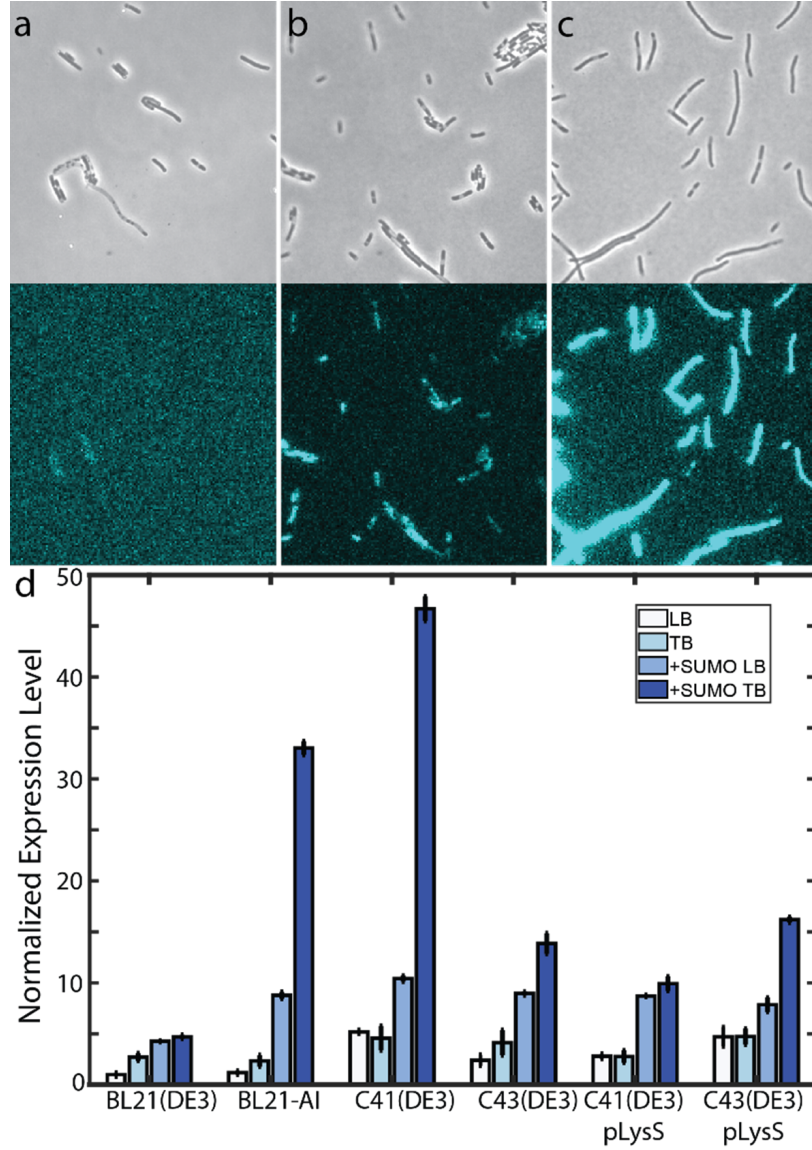
Our AFM results also uncovered a thinning of the membrane in the vicinity of the M protein dimer, which we confirmed with MD simulations. In addition, our MD simulations indicated the propensity of individual M protein dimers to induce local membrane curvature. The line tension resulting from the thinning together with the induced curvature could potentially drive the aggregation viral proteins and the subsequent budding of the virion. To test this idea, we constructed a coarse-grained simulation model of M proteins in a flat membrane and showed that endodomain interactions among M protein dimers were also critical for budding. Together, our work establishes a framework for future studies of viral protein and sheds light on the structure and interactions of M protein dimers within a lipid membrane, yielding insights into important physical processes that drive viral assembly and budding.

## 2.2 Results

### 2.2.1 Expression and high-yield purification of SARS-CoV-2 M protein

To assess the level of expression of SARS-CoV-2 M protein in *E. coli*, we constructed a C-terminal fusion of M with the blue fluorescent protein mCerulean3. This fusion protein is borne on the high-copy number plasmid pUC57 and expressed from the inducible promoter PlacT7. Upon transformation into the standard expression strain BL21(DE3), we observed extremely low expression regardless of the degree of induction with isopropyl  $\beta$ -d-thiogalactopyranoside (IPTG) (Figure 2.1A). Previous reports have demonstrated that expression of M protein from the original SARS-CoV virus could be substantially enhanced by N-terminal fusion of M with a SUMO tag (Zuo et al. 2005). Following this same strategy resulted in enhancement of expression by  $\approx 4$  to 5 times (Figure 2.1B and D). We further optimized yields by quantifying expression of M-mCerulean3 and SUMO-M-mCerulean3 in various strains of *E. coli*, particularly the Walker strains, C41(DE3) and C43(DE3), which have been specifically selected for high expression of membrane-bound and other difficult proteins. We measured expression in these strains grown in both lysogeny broth (LB) and terrific broth (TB), finding that expression of SUMO-M-mCerulean3 in C41(DE3) grown in TB is  $\approx 45$  to 50 times higher than that of the standard use of BL21(DE3) grown in LB.

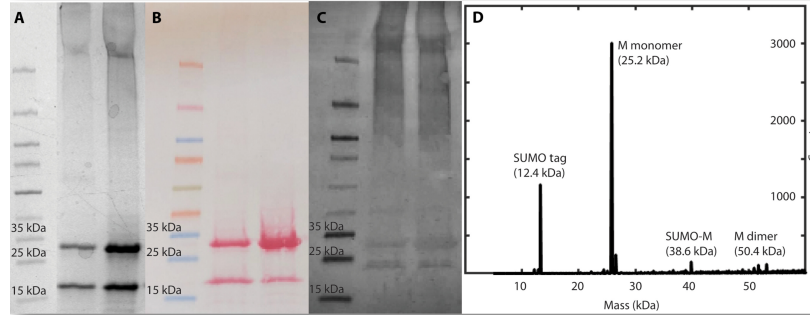
Purification of SUMO-M from C41(DE3) grown in TB using standard immobilized metal affinity chromatography (IMAC) resulted in high yields, typically on the order of 50 to 100 mg per liter of culture. Following purification, we refolded the protein by dialysis and removed the 6xHis tag and SUMO tag by digestion with Ulp1 SUMO protease, resulting



**Figure 2.1: Membrane protein expression in *E. coli*.**

(A to C) Fluorescence microscopy of M-mCerulean3 expression in terrific broth (TB) medium, with phase contrast image shown above and mCerulean3 fluorescence channel shown below. (A) M-mCerulean3 expressed in BL21(DE3). (B) SUMO-M-mCerulean3 expressed in BL21(DE3). (C) SUMO-M-mCerulean3 expressed in C41(DE3). All strains were imaged with identical microscope and camera settings, and images are displayed with identical lookup tables. Scale bars, 10.0  $\mu\text{m}$ . (D) Quantitative comparison of M-mCerulean3 expression in various *E. coli* strains and media [white, lysogeny broth (LB); light blue, TB; SUMO-M-mCerulean3 in LB, medium blue; SUMO-M-mCerulean3 in TB, dark blue]. Expression levels are normalized to expression of M-mCerulean3 in BL21(DE3).





**Figure 2.2: Verification of purified M protein.**

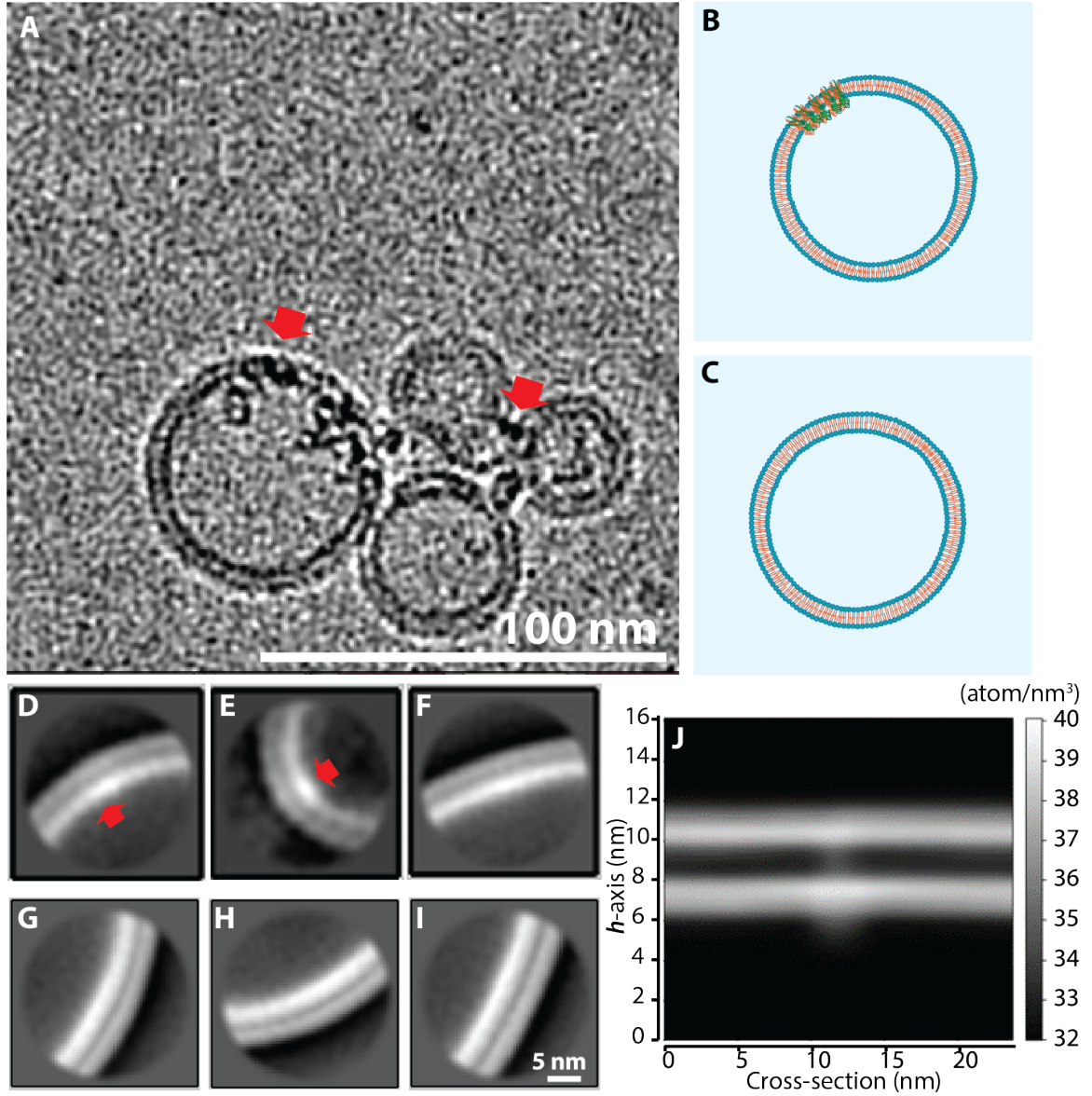
(A) SDS–polyacrylamide gel electrophoresis (PAGE) of purified M sample. Lane 1, ladder; lane 2, after digestion with 0.5 mg SUMO protease per milligram of M; lane 3, digestion with 2× amount SUMO protease. (B) Ponceau all-protein stain of protein transferred to blot from PAGE gel. Lanes same as (A). (C) Western blot of (A). (D) MALDI mass spectrum of purified M sample.

in pure, full-length, native M protein. We quantified the purity by SDS–polyacrylamide gel electrophoresis (PAGE), Western blot, and MALDI (Figure 2.2) and quantified endotoxin levels by Limulus amoebocyte lysate (LAL) endotoxin assay [ $0.35 \pm 0.05$  endotoxin units (EU)/mg]. SDS-PAGE (Figure 2.2A) and MALDI mass spectrometry (Figure 2.2D) show residual cleaved 6xHis-SUMO tag (12.4 kDa), which was subsequently removed by further IMAC purification. Furthermore, Western blot shows (Figure 2.2 B and C), in addition to a large band at the predicted M monomer molecular weight of 25.2 kDa, a large streak of high-molecular weight multimers and aggregates, as well as a small amount of lower-molecular weight products that may be degradation products. However, the majority of detectable products are M monomer and aggregates thereof. We emphasize that, while *E. coli* is unable to perform posttranslational modifications, such as glycosylation that is predicted to occur posttranscriptionally (Gong et al. 2021), it is known for the related SARS-CoV virus that these modifications to M protein are not required for viral assembly

and do not substantially impact protein-protein interactions (Oostraa et al. 2006)-(Laude et al. 1992).

### **2.2.2 Cryo-EM of lipid bilayer membrane reconstituted with M protein in LUVs**

Next, to facilitate studies of the M protein in its physiologically relevant context, we developed a procedure for M protein insertion into lipid membranes of large unilamellar vesicles (LUVs) with a phospholipid composition mimicking that of a typical endoplasmic reticulum-Golgi intermediate compartment (ERGIC) bilayer membrane [described in detail in Materials and Methods and in the Supplementary Materials (Fig. 3.1)]. We then used cryo-EM and AFM to verify the presence of the inserted M protein and its orientation, as described below. First, cryo-EM studies of LUVs with and without M protein were done to confirm the presence of M proteins in the bilayer membranes. A representative cryo-EM image of the M protein embedded in bilayer vesicles is shown in (Figure 2.3A), with the schematic of LUVs with (Figure 2.3B) and without (Figure 2.3C) M protein embedded. (Figure 2.3D to F) shows the typical two-dimensional (2D) class average images of cryo-EM observations from LUVs reconstituted with M protein at a protein-to-lipid mass ratio of 0.015. Two types of 2D class averages were observed in the cryo-EM images. (Figure 2.3D and E) shows the M protein as bright spots on the bilayer corresponding to higher electron density of the M protein embedded in the membrane, while, in the image in (Figure 2.3F), there is no M protein embedded. Most frequently observed spots were  $\approx 4$  to 5nm wide, which can be associated with single dimers or possibly some higher order



**Figure 2.3: Typical cryo-EM micrographs of LUVs reconstituted with M proteins.**

(A) Red arrows indicate M proteins inserted in the vesicle membrane. (B and C) Schematics of LUVs with and without M reconstitution, respectively. (D) Representative two-dimensional (2D) class average image of areas without M protein in reconstituted LUVs. (E and F) Two most representative cryo-EM 2D class average images from LUV bilayer membranes reconstituted with M protein mass ratio,  $R_{m,M/L} = 0.015$ . Arrows, location of M protein as bright spots. (G to I) 2D class average cryo-EM images of blank LUV bilayer membrane. Scale bar, 5.0 nm. (J) Atom number density of the short conformation M protein embedded in a bilayer membrane obtained from all-atom MD.

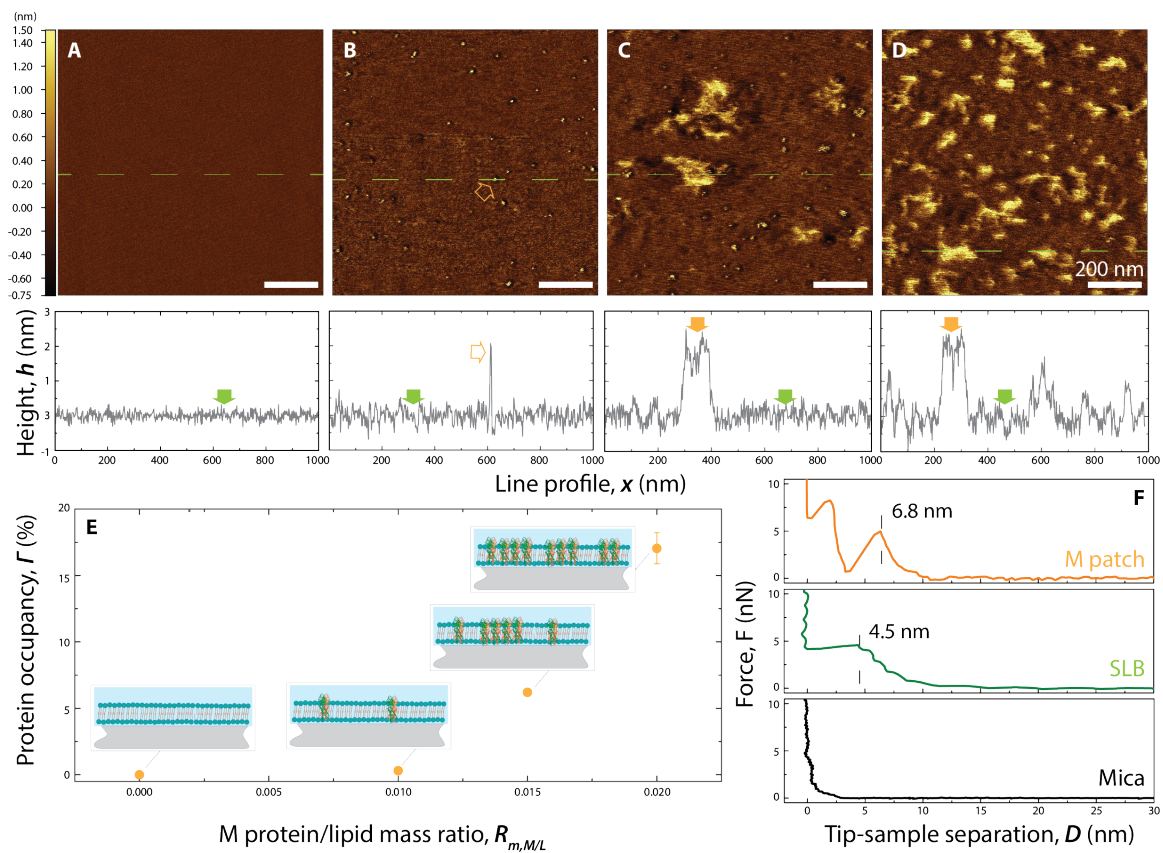
oligomers. We also observed that these spots corresponding to the M protein show a higher intensity on the inner leaflet of the vesicle, indicating that they are likely inserted with the larger and denser C-terminal facing the inside of vesicles. In contrast with regions reconstituted with M proteins, the regions without M proteins show homogeneous intensity from the corresponding uniform electron density throughout the bilayer membrane. In the bottom row as a negative control, blank vesicle samples without M protein were also imaged with cryo-EM. (Figure 2.3G to I) shows typical 2D class average images with a homogenous electron density observed for the LUVs without M protein. Last, (Figure 2.3J) shows a simulation of these cryo-EM micrographs with M protein calculated from the MD simulations of the short form of M protein embedded in a membrane by projecting the atom number density onto a plane perpendicular to the surface of the membrane using a bin size of 0.95 nm by 0.95 nm. Schematics for this MD simulation are shown in (Figure 3.3), with details provided in Materials and Methods. Two different viewing angles are averaged over in creating this simulated image, where lower-resolution density maps for both the long and short forms are shown in (Figure 3.4).

### 2.2.3 AFM topography of SBL reconstituted with embedded M protein

While the cryo-EM results presented above indicated the presence of the M protein in the lipid bilayer, the small molecular weight of 50.2 kDa for the dimer limits high-resolution analysis of the M protein–lipid membrane interaction. To confirm the cryo-EM results, we used the AFM with the height resolution of 0.01 nm and a lateral resolution 0.8 nm, which is ideally suited for understanding the inserted protein morphology, orientation, oligomerization at larger concentration, and the protein-lipid bilayer interaction (Gui et al.

2015; Chen et al. 2020). (Figure 2.4A) shows a 1  $\mu\text{m}$ -by-1  $\mu\text{m}$  AFM topography of a typical ERGIC bilayer membrane generated from LUVs without M protein reconstitution. The height profile along the green dashed line is shown in the bottom panel of (Figure 2.4A). The supported bilayer (SBL) on the mica substrate is smooth with root mean square roughness measured to be 0.3 nm. (Figure 2.4B) shows the typical topography of a 1  $\mu\text{m}$ -by-1  $\mu\text{m}$  surface of a SBL generated from LUVs reconstituted with M protein at a protein-to-lipid mass ratio of 0.01. In contrast to the smooth surface of the SBL in (Figure 2.4A), scattered single-particle protrusions were observed in the AFM image. These can be associated with the embedded protein in the SBL. The line profile (green dashed line) passing through one of the M proteins shows that the height of the feature above the membrane is 2.4 nm, consistent with the C-terminal height of M protein from literature (Zhang et al. 2022; Dolan et al. 2022) and simulations presented below. (Figure 2.4C and D) shows the SBL from LUVs reconstituted at higher protein-to-lipid mass ratios of 0.015 and 0.02, respectively. The topography in (Figure 2.4C) shows a combination of scattered single particles and patches, while patches are mostly observed in (Figure 2.4D). The line profiles show that the patches have the same height as the single particles associated with the M protein. (Figure 2.4E) shows the percent of protein-occupied area in the SBL membrane as a function of protein-to-lipid ratio. The area occupied by the protein is calculated from the sum of the single protein and patch areas. The M protein-occupied area increases from 0.3 to 6.2 and 17.1%, as the protein ratio increased from 0.01 to 0.015 and 0.02, respectively.

In addition to examining the planar surface features of the SBL with and without M protein, we also characterized the depth profile using AFM indentation studies. (Fig-



**Figure 2.4: AFM height image of SBL reconstituted with M protein at increasing protein-to-lipid mass ratios,  $R_{m,M/L}$ , imaged on mica substrates.**

(A to D) Height images and representative single line profiles (green dashed lines in images) collected from supported bilayer (SBL) fabricated from LUVs that were reconstituted with M protein at  $R_{m,M/L} = 0, 0.01, 0.015$ , and  $0.02$ . (E) Area occupation percentage of M protein in SBL as a function of  $R_{m,M/L}$  calculated from (A) to (D). (F) Representative force profiles collected from pristine mica surface (black), SBL (green), and M protein (orange) patches in lipid bilayer; the force profiles were collected during tip engagement. Scale bars: 200 nm (A–D)

ure 2.4F) shows a typical indentation force profile obtained from three different samples: bare mica in the bottom panel; SBL without and with M protein (concentration ratio of  $R_m, M/L = 0.02$ ) in the middle and top panels, respectively. In all three panels of (Figure 2.4F), the horizontal axis is the separation distance between the AFM probe tip and the bare mica surface. The bottom panel shows the typical indentation force curve of the bare mica surface in solution with the tip-rigid surface contact at zero followed by a sharp increase in force by the cantilever bending. The middle panel corresponds to the SBL membrane without M protein. Here, the AFM tip comes in contact with the hydration layer at a separation around 7.5 nm with the applied force leading to deformation of the membrane. With increasing applied force, the tip punctures the membrane surface at  $D = 4.5 \pm 0.1$  nm. This corresponds to the membrane thickness without M protein. With further increase in force, the AFM tip penetrates the membrane until it contacts the rigid mica at 0 nm. The top panel shows the indentation force profile obtained from the patch-like structures from the M protein on the SBL. The force showed a discontinuity at 6.1 nm which corresponds to the tip-protein surface interaction. Large forces lead to penetration of the tip into the M protein patch. The discontinuity at 3.2 nm indicates potential disruption of a specific strong interaction region between neighboring M proteins. From the slope of the approach curves in (Figure 2.4F), the Young's modulus can be calculated to be 9.5 MPa on the lipid bilayer and 41.5 MPa on the protein patches, respectively. (Figure 3.2) shows the average elastic moduli extracted from nanoindentation force profiles for SBL and the M protein patches. The difference in the Young's modulus indicates that the patches are a different type of material from the lipid bilayer.



### 2.2.4 Formation of M protein aggregate patches in the SBL

From (Figure 2.4B), we observed that M proteins in SBL start as single-M protein assemblies at an M protein-to-lipid mass concentration  $R_{m,M/L} = 0.01$ . From the size (length, width, and height), they can be identified as M protein dimers. With the increase of M protein concentration in the SBL, the dimers aggregate to form patches of M proteins as observed in (Figure 2.4C and D). As seen in (Figure 2.4E), the area of the patches grows nonlinearly with the M protein concentration. The M protein-occupied area increases from 0.3 to 6.2 and 17.1%, as the protein-to-lipid mass ratio  $R_{m,M/L}$  increased from 0.01 to 0.015 and 0.02. Thus, the M protein incorporation into the membrane is facilitated by the presence of M proteins already in the membrane. Notice that the height of patch-like assemblies is the same as the isolated single-M protein dimers observed at the lower M concentration. This indicates that the M proteins assemble in a side-by-side manner inside the bilayer, as demonstrated by schematics shown in (Figure 2.4E); these patches are higher order oligomers.

### 2.2.5 Morphological statistics of the isolated M protein particle

We next characterized the properties of single-M protein particles embedded in the SBL for protein-to-lipid mass ratio,  $R_{m,M/L} = 0.01$  (Figure 2.5). (Figure 2.5A) shows a high-resolution 250 nm-by-250 nm surface morphology of a typical single isolated M protein at the concentration ratio of  $R_{m,M/L} = 0.01$ . In (Figure 2.5B), the characterization parameters the length (L) and width (W) are shown. The boundary of the protein particle is identified as an edge higher than the planar membrane height. The planar membrane



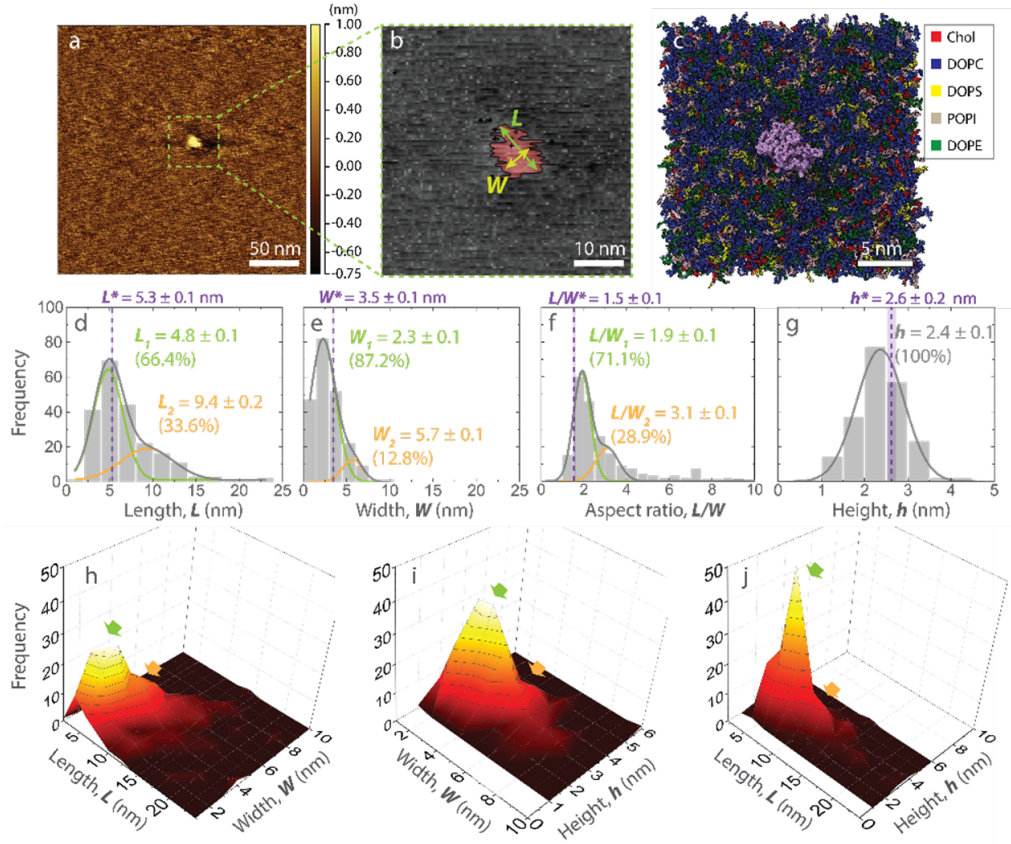


Figure 2.5: **Size and morphology statistics of M protein quantified from AFM images of SBL.**

(A) Representative AFM height image of individual M protein reconstituted in SBL at  $R_m, M/L = 0.01$ . (B) Zoomed-in image of (A), highlighting the protein with red color and lipid bilayer with gray color. The protruded portion of a single-M protein particle above lipid bilayer (left) is detected by setting the height of top surface of SBL as a threshold and taking the above portion. (C) Top view of the short conformation M protein (C terminus in purple) embedded in a multicomponent membrane after 1  $\mu s$  into an all-atom MD simulation. Dimensional parameters—length ( $L$ ), width ( $W$ ), and height ( $h$ )—were measured for detected protein particles. (D to G) Histograms of  $L$ ,  $W$ ,  $L/W$ , and  $h$  of protein protruding above lipid membrane. Solid lines are Gaussian fits. Vertical dashed lines mark corresponding values from MD simulation for short-form M protein reconstituted with C terminus exposed above SBL. (H to J) 3D histograms each showing combined statistical survey of two out of the three dimensions. Green and blue arrows indicate dimer and higher order of oligomers, respectively. Scale bar: 50 nm (A).

height is found by averaging the membrane height far away from the particle (gray region, Figure 2.5B). The particle thus identified is shown in red in (Figure 2.5B). The length (L) of the particle is defined as the maximum distance between any two points on the perimeter of the protein particle. The particle width (W) is defined as the maximum distance between two points on the perimeter, in the direction perpendicular to length. The particle height (h) is the height of the highest point on the protein surface. For comparison, (Figure 2.5C) shows the final frame of an all-atom MD simulation for the short conformation M protein embedded in a multicomponent membrane (see fig. S2 for schematics). The analysis of the AFM data is repeated for 214 samples, and the histogram of the results for L, W, h, and the ratio L/W are shown in (Figure 2.5D to G). Corresponding size classifications from MD simulations are shown in (Figure 2.5D to G), averaged over the last 100 ns of the simulation, through the purple dashed line see (Figure 3.6,3.7,3.8) for more information). Protein length and width were obtained by fitting an ellipse to the projection of the protein above the membrane, while height was based on the difference between the top of the M protein C-terminal and the phospholipid heads adjacent to the protein. The histogram in (Figure 2.5D) shows a bimodal distribution, which, when fit to two normal distributions, leads to mean particle lengths,  $L1 = 4.8 \pm 0.1$  nm and  $L2 = 9.4 \pm 0.2$  nm. (Figure 2.5E) shows the width distribution, again seen to be bimodal, which, when fit to two normal distributions, leads to mean values of  $W1 = 2.3 \pm 0.1$  nm and another at  $W2 = 5.7 \pm 0.1$  nm. In (Figure 2.5F), the aspect ratio L/W again shows two normal distributions: one at  $L/W1 = 1.9 \pm 0.1$  and the other at  $L/W2 = 3.1 \pm 0.1$ . (Figure 2.5G) shows the histogram

obtained from the height of the M protein particles. The values can be fit to one normal distribution with a mean height  $h = 2.4 \pm 0.1$  nm.

### 2.2.6 Confirmation of M protein insertion in the bilayer membrane

High-resolution cryo-EM studies have shown that SARS-CoV-2 membrane protein dimers exist in a conformational equilibrium between the long form and the short form, which are 7.2 and 8.6 nm in height, respectively (Zhang et al. 2022). Comparing these heights with those in (Figure 2.5G), AFM shows that the height of exposed protein particles is much shorter (2.4 nm). This is in agreement with the all-atom MD simulations, where the height ranges from 2.6 to 2.8 nm depending on the protein conformation (short/long, see (Figure 3.6B)). The height measured by AFM is more consistent with the short-form height from simulations; a more thorough discussion of these height fluctuations can be found in the Supplementary Materials (Figure 3.5). The slight differences between AFM and MD can be attributed to the protein deforming slightly as the AFM tip makes contact. Meanwhile, the lateral dimensions of isolated protein particles observed from AFM in (Figure 2.5D and E) are very similar to the literature reported values of 5.0 and 5.7 nm for long and short forms, respectively Zhang et al. (2022). A much smaller height and similar lateral dimensions indicate that the M proteins are inserted into the bilayer membranes instead of absorbing on top of the membranes. This is also supported by our cryo-EM observations in (Figure 2.3D and E). The membrane reconstituted with M protein shows a higher electron density compared with blank membrane. This indicates that a large portion of the protein is buried inside the bilayer membrane.

### 2.2.7 Orientation of reconstituted M protein with C terminus facing the vesicle interior

In (Figure 2.3D and E), we observed that the bright spots corresponding to reconstituted proteins show a higher electron density in the inner leaflets. This indicates that the C terminus, which is larger than the N terminus and contains a higher density of electron-rich “heavy” atoms, is inserted facing the inside of the vesicles. This is further confirmed in AFM measurements and also matches the simulated cryo-EM image from MD in (Figure 2.3J). This stems from averaging the protein over different angles, because the N terminus of the protein on the outer leaflet ranges from small to large arc length along the membrane. Alternatively, the C-terminal domain of the protein along the inner leaflet is largely independent of the viewing angle and will have greater arc length than the N-terminal domain (see (Figure 3.4) for viewing angle differences). In (Figure 2.5G), the height of reconstituted M protein exposed above bilayer membrane show a single distribution with an average height of 2.4 nm, which is consistent with C-terminal size observed in high-resolution cryo-EM images and our simulation predictions, indicating that all M proteins in the SBLs observed in AFM are oriented with their C termini facing the AFM probe. From these results, we can deduce that the M proteins were reconstituted in LUVs with their C terminus encapsulated inside the vesicles. This is because, during the formation of the SBL on mica, the LUVs adsorb to the mica surface with their outer leaflets. Subsequently, the LUVs rupture, exposing their inner leaflets facing upward and accessible to the AFM probe.

### 2.2.8 Reconstituted M protein dimensions and membrane thickness consistent with the short form

In (Figure 2.5H to J), the three-dimensional histograms analyze the relationships between protein length (L) and width (W), width and height (h), as well as length and height, respectively. In (Figure 2.5H) from the L and W, we can observe two distinct populations with their individual lengths and widths. From (Figure 2.5I) plot of the W and h, we can observe that both W populations have the same height above the membrane. From (Figure 2.5J), the same is observed for the two height populations. Combining the above information with results shown in the previous two-dimensional histograms (Figure 2.5D to G), the most frequently observed M protein population in the SBL have a length, width, and height of  $4.8 \pm 0.1$  nm,  $2.3 \pm 0.1$  nm, and  $2.4 \pm 0.1$  nm, respectively. These dimensions show a good agreement with the short form of M protein dimer both from published cryo-EM studies of M protein in detergent micelles (Zhang et al. 2022) and our simulation predictions [averages shown with the purple lines in (Figure 2.5D to G); see (Figure 3.6,3.7,3.8) for the time evolution of these quantities]. The slight discrepancy between the short conformation MD dimensions and AFM results likely stem from the original MD protein structures not including every residue and deformations in the protein from the AFM tip. The second population observed has a much larger size with a length, width, and height of  $9.4 \pm 0.2$  nm,  $5.7 \pm 0.1$  nm, and  $2.4 \pm 0.1$  nm, respectively. These dimensions correspond to that of a multimer of M protein dimers.

### 2.2.9 Reduction of membrane thickness around M protein

(Figure 2.6A) shows a typical isolated M protein particle in a magnified view, with color scale corresponding to the membrane thickness shown to the right. We observed that the membrane thickness in proximity to the single-M protein particle is reduced. The height of the membrane around each protein particle is measured along two perpendicular axes, shown in (Figure 2.6A) as green dashed lines. A similar view from above of the M protein short conformation obtained through all-atom MD simulations can be seen in (Figure 2.6B), where red regions represent a thicker membrane and blue thinner (see (Figure 3.9) for unshifted radial plot). The circular cross section of the protein shown in gold arises because of the protein rotation in the membrane during the simulation shown in (Figure 3.11). (Figure 2.6C) shows an exaggerated diagram of the membrane thinning near the protein and a corresponding side view of the thickness after 1  $\mu$ s of MD simulation is displayed in (Figure 2.6D). The change in membrane height from the edge of the protein (0 nm) along the dashed lines for 25 such proteins is shown in (Figure 2.6E). In addition, the blue points represent shifted membrane thickness values from (Figure 2.6B), based on the distance from the edge of the protein. These distances extend beyond half the box size as we include data from points along the diagonal, giving us a range from 0 to 16.63 nm. The mean change in height with distance from the protein edge is shown by the solid green line. We observed that the bilayer membrane is compressed until the height returns to the normal thickness of 12 nm from the edge of the protein. The maximum membrane compression at the edge of the protein has a mean value of 0.5 nm. It is to be emphasized that this AFM-determined thinning profile matches the MD short-form thinning profile, as

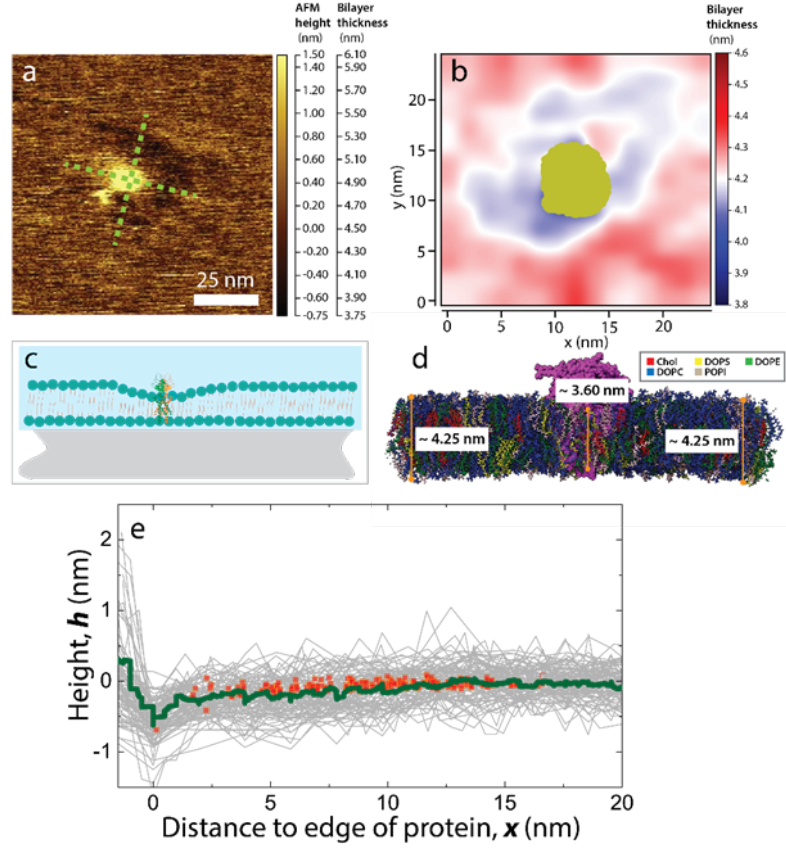


Figure 2.6: **Analysis of membrane thickness around individual M protein particles at  $R_m, M/L = 0.01$  and through MD simulation.**

(A) Example of individual M protein particles reconstituted in SBL. Height profiles were taken along the green dashed lines starting from the center of the protein. (B) View from above of membrane thickness in nm from 500 ns to 1  $\mu$ s of an all-atom MD simulation for the short conformation M protein embedded in a multicomponent membrane. (C) Schematics of membrane thinning around M protein particles. (D) Side view of the MD simulation along the y axis, with the protein in purple, where average thickness is shown dependent on distance from the protein. (E) Solid green line shows the experimentally measured averaged height profile of membrane around individual M protein particles. Inset shows around 100 height profiles obtained from 25 individual protein particles. Here,  $x = 0$  nm is defined by the boundary between the M protein particle and lipid membrane. The blue points represent vertically shifted values of membrane thickness obtained through MD shown in (B). Scale bar: 25 nm (A).

opposed to the long form (long form shown in (Figure 3.9) and agrees with observations of Neuman et al. (Neuman et al. 2011b) for SARS-CoV, where the short conformation is found in thinner regions of membrane as opposed to the long form.

### 2.2.10 Membrane compression, M protein aggregation, and spontaneous membrane curvature

The membrane thickness reduction around M protein observed in (Figure 2.6) is likely due to the height mismatch between the hydrophobic transmembrane domain and the thickness of the lipid bilayer. From the literature (Mahtarin et al. 2022) and as confirmed with our all-atom MD (Figure 3.6D), the transmembrane region of the M protein dimer is around 4 nm. The experimentally measured thickness of the lipid bilayer is 4.5 nm as shown in (Figure 2.4F). Thus, the protein hydrophobic region is  $\approx 0.5$  nm smaller than the membrane thickness. This hydrophobic mismatch leads to compression of the bilayer in the immediate proximity of the M protein that we show in (Figure 2.6C to E). The thinning of the membrane in the vicinity of the protein results in an energetic cost due to bend and tilt deformations in the lipid layers, which gives rise to an effective line tension (Kuzmin et al. 2005). Using our measurements of the mechanical moduli of the membrane, informed by estimates from literature (Doktorova et al. 2017) and the magnitude of the thinning, we can estimate the effective line tension to be at least about 0.35 pN (see the Supplementary Materials for details). This value is consistent with estimates and measurements of line tension for lipid domains with comparable height mismatches (Kuzmin et al. 2005),(?), which control the growth and size of lipid ordered domains (García-Sáez et al. 2007; Usery et al. 2017), suggesting that the line tension can facilitate the aggregation of M into patches.



Our coarse-grained simulation (see the next section) shows that this patch formation leads to the spontaneous curvature of the membrane. In addition, all-atom MD simulations of both M protein conformations show the propensity to individually induce membrane curvature, as seen in (Figure 3.10). However, the attractive interaction between the C-terminal endodomains is essential for the membrane budding (Figure 2.7). Our coarse-grained ERGIC simulation reveals that SARS-CoV-2 nucleocapsid (N) proteins and RNA can effectively introduce membrane curvature as well, providing insight into the mechanism underlying virus budding and encapsulation (Figure 3.12).

### 2.2.11 Coarse-grained MD simulation of M protein-induced budding

In general, our understanding of virus assembly is restricted by the limitations in the spatial and temporal resolution of experimental images, which hinder our ability to observe intermediate structures and states (Garmann et al. 2019), (Tresset et al. 2017), (Zandi et al. 2020), (Garmann et al. 2014). To address this challenge, coarse-grained simulations have been extensively used, successfully improving our understanding of virus assembly processes (Gelbart & Knobler 2009), (Hagan 2014), (Perlmutter et al. 2013), (Li et al. 2017), (Dykeman et al. 2014), (Adhikari et al. 2020), (Panahandeh et al. 2022). Here, we also perform a series of coarse-grained simulations by modeling M proteins and a flat lipid bilayer to illustrate how the inserted M proteins can give rise to virus budding, as shown in (Figure 2.7)A). Each M protein is built from three hard particles (M1, M2, and M3), where M1 and M2 correspond to the transmembrane domain and M3 denotes the endodomain. The membrane is modeled as a triangular lattice with hard particles (M0) occupying the

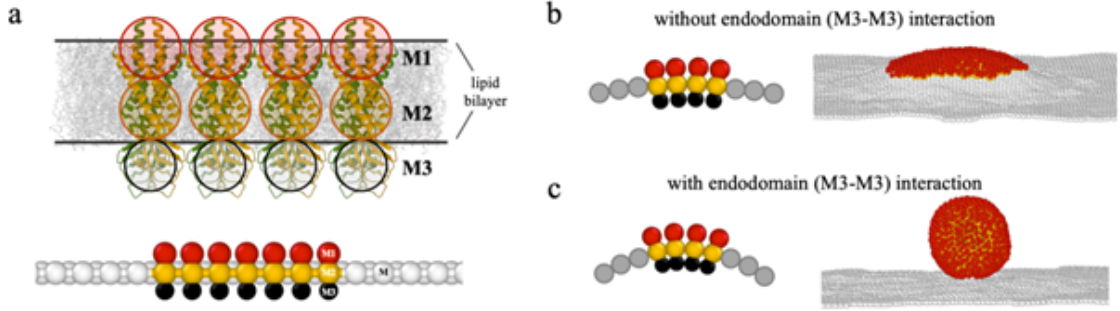


Figure 2.7: **Coarse-grained simulations of M proteins embedded in a flat membrane.**

(A) Schematic illustration of M proteins embedded in lipid bilayers. The top figure shows the actual size of the lipid bilayer and M proteins [Protein Data Bank (PDB): 7VGR (Zhang et al. 2022)], and the bottom figure shows the coarse-grained model used in the simulation. (B) Schematic figure (left) and simulation snapshot (right) for the case that the M proteins have no endodomain interactions ( $\epsilon_{M3-M3} = 0$ ). The M proteins aggregate and form a small bump due to the line tension, but no further budding is observed in the simulation. (C) Schematic figure (left) and simulation snapshot (right) for the case that M proteins have endodomain interactions ( $\epsilon_{M3-M3} = 8kBT$ ). The interaction helps M proteins overcome the energy barrier and the membrane buds into a spherical shell. For each M protein, the domain particles have a diameter size of  $1.0a$  (M1),  $1.0a$  (M2), and  $0.8a$  (M3), where  $a$  is the system length unit corresponding to 3 nm. For both simulations, the transmembrane domain M2 interacts with M2 with strength  $\epsilon_{M2-M2} = 4kBT$ . The membrane is represented as a triangular lattice where rigid particles (M0) occupy the vertices due to the incompressibility of the lipid membrane. The M0 particles interact with M with the strength  $\epsilon_{M-M} = 4kBT$ . The rest of particles have excluded volume repulsion defined through a cutoff Lennard-Jones potential.

vertices, on account of the incompressibility of the lipid membrane. We note that the thickness of the membrane is implicitly taken into account through the stretching and bending moduli of the triangular network (Li & Zandi 2022); see Materials and Methods for more details. Because most of the experiments and all-atom simulations here are performed on the flat membrane, we also use a flat membrane to investigate the role of membrane protein on curvature. The simulation results are shown in (Figure 2.7B and C), where M proteins interact with each other through Lennard-Jones potentials. (Figure 2.7B) shows that trans-membrane domain interaction (M2-M2) cannot give rise to M protein-induced membrane curvature, leading to the membrane budding even if the budding particle has lower energy. This is mainly due to the presence of the energy barrier required to bend the membrane. Consistent with the results for a curved ERGIC membrane (Li & Zandi 2022), we find that an attractive endodomain interaction is necessary under all the conditions examined to overcome the bending energy penalties. (Figure 2.7C) shows that, as we increase the M protein endodomain interaction (M3-M3), the membrane with the embedded M proteins can bud and form a spherical shell. We note that N proteins and RNA can also effectively provide the M protein endodomain interactions and help M proteins bud (Li & Zandi 2022); see (Figure 3.12). The budding prerequisites could indicate the mechanism of how coronavirus can encapsidate N proteins and RNA complex into an infective particle (Li & Zandi 2022).

## 2.3 DISCUSSION

This paper reports an expression and purification protocol for SARS-CoV-2 M protein using SUMO tags, which enables us and other members of the scientific community

to carry out many new experiments to better understand viral protein function and the life cycle of SARS-CoV-2. We have optimized the method to obtain a very high yield of 50 to 100 mg of protein per liter of culture. Furthermore, we developed a method to reliably insert M protein into phospholipid bilayers in vitro, and our cryo-EM images confirm that expressed M protein dimers are inserted homogeneously oriented into LUVs. We can precisely control the M protein concentration by adjusting the protein-to-lipid mass ratio used to create the vesicles. The production of LUVs with adjustable concentrations of inserted M protein constitutes a new experimental platform for studying the behavior and binding properties of vesicles with realistic densities of M protein.

To create a comprehensive picture of the structure and orientation of M protein in the membrane and the proteins' effect on the surrounding membrane, we deposited the M protein-containing vesicle as a supported membrane on a mica surface that allows high-resolution AFM imaging. This reveals the SBL retains a mixture of homogeneously oriented M protein dimers and multimers. This protocol results in the M protein C terminus being exposed on the surface of the suspended bilayer, allowing measurements of M protein binding to other SARS-CoV-2 structural proteins and viral RNA. In addition to AFM imaging, we performed atomistic MD simulations of the M protein in the suspended membrane. Our AFM and MD results indicate that the M protein is oriented perpendicular to the plane of the membrane, with the top of the C terminus extending 2.4 nm above the surface of the membrane, while the N terminus is virtually flush with the membrane surface on the lower leaflet. The membrane-inserted dimer has an elliptical cross section with major and minor diameters of 4.8 and 2.3 nm, respectively. In addition, a second population of embedded

protein is found with the same height but with diameters of 9.4 and 5.7 nm, corresponding to aggregates of dimers. This orientation of the M protein with the C-terminal endodomain extended above the membrane facilitates its proposed role as an anchoring site for other structural proteins.

A particularly interesting result from our study is the structure of the membrane in the immediate vicinity of the M protein dimer. Both the AFM images and the atomistic MD show a region of decreased membrane thickness surrounding the protein, on the order of 0.5 nm, which gradually returns to the normal thickness over a distance of 12 nm. This thinning is likely due to the mismatch in the width of the hydrophobic band in the trans-membrane region of the M protein dimer ( $\approx 4$  nm) and the hydrophobic lipid region of the membrane ( $\approx 4.5$  nm), which induces a line tension in the membrane estimated to be  $\approx 0.35$  pN. This line tension can drive aggregation of M protein dimers, even in the absence of specific inter-dimer attraction. Our membrane is suspended on a rigid surface, so we cannot directly observe via AFM whether the M protein dimers cause membrane bending, but our atomistic MD simulations show that the dimer induces a persistent fold in the membrane (which may manifest as a concave indentation for a simulation of a larger membrane patch). To connect the molecular scale M protein properties to the mesoscale formation of a coronavirus from all its constituent parts, we constructed a coarse-grained model of a membrane containing 14% M protein dimers. We then carried out MD simulations on this model to understand the combined effects of the endodomains, line tension, and spontaneous curvature generation in viral budding. Our simulations revealed that budding and spherical virion formation occur only if M proteins have an intrinsic curvature and interact with each other

via the endodomains. The experimental and simulation techniques developed in this study to date open the door to rapid progress toward fully elucidating the function of M protein. A full understanding of the role of M proteins in driving the budding and release of the newly formed SARS-CoV-2 virion from the cell can pave the path for achieving alternate ways of disrupting their formation.

## 2.4 MATERIALS AND METHODS

### 2.4.1 Strains and media

For protein expression and purification experiments, we used the commonly used protein expression strains BL21(DE3) (Thermo Fisher Scientific) ([Studier & Moffatt 1986](#)), ([Grodberg & Dunn 1988](#)), BL21-AI (Thermo Fisher Scientific), C41(DE3) ([Miroux & Walker 1996](#)), C43(DE3) ([Miroux & Walker 1996](#)), C41(DE3) pLysS ([Miroux & Walker 1996](#)), ([Studier 1991](#)), and C43(DE3) pLysS ([Miroux & Walker 1996](#)), ([Studier 1991](#)). Cells were grown in LB Miller [10 g of tryptone (Bacto, Thermo Fisher Scientific), 5 g of yeast extract (Bacto, Thermo Fisher Scientific), 10 g of NaCl (MilliporeSigma), and ultrapure H<sub>2</sub>O (Milli-Q IQ 7000, MilliporeSigma) to 1 liter of medium ([Bertani 1951](#)), ([Low 1974](#))] or modified TB [20 g of tryptone, 24 g of yeast extract, 4 ml of glycerol (Thermo Fisher Scientific), 0.017 M KH<sub>2</sub>PO<sub>4</sub> (MilliporeSigma), 0.072 M K<sub>2</sub>HPO<sub>4</sub> (MilliporeSigma), and ultrapure H<sub>2</sub>O to 1 liter of medium ([Tartoff & Hobbs 1987](#))].

### 2.4.2 Plasmid construction

Plasmids encoding and expressing M with an N-terminal 6xHis tag and with and without SUMO tag fusions were designed in VectorNTI software (Thermo Fisher Scientific) and synthesized de novo by GENEWIZ/Azenta Gene Synthesis service. Constructs were synthesized to be expressed from a T7lac promoter ([Dubendorf & Studier 1991](#)) a consensus Shine Dalgarno ribosomal binding site and were cloned into plasmid pUC57-kan-GW. A fluorescently tagged variant, M-mCerulean3 ([Markwardt et al. 2011](#)), was designed and produced similarly with C-terminal fluorescent fusion.

### 2.4.3 Cell growth and protein expression and purification

Cells were grown identically for all experiments. First, a seed culture was started by inoculation of 25 ml of medium containing kanamycin ( $25\mu\text{g}/\text{ml}$ ) in a baffled 125 ml of Erlenmeyer flask from freezer stock and allowed to grow overnight at  $37^\circ\text{C}$  in a C76 shaking water bath (New Brunswick Scientific). This culture was used to inoculate 1 liter of medium containing kanamycin ( $25\mu\text{g}/\text{ml}$ ) in a 2-liter baffled Erlenmeyer flask and grown at  $37^\circ\text{C}$  in a C76 shaking water bath until the optical density at 600 nm ( $OD_{600}$ ; SmartSpec Plus spectrophotometer, Bio-Rad) reached  $\approx 0.5$  to 0.6. At this point, the flask was removed from the water bath and transferred to a shaker (Dura-Shaker, VWR) at room temperature. After cooling to room temperature, the appropriate concentration of inducer to effect full induction was added (0.4% w/v l-arabinose for BL21-AI or 1 mM IPTG for all others), and the culture allowed to grow overnight to saturation.

#### 2.4.4 Quantification of protein expression and microscopy

To quantify expression of M protein in various *E. coli* strains and media, cells transformed with the indicated plasmids expressing variants of M-mCerulean3 were grown as described above. Two milliliters of culture was then collected and centrifuged (10,000g, Eppendorf 5424), and the supernatant medium was aspirated away. Cells were then washed with phosphate-buffered saline (PBS) three times and lastly resuspended in fresh PBS. Cells were diluted with PBS such that  $OD_{600} < 1.0$  and distributed to the wells of a 48-well plate (Corning Costar). The plate was placed in a CLARIOstar microplate reader (BMG Labtech), and the  $OD_{600}$  and mCerulean3 fluorescence were measured (excitation at 430/20 nm and emission at 480/20 nm). Expression levels reported in (Figure 2.1) are calculated as mCerulean3 fluorescence divided by  $OD_{600}$ . Each measurement was performed with three technical replicates over three experimental/biological replicates. Fluorescence microscopy was performed on a Nikon Ti-E inverted fluorescence microscope with an Andor DU-897 Ultra camera. For imaging of M-mCerulean3 fluorescence, cells were imaged with a 100 $\times$  total internal reflection fluorescence objective and illuminated with highly inclined 457-nm laser light from a Milles Griot argon laser fitted on a Nikon LU-4A laser launch and using a Chroma filter cube with Z457/10 $\times$  excitation and ET485/30m emission filters.

#### 2.4.5 Protein purification

After growth to saturation as described above, the 1 liter of culture was spun down for 30 min at 4300g in an Eppendorf 5910Ri refrigerated centrifuge maintained at 4°C, and the resulting supernatant was discarded. Cells were resuspended in 10 ml of lysis



buffer [50 mM NaH<sub>2</sub>PO<sub>4</sub>, 300 mM NaCl, and 10 mM imidazole (pH 8.0)] with lysozyme (1 mg/ml; MilliporeSigma) and benzonase nuclease (50 U/ml; MilliporeSigma) and incubated at 4°C for 30 min on a nutating shaker. The cells were flash-frozen in liquid nitrogen, followed by incubation in a room temperature water bath until thawed. This slurry was then centrifuged at 10,000g at 4°C for 30 min. The supernatant clarified cell extract (soluble fraction) was collected and stored at −80°C until purification. The pelleted cell debris was then washed three times in inclusion body wash buffer (PBS with 25% w/v sucrose, 5 mM EDTA, and 1% Triton X-100) followed by centrifugation at 20,000g for 30 min at 4°C. After the final wash, inclusion bodies were resuspended in denaturing lysis buffer [8 M Urea (MilliporeSigma), 10 mM Tris (Thermo Fisher Scientific), 100 mM NaH<sub>2</sub>PO<sub>4</sub> (MilliporeSigma), 50 mM 3-(cyclohexylamino)-1-propanesulfonic acid (CAPS; MilliporeSigma), 0.3% N-lauroyl sarcosine (MilliporeSigma), and 1 mM dithiothreitol (DTT; MilliporeSigma) (pH 11.0)] with nutation at room temperature for 30 min. This insoluble fraction was then stored at −80°C until purification.

For purification, samples were processed using a Bio-Rad NGC Quest 10 Plus fast protein liquid chromatography apparatus with a 5-ml EconoFit Nuvia IMAC affinity column (Bio-Rad). To purify protein from the soluble fraction, the column was first equilibrated with Wash Buffer A (50 mM NaH<sub>2</sub>PO<sub>4</sub>, 300 mM NaCl, and 20 mM imidazole). The sample was loaded onto the column and washed with five column volumes of Wash Buffer A. The column was then washed with a linearly increasing mixture of Wash Buffer A mixed with Wash Buffer B (50 mM NaH<sub>2</sub>PO<sub>4</sub>, 300 mM NaCl, and 250 mM imidazole) from 3 to 100% Wash Buffer B composition over 10 column volumes. Last, the column was washed with five col-

umn volumes of Wash Buffer B. Throughout, the composition of the eluate was monitored by measuring the absorption at 280 nm, and those samples containing the protein of interest were collected and pooled for further processing. To purify protein from the insoluble fraction, the procedure was identical using Denaturing Wash Buffer A [50 mM NaH<sub>2</sub>PO<sub>4</sub>, 300 mM NaCl, 50 mM CAPS, 0.3% N-lauroyl sarcosine, and 20 mM imidazole (pH 11.0)] and Denaturing Wash Buffer B [50 mM NaH<sub>2</sub>PO<sub>4</sub>, 300 mM NaCl, 50 mM CAPS, 0.3% N-lauroyl sarcosine, and 300 mM imidazole (pH 11.0)]. All purified SUMO-fused proteins eluted as a single peak.

#### **2.4.6 Protein refolding and dialysis**

To refold denatured proteins purified from the insoluble fraction, samples were dialyzed in regenerated cellulose dialysis membrane [3.5-kDa molecular weight cutoff (MWCO), Spectra/Por] over 2 days at 4°C in 10× volume of refolding buffer [20 mM tris and 10% glycerol (pH 8.0)] with periodic replacement of buffer at least four times over the course of dialysis.

#### **2.4.7 Cleavage of SUMO tags**

To remove the cleaved 6xHis-SUMO tags and SUMO protease, protein digests were mixed with 1 ml of HisPur Ni-NTA resin (Thermo Fisher Scientific) and placed on a rotating mixer maintained at 4°C overnight. After incubation, the samples were applied to an empty Poly-Prep chromatography column (Bio-Rad) and the purified proteins collected as the flow-through. If higher concentrations were required for a given application, then

these samples were concentrated using a Pierce PES protein concentrator with appropriate MWCO (Thermo Fisher Scientific).

#### **2.4.8 Quantification of yields and endotoxin levels**

During purification, absorbance at 280 nm was measured by the NGC Quest 10 Plus optical detector and yields calculated using extinction coefficients estimated using the ExPASy ProtParam online tool based on protein sequence ([Wilkins et al. 1999](#)). Estimated concentrations were verified using an RC DC protein assay (Bio-Rad). Briefly, 5 $\mu$ l of protein samples and standards (Protein Standards I and II, Bio-Rad) were added to 25 $\mu$ l of reagent A and mixed in the wells of a 96-well plate (Corning Costar). To this, 200 $\mu$ l of reagent B was added to each well, mixed, and allowed to incubate at room temperature. After 15 min, the absorbance at 750 nm was measured using a CLARIOstar microplate reader, and the concentrations of the samples determined by comparison to linear regression of 2 $\times$  dilutions of the protein standards.

Endotoxin levels of purified samples were determined using a ToxinSensor chromogenic LAL endotoxin assay kit (GenScript) according to the manufacturer's instructions. Samples (16.67 $\mu$ l), endotoxin standards, and water as a negative control were added to the wells of a 96-well plate (Corning Costar). To this, 16.67 $\mu$ l of LAL was added to each well and mixed. The plate was incubated at 37°C for 10 min. Reconstituted chromogenic substrate (16.67 $\mu$ l) was added, mixed, and incubated at 37°C for 6 min. Then, 83.3 $\mu$ l of color stabilizer 1 was added and mixed, followed by 83.3 $\mu$ l of color stabilizer 2, and lastly 83.3 $\mu$ l of color stabilizer 3. The absorbance of each sample was measured at 545 nm in a CLARIOstar

microplate reader, and the endotoxin levels of the samples were determined by comparison to linear regression using  $2\times$  dilutions of the endotoxin standards.

#### **2.4.9 SDS-PAGE and Western blots**

To perform SDS-PAGE, 100 ng of each purified protein was mixed with an appropriate volume of  $2\times$  Laemmli sample buffer (Bio-Rad) with 5 mM  $\beta$ -mercaptoethanol (MilliporeSigma). Samples were heated at  $95^{\circ}\text{C}$  for 5 min to denature and reduce the proteins and were then separated by electrophoresis through 4 to 15% Mini-PROTEAN TGX Stain-Free gel (Bio-Rad) with tris/glycine/SDS running buffer. Gels were subsequently stained with Oriole fluorescent gel stain (Bio-Rad) and visualized on a UV transilluminator 2000 (Bio-Rad). For Western blot, the total protein concentration was determined using the bicinchoninic acid protein assay kit (Pierce/Thermo Fisher Scientific, Rockford, IL). Protein samples ( $25\mu\text{g}$ ) were mixed with  $4\times$  lithium dodecyl sulfate (LDS) sample buffer (Life Technologies, CA, USA) and  $10\times$  reducing agent (Invitrogen, Carlsbad, CA) and heated for 5 min at  $100^{\circ}\text{C}$ . Samples were loaded in a 4 to 12% SDS-PAGE gel (NUPAGE, Invitrogen) for electrophoretic separation and subsequently electro-transferred to polyvinylidene difluoride membrane (Invitrogen, USA). Membranes were blocked with 5% bovine serum albumin in TBST (tris-buffered saline with Tween 20) for 1 hour at room temperature and, afterward, incubated overnight at  $4^{\circ}\text{C}$  in primary antibodies. Membranes were then washed with TBST (three times for 15 min) and incubated with anti-rabbit horseradish peroxidase (111 – 036 – 045; Jackson ImmunoResearch; 1 : 5000) secondary antibody. Membranes were then imaged with a Bio-Rad ChemiDoc system (Bio-Rad, Hercules, CA) and analyzed using ImageJ software (National Institutes of Health). The following primary antibodies

were used: SARS-CoV-2 membrane protein antibody (ProSci, 9165; 1 : 1000), SARS-CoV-2 envelope protein antibody (ProSci, 9169; 1 : 1000), SARS-CoV-2 N protein (RayBiotech, QHD43423; 1 : 1000), and SARS-CoV-2 spike protein antibody (Novus biologicals LLC, NB100-56578; 1 : 1000).

#### 2.4.10 Protein reconstitution into vesicles

Monodisperse LUVs were extruded using a vesicle extruder. Lipids with composition mimicking that of the ERGIC were dissolved in chloroform with a solid concentration of 5 mg/ml. The molar ratio of different lipids (Avanti Polar Lipids Inc., Alabaster, AL) are 1-palmitoyl-2-oleoyl-glycero-3-phosphocholine:1-palmitoyl-2-oleoyl-sn-glycero-3-phosphoethanolamine:1-palmitoyl-2-oleoyl-sn-glycero-3-phosphoinositol (POPI):1-palmitoyl-2-oleoyl-sn-glycero-3-phospho-l-serine:cholesterol = 0.45:0.2:0.13:0.07:0.151.

The chloroform solution was dried in a glass vial with a gentle  $N_2$  gas stream and then vacuumed overnight at  $-30$  in Hg at room temperature. The dried lipid mixture was hydrated with biologically relevant buffer (150 mM NaCl and 20 mM HEPES, pH 7.2) with a 30-s vortex, followed by 10 freeze-thaw cycles in dry ice and at  $37^\circ\text{C}$ . After the final thawing step, the aqueous solution was passed 11 times through a polycarbonate membrane with 100-nm pores (Nuclepore Track-Etch membrane, Whatman, Chicago, IL).

The size and zeta potential of the extruded membrane were measured using dynamic and phase analysis light scattering (PALS) techniques with a 90 Plus PALS machine (Brookhaven Instruments, Holtsville, NY).

To reconstitute the M proteins into LUVs, the concentrated stock solution of *n*-dodecyl- $\beta$ -*D*-maltoside (Avanti Polar Lipids Inc., Alabaster, AL) was added to the freshly

extruded LUV solution (5 mg/ml) to reach a final concentration of 80 mM. M protein stabilized by Triton X-100 (stock solution: 1 wt% Triton X-100 per every M protein at 2 mg/ml) was then added to the LUV solution at a mass ratio of M/lipid = 1/100, 1/67, and 1/50 after 30 min of incubation. The solution was allowed another 30 min of incubation before the addition of 40 mg of wet BioBeads (Bio-Rad, Hercules, CA) per milliliter of LUV solution to slowly remove the detergent over 3 hours.

The second and third doses of BioBeads at 40 mg/ml were also added at 3-hour intervals. A final dose of BioBeads was added at a ratio of 380 mg/ml and incubated for another 3 hours. The M-reconstituted LUVs were separated from the solution via centrifugation at 13,000 rpm for 30 min in a microcentrifuge tube.

#### **2.4.11 Preparation of SBL samples**

To prepare SBL on mica for AFM imaging, 75 $\mu$ l of LUV solution collected from the bottom of the microcentrifuge tube was added onto freshly cleaved pristine mica and incubated for 20 min to 1 hour for different coverage. The sample was then gently rinsed with 5 ml of buffer. Caution was taken not to expose the SBL sample directly to air. The samples were kept submerged in aqueous buffer until imaging in an AFM fluid cell. Imaging was performed within 1 hour of the sample preparation.

#### **2.4.12 AFM imaging, force profile collection, and analysis**

AFM height images were obtained inside a fluid cell in tapping mode using a MSNL cantilever (Bruker, Camarillo, CA). Images were taken in 512  $\times$  512 line resolution over areas ranging from 250 nm by 250 nm to 1 $\mu$ m by 1 $\mu$ m. The effective tip size was

calibrated with 10-nm gold nanoparticles (Ted Pella, Redding, CA) using our previously reported protocol 2. Membrane structure was probed with nanoindentation to obtain force profiles. The force profiles were collected in force volume mode with a  $16 \times 16$  sampling resolution over the designated area of interest of 500 nm by 500 nm. To probe the continuum elastic property of the membrane and protein patches, an MLCT probe (Bruker, CA) with a larger calibrated tip radius of 23 nm was used. The area of interest was identified from prior tapping mode imaging. The spring constant of the larger tip radius cantilevers was calibrated to be 0.15 N/m using its thermal oscillation spectrum.

AFM image analysis were performed in Mountains SPIP software (Digital Surf, France). Protein particles were identified by setting the height of SBL top surface as the threshold and only identifying out protrusions above the membrane threshold. The software-identified protrusions were then manually examined to exclude noise and false identification of the protein particles. The false identifications by the software can be ruled out as they are 1- to 2-pixel-wide noise peaks that are much smaller than the proteins that are greater than 40 pixels depending on whether they are isolated protein dimers or large patches.

#### **2.4.13 Cryo-EM data collection and analysis**

Cryo-samples were prepared using a Vitribot Mark IV (Thermo Fisher Scientific). For each sample, Quantifoil grid R2/1 was glow-discharged before a  $3.5 \mu\text{l}$  aliquot was applied and blotted for 4 s at 95% humidity at room temperature. The samples were then plunged into liquid ethane and preserved in liquid nitrogen. The cryo-EM micrographs were obtained using a Talos Arctica (FEI) operating at 200 kV equipped with a Falcon 4i Direct Electron Detector (Thermo Fisher Scientific). Imaging was performed at  $150,000\times$

magnification with a resolution of  $0.95 \text{ \AA}/\text{pixel}$  and a defocus range of  $-1.5$  to  $-2.5 \mu\text{m}$ . About 150 to 250 images were collected for each sample using automated data acquisition with EPU software (Thermo Fisher Scientific). Images were motion-corrected within EPU. The dataset was processed using Relion software. Approximately 500 areas of interest were manually picked to generate the 2D class average images.

#### 2.4.14 All-atom MD simulation

All-atom MD simulations were performed using the CHARMM36m force field with the MD package GROMACS, version 2022.3 (Bauer et al. 2022),(Huang et al. 2017). The CHARMM-GUI input generator was used to set up the simulated systems with periodic boundary conditions and supplied the six steps used for equilibration(Jo et al. 2007),(Jo et al. 2008),(Brooks et al. 2009),(Jo et al. 2009),(Wu et al. 2014),(Lee et al. 2016),(Lee et al. 2019),(Lee et al. 2020),(Park et al. 2021). After equilibration, each system was simulated for 1 ms with a time step of 2 fs in the NPT ensemble. System temperature was maintained at  $303.15K$  using the Nose-Hoover thermostat (Hoover 1985),(Nosé 1984), with the pressure maintained semi-isotropically at 1 bar in the xy dimensions and separately in the z dimension using the Parrinello-Rahman barostat(Nosé & Klein 1983),(Parrinello & Rahman 1981). The coordinates for each simulation were saved once every 50,000 time steps or every 0.1 ns for a total of 10,000 frames.

Three different membrane systems were simulated: one for each form of the M protein dimer (short versus long) and one as a control only consisting of a membrane. For the simulations involving an M protein dimer, the protein was inserted into the membrane with an orientation and depth that matched other studies (Zhang et al. 2022),(Dolan



et al. 2022). Figure S2 shows the schematics for these membrane simulations. In the case of the short form, only residues 9 to 204 were considered [Protein Data Bank (PDB): 7vgs], while residues 9 to 206 were included in the long form structure (PDB: 7vgr)(Zhang et al. 2022). The membrane was composed of 15% cholesterol, 45% 1,2-dioleoyl-sn-glycero-3-phosphocholine (DOPC), 20% 1,2-Dioleoyl-sn-glycero-3-phosphoethanolamine (DOPE), 7% 1,2-dioleoyl-sn-glycero-3-phospho-L-serine (DOPS), and 13% 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphoinositol (POPI) in both leaflets, with the solvent consisting of NaCl at a concentration of 0.15 M and transferable intermolecular potential with 3 points (TIP3P) water. Each simulation, depending on the inclusion of a protein, consisted of a 25 nm-by-25 nm membrane in the xy plane with at least 5 nm of solvent above and below the protruding protein, yielding a total unit cell thickness of  $\approx 18$  nm in the z dimension. For runs involving an M protein, the final trajectory was reoriented frame by frame such that the protein was centered in the box for all frames. While each trajectory was fitted to eliminate protein translation across the membrane, this was not the case in the direction normal to the membrane and for the rotation of the protein. To analyze the processed trajectory, the python library MDAnalysis was used(Gowers et al. 2016),(Michaud-Agrawal et al. 2011). Images of these systems were generated using ChimeraX (Pettersen et al. 2004). In addition, all-atom simulations for each conformation of the M protein dimer in solvent were performed to compare protein stability with the membrane simulations. These systems involved a box size of approximately 15(16) nm by 15(16) nm by 15(16)nm for the short(long) conformation. The same process and conditions as the membrane simulations were used, with slight deviations in the equilibration steps due to the lack of a membrane.

### 2.4.15 Coarse-grained simulation

The coarse-grained simulation is performed using HOOMD-blue package (Anderson et al. 2020) and BondFlip plugin (Li & Zandi 2022). For the membrane, we initially assign a 100a-by-84a triangular lattice plane and place M proteins in the middle disk area with a radius of 15a. The dynamics of the membrane is modeled through Langevin integrator with  $dt = 0.001$  s at room temperature. In every 100 steps, we apply Monte-Carlo bond flip to change the local connectivity of the membrane, where each bond has stretching and bending energy with spring constant  $20kBT/2a$  and bending rigidity  $20kBT$  (Li & Zandi 2022). We note that each simulation is run for 10,000 s, and a height constraint  $[h(t) = 0]$  is applied for the region with radial distance  $r > 35a$  through the simulation time. The simulation is performed on NVIDIA GeForce RTX 3090, and the visualization is performed with OVITO software (Stukowski 2010).

## 2.5 Acknowledgement

### 2.5.1 funding

This work was supported by the University of California Office of the President UC Multicampus Research Programs and Initiatives, grant M21PR3267 (M.E.C., A.G., U.M., R.Z., and T.E.K.); National Science Foundation RAPID, grant 2034794 (U.M., R.Z., and T.E.K.); National Science Foundation, NSF DMR-2131963 (R.Z. and S.L.); National Science Foundation, NSF-CREST: Center for Cellular and Biomolecular Machines at UC Merced, NSF-HRD-1547848 and NSF-HRD-2112675 (A.G.); National Institutes of Health, NIH G-RISE, T32GM141862 (J.M. and A.G.); National Science Foundation, Center for Engineering

Mechanobiology, grant CMMI-1548571 (A.G.); and National Science Foundation, Pinnacles Computing Cluster, NSF-ACI-2019144 (J.M.).

### **2.5.2 Author contributions**

Conceptualization: Y.Z., S.A., J.M., A.G., U.M., R.Z., and T.E.K. Methodology: S.A., Y.Z., J.M., S.L., M.W., P.M., M.E.C., A.G., R.Z., and T.E.K. Investigation: Y.Z., S.A., J.M., S.L., M.W., P.M., U.M., R.Z., and T.E.K. Resources: S.A., P.M., M.W., M.E.C., U.M., R.Z., and T.E.K. Formal analysis: Y.Z., S.A., J.M., S.L., M.E.C., A.G., U.M., and R.Z. Software: J.M., S.L., M.E.C., and R.Z. Visualization: S.A., J.M., S.L., P.M., A.G., U.M., R.Z., and T.E.K. Writing—original draft: Y.Z., J.M., S.L., M.E.C., A.G., U.M., R.Z., and T.E.K. Writing—review and editing: J.M., M.E.C., A.G., U.M., R.Z., and T.E.K. Supervision: M.E.C., A.G., U.M., R.Z., and T.E.K. Project administration: A.G., U.M., R.Z., and T.E.K. Funding acquisition: A.G., U.M., R.Z., and T.E.K.

## **Data Availability**

All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials. Requests for resources and reagents should be directed to and will be fulfilled by the lead contact, T.E.K. [thomas.kuhlman@ucr.edu](mailto:thomas.kuhlman@ucr.edu)

## Chapter 3

# SARS-CoV-2 Structural Proteins - Supplemental Information

### 3.1 Determination of Detergent Conditions for M Protein Insertion

89 The proper function of transmembrane M protein requires insertion into the bilayer membrane with the correct orientation. In this study, the protein insertion utilized a detergent assisted procedure and was optimized for M protein in terms of the type and concentration of detergent used. (Figure 3.1) shows the key results obtained during the optimization process. The as-extruded lipid vesicles were measured to exhibit an intensity-averaged hydrodynamic size of 123 nm in DLS (Fig. 3.1A). As shown in (Fig. 3.1B), as detergents are gradually introduced into the vesicle solution, the lipid vesicles undergo a three-stage evolution, which has been well established in previous studies (Rigaud &

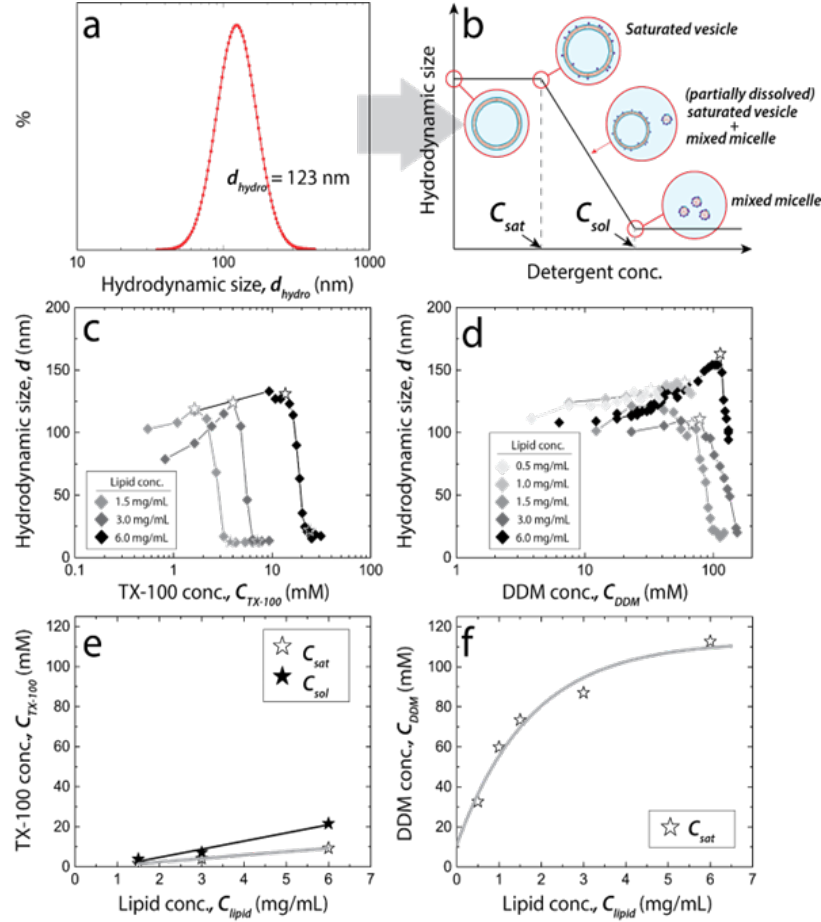


Figure 3.1: **Dynamic light scattering (DLS) study of vesicle size as a function of detergent concentration.**

(A) Intensity distribution vs. hydrodynamic size measured with DLS for as-extruded lipid vesicles in biological buffer. (B) Schematics of the three stages of vesicle dissolution with increasing detergent concentration. (C) and (D) are hydrodynamic size of vesicles measured as functions of TX-100 and DDM detergent concentrations, respectively. The empty and filled stars represent the detergent-saturation concentration,  $C_{sat}$  and vesicle dissolution concentration,  $C_{sol}$ . (E) and (F) are  $C_{sat}$  and  $C_{sol}$  observed at different lipid concentrations for TX-100 and DDM, respectively.

[Lévy 2003](#); [Paternostre et al. 1988](#)). In the first stage, the detergent molecules insert themselves into the lipid vesicle bilayer structures without dissolving the vesicles. In this stage, with more detergent addition, the detergent concentration in the bilayer membrane increases gradually without decreasing the size and number of vesicles. At the saturation concentration,  $C_{sat}$ , the vesicle becomes saturated with detergent. Increasing the detergent concentration beyond saturation leads to the second stage where further detergent addition results in the partial dissociation of the original vesicles. Here, the hydrodynamic size of vesicles significantly decreases with detergent addition, as they are being depleted of lipids. The dissociated lipid molecules form mixed micelles with the detergents. At the end of the second stage, all the vesicles are dissociated completely at the vesicle dissolution concentration,  $C_{sol}$ . In this third stage, only mixed micelles exist. Further addition of detergents only increases the detergent composition in the mixed micelles.

In this study, two non-ionic detergents, TX-100 and DDM were tried. Unlike their ionic counterparts, these detergents are mild and do not disrupt protein folding ([Rigaud & Lévy 2003](#)). (Figs. [3.1C,D](#)) show the hydrodynamic size change with detergent concentration for TX-100 and DDM, where the three stages described in (Fig. [3.1B](#)) were observed for most cases. TX-100 was found to saturate and dissolve vesicles at much lower detergent concentrations compared with DDM. Fig. ([3.1E](#)),f show the saturation and dissolution concentrations as a function of the total lipid concentration in solution for TX-100 and DDM. In both cases, saturation and dissolution concentrations were found to increase with increased lipid concentration. For the concentrations studied TX-100 shows a very gradual linear increase while DDM the saturation concentration increases rapidly and then reaches a

plateau at higher lipid concentrations. Thus, DDM provides much better control and higher tolerance and therefore for the M protein insertion into the lipid bilayers. The DDM concentration used for M protein insertion was 80% of Csat. A similar detergent (octylglucoside) was reported to demonstrate optimum transmembrane protein function and unidirectional insertion in this concentration regime (Rigaud et al. 1988). The unidirectional insertion has been verified by the CryoEM and AFM results as reported below. After M protein incubation, the detergent removal rate was also optimized. The detergent was removed via hydrophobic adsorption, where hydrophobic polystyrene resin (Allen et al. 1980) (SM-2) bio beads at a 40 mg/mL were added into the solution. This lower bead concentration in contrast to the usual 80 mg/mL lead to a slower detergent removal rate, which facilitated a higher efficiency in the protein insertion.

### 3.2 Determination of Elastic Moduli from AFM Measurements

(Figure 3.2) shows the average elastic moduli extracted from nanoindentation force profiles for SBL and the M protein patches. SBL exhibits a Young’s modulus of 9.5 MPa, which agrees with literature values (Drabik et al. 2020). The M protein patches demonstrated a much higher modulus value of 41.5 MPa, indicating the patches are indeed a different type of material from the lipid bilayer.

AFM nanoindentation was performed in force volume mode using a MLCT probe (Bruker, CA). The spring constant was calibrated using a thermal tune,  $k = 0.15$  N/m.

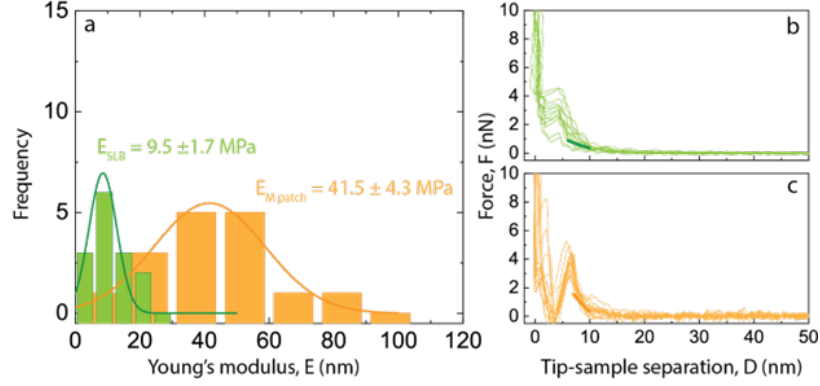


Figure 3.2: **Elastic modulus of SBL and M protein patches obtained from nanoindentation technique.**

(A) Histogram of Young's modulus obtained from analyzing the AFM nanoindentation force profiles using Hertzian contact model. (B) and (C) are the force profiles obtained from SBL and protein patches, respectively. The thick solid lines represent the force profile calculated using the mean elastic modulus values from (A).

Tip radius was calibrated with 10 nm gold colloids (Ted Pella, CA),  $R = 23$  nm. The force profiles were collected over a  $1 \mu\text{m} \times 1 \mu\text{m}$  area at a  $16 \times 16$  resolution. To extract the Young's modulus of SBL and the protein patches, a linearized Hertzian contact model was utilized. According to Hertzian model, the relationship between measured force and sample indentation can be expressed as follow:

$$F = \frac{4}{3} \frac{ER^{1/2}}{(1 - \sigma^2)} \delta^{3/2}. \quad (3.1)$$

Here,  $F$  is the force measured from AFM (Hook's law,  $F = \text{cantilever deflection} \times \text{calibrated spring constant}$ ).  $R$  is the calibrated tip radius of AFM probe,  $\sigma$  is the Poisson's ratio. Here, both  $\sigma_{\text{SBL}}$  and  $\sigma_{\text{Mprotein}}$  are taken to be 0.5, as most biological materials such as lipids and proteins are considered incompressible.  $\delta$  is the indentation depth. The



acquisition of indentation depth is relatively complicated considering it requires the identification of exact point of tip-sample contact. However, the tip-sample separation,  $D$ , can be easily obtained as raw data and the relationship between  $D$  and  $\delta$  is simply:

$$D = c - \delta. \quad (3.2)$$

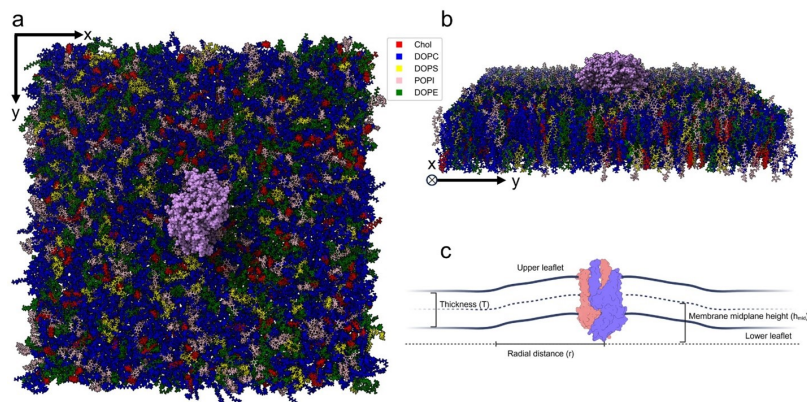
Here,  $c$  is a constant. To simplify the analysis, the Hertzian equation was linearized (Guo & Akhremitchev 2006) as follows:

$$F^{2/3} = C - D * \left[ \frac{4}{3} \frac{ER^{1/2}}{(1 - \sigma^2)} \right]^{2/3} \quad (3.3)$$

Here,  $C$  is another constant. Therefore, by obtaining the slope of the linear relationship between  $F^{2/3}$  and  $D$ , we can extract the elastic modulus. Here, to rule out the substrate effect and make sure the  $E$  is extracted from the linear elastic region, only the initial 20% of force curve was used to be analyzed by the Hertzian model.

### 3.3 Atomistic Molecular Dynamics (MD) Simulations of M-protein Stability

Two interconvertible forms of the M protein homodimer have been reported for SARS-CoV and SARS-CoV-2 (Zhang et al. 2022; Neuman et al. 2011b). To investigate the relative stability of these forms we performed MD simulations of both the long and short forms based on Cryo-EM structures (Zhang et al. 2022). We ran these simulations in two different environments, embedded within a membrane bilayer (as shown in Figure 3.3 ) and



**Figure 3.3: Initial Simulation Schematic for M Protein Embedded in a Multi-component  $\sim 25$  nm x 25 nm Membrane.**

(A) A view from above looking down on the lower leaflet, with the elliptical shape of the C-terminal for the short form shown in purple. (B) A view from the side. (C) Diagram of a long form M protein dimer embedded in a membrane, where the differing protein colors refers to the monomers within the dimer (created with BioRender.com).

in an environment of just water and counter ions.

The latter “solvent” simulations will help determine if the relative stability of the long and short M protein dimer conformations is a result of interactions with the membrane or are intrinsic to the structures themselves. While other M protein dimer simulations using experimentally determined structures have found these conformations to be stable in the membrane, the periodic box size for those simulations was either very small or the long form was not considered (Zhang et al. 2022; Dolan et al. 2022). For the long conformation, a small box allows the protein to interact with itself, possibly stabilizing it if it does need interactions from other proteins for stability in a planar membrane (as opposed to a micelle from Zhang et al. (2022)). From these simulations, the long form in solvent changes most drastically from its initial configuration. The changes in the structural conformation of the short form in solvent and membrane are within the level typical of thermal fluctuations

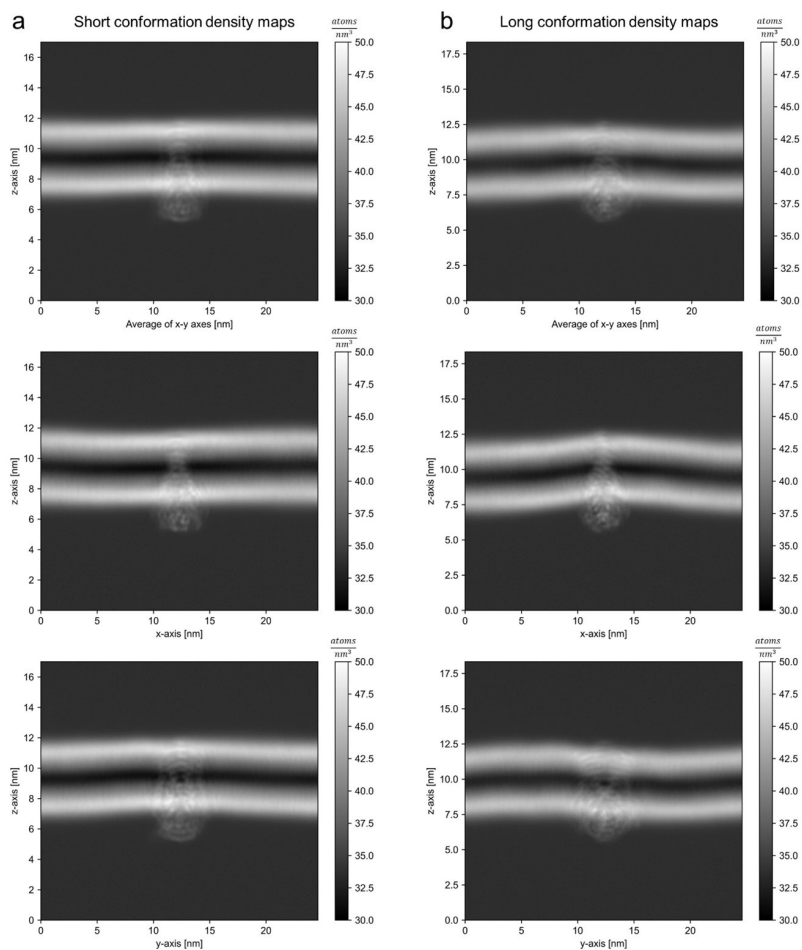
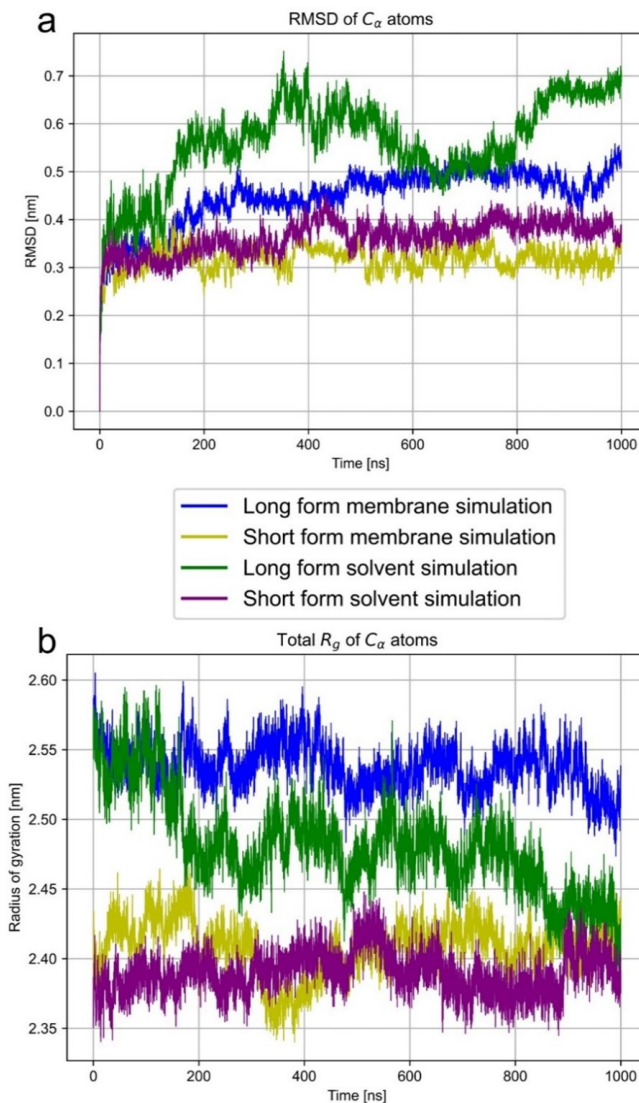


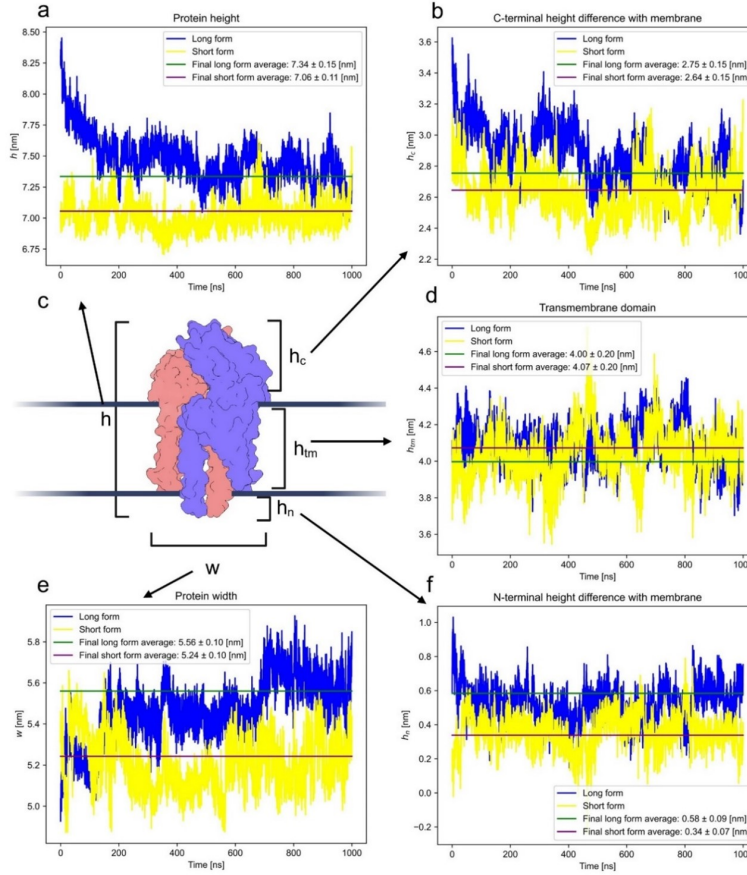
Figure 3.4: Generated using the `gmx densmap` command from GROMACS, (a) and (b) represent atom number densities for the short and long forms respectively. The whole simulation temporally and spatially is considered, where every atom other than hydrogen is accounted for, with a bin size of 0.095 nm. Each row represents a different viewing angle, with the top an average of the below two.



**Figure 3.5: RMSD and  $R_g$  of Four Different M Protein Simulations**

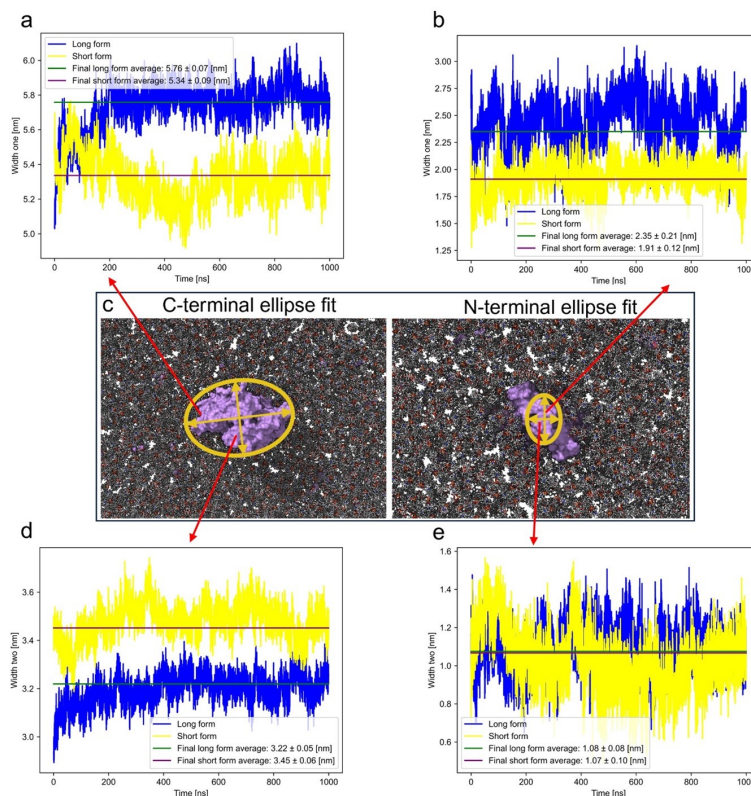
The long/short (blue/yellow) form membrane simulations and long/short (green/purple) form solvent simulations. **(A)** Root mean square distance (RMSD) represents the difference of each  $C_{\alpha}$  atom at the corresponding time with its original position squared, summed over every  $C_{\alpha}$  atom, and then divided by the number of  $C_{\alpha}$  atoms and taken to the half power. **(B)** Total radius of gyration  $R_g$ , defined as the radial distance to a point in which the moment of inertia would match the actual inertia if the body was condensed there, representing how compact the protein is. RMSD and  $R_g$  were calculated using the GROMACS commands `gmx rms` and `gmx gyrate` respectively.

( 0.35 nm) as shown in Figure 3.5 A. In contrast, the solvent-phase long form RMSD is nearly double that observed for the short form solvent-phase simulation after 1 microsecond of simulation (0.7 vs 0.35 nm). In the membrane, the long form structure has a final RMSD that is approximately 30% larger than the short form in membrane, where impact on protein characteristics can be seen in Figure 3.6 ,Figure 3.7 ,Figure 3.8 We see a similar result for the radius of gyration ( $R_g$ ), where the long form in solvent's  $R_g$  decreases over the course of the simulation to match the  $R_g$  observed for the short form in both the membrane and solvent. The  $R_g$  also decreases in the long form membrane simulation, however, it does not drop to a level comparable to the solvent simulation. These results suggest that the long form in solvent is not at a minimum energy configuration initially, unlike the short form, which undergoes little change throughout the simulation. The N-terminal region of the protein curls up without the membrane present to prevent it for the long form. On the other hand, when embedded in the membrane, the alpha helices are split apart by neighboring lipids. Of particular importance is the W31 residue mentioned in Zhang et al. (Zhang et al. 2022), which rotates away from its duplicate due to interactions with nearby DOPC. In comparison to the simulations Zhang et al. ran, these were not performed with the WYF parameter (Liu et al. 2021). However, since the nearby DOPC lipids interact with W31 throughout the simulation, including the WYF parameter would only increase this interaction and most likely would not prevent the alpha helices from splitting apart (Khan et al. 2019). This supports the suggestion that the long form of the M protein dimer



**Figure 3.6: Long and short conformation dimensions obtained throughout MD simulations.**

These calculations were performed with MDAnalysis. **(A)** Protein height, determined by finding the difference between an average of the five highest and lowest  $C_\alpha$  atoms throughout the simulation. **(B)** The difference in height between an average of the five furthest away C-terminal  $C_\alpha$  atoms and the phospholipid heads within 2 nm of the protein. **(C)** Schematic showing the different characteristic quantities (created with BioRender.com). **(D)** Transmembrane domain height, determined by subtracting **(B)** and **(F)** from **(A)**. **(E)** Width of the protein, the maximum distance between  $C_\alpha$  atoms of like residues. **(F)** The height difference between the N-terminal and the membrane, calculated similarly to **(B)**. Averages were taken over last 100 ns, where the corresponding error is standard deviation.



**Figure 3.7: Dimensions of the M Protein When Fitted to an Ellipse**

(A)-(B) and (D)-(E) were obtained by calculating the inertia tensor of every protein atom above the height of the corresponding leaflet, diagonalizing it, and then assuming a uniform protein atom density in the shape of an ellipse. (A)-(B) correspond to the long axis of the C-terminal and N-terminal respectively, while (D)-(E) is the short axis of the C-terminal and N-terminal. (c) A schematic of the principal axes of the protein, with the ellipse fit for each terminal shown for the M protein (created with BioRender.com). Averages were taken over last 100 ns, where the corresponding error is standard deviation.



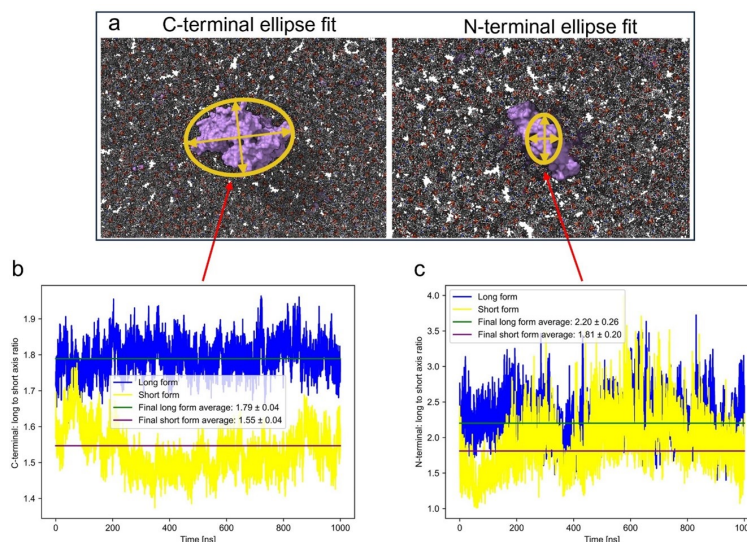


Figure 3.8: **Long to Short Axis Ratios from M Protein Ellipse Fit**

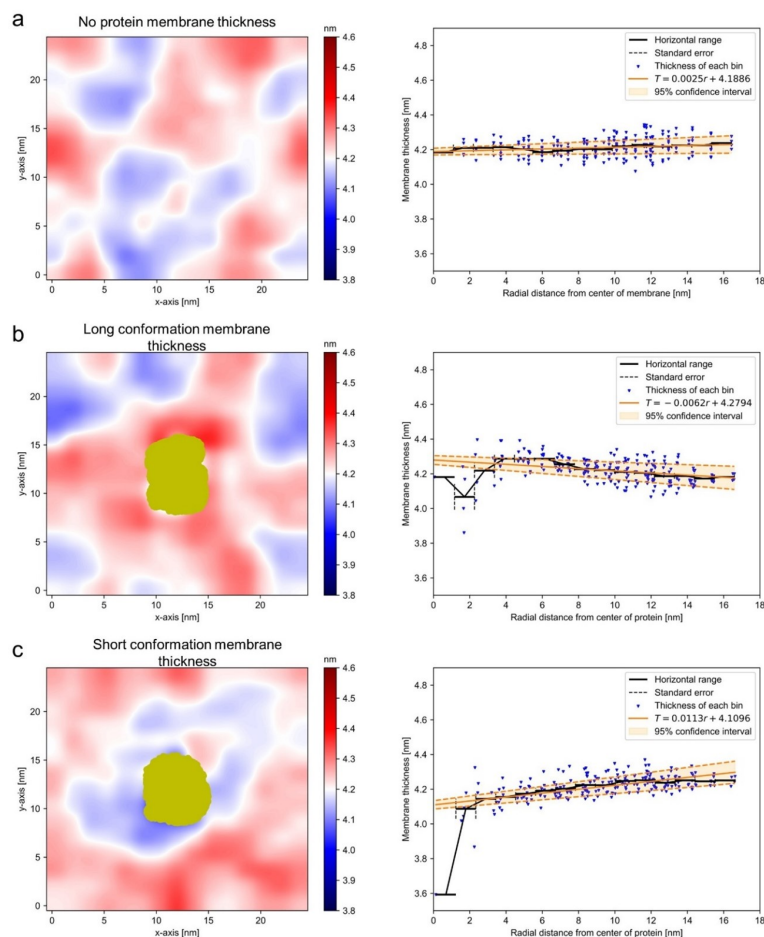
(A) Repeating the schematic shown in Fig. 3.7. (B) Ratios of the C-terminal fit, Fig. 3.7A and 3.7D, and (C) ratio of the N-terminal fit, Fig. 3.7B and 3.7E. Averages were taken over last 100 ns, where the corresponding error is standard deviation.

needs both the membrane and possibly other proteins to stabilize its structure (Neuman et al. 2011b).

### 3.4 M Protein's Impact on Membrane Thickness Through MD Simulations

According to these MD simulations, the short and long conformations alter the thickness of the membrane adjacent to the protein differently. Average values of the thickness, calculated as the difference in height between the phosphate groups in the lower and upper leaflets over the last 500 ns of the simulation, can be seen in Figure 3.9 . At each point in time every phospholipid head was placed into a bin (the membrane was split into 15 x 15 bins), where the average upper leaflet height and lower leaflet height for each bin were





**Figure 3.9: The Average Membrane Thickness is Shown for the Last 500 ns of Each Simulation**

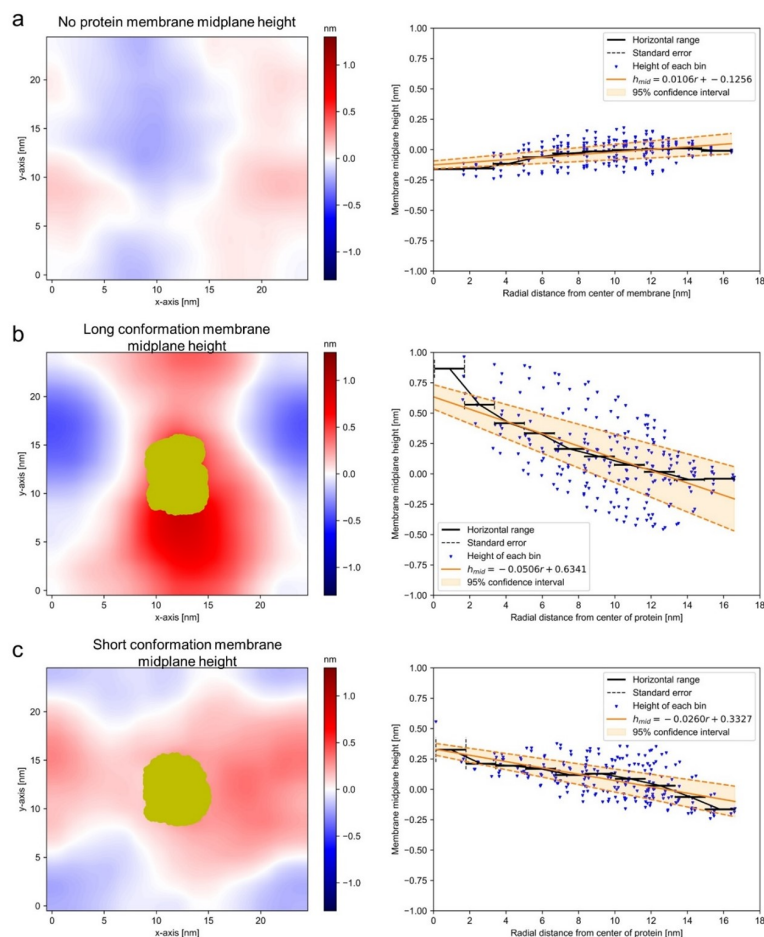
(A) – (C) Visual from above in nm, with a radial plot from the center of the box or center of the protein shown to the right. Each blue dot represents the average thickness at a given bin, orange defines the 95% confidence interval for a linear line of best fit, and black is a radial average of average thickness. Every atom in the protein at every frame averaged over is shown in gold. (C) Shows the non-shifted radial plot used in Fig. 3.7e.

subtracted from each other to determine the thickness. This process was performed using MDAnalysis. Based on the line of best fit and the corresponding 95% confidence intervals, the long and short form induce statistically significant changes in membrane thickness that depend on the radial distance from the protein. In contrast, the membrane-only simulation thickness appears to be independent of the distance from the center, hovering around  $\sim 4.20$  nm. Thus, the membrane-only system experiences thickness fluctuations, while simulations with M protein exhibit persistent thickness dependencies radially around the embedded protein dimer.

Combining these results with the radially binned averages represented with the black line of each bottom plot, the long form of the protein is surrounded by a region of thicker membrane, while the short form is surrounded by a region of thinner membrane. This agrees with Neuman et al. ([Neuman et al. 2011b](#)), and their study of M proteins during the formation of viral particles, in which they found long form in thicker regions and short in thinner. These changes are possibly induced by the different shape and/or different membrane-facing surface areas between the two conformations. Alternatively, the different M protein forms may simply stabilize different membrane thicknesses arising due to membrane fluctuations.

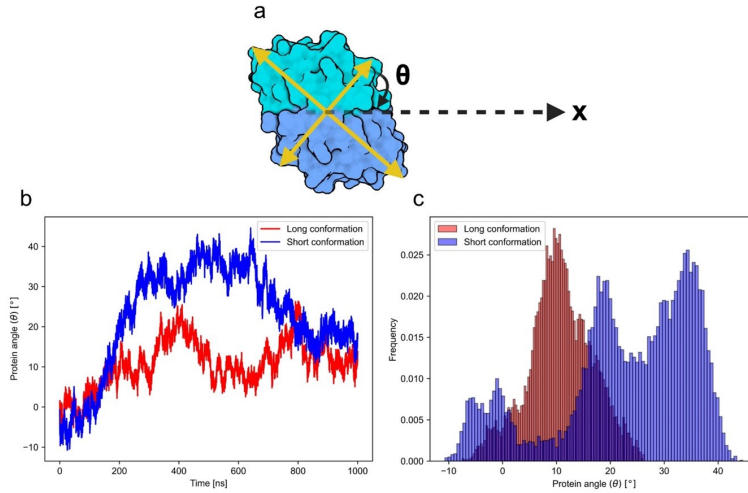
### 3.5 Protein Rotation in MD Simulations

As seen in the superposition of protein cross sections throughout the simulation in Fig [3.7B](#), Fig [3.9](#), and Figure [3.10](#) , the short M protein dimer conformation appears to have a more circular cross section than the long form. Our analysis shows this is a



**Figure 3.10: The Average Membrane Midplane Height is Shown with Respect to the Average Value Along the Boundary**

(A) – (C) Visual from above in nm, with radial plots to the right showing the average heights of each bin (blue) in comparison to either the radial distance from the center of the membrane or the radial distance from the center of the protein. Every atom in the protein at every frame averaged over is shown in gold.



**Figure 3.11: Protein Rotation for Both Forms Relative to Short Axis of the Corresponding M Protein Throughout MD Simulations**

(A) Diagram displaying rotation in the x-y plane, where  $\theta$  defines the angle between the short axis of the protein and the x-axis. The yellow lines represent projections of the long and short axis of the protein (created with BioRender.com). (B) Time evolution of the long/short (red/blue) form orientation ( $\theta$ ) with respect to the x-axis. (C) Shows the frequency of angles throughout the simulation for both short and long forms.

result of the short form's increased rotation in comparison to the long form. This rotation can be measured through the calculation of the protein's inertia tensor, and subsequent computation of the dot product between the x-axis and an inertia eigenvector. Performing this analysis on both the short and long form leads to the results shown in Figure 3.11. As seen in Fig. 3.11(B-C), the short conformation rotates over a wider angular range than the long conformation. The range of rotation is twice as large for the short form than the long form with several peaks in the frequency plot. For the case of the long form, the protein long axis is aligned with the trough of the membrane, which could be an explanation for the bend's formation: that different axes for the protein induce different levels of curvature. Furthermore, the bending aligned with respect to the long axis of the protein can imply a

possible process for the formation of higher order oligomers as other M protein dimers align in this trough. However, it is unknown if this is the preferred curvature profile of M protein or if it is impacted by the small membrane patch size and periodic boundary conditions. Imaging studies of the M protein for SARS-CoV have shown line-like formations across the virion ([Neuman et al. 2011b](#)),([Neuman & Buchmeier 2016](#)).

### 3.6 Impact of M Protein on Membrane Curvature in MD Simulations

To decouple the height of the membrane from the thickness, membrane midplane height is analyzed over the last 500 ns of the simulation. The midplane is defined as the average surface between the lower and upper leaflet shown in Fig 3.3C. As seen in Fig 3.10, while the membrane-only simulation has slight height variation, it is clearly not as substantial as either the short or long form runs. Thus, it can be assumed that these represent membrane fluctuations that persist on the scale of 500 ns.

From the line of best fit, with the given 95% confidence intervals for each simulation, the long conformation of the M protein dimer seems to have the greatest effect on membrane deformation. Additionally, since the confidence interval does not include the horizontal (slope = 0), this radial dependence of average membrane midplane height appears significant.

The shape of the membrane curvature can be determined from the radial average. While the deformation of the membrane-only simulation matches that of height fluctuations, both membranes containing either the long or short M protein conformations seem to adopt

persistent curvature, in the form of a valley-like fold in the membrane enclosing the protein dimer. This shows that the M protein induces curvature in the membrane. Due to the small size of the lipid membrane patch simulated and the periodic boundary conditions imposed, this is only a qualitative confirmation of the membrane curvature induced by the protein. As discussed in (Neuman et al. 2011b) for SARS-CoV, the long conformation appears to induce more curvature than the short conformation, agreeing with our observed membrane deformation.

### 3.7 Coarse-grained Molecular Dynamics Simulation of M Protein Effects on Membrane Curvature

The role of M proteins in the induction of membrane curvature and budding in ERGIC was previously explored (Li & Zandi 2022). ERGIC was modeled as a sphere built from a triangular lattice as shown in Figure 3.12, which shows how the M-M interaction can result in the budding of the region of the ERGIC membrane where the M proteins are embedded. The ERGIC is shown as a gray vesicle in Figure 3.12 and the three regions of M protein are depicted with three hard spheres (M1 in red, M2 in orange, and M3 in black), where M1 and M2 correspond to the transmembrane domain, and M3 denotes the endodomain. Fig. 3.12A shows that in the absence of endodomain interaction, the membrane does not curve. However, the attractive interaction between endodomains results in the membrane curvature and budding as illustrated in Fig. 3.12B. However, in the absence of interaction between endodomains, the attractive interaction between N proteins

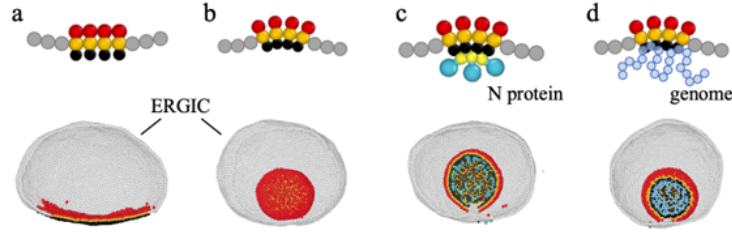


Figure 3.12: **Coarse-grained Simulations of M Proteins Embedded in the ER-Golgi Intermediate Compartment (ERGIC)**

(A) M proteins embedded in the bottom of ERGIC have no endodomain interactions, and no budding is observed in this scenario. (B) M proteins have endodomain interactions, proteins bud into ERGIC and form a spherical shell in 4000s. (C) M proteins have no endodomain interactions, a budding process happens when N proteins are added. Note that N proteins interact with M proteins endodomain and effectively introduce endodomain interactions. (D) M proteins have no endodomain interactions, a budding happens when a short RNA genome (Length = 400a) is added. The M proteins interact with the genome through screened electrostatic interactions, which effectively cause endodomain interactions.

(blue and yellow beads) and M proteins, see Fig. 3.12C, or between RNA and M proteins (Fig. 3.12D) can induce curvature and budding. Since most of the experiments and all-atom simulations in this paper are performed on the flat membrane, for the purpose of this paper, we simulated the role of M proteins embedded in a flat membrane, see the main text. Previously, it was found that the curvature of ERGIC helps with overcoming the energy barriers and leads to budding. In this paper, we found that M proteins can curve even the flat membranes and induce budding if the endodomains interact with each other.

### 3.8 Estimation of Line Tension from Membrane Thinning

For a given height mismatch,  $\delta$ , between a lipid ordered domain of monolayer thickness  $h_r$  and the surrounding lipid disordered membrane of thickness  $h_s$ , the line tension

generated due to the elastic cost of bend and tilt deformation is given by (Kuzmin et al. 2005)

$$\gamma = \left(\frac{\delta}{h}\right)^2 \frac{\sqrt{B_s K_s} \sqrt{B_r K_r}}{\sqrt{B_s K_s} + \sqrt{B_r K_r}}, \quad (3.4)$$

where  $B_s, r$  and  $K_s, r$  are the bending and tilt modulus of the soft and rigid phases respectively, and  $h = (h_r + h_s)/2$ . Since the stiffness of the M protein region (41.5 MPa) is much higher than that of the lipid layer (9.5 MPa), we can take  $B_r \gg B_s$  and  $K_r \gg K_s$ , simplifying our expression to

$$\gamma = \left(\frac{\delta}{h}\right)^2 \sqrt{B_s K_s}. \quad (3.5)$$

We can estimate the bending modulus  $B_s$  from the measured Young's modulus ( $E$ ) via (Deserno 2015)

$$B_s = \frac{E h_s^3}{12(1 - \sigma^2)}. \quad (3.6)$$

With  $\sigma = 0.5$  (incompressible) the bending modulus is approximately  $3 k_B T$  and we can take  $K_s = 3 k_B T / \text{nm}^2$ , which is consistent with lower values reported from simulations and experiments (Doktorova et al. 2017). We can use the line tension expression for a monolayer assuming the height mismatch is evenly distributed among the two leaflets. Using these moduli along with  $\delta = 0.25$  nm and  $h = 2.125$  nm, the simplified expression for the line tension gives  $\gamma = 0.35$  pN (accounting for both leaflets). If we assume the deformation is entirely confined to one leaflet ( $\delta = 0.5$  nm), the line tension increases to  $\gamma = 0.66$  pN.



## Chapter 4

# Optimized Assembly of SARS-CoV-2 Virus-Like Particles and Characterization of Vesicle-Associated RNA

**Author Contributions:** This chapter includes a manuscript currently in preparation for publication. Sara Anbir led experimental design, conducted the experiments, and performed data analysis and interpretation. Sandeep Sharma performed dynamic light scattering (DLS) experiments. Yuanzhong Zhang assisted with atomic force microscopy (AFM) experiments. Hina Singh performed RNA extraction. Marcus Kaul and Thomas E. Kuhlman provided research supervision.

**Abstract** Virus-like particles (VLPs) provide a safe and tractable platform to study viral assembly and vesicle-associated processes under biosafety level-2 conditions. Previous studies have shown that coordinated expression of the SARS-CoV-2 structural proteins membrane (M), nucleocapsid (N), envelope (E), and spike (S), particularly when expressed at ratios reflecting viral mRNA abundance in infected cells, yields particles that closely resemble native virions in size and composition ([Swann et al. 2020](#); [Gourdelier et al. 2022](#)). In this study, mammalian cells were transfected with all four SARS-CoV-2 structural proteins using these optimized expression ratios. VLP formation was validated by Western blotting, atomic force microscopy (AFM), and dynamic light scattering (DLS), confirming the production of particles with physical characteristics consistent with authentic SARS-CoV-2 virions ([Lyonnais et al. 2021](#)).

RNA extracted from purified VLP- and exosome-associated fractions exhibited a size distribution dominated by small RNAs (20–40 nucleotides) and lacked detectable ribosomal RNA. This RNA profile is consistent with selective vesicular RNA packaging rather than cellular RNA contamination or degradation, in line with known properties of coronavirus assembly and extracellular vesicle biology ([Masters 2006](#); [Kim et al. 2021](#)). These findings establish a validated experimental platform for subsequent high-throughput RNA sequencing aimed at identifying and characterizing RNA species associated with SARS-CoV-2 VLPs, with the goal of elucidating selective RNA packaging mechanisms and their potential relevance to virus–host interactions and VLP-based applications.

## 4.1 Introduction

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) is an enveloped, positive-sense single-stranded RNA virus belonging to the *Betacoronavirus* genus and the causative agent of coronavirus disease 2019 (COVID-19). The SARS-CoV-2 virion is organized around four major structural proteins: membrane (M), nucleocapsid (N), envelope (E), and spike (S), which collectively coordinate virion assembly, genome packaging, and host cell entry ([Masters 2006](#); [Hardenbrook & Zhang 2022](#)). Assembly occurs primarily at the endoplasmic reticulum–Golgi intermediate compartment (ERGIC), where M serves as the central scaffold interacting with all other structural components, E contributes to membrane curvature and particle release, and N condenses the viral RNA genome into a ribonucleoprotein (RNP) complex that is encapsidated during budding ([Masters 2006](#); [Bracquemond & Muriaux 2021](#)). Although spike is dispensable for particle assembly, its incorporation is essential for infectivity.

Prior work from our group characterized the membrane-remodeling activities of M and E at the molecular level using in vitro reconstitution. Purified M protein dimers inserted into large unilamellar vesicles (LUVs) were shown by AFM, cryo-EM, and molecular dynamics simulations to adopt exclusively the compact conformation, induce localized membrane thinning, and drive assembly and budding through line tension and endodomain interactions ([Zhang et al. 2024](#)). Subsequent reconstitution of E protein into ERGIC-mimetic bilayers resolved its full oligomeric spectrum and demonstrated that co-reconstitution of M, E, and N–RNA ribonucleoprotein complexes is sufficient to trigger inward vesicle budding in a minimal system. While these in vitro approaches reveal the physical principles

underlying M- and E-driven membrane remodeling, they rely on defined artificial lipid environments and lack the full complement of host cell factors present during natural infection. The present study takes a complementary approach by harnessing the cellular machinery of mammalian cells to express all four structural proteins simultaneously at physiologically relevant ratios, producing VLPs that more faithfully recapitulate authentic SARS-CoV-2 virion composition and morphology.

The stoichiometry of structural protein expression is a critical determinant of VLP assembly efficiency and fidelity. Previous studies have reported conflicting conclusions regarding which combination of structural proteins is sufficient for particle formation, and whether M alone, M and E together, or the inclusion of N is required ([Swann et al. 2020](#); [Plescia et al. 2021](#)). Importantly, recent work demonstrated that expressing M, N, E, and S at ratios reflecting viral mRNA abundance in productively infected cells substantially improves VLP yield, nucleocapsid incorporation, and structural fidelity compared to equal expression ([Gourdelier et al. 2022](#); [Kim et al. 2021](#)). Because N protein is the primary determinant of RNA packaging specificity and its expression level relative to M and E directly governs encapsidation efficiency ([Lu et al. 2021](#); [Rodén et al. 2022](#)), systematic evaluation of plasmid ratio conditions is essential for producing biologically representative VLPs.

A key challenge in characterizing VLP-associated RNA is the co-release of extracellular vesicles (EVs) from transfected producer cells, which overlap with VLPs in size and density and co-purify under standard ultracentrifugation conditions ([Kim et al. 2021](#)). To ensure that detected RNA reflects genuine VLP-associated cargo rather than surface-

adsorbed or EV-derived material, VLP preparations in this study were treated with RNase H prior to RNA extraction, selectively removing RNA present outside intact particles. The resulting RNA was characterized by microfluidic capillary electrophoresis (Agilent Bioanalyzer) to establish size distribution and integrity.

In this chapter, SARS-CoV-2 VLPs are produced in HEK293T cells under multiple M:N:E:S plasmid ratio conditions and validated by Western blotting, dynamic light scattering (DLS), and atomic force microscopy (AFM). RNA isolated from purified VLP fractions is then analyzed to characterize the size distribution of encapsulated nucleic acid species. Together, these experiments establish an optimized and validated VLP platform and lay the groundwork for future sequencing-based studies of selective RNA packaging during coronavirus assembly.

## **4.2 Materials and Methods**

### **4.2.1 Cells and culture conditions**

The HEK293T human embryonic kidney cell line was obtained from ECACC (Sigma-Aldrich, Germany) and maintained in Dulbecco's Modified Eagle Medium (DMEM; Gibco) supplemented with 10% heat-inactivated fetal bovine serum (FBS; Thermo Fisher Scientific, USA), 50 U/mL penicillin and 50  $\mu$ g/mL streptomycin (Penicillin–Streptomycin; Gibco), 1 mM sodium pyruvate (Gibco), and 25 mM HEPES buffer (Gibco). Cells were cultured at 37 °C in a humidified incubator with 5% CO<sub>2</sub>.

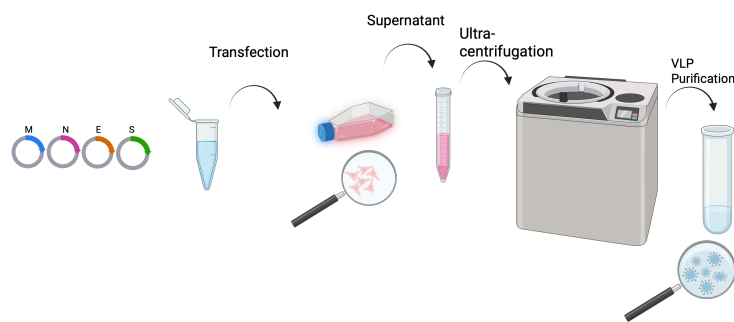


Figure 4.1: Schematic overview of the SARS-CoV-2 virus-like particle (VLP) production workflow. HEK293T cells were transiently transfected with plasmids encoding the four structural proteins (M, N, E, and S) at optimized expression ratios. VLPs were harvested from the culture supernatant, purified by ultracentrifugation through a sucrose cushion, and validated by Western blotting, dynamic light scattering (DLS), and atomic force microscopy (AFM).

#### 4.2.2 DNA plasmid and cloning

Plasmids encoding the SARS-CoV-2 structural proteins membrane (M), envelope (E), nucleocapsid (N), and spike (S) were obtained from Addgene and used for virus-like particle (VLP) production. The pcDNA3.1 SARS-CoV-2 M plasmid (Addgene plasmid #158078; RRID:Addgene\_158078) and pcDNA3.1 SARS-CoV-2 E plasmid (Addgene plasmid #158080; RRID:Addgene\_158080) were gifts from Jeremy Luban. The pcDNA3.1 SARS-CoV-2 N plasmid (Addgene plasmid #158079; RRID:Addgene\_158079) was also obtained from Jeremy Luban([Yurkovetskiy et al. 2020](#)). The spike expression plasmid pcDNA3.1-SARS2-Spike (Addgene plasmid #145032; RRID:Addgene\_145032) was a gift from Fang Li. These plasmids were co-transfected at optimized ratios to generate SARS-CoV-2 VLPs([Shang et al. 2020b](#)).

In addition to individual structural protein expression plasmids, a plasmid encoding membrane, envelope, and spike proteins (pCMV-Spike-VLP; Addgene plasmid #173922; RRID:Addgene\_173922), provided by Andreas Stuermer, was used to directly drive VLP formation in the absence of nucleocapsid protein. This construct enabled assessment of VLP size, morphology, and overall particle quality when N was excluded, allowing direct comparison with VLPs generated using all four structural proteins.

To monitor structural protein expression dynamics and determine the optimal time point for maximal VLP production following transfection, fluorescently tagged variants of M and N were employed. The membrane protein was expressed as a C-terminal fusion with the fluorescent protein mCerulean, and the nucleocapsid protein was expressed as a fusion with the fluorescent protein mScarlet. These tagged constructs enabled real-time

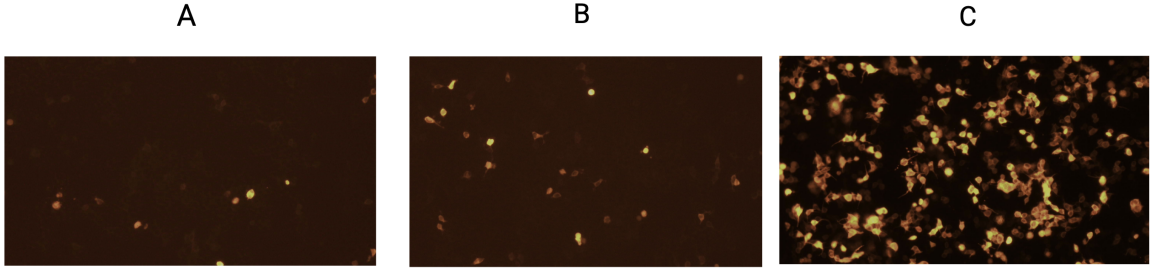


Figure 4.2: **Expression level of Nucleocapsid protein**

(A) Expression level of the Nucleocapsid protein in HEK293T cells transfected with a plasmid encoding Nucleocapsid fused to the mScarlet fluorescent tag, measured 12 hours post-transfection. (B) Expression level at 24 hours post-transfection. (C) Expression level at 48 hours post-transfection.

assessment of expression levels post-transfection and were used to identify the time window corresponding to peak protein expression and maximal VLP release as shown in Figure

. Unless otherwise specified, tagged constructs were used for expression optimization and imaging-based analyses rather than bulk VLP production.

### 4.2.3 DNA transfection

HEK293T cells were seeded into 6-well plates or 10 mL culture dishes 24 h prior to transfection under standard growth conditions. Transient transfection was performed using Lipofectamine 3000 (Thermo Fisher Scientific) according to the manufacturer's protocol. A total of 1.4  $\mu$ g of plasmid DNA was used per well of a 6-well plate.

Based on previously reported optimization of SARS-CoV-2 VLP production, multiple plasmid ratio conditions were evaluated([Gourdelier et al. 2022](#)). Co-transfections were performed using membrane (M), nucleocapsid (N), envelope (E)([Yurkovetskiy et al. 2020](#)),



and spike (S)([Shang et al. 2020b](#)) plasmids at ratios previously shown to enhance VLP yield. The conditions tested included M:N:E at 3:3:3, M:N:E at 3:12:2, and the optimized four-plasmid condition M:N:E:S at 3:12:2:5, which corresponds to the relative viral RNA abundance reported in infected cells and has been associated with maximal VLP production in prior studies([Kim et al. 2021](#); [Swain et al. 2022](#)). For each condition, the total DNA amount per well was maintained at 1.4  $\mu$ g, with plasmid quantities adjusted proportionally to preserve the specified ratios.

Following transfection, cells were incubated for 6 h, after which the medium was replaced with fresh culture medium. Supernatants containing released VLPs were collected between 24 h to 48 h post-transfection.

For cellular protein analysis, adherent cells were washed with phosphate-buffered saline (PBS 1 $\times$ ), harvested, and centrifuged at 1500 rpm for 5 min at 4 °C. Cell pellets were lysed in RIPA buffer (25 mM Tris-HCl pH 7.6, 150 mM NaCl, 1% NP-40, 1% sodium deoxycholate, 0.1% SDS; Thermo Fisher Scientific) supplemented with protease inhibitor cocktail (Sigma). Lysates were clarified by centrifugation at 15,000 rpm for 10 min at 4 °C, and the supernatant was collected for downstream analysis. Protein concentrations were determined using the BCA assay.

Under each experimental condition, cell viability was assessed prior to lysis using trypan blue exclusion to ensure that observed differences in VLP production were not attributable to cytotoxic effects of transfection.

#### 4.2.4 Collected VLPs and purification

Virus-like particles (VLPs) were harvested from the culture supernatant of transfected cells. Cell debris was first removed by centrifugation at 5000 rpm for 5 min at 4 °C. The clarified supernatant was then subjected to ultracentrifugation to pellet VLPs through a 25% (w/v) sucrose cushion prepared in TNE buffer (10 mM Tris-HCl, pH 7.4, 100 mM NaCl, 1 mM EDTA). Ultracentrifugation was performed at either 28,000 rpm or 30,000 rpm for 3 h at 4 °C using Beckman SW41Ti or SW55Ti rotors, respectively, depending on sample volume.

Following centrifugation, VLP-containing pellets were gently resuspended in TNE buffer and allowed to equilibrate overnight at 4 °C. Purified VLP preparations were subsequently stored at 4 °C until further analysis.

#### 4.2.5 DLS Analysis

**Dynamic Light Scattering (DLS).** Hydrodynamic size measurements of VLPs before and after RNase H treatment were performed using a Malvern Zetasizer Nano ZSP at 25 °C. Samples were prepared in TNE buffer at a final concentration of 0.1 mg mL<sup>-1</sup>, and measurements were analyzed using ZS Xplorer software (version 3.00). Prior to RNA extraction, VLP samples were treated with RNase H to remove RNA present outside the particles so that the extracted material would represent RNA incorporated within the VLPs. DLS measurements were then repeated to verify that the treatment did not significantly alter VLP size. In both conditions, VLPs exhibited a hydrodynamic diameter in the range

of 100–146 nm, consistent with the dimensions of native SARS-CoV-2 virions, confirming that RNase H treatment preserves particle integrity (Figure 4.3, Figure 4.4).

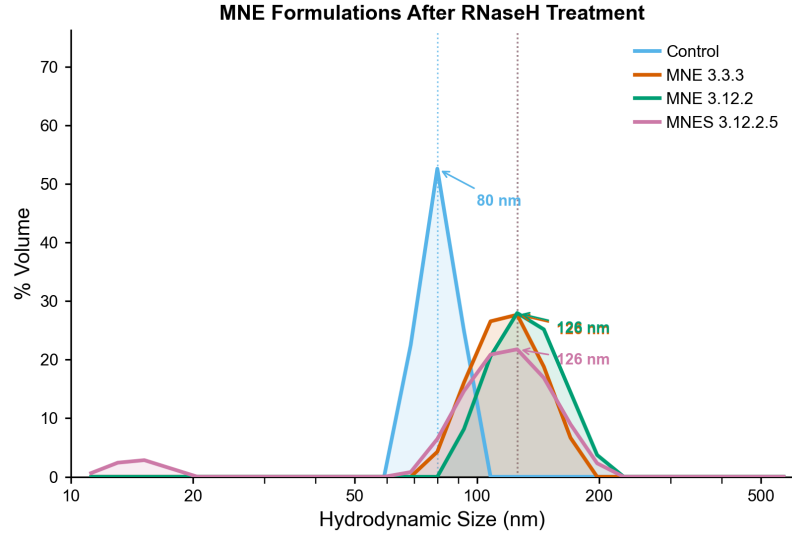


Figure 4.3: Dynamic light scattering size distribution of SARS-CoV-2 VLPs (M:N:E expression condition) after RNase H treatment. Hydrodynamic diameter measurements were performed using a Malvern Zetasizer Nano ZSP in TNE buffer at 25 °C. VLPs exhibit a hydrodynamic diameter in the range of 100–146 nm, consistent with native SARS-CoV-2 virion dimensions. RNase H treatment was applied prior to measurement to remove surface-associated RNA, ensuring that the size distribution reflects intact VLP particles.

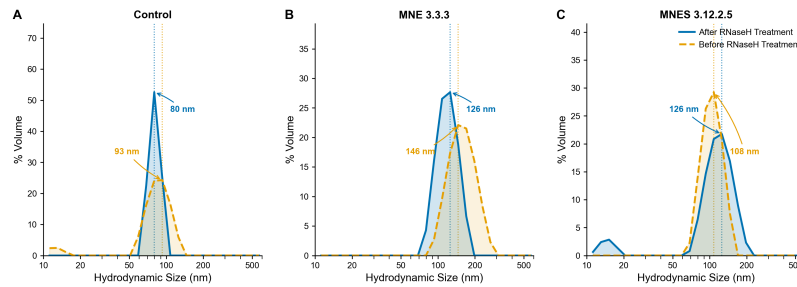


Figure 4.4: Comparison of VLP hydrodynamic size distributions before and after RNase H treatment. In both conditions, VLPs maintain a hydrodynamic diameter in the range of 100–146 nm, consistent with native SARS-CoV-2 virion dimensions. DLS measurements demonstrate that RNase H treatment does not significantly alter VLP size or integrity, confirming that changes in RNA content following treatment reflect removal of externally associated RNA rather than disruption of the particles themselves.

#### 4.2.6 Atomic Force Microscopy

**Atomic Force Microscopy (AFM).** Atomic force microscopy was used to directly visualize purified SARS-CoV-2 VLPs and assess their morphology and size at the single-particle level (Baró & Reifemberger 2012). Freshly cleaved muscovite mica substrates were functionalized with (3-aminopropyl)triethoxysilane (APTES; Sigma-Aldrich) to render the surface positively charged and promote electrostatic adsorption of negatively charged VLP membranes. Briefly, mica surfaces were incubated with 0.1% (v/v) APTES in ethanol for 10 min at room temperature, rinsed thoroughly with ethanol and then ultrapure water, and dried under a gentle stream of nitrogen gas.

Purified VLP samples were diluted to a concentration of approximately  $10\text{--}50\ \mu\text{g mL}^{-1}$  in TNE buffer and deposited onto APTES-functionalized mica by incubating  $20\ \mu\text{L}$  of sample on the surface for 20 min at room temperature. Unbound material was removed by gentle washing with ultrapure water, and the surface was allowed to dry under ambient conditions prior to imaging. AFM imaging was performed in tapping mode (intermittent contact mode) in air using a NanoScope Multimode atomic force microscope (Bruker) equipped with silicon cantilevers (nominal spring constant  $20\text{--}80\ \text{N m}^{-1}$ ; resonance frequency approximately 300 kHz). Scan areas of  $2 \times 2\ \mu\text{m}$  to  $5 \times 5\ \mu\text{m}$  were acquired at a scan rate of 0.5–1 Hz with 512 lines per image.

Height and diameter measurements of individual VLP particles were extracted from topographic images using NanoScope Analysis software (Bruker). A minimum of 50 individual particles were analyzed per sample to obtain representative size distributions. Particle heights were used as the primary size metric, as lateral dimensions are subject to

tip convolution artifacts in AFM imaging. The measured particle heights were compared with DLS-derived hydrodynamic diameters to provide a complementary assessment of VLP dimensions.

AFM topographic imaging of VLPs deposited on APTES-functionalized mica revealed abundant, well-isolated, dome-shaped particles distributed across the substrate surface (Figure 4.5). Individual particles appeared as smooth, rounded protrusions with a morphology consistent with intact, membrane-enveloped vesicles. To quantify particle dimensions, Gaussian fits were applied to horizontal and vertical cross-sectional line profiles drawn through the center of individual particles. The full-width at half-maximum (FWHM) of these profiles yielded values of approximately 88 nm in the horizontal direction and 73 nm in the vertical direction (Figure 4.5), corresponding to an average lateral diameter of approximately 80–100 nm. A representative single-particle line profile confirmed these dimensions, with a Gaussian fit returning an FWHM of 83 nm and a peak height of approximately 17.5 nm above the local baseline (Figure 4.6). These lateral dimensions are in good agreement with the hydrodynamic diameters measured by DLS and with the reported size range of authentic SARS-CoV-2 virions (Lyonnais et al. 2021).

The measured particle heights (approximately 10–18 nm) were substantially smaller than the lateral FWHM values, as is characteristic of enveloped virus particles imaged by AFM in air, where partial dehydration and substrate-induced flattening reduce the apparent height well below the solution-phase hydrodynamic diameter (Lyonnais et al. 2021; Baró & Reifemberger 2012). Three-dimensional surface rendering of the topographic data further confirmed the dome-shaped morphology of individual VLPs, with a representative particle

displaying a lateral diameter of approximately 100 nm and a height of approximately 8.7 nm (Figure 4.7), consistent with a partially flattened spherical vesicle adsorbed onto the mica surface.

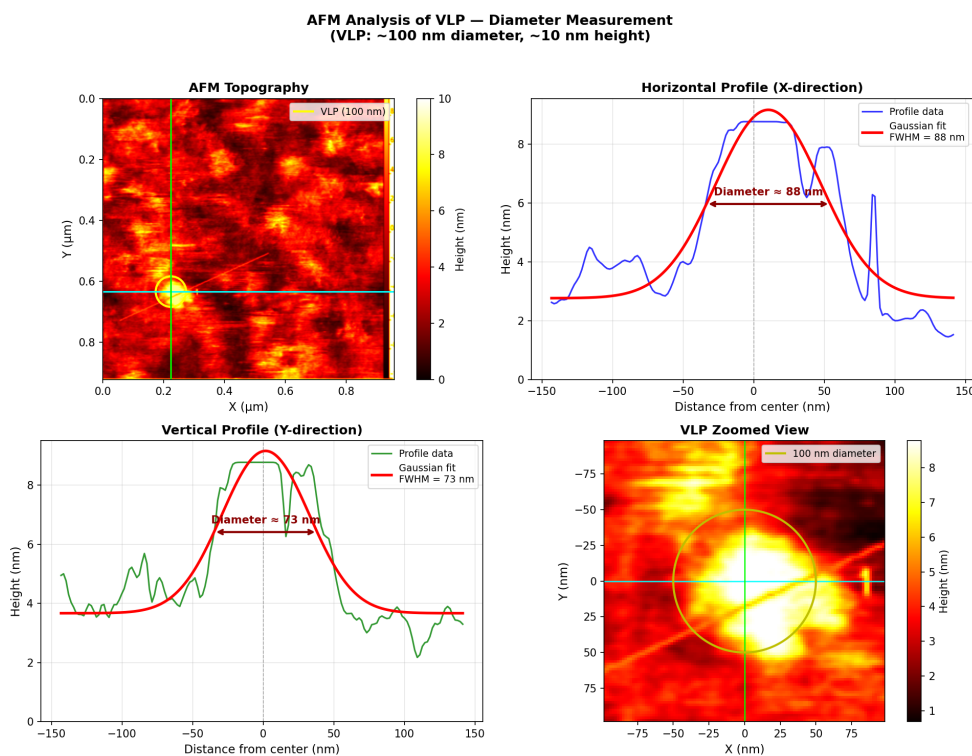


Figure 4.5: AFM diameter analysis of a SARS-CoV-2 VLP produced under the MNE (3:12:2) plasmid ratio condition. **(Top left)** AFM topographic image showing a representative VLP (~100 nm lateral diameter, ~10 nm height) adsorbed on APTES-functionalized mica. **(Top right)** Horizontal line profile (X-direction) across the particle center with Gaussian fit; FWHM = 88 nm. **(Bottom left)** Vertical line profile (Y-direction) with Gaussian fit; FWHM = 73 nm. **(Bottom right)** Zoomed topographic view of the same particle with a 100 nm diameter scale indicator.

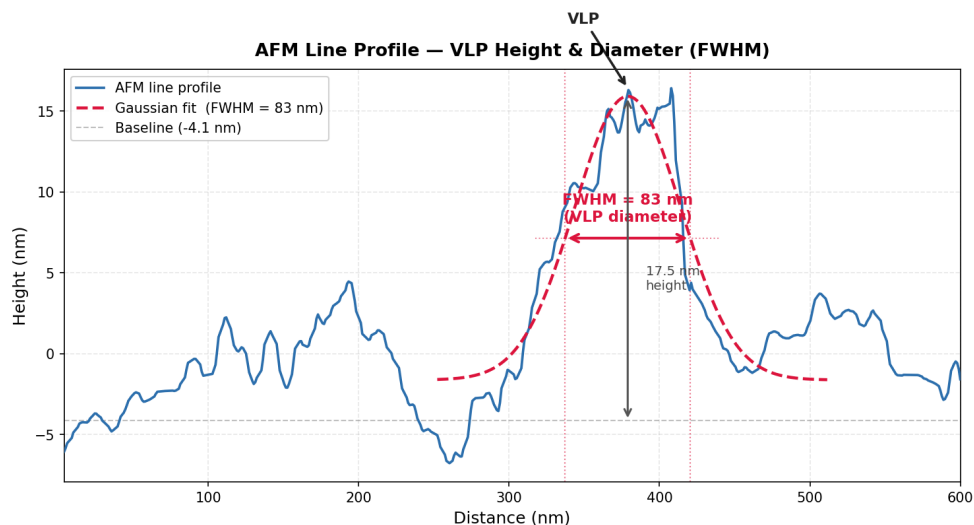


Figure 4.6: AFM line profile of a representative VLP from the MNE (3:12:2) condition. The blue trace shows the raw height profile; the red dashed curve is a Gaussian fit (FWHM = 83 nm). The particle height above the local baseline (−4.1 nm) is approximately 17.5 nm, reflecting partial flattening upon adsorption to the mica substrate.

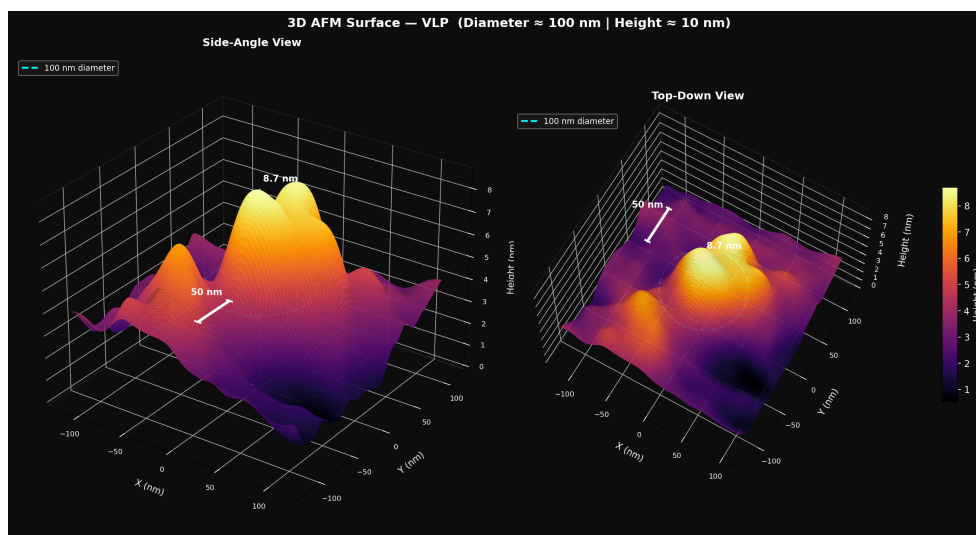


Figure 4.7: Three-dimensional AFM surface rendering of a SARS-CoV-2 VLP from the MNE (3:12:2) condition. **(Left)** Side-angle view showing the dome-shaped morphology; the 50 nm scale bar indicates the particle radius ( $\sim 100$  nm diameter), with a maximum height of approximately 8.7 nm. **(Right)** Top-down view confirming the roughly circular footprint of the adsorbed particle.

#### 4.2.7 Western Blot

**Western Blotting.** To confirm the expression and incorporation of SARS-CoV-2 structural proteins into purified VLP fractions, Western blot analysis was performed. Cell lysate samples containing 20  $\mu$ g total protein or equivalent volumes of VLP preparations were resolved by SDS-PAGE using 4–20% Mini-PROTEAN TGX gradient gels (Bio-Rad). Following electrophoresis, proteins were transferred to nitrocellulose membranes (0.45  $\mu$ m pore size; Thermo Fisher Scientific). Membranes were blocked with 5% bovine serum albumin (BSA) in Tris-buffered saline with 0.1% Tween-20 (TBST) for 1 h at room temperature and subsequently incubated overnight at 4 °C with primary antibodies. Primary antibodies used were: anti-N protein rabbit polyclonal antibody (Rockland, lot 46527), anti-M protein rabbit polyclonal antibody (Fabgenix, lot AC4647827), mouse anti-S protein (GTX632604, Genetex; 1:1000), and anti-GAPDH antibody conjugated to horseradish peroxidase (HRP; G9295, Sigma-Aldrich; 1:25,000). Following three washes with TBST, membranes were incubated with secondary antibodies for 1 h at room temperature: goat anti-rabbit IgG conjugated to Alexa Fluor 680 (Invitrogen; 1:2000) and goat anti-mouse IgG conjugated to Alexa Fluor 647 (Invitrogen; 1:5000). Fluorescent detection was performed using a Chemi-Doc imaging system (Bio-Rad). Band intensities were quantified using ImageJ software ([Schneider et al. 2012](#)).



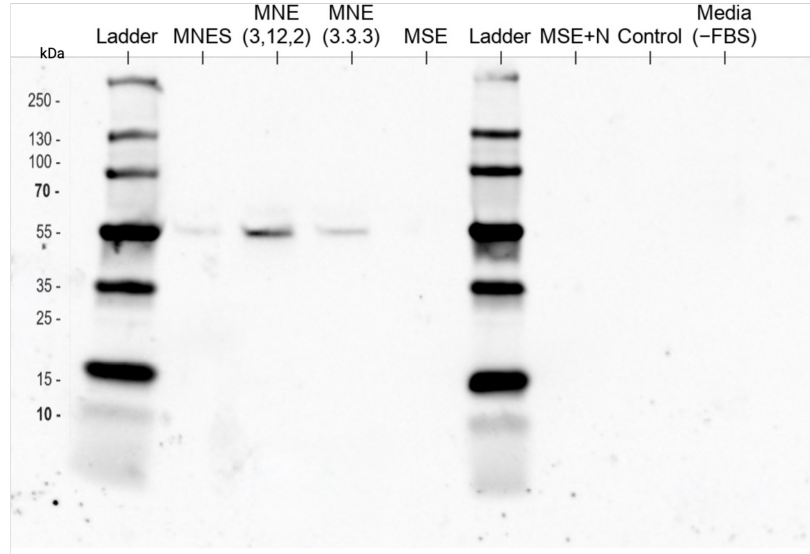


Figure 4.8: **Western blot analysis of nucleocapsid (N) protein incorporation into purified VLPs under different plasmid transfection conditions.** Purified VLP fractions and controls were resolved by SDS-PAGE and immunoblotted with an anti-N antibody. **MNES (3:12:2:5)**: VLPs produced by co-transfection of M, N, E, and S expression plasmids at a 3:12:2:5 molar ratio. **MNE (3:12:2)**: VLPs produced with M, N, and E plasmids at a 3:12:2 molar ratio, showing the strongest N incorporation (~50 kDa). **MNE (3:3:3)**: VLPs produced with M, N, and E plasmids at an equimolar ratio, showing reduced N incorporation. **MSE**: VLPs produced without the N plasmid; absence of the N band confirms specificity of incorporation. **MSE+N**: VLPs produced by co-transfection of a single polycistronic MSE plasmid with a separate N expression plasmid; no detectable N band in the purified fraction, indicating that this plasmid ratio is incompatible with efficient N packaging into VLPs. **Control**: concentrated conditioned media harvested from untransfected cells processed identically to VLP samples, confirming the absence of non-specific background. **Media (-FBS)**: serum-free conditioned media control. Molecular weight markers (kDa) are indicated. All blots are representative of  $n = 3$  independent experiments.

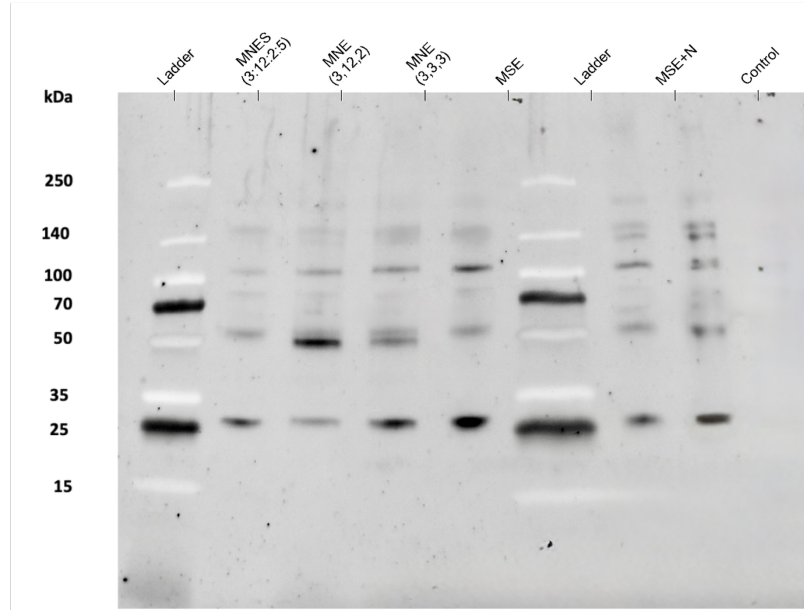


Figure 4.9: **Western blot analysis of membrane (M) protein incorporation into purified VLPs under different plasmid transfection conditions.** Purified VLP fractions and controls were resolved by SDS-PAGE and immunoblotted with an anti-M antibody. **MNES (3:12:2:5)**: VLPs produced by co-transfection of M, N, E, and S expression plasmids at a 3:12:2:5 molar ratio. **MNE (3:12:2)**: VLPs produced with M, N, and E plasmids at a 3:12:2 molar ratio, showing a prominent M monomer band at ~25 kDa. **MNE (3:3:3)**: VLPs produced with M, N, and E plasmids at an equimolar ratio. **MSE**: VLPs produced with M, S, and E plasmids. **MSE+N** (two lanes): VLPs produced by co-transfection of a single polycistronic MSE plasmid with a separate N expression plasmid, shown as biological replicates. **Control**: concentrated conditioned media harvested from untransfected cells processed identically to VLP samples. M protein is detected as a monomer at ~25 kDa, with additional bands corresponding to SDS-stable dimers (~50 kDa) and higher-order oligomers (~100 kDa), consistent with the known oligomeric behavior of the M protein. Molecular weight markers (kDa) are indicated. All blots are representative of  $n = 3$  independent experiments.

#### 4.2.8 RNA Incorporation in VLP

To isolate RNA species specifically incorporated within VLPs rather than non-specifically adsorbed to particle surfaces, purified VLP preparations were subjected to RNase H treatment prior to RNA extraction. Purified VLPs resuspended in TNE buffer were incubated with RNase H (New England Biolabs) at a final concentration of 5 U per 50  $\mu$ L reaction volume for 30 min at 37 °C. This treatment degrades RNA:DNA hybrids and removes RNA molecules present on the external surface of vesicles, ensuring that subsequently extracted RNA represents material encapsulated within the VLP lumen ([Kim et al. 2021](#)). Following RNase H treatment, VLPs were pelleted by ultracentrifugation at 30,000 rpm for 1 h at 4 °C to remove digestion products and free nucleotides, and resuspended in nuclease-free TNE buffer.

RNA was extracted from RNase H-treated VLP pellets and from matched exosome-enriched fractions using the miRNeasy Mini Kit (Qiagen) according to the manufacturer’s protocol with modifications for low-input samples. Briefly, pellets were lysed directly in QIAzol lysis reagent, vortexed vigorously for 1 min, and incubated at room temperature for 5 min to ensure complete homogenization. Chloroform was added at a 0.2:1 (v/v) ratio, and phase separation was achieved by centrifugation at 12,000  $\times$ g for 15 min at 4 °C. The aqueous phase was transferred to a fresh tube, and 1.5 volumes of 100% ethanol were added to facilitate binding to the RNeasy MinElute spin column. Columns were washed with Buffer RWT and RPE according to the standard protocol, and RNA was eluted in 14  $\mu$ L of nuclease-free water. RNA concentration and purity were assessed using a NanoDrop 2000 spectrophotometer (Thermo Fisher Scientific), with  $A_{260}/A_{280}$  and  $A_{260}/A_{230}$

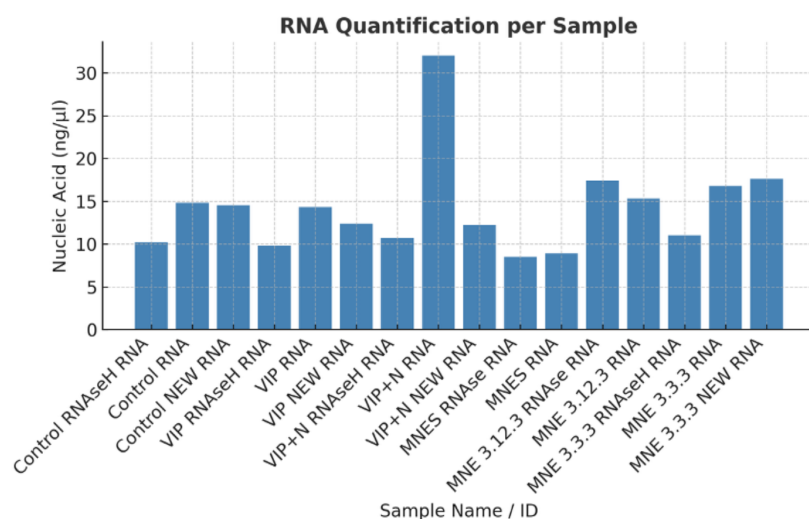


Figure 4.10: **RNA quantification of VLP samples following RNA extraction.** Nucleic acid concentration ( $\text{ng}/\mu\text{l}$ ) was measured across all VLP conditions after RNA extraction. Samples include MSE, MSE+N, MNES, MNE (3:12:2), and MNE (3:3:3) preparations, with RNaseH-treated replicates included as negative controls to confirm RNA identity. All samples yielded detectable RNA, with concentrations ranging from approximately 8 to 32  $\text{ng}/\mu\text{l}$ . The MSE+N condition showed the highest RNA yield ( $\sim 32 \text{ ng}/\mu\text{l}$ ). NEW RNA labels indicate independently prepared biological replicates.

ratios recorded to evaluate sample purity. RNA integrity and size distribution were subsequently characterized using an Agilent 2100 Bioanalyzer with the Small RNA chip (Agilent Technologies), which provides resolution in the 6–150 nucleotide range with high sensitivity appropriate for the low-abundance RNA populations expected from VLP preparations (Gourdelier et al. 2022). Electropherogram profiles were examined for the presence or absence of ribosomal RNA peaks (18S and 28S), which would indicate cellular RNA contamination, and for enrichment of small RNA species in the 20–40 nucleotide range consistent with selective vesicular RNA packaging (Kim et al. 2021).

## 4.3 Results and Discussion

To generate and characterize SARS-CoV-2 VLPs, HEK293T cells were transiently transfected with plasmids encoding the four structural proteins under multiple plasmid ratio conditions. VLP production was assessed by Western blotting of purified particle fractions, and particle size and morphology were determined by dynamic light scattering (DLS) and atomic force microscopy (AFM). Following biophysical validation, VLP-associated RNA was isolated and characterized by Bioanalyzer profiling to establish the size distribution and integrity of encapsulated RNA species.

### 4.3.1 VLP Characterization

Western blot analysis was used to confirm the incorporation of SARS-CoV-2 structural proteins into purified VLP fractions and to evaluate the effect of plasmid ratio on protein packaging efficiency. Immunoblotting with an anti-N antibody revealed that nucle-

ocapsid incorporation is strongly dependent on the plasmid transfection ratio (Figure 4.8). The MNE (3:12:2) condition produced the strongest N incorporation, with a prominent band at approximately 50 kDa, consistent with the expected molecular weight of the nucleocapsid protein. The equimolar MNE (3:3:3) condition yielded comparatively reduced N incorporation, indicating that N expression level relative to M and E is a critical determinant of packaging efficiency (Gourdelier et al. 2022). VLPs produced under the four-component MNES (3:12:2:5) condition also showed detectable N incorporation, demonstrating that spike co-expression does not abrogate nucleocapsid packaging. Importantly, VLPs generated from the MSE construct, which lacks the N expression plasmid, showed no anti-N signal, confirming the specificity of the detected band. Similarly, the MSE+N condition, in which a separate N plasmid was co-transfected with the polycistronic MSE construct, yielded no detectable N in purified VLP fractions, suggesting that the plasmid stoichiometry in this configuration is incompatible with efficient N packaging (Swann et al. 2020). Untransfected cell conditioned media and serum-free media controls produced no signal, confirming the absence of non-specific background.

Immunoblotting with an anti-M antibody demonstrated consistent membrane protein incorporation across all transfection conditions tested (Figure 4.9). A prominent M monomer band was detected at approximately 25 kDa in all VLP-containing fractions. Notably, additional bands at approximately 50 kDa and 100 kDa were also observed, corresponding to SDS-stable M dimers and higher-order oligomers, respectively. This pattern is consistent with the well-established propensity of the coronavirus M protein to form stable homo-oligomeric assemblies that resist SDS denaturation (Hardenbrook & Zhang 2022).

The MNE (3:12:2) condition produced particularly robust M signal in the VLP fraction relative to the cell lysate, suggesting efficient M enrichment during ultracentrifugation. Taken together, the Western blot data confirm that the optimized M:N:E:S ratio of 3:12:2:5 supports incorporation of both N and M into purified VLP fractions, validating these particles as biochemically representative of SARS-CoV-2 virions.

Biophysical characterization of purified VLPs by dynamic light scattering (DLS) confirmed the production of particles with size distributions consistent with native SARS-CoV-2 virions. VLPs produced under the MNE expression condition exhibited a hydrodynamic diameter in the range of 100–146 nm following RNase H treatment (Figure 4.3), in good agreement with published dimensions of authentic SARS-CoV-2 particles (Lyonnais et al. 2021). To assess whether RNase H treatment altered particle integrity, DLS measurements were performed on the same VLP preparations before and after enzymatic treatment (Figure 4.4). The size distributions before and after RNase H treatment were indistinguishable, with both conditions yielding particles in the 100–146 nm range, demonstrating that RNase H digestion removes surface-associated RNA without disrupting the structural integrity of the VLPs. This result validates the use of RNase H pre-treatment as a strategy to specifically isolate internally encapsulated RNA for downstream sequencing-based analysis.

#### 4.3.2 RNA Packaging Analysis

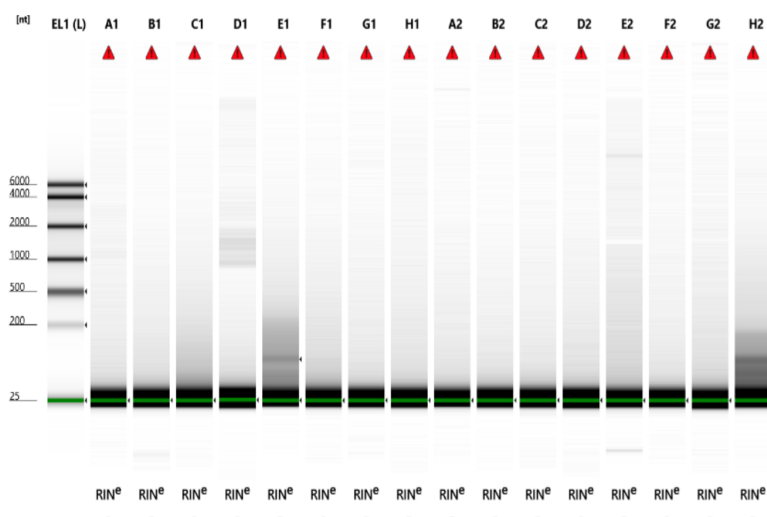
To characterize the RNA species associated with SARS-CoV-2 VLPs, RNA was extracted from purified VLP fractions following RNase H treatment and analyzed by Agilent Bioanalyzer using the Small RNA chip. The electropherogram profiles revealed a size distribution dominated by small RNA species in the approximately 20–40 nucleotide range.

Critically, neither the 18S nor 28S ribosomal RNA peaks characteristic of cellular RNA contamination were detectable in the VLP-associated RNA profiles. The absence of ribosomal RNA bands indicates that the extracted material does not reflect bulk cytoplasmic RNA contamination carried over during VLP purification, but rather represents a distinct subset of RNA species selectively associated with the vesicular fractions (Kim et al. 2021).

The enrichment of small RNAs in the 20–40 nucleotide range is consistent with known properties of extracellular vesicle-associated RNA cargo, which is typically depleted of high-molecular-weight species and enriched for small regulatory RNAs and RNA fragments (Kim et al. 2021). This size profile is also compatible with the encapsulation capacity of VLPs, which lack the nucleocapsid-bound genomic RNA of authentic virions but may selectively package host-derived or viral-derived small RNA species through interactions with the N protein RNA-binding domains or through passive encapsidation during budding (Masters 2006; Swain et al. 2022). The presence of a defined small RNA population, rather than a broad smear indicative of RNA degradation, further supports the conclusion that the RNA profiles obtained reflect genuine vesicular RNA content rather than an artifact of sample processing.

These findings establish a validated experimental framework for identifying and characterizing RNA species selectively packaged into SARS-CoV-2 VLPs. The enrichment of small RNAs and the absence of ribosomal RNA confirm the suitability of these preparations for subsequent high-throughput small RNA sequencing, which will enable systematic identification of the RNA species incorporated during VLP assembly and provide insight





**Figure 4.11: RNA quantification of VLP samples following RNA extraction. RNA Size Distribution of SARS-CoV-2 Virus-Like Particle (VLP)-Associated RNA Species Analysed by Agilent Bioanalyzer Small RNA Chip.** Capillary electrophoresis gel image depicting the size distribution of RNA extracted from purified SARS-CoV-2 VLP fractions following RNase H treatment, analysed across sixteen samples (A1–H1 and A2–H2) using an Agilent Bioanalyzer Small RNA chip. Size markers range from 25 to 6,000 nucleotides (nt), with the lower internal standard indicated by the green fluorescent band at 25 nt. RNA integrity numbers (RIN<sup>e</sup>) were not calculable for any sample (denoted “–”), consistent with the absence of intact ribosomal RNA species and the heterogeneous nature of VLP-encapsidated RNA cargo. The dominant signal observed across samples corresponds to small RNA species in the approximately 20–40 nt range, with neither 18S nor 28S ribosomal RNA peaks detected, confirming the absence of bulk cytoplasmic RNA contamination carried over during VLP purification ([Kim et al. 2021](#)). The enrichment of this discrete small RNA population, rather than a broad degradation smear, is consistent with selective encapsidation of host- or viral-derived small RNA species within VLPs, potentially mediated through interactions with the N protein RNA-binding domains or passive packaging during vesicle budding ([Masters 2006](#); [Swain et al. 2022](#)). Inter-sample variability in higher molecular weight signal (notably D1 and E2) may reflect differences in VLP yield or residual RNA species co-purified within vesicular fractions. Red triangles denote individual sample injection points.

into the mechanisms governing selective RNA packaging in the absence of productive viral replication.

## 4.4 Conclusion

This study demonstrates the successful generation and biophysical validation of SARS-CoV-2 virus-like particles using optimized co-expression of the four structural proteins M, N, E, and S in HEK293T cells. Western blot analysis confirmed that nucleocapsid incorporation into purified VLP fractions is strongly dependent on the relative plasmid transfection ratio, with the MNE (3:12:2) condition yielding the most efficient N packaging and the four-component MNES (3:12:2:5) condition providing a biologically representative particle that includes all major structural components. The consistent detection of M protein across all conditions, together with the identification of SDS-stable M oligomers, further confirms the structural integrity of the assembled particles. DLS measurements established that purified VLPs have a hydrodynamic diameter of 100–146 nm, consistent with the dimensions of authentic SARS-CoV-2 virions, and demonstrated that RNase H pre-treatment — used to remove surface-adsorbed RNA — does not perturb particle size or morphology (Lyonnais et al. 2021; Gourdelier et al. 2022).

RNA profiling of purified VLP fractions by Agilent Bioanalyzer revealed a size distribution enriched for small RNA species in the 20–40 nucleotide range and devoid of detectable ribosomal RNA, indicating that the extracted RNA reflects genuine vesicular cargo rather than cellular contamination. This RNA profile is consistent with the selective RNA packaging properties of extracellular vesicles and coronaviruses and provides the first

characterization of RNA species associated with SARS-CoV-2 VLPs produced under these optimized expression conditions ([Kim et al. 2021](#); [Masters 2006](#)).

Collectively, these results establish a validated experimental platform for systematic investigation of RNA packaging during SARS-CoV-2 VLP assembly. The biochemical and biophysical characterization presented here confirms that these VLPs are structurally and compositionally representative of native virions and that their associated RNA cargo can be reliably isolated for downstream analysis. Future work will employ high-throughput small RNA sequencing to comprehensively identify and quantify the RNA species incorporated into VLPs under different assembly conditions, with the goal of defining the molecular determinants of selective RNA packaging and their potential relevance to virus–host interactions, innate immune modulation, and the development of VLP-based therapeutic and vaccine platforms.

## Chapter 5

# Human Stress Response Specificity through Bioresonance Selectivity

**Author Contributions:** The content of this chapter has been deposited on bioRxiv and submitted to *Nature Communications*, where it is currently under review. Michael Worcester and Shayan Nejad (co-first authors) led sample preparation, microscopy, and data analysis. Devin O'Donnell developed and performed the pump system experiments. Arya Eimagh Naein contributed to image processing, RNA sequencing, and data analysis. Sara Anbir collaborated in Western blot assays, stable transfection of cells with VPC constructs, transduction with MEF2A lentivirus, and data analysis. Surya Arian performed RNA extraction and RT-qPCR. Pratyasha Mishra conducted Western blot assays and data analysis. Siyu Li and Kevin Yang performed p38 signaling pathway modeling. Matthew Guevara and Nina Y. Yuan performed Western blot assays. Seán O'Leary, Marcus Kaul, Roya Zandi, and Thomas E. Kuhlman provided supervision and support. Thomas E. Kuhlman acts as

the corresponding author. Michael Worcester and Shayan Nejad contributed equally to this work.

**Abstract** In all eukaryotes, the mitogen activated protein kinase (MAPK) cascade, a multilayered interconnected network of enzymes, connects external stimuli to gene regulation, dictating cellular fate. However, mechanisms for encoding information in this complex, fluctuating network to activate specific responses remain elusive. Here, we demonstrate that the central human stress regulator protein *p38* MAPK encodes information regarding experienced stresses as different frequency oscillations of its activation state. These oscillations are used to drive specific responses through frequency-dependent resonance of oscillating biochemical phosphorylation reactions between *p38* and downstream targets. These interactions closely mirror those of electronic alternating current (AC) circuits and their components, providing a unique framework through which to understand signal transduction in the MAPK cascade. Finally, we demonstrate how this understanding of bioresonance allows us to induce specific genetic responses simply by exposing cells to sugar to force activation state oscillations of *p38* at predetermined frequencies.

## 5.1 introduction

Enzyme specificity is essential for the adaptability and survival of all living organisms. An example relevant to all eukaryotes is the mitogen activated protein kinase (MAPK) cascade, a complex network of dozens of enzymes responsible for sensing environmental stimuli and propagating the signal to the genome to alter gene expression. The cascade transmits received signals as a current of activating and deactivating phosphoryla-

tion reactions between MAPK enzymes. In the case of environmental stresses, the current starts from the stress receptors, is passed through the MAP3K and MAP2K layers of the cascade, and is ultimately received by the master stress response regulator protein p38 MAPK (Canovas & Nebreda 2021; Trempolec et al. 2013). Once activated by dual phosphorylation at Thr-180 and Tyr-182 by the MAP2K enzymes MKK3 or MKK6 (Raingeaud et al. 1996), p38 must then selectively phosphorylate and activate the correct substrates from a pool of hundreds of different possible protein targets. This initiates specific gene expression responses appropriate to the received stress (Trempolec et al. 2013). Selection of incorrect responses can be catastrophic, as inappropriate p38 activity is implicated in many devastating human diseases, including Alzheimer’s disease (Lloret et al. 2015), amyotrophic lateral sclerosis [ALS, (Holasek et al. 2005), (Tortarolo et al. 2003)], neuroHIV (Medders et al. 2010; K.E. Medders 2011), and cancer (Gui et al. 2020; García-Hernández et al. 2022), among many others (Canovas & Nebreda 2021), (Burton et al. 2021). How can p38, which can exist in an active phosphorylated state or an inactive dephosphorylated state, encode enough information to provide specificity in response selection? In some cases, specificity in p38-target interactions is aided by allosteric control (Rohm et al. 2020) and/or regulation of p38 and substrate availability or localization (Uhlik et al. 2003), (Plotnikov et al. 2011), (Ma Y 2023). However, because p38 regulates hundreds of targets, it is difficult to imagine how to generalize such approaches to form a comprehensive model of p38 selectivity due to the numbers of additional allosteric and localization effectors that would be required. Extracting order from this complex network of interacting components is further complicated by dynamic temporal patterns and oscillations in number and activation state that have been

observed in thousands of proteins within and associated with the MAPK cascade (Sun et al. 2008a). These include the master regulator ERK1/2 (Shankaran et al. 2009),(Waters et al. 2014), the MAP2Ks MKK3 and MKK6 (Dérjard et al. 1995),(Raingeaud et al. 1996),(Han et al. 1996),(Sorkin et al. 2009), and the transcription factors NF- $\kappa$ B (Hoffmann et al. 2002),(Sung et al. 2009)and p53 (Geva-Zatorsky et al. 2010),(Lahav et al. 2004), among many others (Sun et al. 2008a). Because the output of a particular signaling pathway relies on the balance between activation and deactivation by kinase and phosphatase enzymes, the temporal behavior of pathway components has been proposed to contribute to signaling specificity (Hoffmann et al. 2002),(Sung et al. 2009),(Werner et al. 2005). However, specific mechanisms for how such temporal dynamics are interpreted or employed by signaling pathways remain largely obscure.

## 5.2 Result

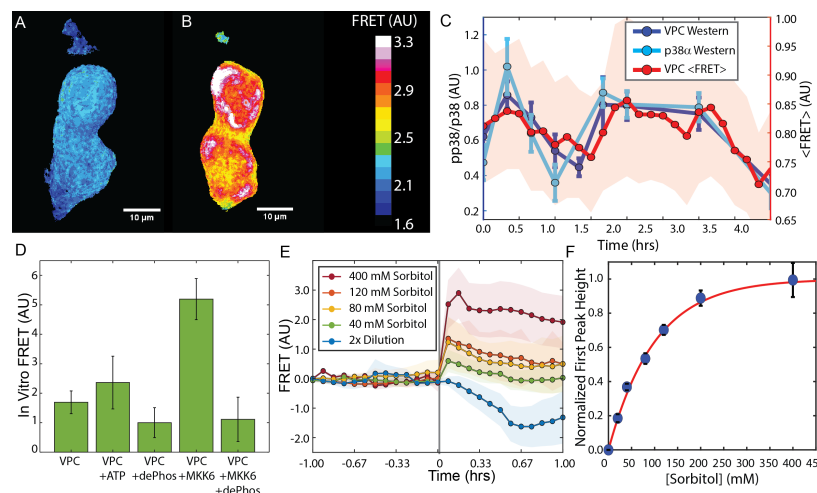
### 5.2.1 VPC directly reports p38 activation state by FRET

Existing tools for studying the flow of information through the MAPK cascade and p38 rely upon indirect measurement of p38 activity. These include kinase translocation reporters (KTRs), which are kinase substrates that report kinase activity through fluorophore translocation into and out of the nucleus upon phosphorylation(Peterson et al. 2022), and Förster resonance energy transfer-based kinase substrates that change their FRET state when acted upon by the kinase of interest. Tomida et al. previously developed the FRET-based p38 substrate PerKy-38 (Tomida et al. 2015), which is based upon the p38 docking domain of the transcription factor protein MEF2A and therefore acts as a MEF2A-like

substrate for p38. However, p38 serves only to phosphorylate PerKy-38 and KTRs, while they are dephosphorylated independently by other phosphatases (Wu et al. 2001),(Chen et al. 2007). Hence PerKy-38 and KTRs report more accurately on the activation state dynamics and interactions of p38 substrates than those of p38 itself. To overcome this limitation, we developed VPC (Venus-p38-Cerulean), which directly reports on p38 activation state by FRET, and we validated that the phosphorylation and activity dynamics of VPC are representative of native p38 on the time scales measured here. VPC is a translational fusion of the N and C terminal ends of p38 $\alpha$ , the most ubiquitously expressed isoform of p38(Hale et al. 1999), to the monomeric fluorescent reporters mVenus (Nagai et al. 2002)and mCerulean3 (Markwardt et al. 2011).

When human HEK-293T cells are stably transfected to express low levels of VPC, the VPC displays intracellular localization patterns upon stimulation that are consistent with observations of nuclear and cytoplasmic activity displayed by native p38 upon phosphorylation Figure 5.1 (A,B)(Gong et al. 2010). The dynamics of the phosphorylation state of VPC measured by Western blot and ensemble-averaged FRET recapitulate those of p38 in untransfected HEK-293T cells Figure 5.1C. Furthermore, the FRET signal of VPC purified from unstimulated cells increases when phosphorylated in vitro by the p38 kinase MKK6 (Figure 5.1D, (you need to ref to obtain data)FIGURE S1B, C). Finally, the FRET state of stimulated VPC- expressing cells changes dynamically in time Figure 5.1E and proportionally in amplitude Figure 5.1F in response to osmotic shocks of various strengths.



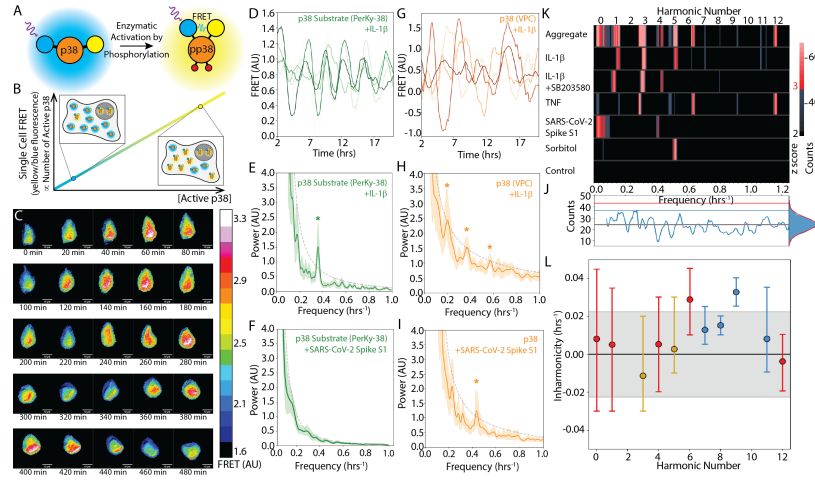


**Figure 5.1: VPC, a direct FRET-based reporter of p38 activation state.**

(A) FRET state of two resting HEK-293T cells expressing VPC (VPC293). (B) The same two cells, immediately after hyperosmotic shock induced by the addition of 120 mM sorbitol. (C) VPC (dark blue, from VPC293 cells) and p38 $\alpha$  (light blue, from untransfected HEK-293T cells) phosphorylation dynamics measured using Western blot and ensemble-averaged FRET (red,  $N \approx 1,500$  cells). Shaded region is the 95% confidence bound for FRET data, while error bars represent standard deviation of three biological Western blot replicates. See Fig. S1 for example Western blots. (D) In vitro FRET of purified VPC protein: as purified (VPC), VPC + 0.25 mM ATP-Mg (VPC +ATP), VPC + 0.25 mM ATP-Mg + 100 U/mL Antarctic Phosphatase (VPC +dePhos), VPC + 0.25 mM ATP-Mg + 200 nM MKK6 kinase (VPC +MKK6), and VPC + 0.25 mM ATP-Mg + 200 nM MKK6 kinase + 100 U/mL Antarctic Phosphatase. Values are means of 15 measurements; error bars represent SEM. (E) Dynamic ensemble-averaged FRET response of VPC293 cells upon hyperosmotic (green: 40 mM, orange: 80 mM, blue: 120 mM sorbitol) or hypoosmotic (red: 2 $\times$  dilution with ultrapure H<sub>2</sub>O) shock. Shock was induced at  $t = 0$ , indicated by the vertical gray line. Points represent the average of three biological replicates; shaded region indicates 95% confidence bounds. (F) Normalized height of the first peak in the VPC FRET signal upon induction of hyperosmotic shock via addition of sorbitol, as in (E). Blue points: mean of three biological replicates; error bars: standard deviation. Red line: best fit to the function  $1 - e^{-b[\text{Sorbitol}]}$ , with best-fit value  $b = 0.01 \pm 0.002 \text{ mM}^{-1}$ . Hyperosmotic shock with 400 mM sorbitol results in cell death. Scale bar: 20  $\mu\text{m}$  (A, B).

### 5.2.2 p38 and the TF-based p38 substrate PerKy-38 exhibit activation state oscillations

By measuring the FRET responses of both p38 and the transcription factor (TF)-based p38 substrate PerKy-38 in response to stimuli, we directly observe at the single cell level and in real time how information is passed between layers of the MAPK cascade. As the FRET level of each individual protein changes discretely when activated Figure 5.2 A, FRET levels measured from individual cells are proportional to the amount of activated PerKy-38 or VPC within those cells Figure 5.2B. When stimulated, the activation states of both p38 and PerKy-38 oscillate in time Figure 5.2C, corresponding to temporal waves in numbers of activated molecules. The activation state of the TF-based p38 substrate PerKy-38 oscillates at a specific frequency when HEK-293T cells stably transfected to express PerKy-38 (henceforth PerKy293) are stimulated with the pro-inflammatory cytokine interleukin-1 $\beta$  (IL-1 $\beta$ ), as reported previously (Tomida et al. 2015). Example PerKy293 FRET time traces are shown in Figure 5.2D. The oscillatory response is eliminated upon treatment of the cells with SB203580 (Figure S2), a p38-specific inhibitor that binds to the p38 ATP-binding pocket and prevents phosphorylation of downstream targets (Young et al. 1997). To quantify the oscillations, we performed Fourier analysis on the FRET time traces of approximately 700 cells recorded over 48 hours. (see Supplementary Materials (SM) I). The ensemble-averaged Fourier spectrum of the SB203580-treated PerKy293 control exhibits a smooth noise background (Figure S2B,C), and quantification of the statistical fluctuations from power-law behavior allows us to assign  $p$ -values and statistical significance to peaks observed in non-control spectra (see SM II, Figure S2D). For PerKy293 cells stim-



**Figure 5.2: PerKy-38 and p38 exhibit phosphorylation state oscillations with stimulus-specific frequencies.**

(A–B) Interpretation of single-cell FRET signals. (A) Activation of a p38 protein by phosphorylation discretely increases the FRET signal from that molecule. (B) Individual cells contain many p38 molecules; hence, the single-cell FRET signal is proportional to the number of active p38 molecules. (C) Oscillating p38 activation state shown for a single VPC293 cell stimulated by hyperosmotic shock with sorbitol. The first image ( $t = 0$ ) is 35 hours after addition of 120 mM sorbitol. Scale bar:  $20 \mu\text{m}$  (C). (D) FRET time traces of five fields of view (FOVs) of PerKy293 cells exposed to 300 ng/mL IL-1 $\beta$ . (E) Ensemble-averaged Fourier spectrum of 70 FOVs of PerKy293 (as in D). Shaded region: 95% confidence bounds; dashed gray line: 99% confidence interval for peak magnitudes. Asterisks indicate statistically significant peaks. (F) Ensemble-averaged Fourier power spectrum of 100 FOVs of PerKy293 cells exposed to 2 nM SARS-CoV-2 Spike S1. Note the absence of statistically significant peaks. (G) FRET time traces of five FOVs of VPC293 cells exposed to 300 ng/mL IL-1 $\beta$ . (H) Ensemble-averaged Fourier spectrum of 66 FOVs of VPC293 cell line (as in G). Asterisks indicate statistically significant peaks; gray dashed line represents the 99% confidence level. (I) Ensemble-averaged Fourier power spectrum of 70 FOVs of VPC293 cells exposed to 2 nM SARS-CoV-2 Spike S1 protein. Note the statistically significant peak at  $f = 0.45 \pm 0.01 \text{ hr}^{-1}$  ( $p = 2.3 \times 10^{-4}$ , marked with an asterisk), with no significant peak at  $f = 0.36 \text{ hr}^{-1}$ . (J) Statistics of local maxima in Fourier spectra of unstimulated VPC293 cells (see SM II). Integration over frequency yields a normal distribution (right), used to quantify  $z$  score;  $z$  = number of standard deviations from the mean. Gray: mean; blue:  $z = 2$ ; red:  $z = 3$ . (K) Oscillatory frequencies in individual cells for different stimuli. Top: aggregated spectrum. Expected frequencies of the  $n$ th harmonic [ $f_n = (n + 1)f_0$ ,  $n = 0, 1, 2, \dots$ , and  $f_0 = 0.09 \text{ hr}^{-1}$ ] are indicated. Blue:  $2 \leq z < 3$ ; red:  $z \geq 3$ . (L) Inharmonicity of identified frequencies. Points: band center of mass; error bars: band width at  $z$  score boundary. Blue:  $2 \leq z < 3$ ; red:  $z \geq 3$ ; yellow:  $z > 10$ . Gray shaded region: interval of  $\pm f_0/4$  (see SM III).

ulated with IL-1 $\beta$ , we find the only statistically significant peak occurring at a frequency of  $f = 0.36 \pm 0.02 \text{ hr}^{-1}$  ( $p = 2.5 \times 10^{-13}$ ), corresponding to a period of approximately 2.8 hours Figure 5.2E, consistent with Tomida *et al.*'s qualitatively determined period of 2–3 hours in HeLa cells (Tomida *et al.* 2015). Conversely, we observe no PerKy-38 activity or oscillations when cells are exposed to SARS-CoV-2 Spike S1 subunit protein Figure 5.2F or when osmotically shocked by the addition of 120 mM sorbitol to the medium. Direct measurement of the p38 activation state in HEK-293T cells stably transfected to express VPC (henceforth VPC293) also shows an oscillatory FRET response when stimulated with interleukin-1 $\beta$  (IL-1 $\beta$ ) Figure 5.2G. However, unlike PerKy-38, p38 oscillations occur regardless of the specific stimulus used. As a control, cells transfected to express a combination of singly labeled Venus-p38 and p38-mCerulean3 (“VPPC293”) show no FRET response (Figure S3A), and we againWe analyzed the noise spectrum of this control (Figure S3B–D, Supplementary Materials II) to allow us to assign  $p$ -values and statistical significance to peaks observed in spectra obtained from VPC293 cells. Similarly to PerKy-38, the ensemble-averaged Fourier spectrum decomposed from VPC293 cells stimulated with interleukin-1 $\beta$  (IL-1 $\beta$ ) displays a statistically significant peak at a frequency of  $0.36 \pm 0.02 \text{ hr}^{-1}$  ( $p = 4.6 \times 10^{-3}$ ), but additionally shows other significant peaks at  $0.18 \pm 0.02 \text{ hr}^{-1}$  ( $p = 2.9 \times 10^{-3}$ ) and  $0.54 \pm 0.02 \text{ hr}^{-1}$  ( $p = 3.1 \times 10^{-3}$ ) (Figure 2H). In contrast, when VPC293 cells are exposed to SARS-CoV-2 Spike S1, the VPC Fourier spectrum displays a statistically significant peak at a frequency of  $0.45 \pm 0.01 \text{ hr}^{-1}$  ( $p = 5.7 \times 10^{-4}$ ; Figure 5.2I). Hyperosmotic shock via the addition of 120 mM sorbitol to the medium results in a peak at a frequency of  $0.54 \pm 0.02 \text{ hr}^{-1}$  ( $p = 2.4 \times 10^{-4}$ ). Direct measurement of the p38 activation state in HEK-293T cells stably

transfected to express VPC (henceforth VPC293) also shows an oscillatory FRET response when stimulated with interleukin-1 $\beta$  (IL-1 $\beta$ ) Figure 5.2G. However, unlike PerKy-38, p38 oscillations occur regardless of the specific stimulus used. As a control, cells transfected to express a combination of singly labeled Venus-p38 and p38-mCerulean3 (“VPPC293”) show no FRET response (Figure S3A), and we again analyzed the noise spectrum of this control (Figure S3B–D, Supplementary Materials II) to allow us to assign  $p$ -values and statistical significance to peaks observed in spectra obtained from VPC293 cells. Similarly to PerKy-38, the ensemble-averaged Fourier spectrum decomposed from VPC293 cells stimulated with interleukin-1 $\beta$  (IL-1 $\beta$ ) displays a statistically significant peak at a frequency of  $0.36 \pm 0.02 \text{ hr}^{-1}$  ( $p = 4.6 \times 10^{-3}$ ), but additionally shows other significant peaks at  $0.18 \pm 0.02 \text{ hr}^{-1}$  ( $p = 2.9 \times 10^{-3}$ ) and  $0.54 \pm 0.02 \text{ hr}^{-1}$  ( $p = 3.1 \times 10^{-3}$ ) Figure 5.2H. In contrast, when VPC293 cells are exposed to SARS-CoV-2 Spike S1, the VPC Fourier spectrum displays a statistically significant peak at a frequency of  $0.45 \pm 0.01 \text{ hr}^{-1}$  ( $p = 5.7 \times 10^{-4}$ ; Figure 5.2I). Hyperosmotic shock via the addition of 120 mM sorbitol to the medium results in a peak at a frequency of  $0.54 \pm 0.02 \text{ hr}^{-1}$  ( $p = 2.4 \times 10^{-4}$ ).

### 5.2.3 p38 activation state oscillates with stimulus-specific frequencies

As the above experiments suggest, when treating VPC293 cells with a variety of stimuli, we observe p38 oscillatory spectra composed of different frequencies. Line spectra of detected frequencies for VPC293 cells stressed with various stimuli are shown in Figure 5.2(J,K). Combining these individual spectra into a single aggregate spectrum (Figure 5.2K, top) shows that every detected frequency from all stimuli tested here is an evenly spaced integer multiple harmonic of a fundamental frequency,  $f_0$ , such that the fre-

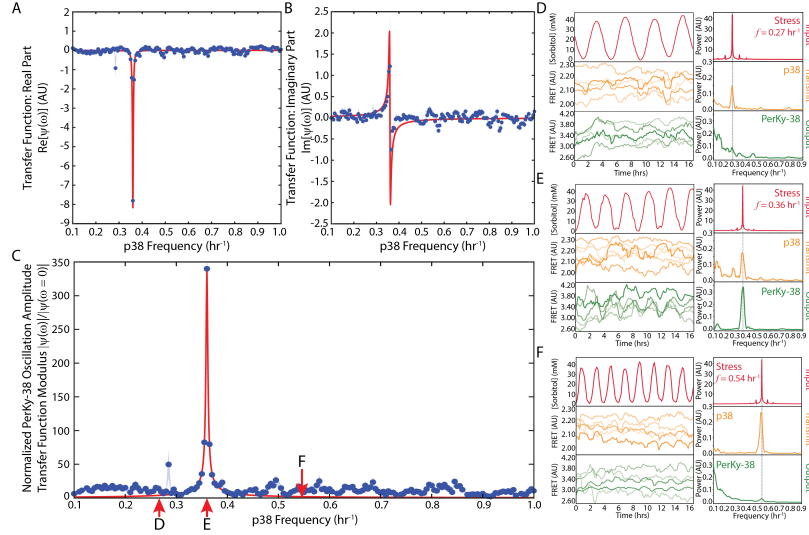
quency of the  $n$ th harmonic is given by  $f = (n + 1)f_0$ , with  $n$  an integer,  $n = 0, 1, 2, \dots$ . Quantitative analysis of these frequencies extracts a best-fit fundamental frequency of  $f_0 = 0.090 \pm 0.002 \text{ hr}^{-1}$  (see Supplementary Materials III and Figures S4A,B), and up to the 12th harmonic is observed in our experiments. Comparison of the distribution of these frequencies with that expected from a collection of randomly distributed frequencies demonstrates that the data are more closely distributed as evenly spaced harmonics than would be expected by chance ( $p = 0.004$ ; see Figure 5.2L, Figure S4C, Supplementary Materials III). This demonstrates that various stimuli result in differential enhancement of stimulus-specific p38 harmonic frequencies. The pro-inflammatory cytokine IL-1 $\beta$  (300 ng/mL) elicits a complex spectrum including a mix of 1st, 3rd, and 5th harmonics, with weaker responses of the 6th, 7th, and 11th (see Figure 5.2K, second row). Addition of SB203580 (10  $\mu$ M) along with IL-1 $\beta$  (300 ng/mL) results in replacement of the 5th with the 4th harmonic, along with weak activation of the 8th (Figure 5.2K, third row). Conversely, stimulation with the multifunctional cytokine Tumor Necrosis Factor (TNF, 200 ng/mL) also results in a complex response, but with strong activation of the 3rd, 6th, and 12th harmonics, and weak activation of the 0th, 1st, 5th, and 9th harmonics (Figure 5.2K, fourth row). Stimulation with SARS-CoV-2 Spike S1 primarily activates the 0th and 4th harmonics (Figure 5.2K, fifth row), and hyperosmotic shock induced by addition of 120 mM sorbitol activates the 5th harmonic with weak activation of the 0th (Figure 5.2K, sixth row).

Accordingly, for the stimuli tested here, we observe a response in PerKy293 cells only when stimulated with the cytokines IL-1 $\beta$  and TNF, both of which activate the 3rd harmonic at  $f_3 = 0.36 \text{ hr}^{-1}$ . In contrast, we do not observe a PerKy-38 response when

stimulated with SARS-CoV-2 Spike S1 or hyperosmotic shock (e.g., Figure 5.2F). The time dependence of the appearance of different frequencies can be complex (see Figure S5, Supplementary Materials IV), although here we focus on the time-averaged response.

#### 5.2.4 p38 activates substrates through bioresonance

The Fourier spectra of p38 and PerKy-38 exhibit precisely the relationship expected between two oscillators coupled by resonance (Lawson et al. 2020), where an oscillating target undergoes large amplitude oscillations only when pumped by a driving force oscillating at the same natural frequency of the target (see SM V, SM VI). Here, p38 serves as the driving oscillator, while PerKy-38 acts as the target, with its natural oscillatory frequency of  $f = 0.36 \text{ hr}^{-1}$  determined by interactions of PerKy-38 with p38, its phosphatases, and other proteins in this HEK-293T cell line (see SM VI). We call this phenomenon bioresonance to distinguish it from other unrelated uses of the term resonance in chemistry and biochemistry. The relationship between the p38 and PerKy-38 spectra can be quantified by the transfer function,  $\psi(\omega)$ , which describes how the output oscillatory spectrum of the PerKy-38 target,  $x(\omega)$ , arises from coupling to the input oscillations of the p38 driver,  $y(\omega)$ , such that  $x(\omega) = \psi(\omega)y(\omega)$  (see SM VII). For resonantly coupled oscillators,  $\psi(\omega)$  is a complex function whose real and imaginary parts describe the expected phase relationships and transfer of energy between the oscillators (see SM VII and Figure S6). We determined the PerKy-38/p38 transfer function by taking  $> 1,000$  ratios between PerKy-38 and p38 Fourier spectra,  $\psi(\omega) = x(\omega)/y(\omega)$ . We find that the PerKy-38/p38 transfer function is exactly predicted by describing the p38/PerKy-38 system as resonantly coupled oscillating



**Figure 5.3: p38 drives target activation through forced resonant oscillations.**

**Real (A) and Imaginary (B) parts of the transfer function,  $\psi(\omega)$ , between p38 [ $y(\omega)$ ] and PerKy-38 [ $x(\omega)$ ],  $\psi(\omega) = x(\omega)/y(\omega)$ ; see SM VII and Fig. S6.** Blue points: data,  $N = 1,092$  ratios of Fourier spectra from VPC293 and PerKy293 cells; blue shaded region: 95% confidence region; red lines: fits to the expected form of the transfer function of a driven simple harmonic oscillator (see SM VII, Eqs. S7.3, S7.7; cf. Fig. S6A,B). (C) Normalized modulus of the transfer function between p38 and PerKy-38, showing PerKy-38 amplitude response as a function of p38 driving frequency (see SM V, Eq. S5.3, and SM VII, Eq. S7.8; cf. Fig. S6C). Fits yield damping parameter  $\gamma = 0.00423 \pm 0.00009$  (SEM)  $\text{hr}^{-1}$  and quality factor  $Q \equiv \omega_0/\gamma = 535 \pm 11$ . (D–F) Driving the p38 MAPK system with periodic hyperosmotic shock. Time series (left) and Fourier spectra (right) of the input sorbitol flow (red, top, measured by rhodamine B fluorescence), p38 response (orange, middle, VPC FRET), and PerKy-38 output (green, bottom, PerKy-38 FRET) when cells are subjected to periodic hyperosmotic stress by oscillatory flow of sorbitol between 0 mM and 40 mM at a frequency of (D)  $0.27 \text{ hr}^{-1}$ , (E)  $0.36 \text{ hr}^{-1}$ , and (F)  $0.54 \text{ hr}^{-1}$ . The corresponding driving frequencies relative to the modulus of the transfer function are indicated by red arrows in (C).

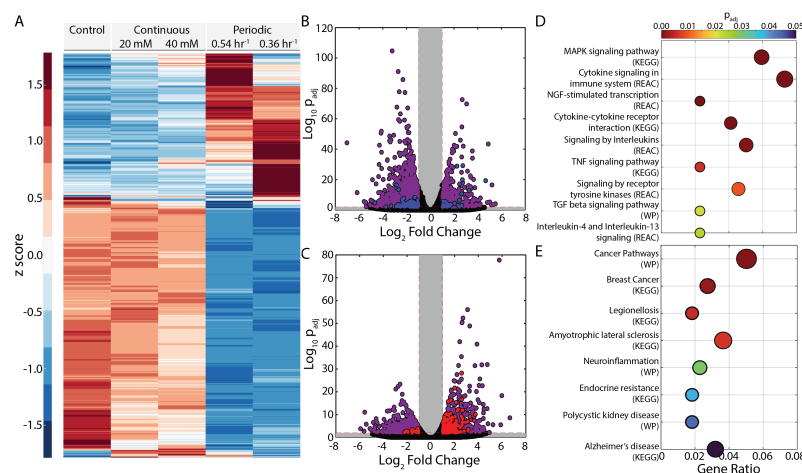


biochemical reactions Figure 5.3A-C; see SM VII, Eq. S7.7). Fitting the transfer function extracts parameters of the resonant interaction, showing that damping of the system is weak [the total rate of spontaneous dephosphorylation,  $\gamma = 0.00423 \pm 0.00009$  (SEM)  $\text{hr}^{-1}$ ; see SM V, Eq. S5.1 and SM VI, Eq. S6.7.2] and that the quality factor,  $Q$ , which measures the magnitude and specificity of amplification upon resonance, is high for such low-frequency oscillations,  $Q = 535 \pm 11$  (see SM V, Eqs. S5.5, S5.6).

### 5.2.5 Activation of specific genetic responses through dynamic bioresonant stimulation

We hypothesized that the different frequencies present in p38 activation state oscillations resonate with other naturally occurring targets with corresponding natural frequencies to selectively drive specific responses to various stimuli. This immediately suggests that it should be possible to spoof the MAPK network to elicit specific genetic responses by periodically stimulating cells to force p38 activation state oscillations of a specific frequency. To test this hypothesis, we designed a pump system (FIGURE S7) to deliver sinusoidally varying concentrations of sorbitol to induce weak periodic osmotic shock with a predetermined frequency (see FIGURES 5.1E,F). Recall that in the case of hyperosmotic shock with sorbitol at constant concentration, where the p38 spectrum includes primarily  $f_5 = 0.54 \text{ hr}^{-1}$ , we do not observe any PerKy-38 response. Conversely, when cells are stimulated by weakly oscillating hyperosmotic shock (sorbitol concentration between 0–40 mM) at PerKy-38’s resonant frequency of  $f = 0.36 \text{ hr}^{-1}$ , both the VPC and PerKy-38 activation state spectra include statistically significant peaks at that frequency Figure 5.3E. Conversely, when VPC293 and PerKy293 cells are driven at off-resonant frequencies, we

observe a signal at that frequency in VPC293 cells, but no or only a very weak response in PerKy293 cells Figure 5.3D,F, consistent with PerKy-38 activation due to bioresonance. RNA sequencing demonstrates that cells driven by periodic osmotic shock exhibit radically altered gene expression profiles when driven at different frequencies, compared to continuously osmotically shocked cells. We collected RNA from cells where, for 48 hours, the sorbitol concentration oscillated between 0–40 mM at a frequency of  $0.36 \text{ hr}^{-1}$  or  $0.54 \text{ hr}^{-1}$ , the two most statistically significant frequencies appearing in the cells' Fourier spectra Figure 5.2K. As controls, we also sequenced RNA from untreated negative control samples, as well as samples exposed continuously to 20 mM or 40 mM sorbitol for 48 hours (the mean and maximum sorbitol concentrations of the oscillations, respectively). Cluster analysis of genes identified as differentially expressed relative to the negative control shows that the negative controls and the samples treated with continuous 20 mM and 40 mM sorbitol exhibit similar overall expression profiles Figure 5.4A, indicating that the low sorbitol concentrations and mild osmotic shock imposed on the oscillatory samples are not, by themselves, sufficient to trigger a significantly altered gene regulatory response. Conversely, the expression profiles of cells driven by periodic stimulation differ dramatically from those treated with continuous stimulation, and from each other depending on the driving frequency Figures 5.4A–C. Enrichment analysis of genes uniquely differentially expressed when driven at different frequencies demonstrates that driving p38 oscillations at distinct frequencies activates different genetic programs. Notably, driving p38 oscillations using only sugar at a frequency of  $0.36 \text{ hr}^{-1}$  induces an inflammatory response typically associated with cytokine exposure. To further investigate this phenomenon, we compared the lists of genes



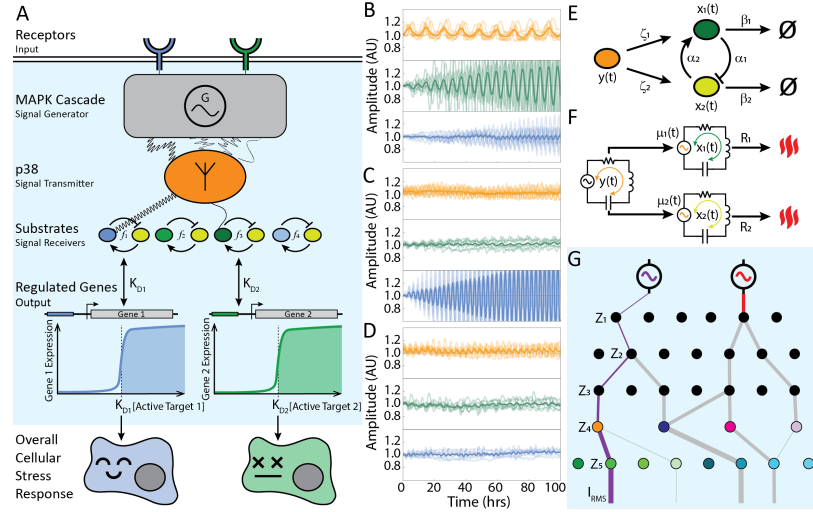
**Figure 5.4: Activation of specific genetic responses through dynamic bioresonant stimulation.**

(A) Clustergram of expression levels of 3,336 statistically significant ( $\text{padj} < 0.01$ ) differentially expressed genes with  $|\log_2(\text{fold change})| > 1$ , as determined by RNA sequencing of untreated negative control samples (Control), samples osmotically stressed with continuously applied sorbitol (20 mM or 40 mM), and samples driven with sorbitol oscillating between 0–40 mM at a frequency of either  $0.54 \text{ hr}^{-1}$  or  $0.36 \text{ hr}^{-1}$ , as in Figure 5.3D–F. All RNA-seq was performed using two biological replicates for each sample. (B–C) Volcano plots of RNA-seq results from samples periodically driven at a frequency of  $0.54 \text{ hr}^{-1}$  (B) and  $0.36 \text{ hr}^{-1}$  (C). Purple points: genes differentially expressed regardless of driving frequency ( $N = 1,144$ ); blue points: genes differentially expressed only at  $0.54 \text{ hr}^{-1}$  ( $N = 184$ ); red points: genes differentially expressed only at  $0.36 \text{ hr}^{-1}$  ( $N = 195$ ). (D–E) Enrichment analysis of genes uniquely expressed at a driving frequency of  $0.36 \text{ hr}^{-1}$  [red points in (C)], showing enriched (D) regulatory pathways and (E) disease-related pathways. Data sources include KEGG, Reactome (REAC), and WikiPathways (WP). Point size increases linearly with the number of intersecting genes (range: 5–20). Terms are filtered for brevity and relevance; full lists of enriched terms are provided in Figures S8–S10.

identified as differentially expressed when cells were driven at  $0.54 \text{ hr}^{-1}$  and  $0.36 \text{ hr}^{-1}$  to a reference set of differentially expressed genes identified from samples subjected to other stimulation frequencies ranging from  $0.09 \text{ hr}^{-1}$  to  $0.90 \text{ hr}^{-1}$ , spaced at intervals of  $0.09 \text{ hr}^{-1}$  (Table S1). From this analysis, we identified 195 genes that are uniquely differentially expressed when cells were driven at  $0.36 \text{ hr}^{-1}$ , and 184 genes differentially expressed when driven at  $0.54 \text{ hr}^{-1}$  (blue and red points in Figure 5.4B,C, respectively. Enrichment analysis of genes uniquely differentially expressed at  $0.54 \text{ hr}^{-1}$  (Table S2) yielded only one statistically significant annotated pathway, corresponding to Ras signaling (KEGG,  $\text{padj} = 0.053$ ). In contrast, enrichment analysis of the genes uniquely differentially expressed in cells driven at  $0.36 \text{ hr}^{-1}$  (Table S3)—a frequency that we found arises naturally in response to the cytokines IL- $1\beta$  and TNF Figure 5.2K—indicates that these genes are associated with MAPK responses to cytokines in general, and specifically with interleukin and TNF signaling. They are also linked to inflammation, transcriptional regulation of cell death genes, and other biological processes known to be regulated by p38 (Figure 5.4D and Figs. S8–S10). Intriguingly, we also find that genes differentially expressed when p38 oscillations are driven at a frequency of  $0.36 \text{ hr}^{-1}$  are significantly associated with a variety of disease states, including cancer ( $\text{padj} = 7.1 \times 10^{-5}$ ) and multiple neurodegenerative disorders such as neuroinflammation, amyotrophic lateral sclerosis (ALS,  $\text{padj} = 6.1 \times 10^{-3}$ ), and Alzheimer’s disease ( $\text{padj} = 5.1 \times 10^{-2}$ ), all of which have previously been linked to inappropriate p38 activity [Figure 5.4E, (Canovas & Nebreda 2021; Lloret et al. 2015; Holasek et al. 2005; Tortarolo et al. 2003; Medders et al. 2010; K.E. Medders 2011; Gui et al. 2020; García-Hernández et al. 2022; Burton et al. 2021)].

### 5.2.6 A theoretical framework describing bioresonance

To demonstrate the generality of bioresonance, we developed a mathematical framework to describe the effect of p38 oscillatory dynamics on downstream signaling pathways (see SM V–SM IX). A schematic is presented in Figure 5.5 . We simulate p38 dynamics by extending a model describing p38 phosphorylation oscillations developed by Tomida *et al.* (Tomida et al. 2015) to include interactions with upstream cascade components oscillating at different frequencies (Sun et al. (2008b); Sorkin et al. (2009); Sun et al. (2008a)). This generates oscillations in the p38 activation state whose spectrum is composed of stimulus-specific frequencies (orange lines in Figure 5.5B–D). These p38 frequencies then resonantly drive oscillations in two downstream targets, each possessing a distinct natural frequency due to their individual interactions with other cellular components. For each case, we employ standard linear models of protein oscillations—such as those used for p53 and Mdm2 (Geva-Zatorsky et al. 2010)—to simulate the system (see Supplementary Materials VI and IX). Time traces of the resulting target oscillations (FIGURE 5.5B–D) illustrate the bioresonance effect, which manifests as the amplification of specific oscillatory modes. target oscillation amplitudes when the natural frequency of the selected target appears in the p38 spectrum. We further show that this specific target activation can be used to selectively control gene expression through standard models of transcriptional regulation by protein transcription factors (see SM VIII, FIGURE S11).



**Figure 5.5: An Integrated Conceptual Framework Describing MAPK Bioresonance and Output.**

(A) When stimuli are received by the cell, receptors activate the MAPK cascade, which we treat as a “black box” signal generator (generator symbol) that sinusoidally drives p38 oscillations (orange oval) at different frequencies. p38 functions as a transmitter (antenna symbol) whose oscillatory frequencies resonantly drive oscillations of downstream targets (green, blue, and yellow ovals), which we model as two-component biological oscillators with natural frequencies  $f_i$ . Resonantly activated substrates then activate or repress genes regulated by those substrates. The combined changes in gene expression lead to the overall correct (blue) or incorrect (green) cellular stress response. (B–D) Simulated time traces of p38 (orange) resonantly interacting with two downstream targets, Substrate A and Substrate B, with natural frequencies  $f_A = 0.18 \text{ hr}^{-1}$  (green) and  $f_B = 0.36 \text{ hr}^{-1}$  (blue). Shown are examples with: (B) p38 driven at  $f = 0.09 \text{ hr}^{-1}$ , which generates harmonics on-resonance with Substrate A; (C) p38 driven at  $f = 0.36 \text{ hr}^{-1}$ , on-resonance with Substrate B; (D) p38 driven at  $f = 0.27 \text{ hr}^{-1}$ , off-resonance with both substrates. Transparent lines: individual simulations; solid lines: mean of 7 simulations simulating one FOV. Fourier spectra are averages of 14 simulated FOVs; see SM IX. (E) A quantitative model of p38 [ $y(t)$ , orange] driving oscillations in two interacting TFs [ $x_0(t)$ , green and yellow]. TFs are phosphorylated with rates  $\zeta_0$ ; TFs react with each other with rates  $\alpha_0$ ; TFs spontaneously dephosphorylate (null symbol) with rates  $\beta_0$ . See SM VI. (F) Equivalent representation of (E) in terms of AC circuit elements, oscillating phosphate current in p38 (orange) inducing chemical potentials,  $\mu_0(t)$ , in downstream TFs, with spontaneous dephosphorylation analogous to resistive heat losses. See SM VI and SM X–XI. (G) Representation of the MAPK cascade as a phosphate current amplifier and distribution network. Phosphate currents ( $I_{\text{RMS}}$ , lines) are amplified or suppressed along paths through the cascade based upon the net frequency-dependent impedance,  $Z$ , of each MAPK element (circles) along each path.

### 5.3 Discussion

Bioresonance provides a unique regulatory mechanism whereby, instead of requiring a distinct enzyme conformation for each biochemical reaction, a single enzyme can selectively control a broad range of downstream targets without alterations to its amino acid sequence or structure. This specificity is achieved by modulating the frequency of oscillations in the enzyme’s activation state. Upon activation, an enzyme governed by bioresonance may bind and act upon all potential substrates at low efficiency; however, each substrate possesses a distinct natural oscillation frequency due to its interactions with other cellular components. Only those substrates whose dynamics resonate with the enzyme’s oscillatory frequency experience selective amplification through bioresonance. As demonstrated in this work, bioresonance alone is sufficient to confer selectivity in gene regulatory responses. Nevertheless, this mechanism does not preclude the involvement of additional layers of regulation, such as allosteric control or spatial localization, which may further refine substrate specificity. Our theoretical description of bioresonance and the flow of a current of phosphates through the MAPK cascade (FIGURE 5.5E) parallels the framework used to describe alternating current (AC) electronic circuits (FIGURE 5.5F), where resonance-based selectivity is a common control mechanism. For instance, in a radio receiver, a specific signal is isolated from a multitude of transmitted signals by tuning the receiver’s internal AC circuit to resonate with the frequency of the desired electromagnetic wave. In AC circuit terminology, this resonance tuning minimizes the frequency-dependent impedance,  $Z(\omega)$ , which is inversely proportional to the circuit’s transfer function,  $Z(\omega) \propto \psi^{-1}(\omega)$  (see SM X). The result is a high-amplitude oscillating current within the receiver, which is then

decoded into audio information. Analogously, the precision and efficacy with which the MAPK signaling network selectively amplifies frequency-specific signals is characterized by the quality factor ( $Q$ ) of the system. Notably, the  $Q$  factor of the p38–PerKy-38 interaction, which operates at ultra-low frequencies in the millihertz to microhertz range, compares favorably to that of conventional electronic components operating at kilohertz to megahertz frequencies. Our results demonstrate that the p38 MAPK system—and potentially other oscillatory systems within the cell—operates via a biochemical analogue of the resonance selectivity mechanism commonly utilized in alternating current (AC) circuits. Numerous components of the MAPK cascade have been shown to oscillate similarly to p38 and its substrates as described here ([Shankaran et al. 2009](#); [Waters et al. 2014](#); [Dérillard et al. 1995](#); [Raingeaud et al. 1996](#); [Han et al. 1996](#); [Sorkin et al. 2009](#); [Hoffmann et al. 2002](#); [Sung et al. 2009](#); [Geva-Zatorsky et al. 2010](#); [Lahav et al. 2004](#); [Sun et al. 2008a](#); [Werner et al. 2005](#); [Peterson et al. 2022](#); [Tomida et al. 2015](#)), supporting the conclusion that the MAPK cascade functions as an oscillating current circuit. In this model, an oscillating electrical current of negatively charged phosphate residues is amplified and distributed to downstream elements at the terminal ends of the cascade, which directly regulate gene expression (see SM X–IX and FIGURE 5.5G).

The magnitude of the phosphate current received by each terminal transcription factor (TF) is governed by the product of the transfer functions—or, equivalently, the net impedance to phosphate current flow—through each bioresonant element along the signaling pathway (FIGURE 5.5G; see SM XI, eqs. S11.6–S11.9). This, in turn, determines the extent to which genes regulated by those TFs are differentially expressed. This correspon-



dence with AC circuits suggests that other approaches used in AC electronic information encoding may also be employed by the cell. For example, heterodyning, where low-frequency information is encoded in the beats of two combined high-frequency oscillators, could be used to frequency-shift oscillations from the short molecular time scales of chemical reactions to longer time scales relevant to cells or organisms. Finally, we have demonstrated that this quantitative framework describing the bioresonant circuit properties of the MAPK cascade enables control over the cascade’s output and resulting gene expression responses using nothing more than sugar applied at different frequencies. We propose that this work unifies a wide range of previously disparate observations regarding the oscillatory behavior of signaling components and stimulus-induced gene expression ([Shankaran et al. 2009](#); [Waters et al. 2014](#); [Dérillard et al. 1995](#); [Raingeaud et al. 1996](#); [Han et al. 1996](#); [Sorkin et al. 2009](#); [Hoffmann et al. 2002](#); [Sung et al. 2009](#); [Geva-Zatorsky et al. 2010](#); [Lahav et al. 2004](#); [Sun et al. 2008a](#); [Werner et al. 2005](#); [Peterson et al. 2022](#); [Tomida et al. 2015](#)) into a coherent and predictive framework for understanding, manipulating, and potentially engineering human gene regulation.

Further molecular-level investigation is required to elucidate the mechanisms that generate stimulus-specific p38 frequencies and to determine the parameters, time delays, or biochemical interactions that set the oscillation frequencies of p38 and its substrates ([Zhang et al. 2017](#); [Bratsun et al. 2005](#)). Additionally, our findings (FIGURE 5.4E) suggest that a number of human diseases associated with aberrant p38 signaling [Canovas & Nebreda \(2021\)](#); [Medders et al. \(2010\)](#); [Gui et al. \(2020\)](#); [García-Hernández et al. \(2022\)](#); [Burton et al. \(2021\)](#); [Atzori et al. \(2001\)](#); [Colié et al. \(2017\)](#); [Darlyuk-Saadon et al. \(2021\)](#); [Gee](#)

et al. (2020); Hensley et al. (1999); Kheiri et al. (2019); Lee & Kim (2017); Munoz & Ammit (2010); Sheng et al. (2001); Sun et al. (2003); Viorel et al. (2024); Zhu et al. (2000) may originate from dysfunction in this bioresonance mechanism.

As we demonstrate that diverse signaling responses can be elicited independently of the chemical or physical identity of the stimulus—provided that the frequency of stimulation is tuned appropriately—these results may have broad implications for pharmaceutical development and for strategies aimed at modulating cellular responses in disease.

## 5.4 Acknowledgement

We thank Nigel Goldenfeld, Edward C. Cox, Hernan Garcia, Joseph Rudnick, Matthew Scott, Zhenyu Jia, and Nathan Gabor for useful discussions.

### 5.4.1 funding

Startup funds were provided by the University of California, Riverside (TEK). Additional support was received from the UCR School of Medicine through the Dean’s Collaborative Seed Grant (TEK, MK, RZ), and the University of California Office of the President via the UC Multicampus Research Programs and Initiatives grant M21PR3267 (TEK, MK, RZ). This work was also supported by the National Science Foundation (NSF DMR-2131963, RZ) and the National Institutes of Health (NIH R01 DA052209, MK).

#### 5.4.2 Author contributions

### Author Contributions

**Conceptualization:** MK, RZ, TEK

**Methodology:** MW, SN, DO, SA, PM, AN, SL, MG, SO, MK, RZ, TEK

**Investigation:** MW, SN, DO, SA, PM, AN, SL, AA, KY, MG, NY, SO, MK, RZ, TEK

**Visualization:** MW, SN, DO, PM, SL, AA, KY, NY, AEN, TEK

**Funding acquisition:** MK, RZ, TEK

**Project administration:** MK, RZ, TEK

**Supervision:** MK, RZ, TEK

**Writing – original draft:** MW, TEK

**Writing – review & editing:** MW, SN, SL, KY, SO, MK, RZ, TEK

### Data and Materials Availability

RNA sequencing data are available from the Sequence Read Archive (SRA, BioProject PRJNA1225323) and the Gene Expression Omnibus (GEO, accession GSE289999). Code used for data analysis and simulation is available at <https://github.com/kuhlman-lab-ucr/p38>. The PerKy-38 construct was provided under a Material Transfer Agreement (MTA) with Dr. Taichiro Tomida and Dr. Atsushi Miyawaki. Plasmids pB510B-1 and pB210PA-1 for piggyBac-mediated stable transfection are restricted from distribution under a purchase

agreement with Systems Biosciences. The plasmid pUC57-~~kan~~-VPC is available upon request from [thomas.kuhlman@ucr.edu](mailto:thomas.kuhlman@ucr.edu) and can be transferred under an MTA with UC Riverside.

## Chapter 6

# Human Stress Response Specificity through Bioresonance Selectivity-Supplementary Information

### 6.1 Materials and Methods

#### 6.1.1 Plasmids, Constructs, and Proteins

The plasmid pB510B-1-PerKy-38 was the kind gift of Dr. Taichiro Tomida <sup>?</sup>. The coding sequence for VPC, expressed from a CMV promoter, was designed using VectorNTI software (Thermo Fisher Scientific), synthesized de novo, and cloned into the high copy number pUC57-kan (GENEWIZ/Azenta). A second construct, VPC-2xStrep, was similarly

designed and synthesized for VPC purification. VPC was then subcloned into the expression vector pB510B-1 (System Biosciences) by digestion of both pUC57-kan-VPC and pB510B-1 with `fcgdtgyhNheI`-HF and `BamHI`-HF (New England Biolabs), gel purification of the VPC band from pUC57-kan (Qiagen QIAquick gel extraction kit), and ligation into digested pB510B-1. The resulting plasmid, pB510B-1-VPC, was verified by nanopore sequencing (Plasmidsaurus).

SARS-CoV-2 Spike S1 subunit protein was purchased from R&D Systems (Minneapolis, MN). IL-1 $\beta$  was purchased from Invitrogen (Waltham, MA). TNF was purchased from InvivoGen (San Diego, CA). Sorbitol was purchased from Fisher Scientific. Kinase inhibitor SB203580 was purchased from Selleck Chemicals (Houston, TX).

### 6.1.2 Cell lines and culturing

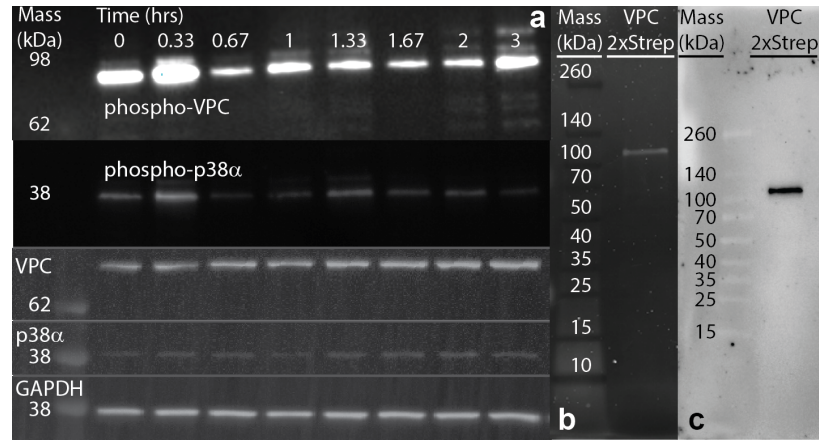
Transient transfection of HEK-293T cells with pB510B-1-PerKy-38, pB510B-1-VPC, or pUC57-kan-VPC-2xStrep was performed using FuGENE HD Transfection Reagent (Promega) at a ratio of 3:1. HEK-293T cells stably transfected with PerKy-38 or VPC were generated by the PiggyBac Transposon Vector System (System Biosciences) according to the manufacturer's directions. Briefly, HEK-293T cells were doubly transfected with the pB510B-1 vectors along with PiggyBac Transposase vector pB210PA-1. Cells were allowed to recover for 48 hours, after which 5  $\mu$ g/ml puromycin was added and cultures were expanded. After 14 days, surviving cells exhibiting blue and yellow fluorescence were collected, and the sequence of the inserted cassette was verified by PCR and sequencing.

For experiments described here, cells were cultured in complete media containing 90% DMEM with high glucose (Gibco), 10% heat-inactivated FBS (Gibco), 200 mM L-

glutamine (100X), 10 mM MEM nonessential amino acids (100X), 100 mM MEM sodium pyruvate (100X), and concentrated puromycin (Gibco).

### **6.1.3 Protein Purification and In Vitro FRET Assay**

HEK-293T cells maintained as above were transiently transfected with pUC57-kan-VPC-2xStrep. Transfected cells were washed with ice-cold Dulbecco's Phosphate Buffered Saline (DPBS), collected in a 15 ml tube (Falcon), and spun for 5 min at room temperature, 300 g. The supernatant was discarded and cells were lysed by addition of 0.5 ml ice-cold lysis buffer [1X RIPA buffer (Cell Signaling Technologies) + complete Protease Inhibitor Cocktail (Cell Signaling Technologies) + Phosphatase Inhibitor Cocktail (Cell Signaling Technologies)]. Contents were transferred to a pre-chilled 1.5 ml Eppendorf tube and incubated on ice for 10 min on a platform rotator. 3 U per  $10^6$  cells Benzonase nuclease (Sigma-Aldrich) was added, followed by another 15 min incubation on ice. The sample was centrifuged at 10,000 g, 4°C for 10 min, and the supernatant collected into a fresh 1.5 ml pre-chilled Eppendorf tube. VPC-2xStrep protein was purified from this lysate using MagStrep Strep-Tactin XT beads (IBA Lifesciences) with a magnetic separator (IBA Lifesciences). After washing twice with wash buffer (100 mM Tris-HCl, 150 mM NaCl, 1 mM EDTA), protein was eluted four times with magnetic beads elution buffer (100 mM Tris-HCl, 150 mM NaCl, 1 mM EDTA, 2.5 mM desthiobiotin). Pure protein was verified by sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS–PAGE) and western blotting (see below and Figure 6.1B,C). VPC protein concentration was determined using the bicinchoninic acid (BCA) protein assay kit (Pierce, Thermo Fisher Scientific, Rockford,



**Figure 6.1: Western blot analysis of p38 $\alpha$  and VPC phosphorylation**

**(A)** Representative immunoblots as a function of time following addition of 300 nM IL-1 $\beta$ , showing (from top to bottom): phospho-VPC, phospho-p38 $\alpha$ , total VPC, total p38 $\alpha$ , and GAPDH as a loading control. **(B)** SDS-PAGE gel of purified VPC-2 $\times$ Strep stained with SYPRO. **(C)** Western blot analysis of purified VPC-2 $\times$ Strep.

IL, USA). In vitro FRET assays were performed using VPC protein purified as described above. Purified VPC was diluted to a working concentration of 600 nM in assay buffer (20 mM HEPES, 100 mM NaCl).

FRET measurements were performed under six experimental conditions: (1) dilution buffer + 2.5 mM ZnCl<sub>2</sub>; (2) VPC (as purified) + 2.5 mM ZnCl<sub>2</sub>; (3) VPC + 0.25 mM ATP-Mg + 2.5 mM ZnCl<sub>2</sub>; (4) VPC + 0.25 mM ATP-Mg + 2.5 mM ZnCl<sub>2</sub> + 100 U ml<sup>-1</sup> Antarctic Phosphatase (New England Biolabs); (5) VPC + 0.25 mM ATP-Mg + 2.5 mM ZnCl<sub>2</sub> + 200 nM MKK6 kinase (Sigma-Aldrich); and (6) VPC + 0.25 mM ATP-Mg + 2.5 mM ZnCl<sub>2</sub> + 200 nM MKK6 kinase + 100 U ml<sup>-1</sup> Antarctic Phosphatase.

For each condition, 100  $\mu$ l of reaction mixture was prepared in wells of a black 96-well optical plate (Thermo Fisher Scientific) and incubated at 37 °C for 1 h. Fluorescence measurements were acquired using a BMG CLARIOstar Plus plate reader, with excitation at 430  $\pm$  10 nm and emission at 480  $\pm$  10 nm (mCerulean3) and 515  $\pm$  10 nm (mVenus).



Each condition was measured using three biological replicates with five technical replicates per experiment. Reported values represent the mean of fifteen total replicates. Error bars correspond to the standard error of the mean (SEM), calculated as follows:

$$z = \frac{\langle x \rangle}{\langle y \rangle}, \quad (6.1)$$

$$\text{SEM}_z = \frac{\sigma_z}{\sqrt{N}}, \quad (6.2)$$

$$\sigma_z^2 = \frac{\sigma_x^2}{\langle x \rangle^2} + \frac{\sigma_y^2}{\langle y \rangle^2} - 2 \frac{\sigma_{xy}}{\langle x \rangle \langle y \rangle}, \quad (6.3)$$

where  $\langle x \rangle$  and  $\langle y \rangle$  are the mean fluorescence intensities,  $\sigma_x$  and  $\sigma_y$  are the corresponding standard deviations, and  $\sigma_{xy}$  is the covariance between donor and acceptor signals.

#### 6.1.4 Western Blot Analysis

For Western blot analysis, cells were grown as described above in 6-well plates (Corning Costar). Cells were washed with phosphate-buffered saline (PBS) and lysed in 100  $\mu$ l ice-cold lysis buffer. Lysates were pulse-sonicated to ensure complete cell disruption and centrifuged at 21,000  $\times g$  for 15 min in a refrigerated centrifuge. Supernatants were collected and stored at  $-80^\circ\text{C}$  until use.

For immunoblotting, 25  $\mu$ g of total protein was mixed with 4 $\times$  lithium dodecyl sulfate (LDS) sample buffer (Life Technologies) and 10 $\times$  reducing agent (Invitrogen), followed by heating at  $100^\circ\text{C}$  for 5 min. Samples were resolved on 4–12% SDS-PAGE

gels (Mini-PROTEAN TGX, Bio-Rad) and subsequently electrotransferred to nitrocellulose membranes (Bio-Rad).

Membranes were blocked for 1 h at room temperature in either 5% bovine serum albumin (BSA; phospho-blot; RPI Corp., 830075) or 5% nonfat dry milk (non-phospho blot; RPI Corp., M17200) prepared in 1× TBST (Tris-buffered saline containing Tween-20). Membranes were washed three times with 1× TBST for 5 min each and incubated overnight at 4 °C with primary antibodies.

Fully activated p38 MAPK, doubly phosphorylated at Thr180 and Tyr182, was probed as described previously (53). Primary rabbit antibodies used were as follows: total p38 (Cell Signaling Technology; 9212S; 1:1000 in 1× TBST + 5% milk), phospho-p38 (ProSci; 9169; 1:1000 in 1× TBST + 5% BSA), and GAPDH (Cell Signaling Technology; 2118S; 1:1000 in 1× TBST + 5% milk).

Following primary incubation, membranes were washed three times for 15 min each in 1× TBST and incubated with goat anti-rabbit horseradish peroxidase (HRP)-conjugated secondary antibody (Invitrogen; 31460; 1:2000 in 1× TBST containing 1% BSA or milk, as appropriate). Membranes were washed three times with 1× TBST for 5 min each and developed using SuperSignal West Dura Extended Duration Substrate (Thermo Fisher Scientific).

Membranes were imaged using a Bio-Rad ChemiDoc imaging system, and band intensities were quantified using Fiji/ImageJ ([Schneider et al. 2012](#)). Reported values represent the mean of three biological replicates, with error bars indicating the standard error

of the mean (SEM). For values reported as ratios (e.g., p38/GAPDH), error propagation was calculated as follows:

$$z = \frac{\langle x \rangle}{\langle y \rangle}, \quad (6.4)$$

$$\text{SEM}_z = \frac{\sigma_z}{\sqrt{N}}, \quad (6.5)$$

$$\sigma_z^2 = \frac{\sigma_x^2}{\langle x \rangle^2} + \frac{\sigma_y^2}{\langle y \rangle^2} - 2 \frac{\sigma_{xy}}{\langle x \rangle \langle y \rangle}, \quad (6.6)$$

where  $\langle x \rangle$  and  $\langle y \rangle$  denote the mean intensities,  $\sigma_x$  and  $\sigma_y$  the corresponding standard deviations, and  $\sigma_{xy}$  the covariance between measurements.

### 6.1.5 Fluorescence Microscopy

Cells were seeded for microscopy at approximately 30% confluency in 35 mm glass coverslip dishes (Cellvis). Cells were maintained at 37 °C and 5% CO<sub>2</sub> for at least 24 h prior to imaging.

Live-cell imaging was performed using a motorized Nikon Ti-E inverted microscope. During imaging, cells were maintained in an Ibidi Silver Line stage-top incubator at 37 °C, 5% CO<sub>2</sub>, and 95% relative humidity. mCerulean3 excitation was provided by a Nikon Intensilight epifluorescence illuminator using a D436/20 excitation filter (Chroma Technology). Donor and acceptor emission signals were acquired simultaneously using a Photometrics Dual-View system equipped with a CFP/YFP filter set and detected on an Andor iXon DU897 Ultra camera.

Images were acquired with 500 ms exposure every 10 min across a field-of-view (FOV) grid programmed into the motorized stage. For a typical experiment, approximately 100 FOVs were imaged over a 48 h period, with an average of 7–8 cells per FOV. Image segmentation and single-cell tracking were performed using CellProfiler ([Stirling et al. 2021](#)).

Time traces corresponding to sinusoidally varying sorbitol concentrations were quantified using imaging media supplemented with the fluorescent dye rhodamine B.

#### **6.1.6 Pump System and Dynamic Stimulation**

Media flow was generated using four New Era NE-9000 series peristaltic pumps equipped with YG4 pump head kits, supporting tubing with a 1/16 inch inner diameter. Food-grade silicone tubing (1/16 inch inner diameter, 1/8 inch outer diameter), sanitized with 70% ethanol prior to use, was used to connect the pumps to the media reservoirs, mixing chamber, experimental dish, and waste bottle.

The experimental dish was sealed with a custom-fabricated glass cover to allow insertion of tubing through a barbed elbow connector. Pump flow rates were controlled individually through RS-232 serial connections using a custom-written Python program incorporating standard serial communication, USB interfacing, and NumPy libraries. This setup enabled generation of sinusoidal sorbitol concentration waveforms over time.

#### **6.1.7 Fourier Analysis**

Precise details of the Fourier analysis implementation are provided in Supplementary Methods I. For ensemble-averaged Fourier spectra, individual fields of view (FOVs), each containing an average of 7–8 cells, were first segmented using CellProfiler [Stirling et al.](#)

(2021). Fluorescent donor and acceptor channels were combined into FRET images using Fiji/ImageJ [Schneider et al. \(2012\)](#). Fourier analysis of single-cell FRET time traces was then performed using a custom-written Python script.

Spectra obtained from each FOV ( $N \geq 50$ ) were subsequently averaged to generate the ensemble-averaged spectra presented in Figures(5.2),(5.3).

### 6.1.8 RNAseq

Total RNA was extracted from HEK-293T cells using the RNeasy Mini Kit (Qiagen) according to the manufacturer's instructions. Negative control samples, samples subjected to continuous exposure to 20 mM and 40 mM sorbitol, and samples exposed to sorbitol oscillating between 0–40 mM at frequencies of  $0.36 \text{ h}^{-1}$  and  $0.54 \text{ h}^{-1}$  were collected with two independent biological replicates. Samples exposed to sorbitol oscillating between 0–40 mM at frequencies of  $0.09 \text{ h}^{-1}$ ,  $0.18 \text{ h}^{-1}$ ,  $0.27 \text{ h}^{-1}$ ,  $0.45 \text{ h}^{-1}$ ,  $0.63 \text{ h}^{-1}$ ,  $0.72 \text{ h}^{-1}$ ,  $0.81 \text{ h}^{-1}$ , and  $0.90 \text{ h}^{-1}$  were collected as single biological samples.

Total RNA samples were submitted to Novogene for messenger RNA (mRNA) sequencing. mRNA was enriched from total RNA using poly(T) oligo-attached magnetic beads. Following fragmentation, first-strand cDNA was synthesized using random hexamer primers, followed by second-strand cDNA synthesis using either dTTP for non-strand-specific libraries or dUTP for strand-specific libraries. Libraries were generated following end repair, A-tailing, adapter ligation, size selection, amplification, and purification. Library quality and concentration were assessed using Qubit fluorometry, real-time PCR, and an Agilent Bioanalyzer for fragment size distribution analysis.

Following library quality control, libraries were pooled based on effective concentration and target sequencing depth and sequenced using 150 bp paired-end reads on an Illumina NovaSeq X platform. Raw sequencing data (FASTQ format) were processed using *fastp*. During preprocessing, adapter sequences, reads containing poly-N stretches, and low-quality reads were removed to generate high-quality clean reads. Quality metrics including Q20, Q30, and GC content were calculated for all samples. All downstream analyses were performed using clean, high-quality reads, with each sample containing more than  $8.0 \times 10^7$  reads.

Clean reads were aligned to the *Homo sapiens* reference genome (hg38) downloaded from NCBI. Reference genome indexing and read alignment were performed using HISAT2 v2.0.5. Gene-level read counts were quantified using featureCounts v1.5.0-p3. Differential expression (DE) analysis was performed using MATLAB, which implements the DESeq differential expression analysis pipeline [Love et al. \(2014\)](#). Raw read counts were normalized using a library size factor calculated as the median ratio of each gene count relative to the geometric mean of that gene across all conditions prior to hypothesis testing.

Genes exhibiting an absolute fold change greater than two relative to the negative control (i.e.,  $|\log_2(\text{fold change})| > 1$ ) and an adjusted  $p$ -value ( $p_{\text{adj}} < 0.01$ ) were considered statistically significantly differentially expressed.

Functional enrichment analysis was performed using g:Profiler (version e112\_eg59\_p19\_25aa4782) with the g:SCS multiple testing correction method, applying a significance threshold of 0.05 ([Kolberg et al. 2023](#)).

## 6.2 Supplementary Text

### 6.2.1 I. Fourier Analysis

The FRET signal measured from individual cells is a continuous function of time,  $x(t)$ . Measurements were acquired at discrete time points with a sampling period of  $T = 10$  min, corresponding to a sampling frequency  $f_s = 1/T$ . This results in a discrete time series of  $N$  FRET measurements,  $x_n = x_0, x_1, \dots, x_{N-1}$ .

The frequency spectrum of the oscillatory signal,  $X(f)$ , was obtained by computing the discrete Fourier transform (DFT) of the sampled signal,

$$X(f) = \sum_{n=0}^{N-1} x_n e^{-i2\pi f n T}. \quad (6.7)$$

The resulting spectrum consists of  $N$  evenly spaced discrete frequency components ranging from  $f = 0$  to  $f = (N - 1)/(NT) = (N - 1)f_s/N$ , such that

$$f = 0, \frac{f_s}{N}, \frac{2f_s}{N}, \dots, \frac{(N - 1)f_s}{N}. \quad (6.8)$$

Because frequency components above the Nyquist limit do not contain independent information, only frequencies up to the Nyquist frequency  $f_{\text{Nyquist}} = f_s/2$  were retained. Accordingly, the analyzed spectrum consisted of the first  $N/2$  frequency components,

$$f = 0, \frac{f_s}{N}, \frac{2f_s}{N}, \dots, \frac{f_s}{2}. \quad (6.9)$$

A limitation of discrete Fourier analysis is the finite frequency resolution imposed by the sampling window. If dominant oscillation frequencies do not coincide with discrete

frequency bins, spectral peaks may be broadened or underestimated. To improve effective frequency resolution, we implemented a procedure in which subsets of the original time series were analyzed to generate spectra sampled at additional frequency points, denoted  $g(f)$ .

Specifically, a family of discrete spectra  $X_j(f)$  was computed using truncated subsets of the original data containing  $j$  data points, where  $j$  ranged from a minimum value  $j_{\min} = N/2 + 1$  to  $j = N$ ,

$$X_j(f) = \sum_{n=0}^{j-1} x_n e^{-i2\pi f n T}, \quad j = j_{\min}, \dots, N. \quad (6.10)$$

Each truncated time series yields a spectrum with a distinct set of discrete frequency bins. The union of these spectra was then combined to generate an enhanced-resolution frequency set,

$$g(f) = \bigcup_{j=j_{\min}}^N X_j(f). \quad (6.11)$$

### 6.2.2 II. Statistical Significance of Fourier Peaks

We found that PerKy-38 and p38 activation state measured by FRET exhibit Brownian noise (i.e., a power-law spectrum  $Af^m$  with  $m = -2$ ; Figures 6.2A and 6.4A). Accordingly, ensemble-averaged Fourier spectra were well described by a power-law function (Figures 6.2C and 6.4C). Deviations of the data from the fitted line in log space were approximately normally distributed in control conditions (Figures 6.2D and 6.4D), enabling the use of normal statistics to compute  $p$ -values and confidence intervals.



To generate the line spectra shown in Figure 5.2K, we computed Fourier spectra for individual fields of view (FOVs) and removed their respective noise backgrounds by subtraction in log space. We then applied a peak-detection algorithm to identify spectral peaks and weighted each peak by the Fourier power of its corresponding frequency bin in the ensemble Fourier spectrum. The resulting set of peaks was then segmented using a watershed procedure based on statistical significance, where the null distribution of peak magnitudes (deviations from the mean) was estimated from unstimulated VPC293 cells Figure 5.2J.

### 6.2.3 III. Analysis of Harmonics

Stimulus line spectra, as shown in Figure 5.2K, display statistically overrepresented frequencies in the Fourier spectra. We analyzed these line spectra to quantify the likelihood that the observed frequencies are harmonically organized with uniform spacing and to determine the fundamental frequency from which the harmonics are derived.

For a given candidate fundamental frequency  $f_0$ , a series of harmonics was generated in the form of a Dirac comb,

$$C_{f_0}(f) = \sum_{n=1}^{\infty} \delta(f - nf_0), \quad (6.12)$$

where  $\delta$  denotes the Dirac delta function. The locations of these harmonics are illustrated by the green tines in Figure 6.5A. For each frequency band in the aggregate

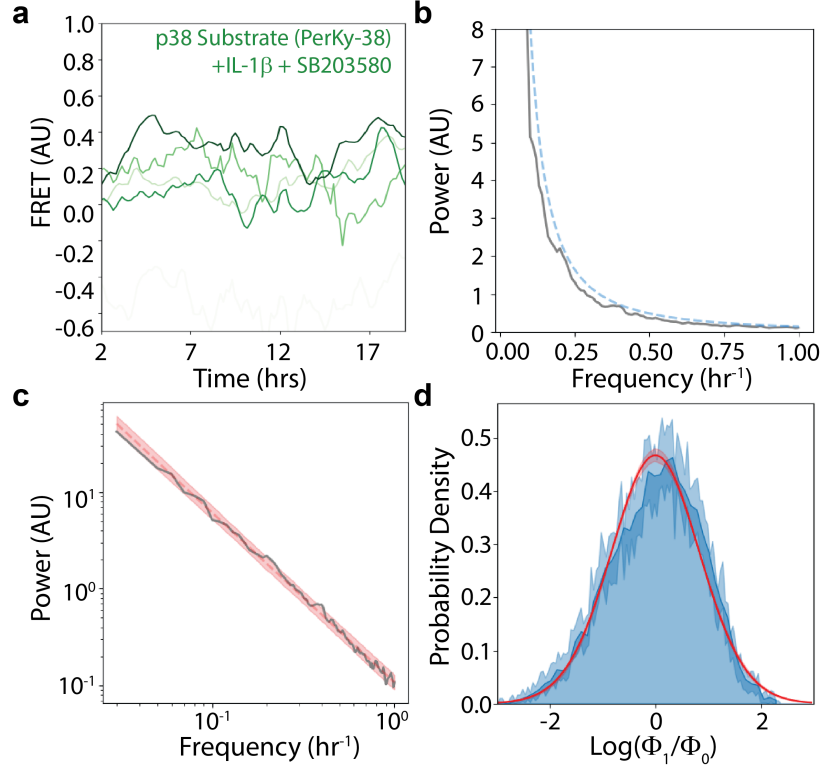
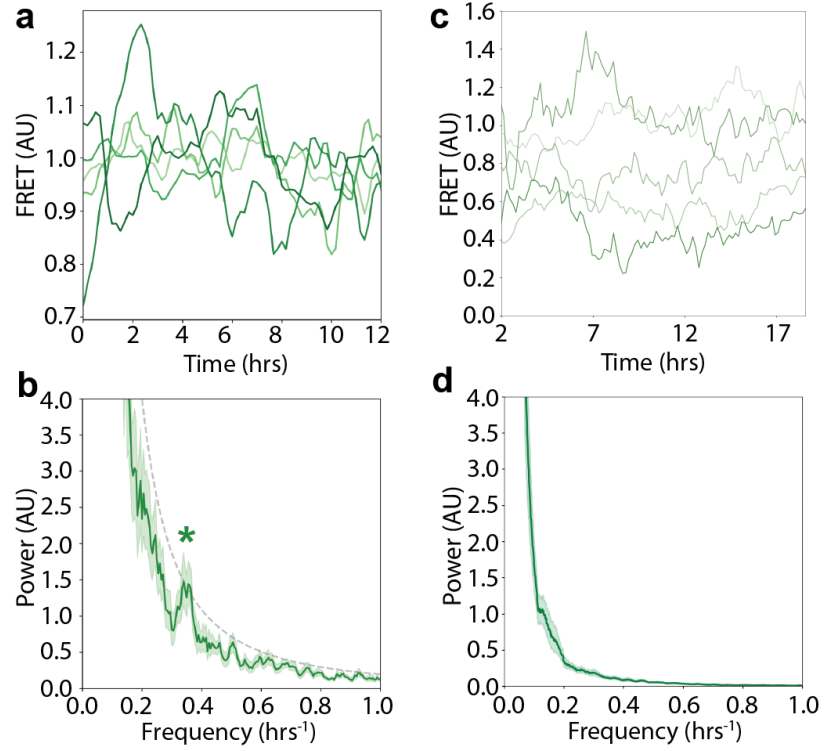


Figure 6.2: **Extended Data Fig. 2. PerKy293+SB203580 Control Statistics**

(A) Representative FRET time traces of PerKy293 cells treated with 300 ng mL<sup>-1</sup> IL-1 $\beta$  and 10  $\mu$ M SB203580. (B) Ensemble-averaged power spectrum of  $N = 450$  PerKy293 cells under the conditions shown in panel (A). The blue dashed line indicates the 99% confidence interval determined using the statistical procedure described below. (C) Log-log plot of the PerKy293 power spectrum under IL-1 $\beta$  and SB203580 treatment. The red dashed line represents a fit to a power-law  $f^{-m}$ , yielding  $m = 1.74 \pm 0.34$ . The red shaded region denotes the best fit  $\pm$  fitting error. (D) Histogram of deviations of the PerKy293 power spectrum,  $\Phi_1$ , from the power-law fit,  $\Phi_0$ . Dark blue bars indicate deviations from the best fit, while the light blue shaded region corresponds to deviations from the best fit  $\pm$  the errors shown in panel (B). The solid red line shows the best fit to a log-normal distribution, and the shaded red region indicates the uncertainty in the log-normal fit.



**Figure 6.3: PerKy293 Response to Continuous Hyperosmotic Shock and TNF Stimulation**

(A) Representative FRET time traces from five fields of view (FOVs) of PerKy293 cells treated with 120 mM sorbitol. (B) Ensemble-averaged power spectrum of PerKy293 cells treated with 120 mM sorbitol. (C) Representative FRET time traces from five FOVs of PerKy293 cells treated with 200 nM TNF. (D) Ensemble-averaged power spectrum of PerKy293 cells treated with 200 nM TNF.

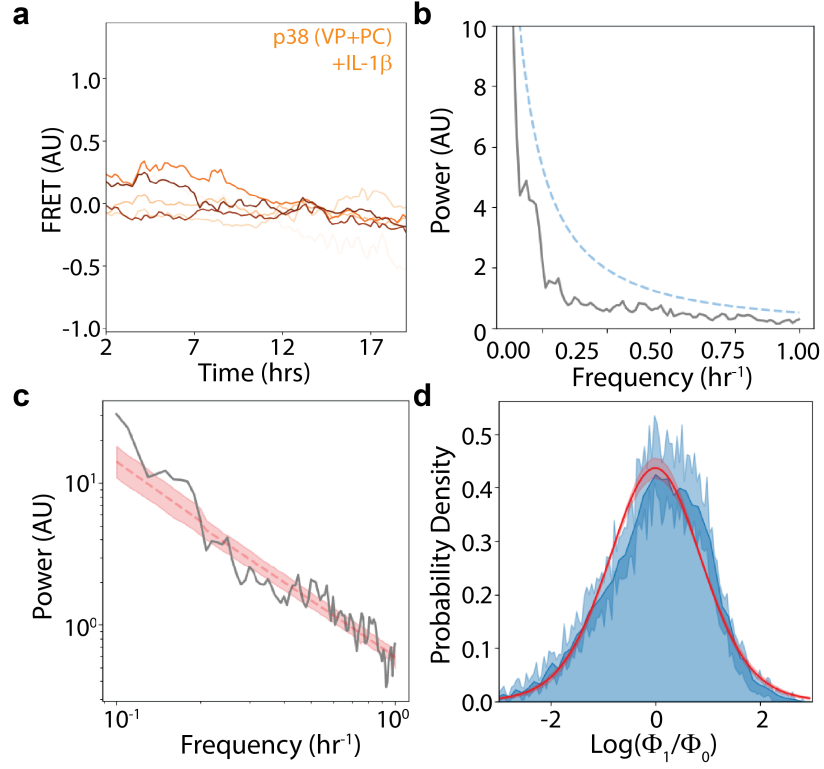


Figure 6.4: **VPPC293 Control Statistics**

(A) Representative FRET time traces from five fields of view (FOVs) of VPPC293 control cells treated with 300 ng mL<sup>-1</sup> IL-1 $\beta$ . (B) Ensemble-averaged power spectrum of VPPC293 cells. The blue dashed line indicates the 99% confidence interval determined using the statistical procedure described below. (C) Log-log plot of the VPPC293 power spectrum. The red dashed line represents a fit to a power-law  $f^{-m}$ , yielding  $m = 1.86 \pm 0.20$ . The red shaded region denotes the best fit  $\pm$  fitting error. (D) Histogram of deviations of the VPPC293 power spectrum,  $\Phi_1$ , from the power-law fit,  $\Phi_0$ . Dark blue bars indicate deviations from the best fit, while the light blue shaded region corresponds to deviations from the best fit  $\pm$  the errors shown in panel (B). The solid red line shows the best fit to a log-normal distribution, and the shaded red region indicates the uncertainty in the log-normal fit.

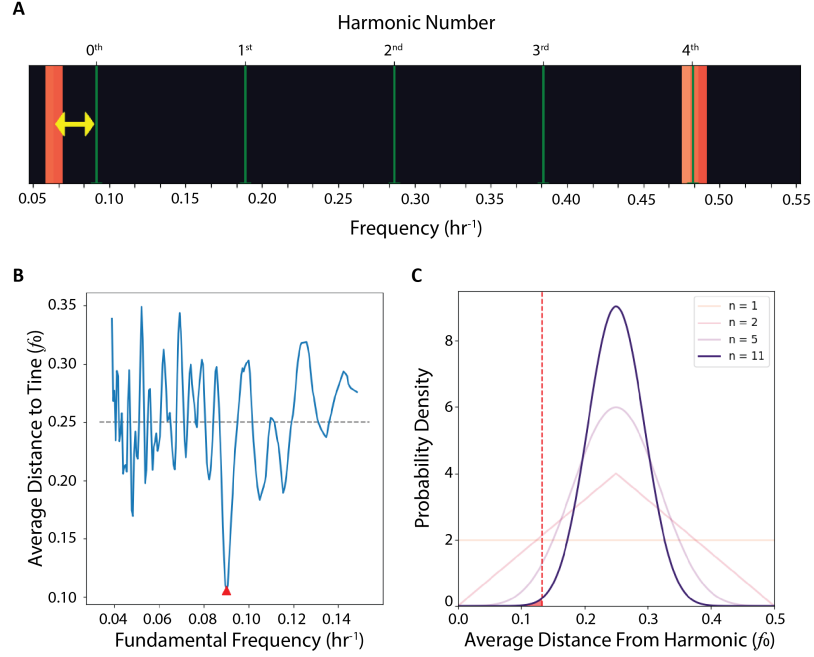


Figure 6.5: **Harmonic Analysis of Stimulus Spectra.**

(A) Distances (yellow) of identified frequency bands (red, as in Figure 5.2K) measured relative to the tines of a Dirac comb with evenly spaced harmonics (green). (B) Mean distance of frequency bands to the nearest tine of the Dirac comb as a function of the fundamental frequency  $f_0$ . The best-fit fundamental frequency corresponds to the minimum of this curve (red), yielding  $f_0 = 0.090 \pm 0.002 \text{ h}^{-1}$ . (C) Distribution of the mean distance of anharmonically (randomly) placed frequency bands from their nearest harmonic, expressed in units of the fundamental frequency  $f_0$ . The distribution follows a Bates distribution on the interval  $[0, 1/2]$  and is shown for different numbers of frequency bands. As 11 frequency bands are observed in the aggregate spectrum (Figure 5.2K, top),  $n = 11$  (purple curve) was used to calculate  $p$ -values, with distances measured from the center of mass of each band (red dashed line).

spectrum shown in Figure 5.2K, we computed the distance to the nearest harmonic. By construction, the distance of any frequency band from the nearest tine of the Dirac comb cannot exceed  $f_0/2$ .

The best-fit fundamental frequency was determined by scanning over all candidate values of  $f_0$  and minimizing the mean distance of spectral bands to their nearest harmonic tine (Figure 6.5B). Using this procedure, we identified a best-fit fundamental frequency of  $f_0 = 0.090 \pm 0.002 \text{ h}^{-1}$ .

However, a series of frequency bands placed at random frequencies may artifactually appear harmonic when evaluated solely by their distance to the nearest harmonic.

Assuming the frequencies where bands appear are random, the probability of finding a single band a given distance from the nearest tine of the Dirac comb is a uniform distribution over the interval  $[0, f_0/2]$ . Consequently, the probability to find  $n$  frequency bands an average distance  $X$  from their closest tine is the Bates distribution on the interval  $[0, 1/2]$ , i.e., the distribution of the mean,  $X$ , of  $n$  independent random variables distributed uniformly over  $[0, 1/2]$ :

$$f_X(x; n) = \frac{n^n}{(n-1)!} \sum_{k=0}^{n-1} (-1)^k \binom{n}{k} (2nx - k)^{n-1} \text{sgn}(2nx - k), \quad (6.13)$$

where  $\binom{n}{k}$  is the binomial coefficient,

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}. \quad (6.14)$$

The sign function is defined as

$$\text{sgn}(2nx - k) = \begin{cases} -1, & 2nx < k, \\ 0, & 2nx = k, \\ 1, & 2nx > k. \end{cases} \quad (6.15)$$

As we observe 11 frequency bands in our aggregated spectrum, we calculate  $p$  values using this distribution with  $n = 11$  and  $x$  measured from the center of mass of the band, as described in the text and shown in Figure 6.5C.

#### 6.2.4 IV. Short Time Fourier Transform (STFT)

Signals that evolve in time can be visualized in the frequency domain through the application of the Short Time Fourier Transform (STFT), where a long time signal is divided into segments of equal length and spacing and then Fourier transformed in steps.

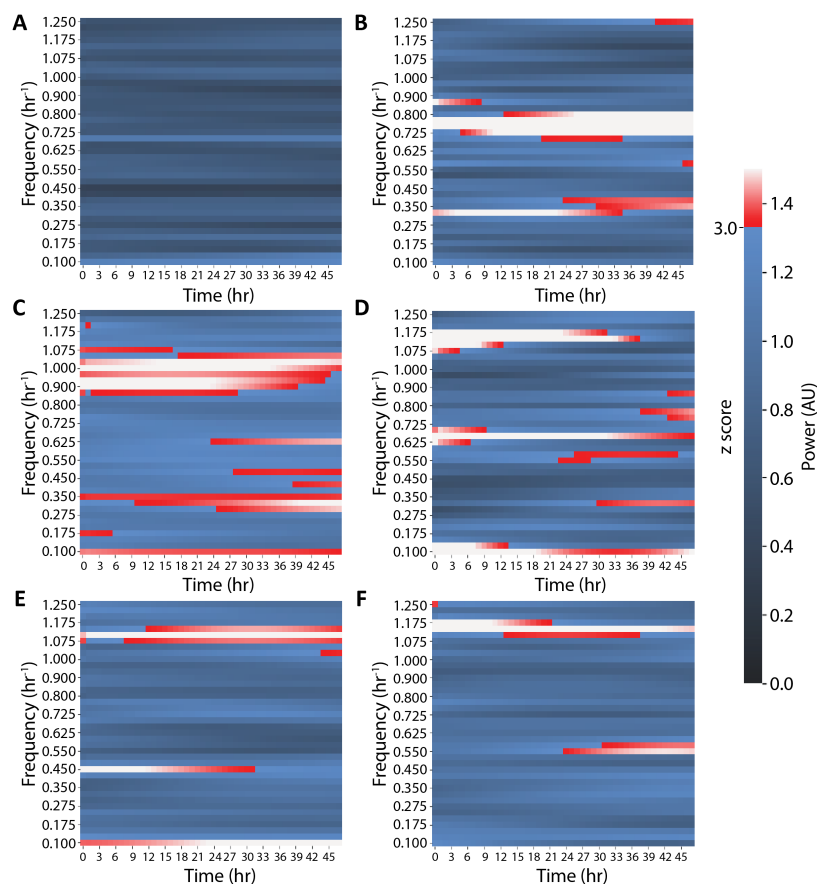
This is accomplished by multiplying the signal,  $x(t)$ , by a window function,  $w(\tau)$ , and performing the Fourier transform while sliding the window function over the length of the signal,

$$\text{STFT}_x(t, \tau, \omega) = \int_{-\infty}^{\infty} x(t) w(t - \tau) e^{-i\omega t} dt. \quad (6.16)$$

In the discrete-time case relevant here, this is performed as

$$\text{STFT}_x(nT, f) = \sum_{n=0}^{N-1} x_n w(nT - \tau) e^{-i2\pi f nT}. \quad (6.17)$$

We employed a Blackman window,



**Figure 6.6: Time Dependence of the Frequency Response of p38 to Various Stimuli.**

Short Time Fourier Transform (STFT) spectrograms illustrating how the frequency response of p38 changes over time when subjected to various stimuli. Note that by adding resolution along the time axis, resolution along the frequency axis is reduced; hence these spectrograms are not directly comparable to the time-independent Fourier spectra presented elsewhere, which are effectively (but not mathematically identical to) averages over the time axis of these spectrograms. (A) Unstimulated VPC293 control. (B) 300 ng mL<sup>-1</sup> IL-1 $\beta$ . (C) 300 ng mL<sup>-1</sup> IL-1 $\beta$  + SB203580. (D) 200 ng mL<sup>-1</sup> TNF. (E) 2 nM SARS-CoV-2 Spike S1. (F) Hyperosmotic shock induced by addition of 120 mM sorbitol.



$$w_n = a_0 - a_1 \cos\left(\frac{2\pi n}{N}\right) + a_2 \cos\left(\frac{4\pi n}{N}\right). \quad (6.18)$$

with  $a_0 = 0.42$ ,  $a_1 = 0.5$ , and  $a_2 = 0.08$ . The spectrograms plotted in Figure S6 represent the squared magnitude of the STFT,

$$\text{spectrogram}\{x_n\}(t, f) = |\text{STFT}\{x_n\}(t, f)|^2. \quad (6.19)$$

It should be noted that, by increasing the time resolution of the Fourier transform by employing the STFT, resolution along the frequency axis is lost according to the Heisenberg–Gabor limit,

$$\sigma_t \sigma_f \geq \frac{1}{2\pi}. \quad (6.20)$$

where  $\sigma_t$  and  $\sigma_f$  are the standard deviations of the time and frequency representations, respectively. Hence the spectrograms in Figure 6.6 are not directly comparable to the time-independent spectra elsewhere, and the specific frequencies with which p38 oscillates are represented with less resolution.

### 6.2.5 V. Forced Oscillations, Resonance, and Resonance Selectivity

A damped linear harmonic oscillator driven by a sinusoidally oscillating force,  $F(t) = F_0 \cos \omega t$ , is described by

$$\ddot{x} + \gamma \dot{x} + \omega_0^2 x = F(t) = F_0 \cos \omega t, \quad (6.21)$$

where  $x(t)$  is the time-dependent position of the oscillator,  $\gamma$  the damping parameter, and  $\omega_0$  is the natural angular frequency ( $\omega_0 = 2\pi f_0$ ) of the forced oscillator.  $F_0$  and  $\omega$  are the coupling strength and angular frequency of the driving oscillator, respectively. Dot notation indicates time derivatives.

The solution is given by

$$x(t) = A \cos(\omega t + \phi), \quad (6.22)$$

where the frequency-dependent amplitude response of the driven oscillator is

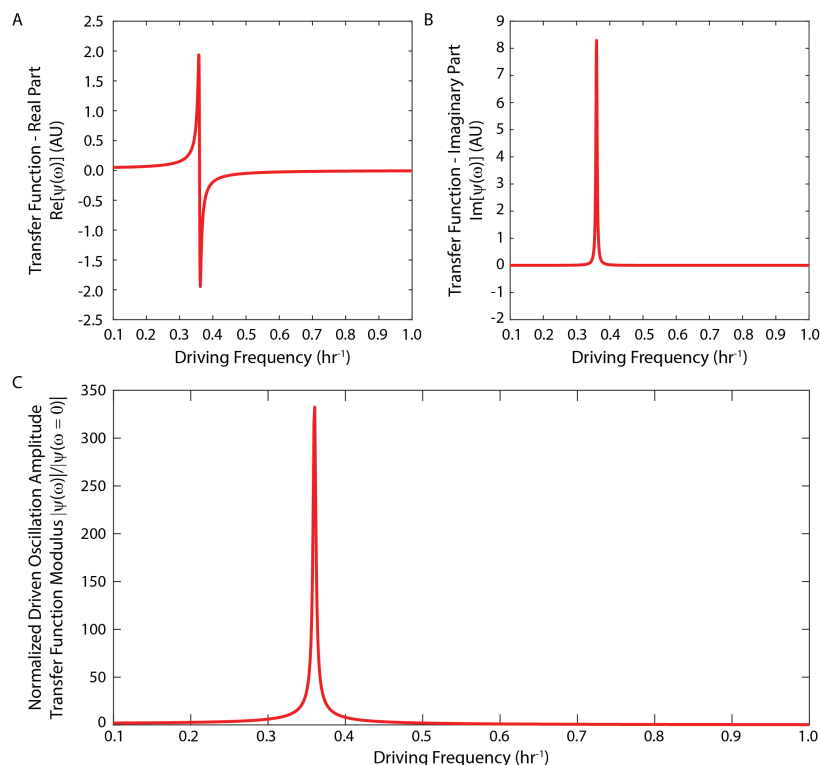
$$A(\omega) = \frac{F_0}{\sqrt{(\omega_0^2 - \omega^2)^2 + (\gamma\omega)^2}}, \quad (6.23)$$

which corresponds to the modulus of the transfer function; see Supplementary Methods VII below, Figure 5.3, and Figure 6.7. Consequently, the amplitude of the driven oscillations peaks when  $\omega = \omega_0$ , corresponding to resonance between driver and target.

The phase difference between driver and target is given by

$$\phi = \phi_x - \phi_y = \tan^{-1} \left( \frac{\gamma\omega}{\omega_0^2 - \omega^2} \right). \quad (6.24)$$

Resonance selectivity occurs when a driving oscillator with angular frequency  $\omega$  is coupled to multiple other oscillatory targets, each oscillating with their own natural angular frequencies  $\omega_{0,j}$ . When the driving frequency is close to the natural frequency of one of the substrates,  $\omega \sim \omega_{0,a}$ , the amplitude of the oscillations of substrate  $a$  is selectively amplified relative to those substrates with natural frequencies not near that of the driver  $\omega_{0,j}$  ( $j \neq a$ ).



**Figure 6.7: Transfer Function of a Forced, Lightly Damped Linear Oscillator.**

(A) Real part, (B) imaginary part, and (C) modulus of the transfer function expected for a forced, lightly damped linear oscillator, 6.37, using the parameters extracted from fitting the p38/PerKy-38 transfer function. Plots are scaled by an arbitrary parameter to allow for direct comparison to Fig5.3. Note that, compared to Fig5.3, the real and imaginary parts are swapped and the imaginary part inverted, exactly as predicted when treating the system as coupled chemical reactions, eqs. S7.7.

The quality factor,  $Q$ , is defined relative to the full width at half maximum (FWHM) of the amplification peak,  $\Delta\omega$ ,

$$Q \equiv \frac{\omega_0}{\Delta\omega}. \quad (6.25)$$

The relative magnitude of the amplification at resonance is characterized by  $Q$ ,

$$\frac{A(\omega = \omega_0)}{A(\omega = 0)} = Q. \quad (6.26)$$

Hence, ideal resonance selectivity with large, selective amplification of specific substrates corresponds to a system with a large  $Q$  factor.

### 6.2.6 VI. Bioresonance

A typical linear model for a two-component biological oscillator, e.g., the p53/Mdm2 system ([Markwardt et al. 2011](#)), is given as

$$\frac{dx_1}{dt} = -\lambda_1 x_2 - \mu_1 x_1, \quad (6.27)$$

$$\frac{dx_2}{dt} = \lambda_2 x_1 - \mu_2 x_2, \quad (6.28)$$

with

$$x_i(t) = x_i(t) - \langle x_i(t) \rangle. \quad (6.29)$$

describing the amplitude of the oscillations of the concentration of component  $x_i(t)$ ,  $x_i(t)$ , from the mean activation level,  $\langle x_i(t) \rangle$ .

In matrix notation,

$$\begin{pmatrix} \dot{x}_1 \\ \dot{x}_2 \end{pmatrix} = \begin{pmatrix} -\mu_1 & -\lambda_1 \\ \lambda_2 & -\beta_2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \quad (6.30)$$

$$\dot{\mathbf{x}}(t) = A \mathbf{x}(t). \quad (6.31)$$

If this system is driven enzymatically by a third component,  $y(t)$ , with coupling parameters  $\zeta_1$  and  $\zeta_2$ , and assuming that the  $x_i$  are in excess relative to  $y$ , then

$$\begin{pmatrix} \dot{x}_1 \\ \dot{x}_2 \end{pmatrix} = \begin{pmatrix} -\mu_1 & -\lambda_1 \\ \lambda_2 & -\beta_2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} \zeta_1 \\ \zeta_2 \end{pmatrix} y(t), \quad (6.32)$$

$$\dot{\mathbf{x}}(t) = A \mathbf{x}(t) + \boldsymbol{\zeta} y(t), \quad (6.33)$$

where the  $\zeta_j$  are the  $v_{\max}$  rates of the enzyme  $y$  reacting with  $x_j$ .

Taking the Laplace transform,

$$(sI - A) \mathbf{X}(s) = \boldsymbol{\zeta} Y(s), \quad (6.34)$$

yields

$$\mathbf{X}(s) = (sI - A)^{-1} \boldsymbol{\zeta} Y(s), \quad (6.35)$$

$$= \frac{\text{adj}(sI - A)}{\det(sI - A)} \boldsymbol{\zeta} Y(s), \quad (6.36)$$

$$= \frac{1}{\det(sI - A)} \begin{pmatrix} s + \beta_2 & \lambda_1 \\ -\lambda_2 & s + \mu_1 \end{pmatrix} \begin{pmatrix} \zeta_1 \\ \zeta_2 \end{pmatrix} Y(s). \quad (6.37)$$

$$\det(sI - A) \mathbf{X}(s) = \begin{pmatrix} s + \beta_2 & \lambda_1 \\ -\lambda_2 & s + \mu_1 \end{pmatrix} \begin{pmatrix} \zeta_1 \\ \zeta_2 \end{pmatrix} Y(s). \quad (6.38)$$

From this we obtain

$$\ddot{x}_1 + (\mu_1 + \beta_2)\dot{x}_1 + (\mu_1\beta_2 + \lambda_1\lambda_2)x_1 = \zeta_1\dot{y} + (\beta_2\zeta_1 + \lambda_1\zeta_2)y. \quad (6.39)$$

$$\ddot{x}_2 + (\mu_1 + \beta_2)\dot{x}_2 + (\mu_1\beta_2 + \lambda_1\lambda_2)x_2 = \zeta_2\dot{y} + (\mu_1\zeta_2 - \lambda_2\zeta_1)y. \quad (6.40)$$

With the identifications

$$\omega_0^2 = \mu_1\beta_2 + \lambda_1\lambda_2, \quad (6.41)$$

$$\gamma = \mu_1 + \beta_2, \quad (6.42)$$

$$F_1(t) = \zeta_1\dot{y} + (\beta_2\zeta_1 + \lambda_1\zeta_2)y, \quad (6.43)$$

$$F_2(t) = \zeta_2\dot{y} + (\mu_1\zeta_2 - \lambda_2\zeta_1)y, \quad (6.44)$$

these become the standard equations of forced, damped linear harmonic oscillators,

6.21,

$$\ddot{x}_1 + \gamma \dot{x}_1 + \omega_0^2 x_1 = F_1(t), \quad (6.45)$$

$$\ddot{x}_2 + \gamma \dot{x}_2 + \omega_0^2 x_2 = F_2(t). \quad (6.46)$$

If the driving enzyme oscillations,  $y(t)$ , are expressed as a Fourier spectrum with a set of frequencies  $\omega_k$ ,

$$y(t) = \sum_k y_k e^{i\omega_k t}, \quad (6.47)$$

then

$$F_1(t) = \sum_k [(i\omega_k \zeta_1 + \beta_2 \zeta_1 + \lambda_1 \zeta_2) y_k] e^{i\omega_k t} = \sum_k A_{1,k} e^{i(\omega_k t + \phi_{1,k})}, \quad (6.48)$$

$$F_2(t) = \sum_k [(i\omega_k \zeta_2 + \mu_1 \zeta_2 - \lambda_2 \zeta_1) y_k] e^{i\omega_k t} = \sum_k A_{2,k} e^{i(\omega_k t + \phi_{2,k})}, \quad (6.49)$$

with

$$\phi_{j,k} = \tan^{-1} \left( \frac{\text{Im}(A_{j,k})}{\text{Re}(A_{j,k})} \right). \quad (6.50)$$

$$A_{j,k} = y_k \sqrt{\zeta_j^2 + \omega_k^2}. \quad (6.51)$$

$$\zeta_1 = \lambda_2 \zeta_1 - \lambda_1 \zeta_2, \quad (6.52)$$

$$\zeta_2 = \lambda_1 \zeta_2 + \lambda_2 \zeta_1. \quad (6.53)$$

The  $F_j(t)$  are therefore also sinusoidal with the same set of frequencies  $\omega_k$ .

Consequently, it is clear that if the driving enzyme  $y(t)$  oscillates, then any two-component biological oscillator coupled to that enzyme will demonstrate all of the familiar phenomena of driven damped oscillators, including resonance and resonance selectivity, as demonstrated above (SM V). As described above in SM V, ideal resonance selectivity requires a large  $Q$  factor,

$$Q = \frac{\omega_0}{\gamma} = \frac{\mu_1 \beta_2 + \lambda_1 \lambda_2}{\mu_1 + \beta_2}, \quad (6.54)$$

that is, a system with small spontaneous degradation rates,  $\beta_j$ , relative to the  $v_{\max}$  enzymatic rates,  $\alpha_j$ . We show below that this is consistent with the p38/PerKy-38 system, i.e.,  $\gamma \ll \omega_0$  (see SM VII), leading to

$$F_j(t) \simeq \zeta_j y(t). \quad (6.55)$$

### 6.2.7 VII. Linear Response Theory and the p38/PerKy-38 Transfer Function

The spectral response of a forced oscillator,  $x_j$ , can be expressed as proportional to the spectrum of the driving input,  $y$ , using the transfer function,  $\chi_j$ , also referred to as the generalized susceptibility or response function, as



$$x_j = \chi_j y. \quad (6.56)$$

where  $x_j$  and  $y$  are the Fourier spectra of  $x_j(t)$  and  $y(t)$  (see SM VI, [6.45,6.55](#), respectively).

For a driven damped simple harmonic oscillator, the transfer function is

$$\chi(\omega) = \frac{1}{\omega_0^2 - \omega^2 + i\gamma\omega}, \quad (6.57)$$

$$\chi(\omega) \equiv \chi' + i\chi''. \quad (6.58)$$

Of particular interest, the imaginary part of the transfer function,  $\chi''$ , describes the time-averaged rate of energy absorption by the driven oscillator  $j$  as work,  $W_j$ ,

$$\frac{dW_j}{dt} = \omega \chi''. \quad (6.59)$$

Here this corresponds to the rate that energy is consumed to phosphorylate the substrate  $x_j$  using ATP.

For weak damping,  $\gamma \ll 1$ , only behavior near resonance is significant,  $\omega \approx \omega_0$ , yielding

$$\chi(\omega) \simeq \frac{1}{\omega_0^2 - \omega^2 + i\gamma\omega}. \quad (6.60)$$

This expected form of the real part, imaginary part, and modulus of the transfer function of a driven, lightly damped linear oscillator is shown in Figure [6.7](#).

Expressing  $\chi_j$  in terms of the spectrum of the driver,  $\chi_j = \chi_j + i\omega\chi_j y$ , yields

$$x_1 = (\chi_1 + i\omega\chi_1) y, \quad (6.61)$$

$$x_2 = (\chi_2 + i\omega\chi_2) y, \quad (6.62)$$

which defines a modified transfer function,  $\tilde{\chi}_j$ ,

$$x_j = \tilde{\chi}_j y. \quad (6.63)$$

with

$$\tilde{\chi}_j = \chi_j + i\omega\chi_j, \quad (6.64)$$

$$\tilde{\chi}_j = \chi'_j + i\chi''_j. \quad (6.65)$$

The real and imaginary parts of  $\tilde{\chi}_1$  we experimentally obtain for the p38/PerKy-38 system are shown in Figure 5.3A,B. Notably, we observe the behavior of the real and imaginary parts switched and the real part inverted compared to that naïvely expected for a driven harmonic oscillator oscillator, 6.60 (compare Figure 5.3A,B to Figure 6.7A,B). Hence, we conclude that for this system  $\gamma \ll \omega_0$  and

$$\tilde{\chi}_1 = i\omega\chi_1, \quad (6.66)$$

$$\tilde{\chi}_1 \simeq -\chi''_1 + i\chi'_1, \quad (6.67)$$

switching the real and imaginary parts and inverting the real part, as experimentally observed (Figure 5.3A,B).

If driven by a sinusoidal oscillator with discrete frequency  $\omega$ , the amplitude of the driven oscillations in PerKy-38 is given by the modulus of the modified transfer function,

$$x_1(t) = |\tilde{\chi}_1| y(t), \quad (6.68)$$

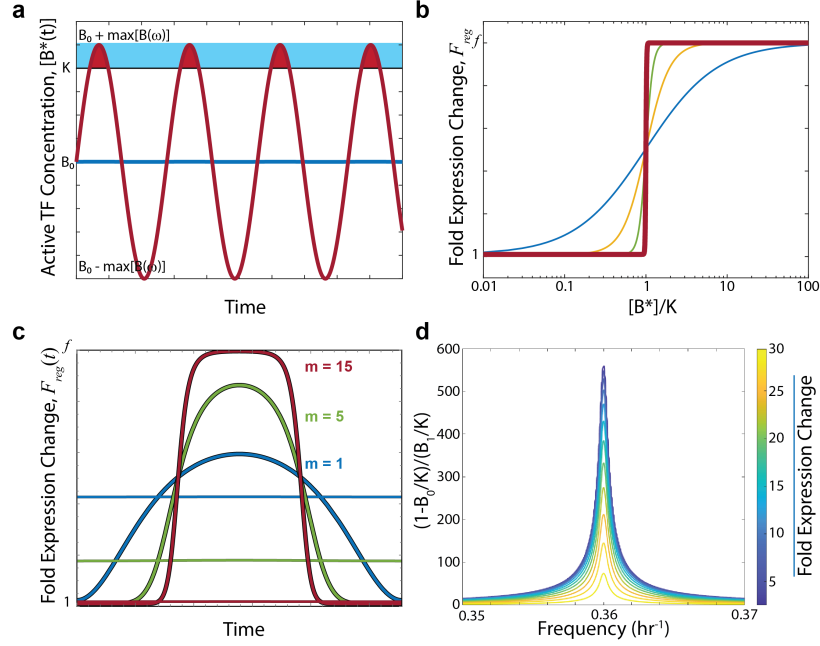
$$= \frac{\omega_0}{\sqrt{(\omega_0^2 - \omega^2)^2 + (\gamma\omega)^2}} y(t), \quad (6.69)$$

$$\simeq \frac{\omega_0}{\sqrt{(\omega_0^2 - \omega^2)^2 + (\gamma/2)^2}} y(t), \quad (6.70)$$

cf., 6.23above. The modulus of  $\tilde{\chi}_1$  for the p38/PerKy-38 system is shown in Figure 5.3C, yielding a best-fit value of  $\gamma = 0.00423 \pm 0.00009$  (SEM)  $\text{h}^{-1}$ . This corresponds to a quality factor,  $Q \equiv \omega_0/\gamma$ , of  $Q = 535 \pm 11$ . Since measurement by fluorescence also introduces an unknown scaling due to fluorescence units,  $\omega_0$  cannot be directly determined from these data by fitting.

### 6.2.8 VIII. Bioresonant Transcriptional Regulation

We next address whether bioresonance can provide specificity in gene regulation. A simple model of the output of a genetic circuit regulated by the binding and unbinding of transcription factors (TFs) at some cis-regulatory region can be described using statistical mechanics (Bintu et al. 2005a),(Bintu et al. 2005b). As the oscillations we observe have



**Figure 6.8: Bioresonant Transcriptional Regulation**

(A) Concentration of active resonant (red) and off-resonant transcription factor (blue line) over time, 6.74. The dissociation constant for binding to a cis-regulatory region,  $K$ , is indicated by a horizontal black line. The duty cycle,  $D$  6.85, is calculated as the ratio of the red to blue shaded areas over one period. (B) The regulatory function  $F_{\text{reg}}$  6.73 plotted as a function of active transcription factor concentration for various values of the Hill coefficient,  $m$ . Blue,  $m = 1$ ; orange,  $m = 3$ ; green,  $m = 10$ ; red,  $m = 100$ . The “duty cycle” approximation corresponds to large values of  $m$  resulting in a switch-like step function, e.g., the red curve. (C) The regulatory function  $F_{\text{reg}}$  plotted as a function of time (i.e., the integrand of 6.76 for various values of the Hill coefficient,  $m$ , for resonant (thick lines) and off-resonant (thin lines) transcription factors. The ratio of areas under the thick versus thin lines represents the enhancement of gene expression when a transcription factor is on-resonance versus off-resonance. Parameter values used:  $f = 100$ ,  $B_0/K = 0.73$ ,  $B_1/K = 0.001$ ,  $f_0 = 0.36 \text{ h}^{-1}$ ,  $\gamma = 0.004$ . (D) The gene expression fold change as a function of the driving frequency and parameter combination  $(1 - B_0/K)/(B_1/K)$  under the duty cycle approximation, 6.84 and 6.85. Note the narrow range of the  $x$ -axis, indicating that gene expression is enhanced only very close to the resonant frequency of the transcription factor. Parameter values used:  $f = 100$ ,  $f_0 = 0.36 \text{ h}^{-1}$ ,  $\gamma = 0.004$ .

periods of multiple hours, we assume that on the short time scales of TF binding and unbinding the system can be treated as in equilibrium.

Consider a TF  $B$  that can exist in either an active ( $B^*$ ) or inactive ( $B$ ) state. When active, it binds to a cis-regulatory region with dissociation constant  $K$  and Hill coefficient  $m$ , which quantifies the degree of TF binding cooperativity. When bound, the TF interacts with RNA polymerase to a degree quantified by the capacity  $f$ . If RNA polymerase binds at the promoter with dissociation constant  $K_P$ , the probability to find RNA polymerase bound at the promoter is described by (Bintu et al. 2005a),(Bintu et al. 2005b)

$$P_{\text{RNAP bound}} = \frac{1}{1 + K_P [\text{RNAP}] F_{\text{reg}}}. \quad (6.71)$$

where  $F_{\text{reg}}$  is given by

$$F_{\text{reg}} = \frac{1 + f \cdot (B^*/K)^m}{1 + (B^*/K)^m}. \quad (6.72)$$

The rate of transcription, and therefore the rate of gene expression, is proportional to this probability Bintu et al. (2005b,a). With the standard approximation that the promoter is weak,  $[\text{RNAP}] K_P \ll 1$ ,

$$P_{\text{RNAP bound}} \simeq [\text{RNAP}] K_P F_{\text{reg}} = [\text{RNAP}] K_P \frac{1 + f \cdot (B^*/K)^m}{1 + (B^*/K)^m}. \quad (6.73)$$

In the MAPK cascade, transcription factors are activated and deactivated by phosphorylation and dephosphorylation reactions, resulting in oscillations of the transcription factor activation state (see SM VI). Consequently, the amount of active transcription factor at any time  $t$  is given by

$$B^*(\omega, t) = B_0 + B \cos(\omega t + \phi). \quad (6.74)$$

with  $B_0$  the mean level of activation,  $\omega$  the angular frequency, and  $\phi$  an arbitrary phase shift. The amplitude of oscillations,  $B(\omega)$ , is given by 6.23,

$$B = \frac{B_1}{\sqrt{(\omega_0^2 - \omega^2)^2 + (\gamma\omega)^2}}, \quad (6.75)$$

with  $B_1$  a scaling parameter. For such a system,  $B = 2B_1/\gamma \leq B_0$ , such that the transcription factor concentration always remains positive.

The time-averaged rate of gene expression is obtained as

$$\text{Gene Expression Rate} \propto \langle P_{\text{RNAP bound}} \rangle \simeq \frac{1}{T} [\text{RNAP}] K_P \int_0^T \frac{1 + f \cdot (B^*(\omega, t)/K)^m}{1 + (B^*(\omega, t)/K)^m} dt. \quad (6.76)$$

For bioresonance to work as a transcriptional regulatory strategy, it must be the case that Figure 6.8A

$$B_0, B < K \sim B_0 + B. \quad (6.77)$$

For an off-resonant transcription factor,  $B \simeq 0$ , and the integrand of 6.76 loses time dependence. The integral in 6.76 can be performed trivially to yield

$$\text{Gene Expression Fold Change}_{\text{OFF}} \equiv \text{GEFC}_{\text{OFF}} = \frac{1 + f \cdot (B_0/K)^m}{1 + (B_0/K)^m} \frac{1 + (b_0)^m}{1 + f \cdot (b_0)^m}, \quad (6.78)$$

i.e., off-resonance transcription factors controlled by the same master regulator are activated to a degree quantified by  $b_0 = B_0/K$ .

For an on-resonance transcription factor, 6.76 can be evaluated analytically using, e.g., Cauchy's residue theorem. However, the result quickly becomes unwieldy for  $m > 1$ . We demonstrate below that there must be some minimum degree of cooperativity characterized by a Hill coefficient  $m^* \gg 1$  for bioresonance to specifically activate genetic responses.

For  $m = 1$ , we obtain from integrating 6.76

$$\text{GEFC} = \frac{1 + fb_0(1 - b_1 + b_0^2)}{1 + b_0(1 - b_1 + b_0^2)} - f \cdot \frac{1 - (1 - b_1 + b_0^2)}{1 + b_0(1 - b_1 + b_0^2)}. \quad (6.79)$$

$$= \frac{1}{1 + f \cdot \delta b_0 / (1 + b_0)} - f \cdot \frac{1 - \delta}{1 + b_0}, \quad (6.80)$$

with

$$1 - b_1 + b_0^2 \leq 1. \quad (6.81)$$

where we have again absorbed  $K$  into the definitions of  $b_0$  and  $b_1$ . Note that Eqs. 6.79 and 6.80 reduce to Eq. 6.78 when  $b_\omega \neq 0 \simeq 0$  and  $\omega \neq 0 \simeq 1$ . On resonance,  $\omega = \omega_0$ , and

$$\text{GEFC}_{\text{ON}} = \frac{1}{1 + f \cdot \delta b_0 / (1 + b_0)} - f \cdot \frac{1 - \delta}{1 + b_0}, \quad \delta = 1 - \frac{2b_1}{1 + b_0^2}. \quad (6.82)$$

For any reasonable choice of parameters 6.77, it is therefore the case when  $m = 1$  that

$$\text{GEFC}_{\text{ON}} = \text{GEFC}_{\text{OFF}}, \quad (6.83)$$

i.e., bioresonance cannot provide specificity in transcriptional regulation when TF binding is non-cooperative ( $m = 1$ ).

Conversely, for  $m \gg 1$ , we consider a duty cycle approximation of 6.76. If TF binding at the cis-regulatory region is highly cooperative (i.e.,  $m \gg 1$ ),  $F_{\text{reg}}$  can be approximated as a step-function switch; that is, expression is at its maximum level if  $B^* > K$ , whereas expression is completely off if  $B^* < K$  Figure 6.8B. With this approximation, the time-averaged gene expression fold change is

$$\text{GEFC} \simeq f \cdot D, \quad (6.84)$$

where  $D(\omega)$  is the TF's duty cycle, or the fraction of time when  $B^* > K$  and the TF is bound. For sinusoidal oscillations of  $B^*$  as given by 6.74,  $D(\omega)$  can be defined piecewise as

$$D = \begin{cases} 0, & 1 - \frac{b_0}{b} > 1, \\ \frac{1}{2} - \frac{1}{\pi} \sin^{-1} \left( \frac{1 - b_0/b}{1} \right), & -1 \leq \frac{1 - b_0}{b} \leq 1, \\ 1, & 1 - \frac{b_0}{b} < -1. \end{cases} \quad (6.85)$$

This results in an extremely sharp enhancement of gene expression only for on-resonance transcription factors when  $\omega = \omega_0$ , with zero gene expression otherwise, as shown in Figure 6.8D.



Consequently, we conclude that for on-resonance transcription factors to lead to a greater effect on gene expression than off-resonant transcription factors, nonlinearity in transcription factor binding must be present to suppress the level of gene expression when  $B^* < K$  and amplify the effect when  $B^* > K$  Figure 6.8B. That is, transcription factor binding at cis-regulatory regions must be highly cooperative for bioresonance to provide specificity in gene regulation Figure 6.8C, which is consistent with the experimentally observed generality of high cooperativity in eukaryotic transcriptional regulation (Martin et al. 2023).

### 6.2.9 IX. Simulations of p38 bioresonance selectivity

For simulations of p38 activity and bioresonance selectivity shown in Figure 5.5, we use the set of nonlinear differential equations describing the p38 (MAPK) phosphorylation state developed by Tomida et al. (Tomida et al. 2015),

$$\frac{d \text{MAP2K}}{dt} = k_0 S(t) [1 - \text{MAP2K}] - k_1 \text{MAP2K}, \quad (6.86)$$

$$\frac{d \text{MAPK}}{dt} = k_2 \text{MAP2K} [1 - \text{MAPK}] - \frac{k_3 \text{MKP1} \text{MAPK}}{k_4 + \text{MAPK}}, \quad (6.87)$$

$$\frac{d \text{MKP1RNA}}{dt} = k_5 \text{MAPK} [1 - \text{MKP1RNA}] - k_6 \text{MKP1RNA}, \quad (6.88)$$

$$\frac{d \text{MKP1}}{dt} = k_7 \text{MKP1RNA} [1 - \text{MKP1}] - k_8 \text{MKP1}. \quad (6.89)$$

Parameter values are as given in (Tomida et al. 2015), multiplied by a factor of 0.1175 to yield the experimentally observed p38 fundamental frequency of  $0.09 \text{ h}^{-1}$ :

$$k_2 = 1.06 \text{ h}^{-1},$$

$$k_3 = 1.13 \text{ h}^{-1},$$

$$k_4 = 7.05 \times 10^{-4},$$

$$k_5 = 0.38775 \text{ h}^{-1},$$

$$k_6 = 0.353 \text{ h}^{-1},$$

$$k_7 = 1.41 \text{ h}^{-1},$$

$$k_8 = 0.141 \text{ h}^{-1}.$$

Note that the first equation for the time evolution of [MAP2K] is decoupled from, and can be solved independently of, the others. Tomida et al. (Tomida et al. 2015) used  $S(t) = 1$  when p38 is activated by MAP2K and  $S(t) = 0$  otherwise. Instead, we treat [MAP2K] as a driving oscillator representing p38's interactions with oscillating components of the phosphorylation cascade in the equation for the time evolution of [MAPK] (p38), setting

$$\text{MAP2K}(t) = A + B \sin(2\pi ft),$$

with  $A = 0.2857$  and  $B$  a random number drawn from a uniform distribution on  $[0, 1]$ .

To generate the simulated FRET time traces and ensemble-averaged Fourier spectra shown in Figure 5.5C–D, we solve the above equations using a fourth-order Runge–Kutta method with  $f = 0.09 \text{ h}^{-1}$  Figure 5.5B,  $f = 0.36 \text{ h}^{-1}$  Figure 5.5C, or  $f = 0.27 \text{ h}^{-1}$  Figure 5.5D to yield solutions  $[\text{MAPK}_f(t)]$ . To simulate resonance selectivity, we model two

substrates, Substrate A and Substrate B, as driven two-component linear oscillators with repressors  $R_i$ , as described above in SM VI:

$$\begin{pmatrix} \dot{A} \\ \dot{R}_A \end{pmatrix} = \begin{pmatrix} -\alpha_{A1} & -\alpha_{A1} \\ \alpha_{A2} & -\beta_{A2} \end{pmatrix} \begin{pmatrix} A \\ R_A \end{pmatrix} + \begin{pmatrix} \zeta_1 \\ \zeta_2 \end{pmatrix} \text{MAPK}_f(t), \quad (6.90)$$

$$\begin{pmatrix} \dot{B} \\ \dot{R}_B \end{pmatrix} = \begin{pmatrix} -\alpha_{B1} & -\alpha_{B1} \\ \alpha_{B2} & -\beta_{B2} \end{pmatrix} \begin{pmatrix} B \\ R_B \end{pmatrix} + \begin{pmatrix} \zeta_1 \\ \zeta_2 \end{pmatrix} \text{MAPK}_f(t). \quad (6.91)$$

We set  $\zeta_2 = 0$  and use

$$\alpha_{A1} = 1.6 \text{ h}^{-1},$$

$$\alpha_{A2} = 0.8 \text{ h}^{-1},$$

$$\beta_{A1} = \beta_{A2} = 0.004 \text{ h}^{-1},$$

$$f_{A0} = \frac{\sqrt{\alpha_{A1}\alpha_{A2} + \beta_{A1}\beta_{A2}}}{2\pi} = 0.18 \text{ h}^{-1},$$

$$Q_A = \frac{\alpha_{A1}\alpha_{A2} + \beta_{A1}\beta_{A2}}{\alpha_{A1} + \alpha_{A2}} = 141,$$

and

$$\alpha_{B1} = 3.2 \text{ h}^{-1},$$

$$\alpha_{B2} = 1.6 \text{ h}^{-1},$$

$$\beta_{B1} = \beta_{B2} = 0.004 \text{ h}^{-1},$$

$$f_{B0} = \frac{\sqrt{\alpha_{B1}\alpha_{B2} + \beta_{B1}\beta_{B2}}}{2\pi} = 0.36 \text{ h}^{-1},$$

$$Q_B = \frac{\alpha_{B1}\alpha_{B2} + \beta_{B1}\beta_{B2}}{\alpha_{B1} + \alpha_{B2}} = 224.$$

### 6.2.10 X. Resonance and Impedance in Alternating Current (AC) Circuits

Series resistor–inductor–capacitor (RLC) electronic circuits, when driven by a sinusoidal voltage  $V(t)$  at some angular frequency  $\omega$ , generate an oscillating alternating current (AC),  $I(t) = \dot{Q}(t)$ , described by

$$\ddot{Q}(t) + \frac{R}{L}\dot{Q}(t) + \frac{1}{LC}Q(t) = \frac{1}{L}V(t). \quad (6.92)$$

where  $R$  is the resistance,  $L$  the inductance, and  $C$  the capacitance. With the identifications

$$x = Q(t), \quad (6.93)$$

$$\omega_0 = \frac{1}{\sqrt{LC}}, \quad (6.94)$$

$$\gamma = \frac{R}{L}, \quad (6.95)$$

$$F(t) = \frac{V(t)}{L} = \frac{V_0}{L} \cos(\omega t), \quad (6.96)$$

$$I(t) = \dot{Q}(t), \quad (6.97)$$

eq. S10.1 becomes

$$\ddot{x} + \gamma\dot{x} + \omega_0^2 x = F(t), \quad (6.98)$$

the general equation of a driven damped linear harmonic oscillator, 6.21.

The solution, 6.22, can be expressed in terms of the complex frequency-dependent impedance of the circuit,  $Z$ ,

$$Z = R + iX = e^{i\phi_Z} \sqrt{R^2 + X^2}, \quad (6.99)$$

where  $R$  is the resistance of the circuit and  $X(\omega)$  the reactance, which also depends on  $L$  and  $C$ . In the language of SM VII, the impedance is the inverse of the modified transfer function of the circuit,

$$Z = \tilde{\chi}^{-1}. \quad (6.100)$$

The current in the driven circuit is therefore given in terms of the driving voltage as in 6.22,

$$I(t) = \psi V(t) = \frac{V(t)}{Z} = \frac{V(t)}{\sqrt{R^2 + X^2}} e^{-i\phi_Z} = \frac{V(t) e^{-i\phi_Z}}{\sqrt{(\omega_0^2 - \omega^2)^2 + (\gamma\omega)^2}}, \quad (6.101)$$

i.e., the AC equivalent of Ohm's law, with  $\phi_Z$  the phase shift between the current and the driving voltage.

### 6.2.11 The MAPK Cascade as Oscillating Phosphate Current Amplifier and Distribution Network

The oscillating components of the MAPK cascade carry an oscillating current composed of phosphates being added and removed from components by enzymatic chemical reactions. Comparison of 6.92 to 6.45 leads to identifying the phosphate charges as  $x_j(t)$

and  $y(t)$ , with oscillating phosphate currents  $\dot{x}_j(t)$  and  $\dot{y}(t)$ . The  $\zeta_j(t)$  in 6.55 therefore represent driving chemical potentials induced by the upstream  $\dot{y}(t)$  currents to drive the  $\dot{x}_j(t)$  currents,

$$\zeta_j(t) \simeq j y(t). \quad (6.102)$$

which is recognized as Ohm's law,

$$V(t) = R \dot{Q}(t). \quad (6.103)$$

The net chemical potential is determined as the superposition of all input potentials (Figure 5.5).

Consequently, each component in the MAPK cascade is an oscillating phosphate current circuit element with complex impedance equal to the reciprocal of its transfer function,  $Z_j = \tilde{\chi}_j^{-1}$ . Each component is coupled to other components, such that the current in that element is

$$\dot{x}_j(t) = \kappa_j \sum_k \tilde{\chi}_j(\omega_k) i\omega_k \zeta_j y_k(t) = \frac{1}{Z_j} \sum_k i\omega_k \zeta_j y_k(t). \quad (6.104)$$

with the complex impedance  $Z_j(\omega)$  encapsulating the resonant response of the element,

$$Z_j(\omega) = e^{i\phi_{Z_j}} \sqrt{(\omega_{0,j}^2 - \omega^2)^2 + (\gamma_j \omega)^2}, \quad (6.105)$$

$$\phi_{Z_j} = \tan^{-1} \left( \frac{\gamma_j \omega}{\omega_{0,j}^2 - \omega^2} \right). \quad (6.106)$$

A component of the MAPK cascade interacting with a single upstream driver will carry oscillating phosphate current given by Ohm's law,

$$\dot{x}_j(t) = \tilde{\chi}_j y(t) = \frac{i\omega\zeta_j}{Z_j} y(t), \quad (6.107)$$

which corresponds to a current gain of

$$\text{Gain} = \frac{\dot{x}_j(t)}{y(t)} = \frac{\zeta_j}{Z_j}. \quad (6.108)$$

As we find experimentally  $\gamma \ll 1$  and  $Q \gg 1$ , the gain will only be significant near resonance ( $\omega \approx \omega_{0,j}$ ), yielding

$$\text{Gain}_j \simeq \frac{\zeta_j}{\gamma_j}. \quad (6.109)$$

Gain greater than one, corresponding to the component acting as a phosphate current amplifier, therefore requires  $\zeta_j > \gamma_j$ , that is, the  $v_{\text{max}}$  rate of the driving enzyme phosphorylating the targets must be greater than the sum of the spontaneous dephosphorylation rates of the targets. We indeed find this to be the case for the p38/PerKy-38 system (see SM VII above).

For multiple components in series, the driving current  $y(t)$  in 6.107 will itself have been generated in response to a further upstream component. Thus,  $N$  components in series generate a net gain

$$\text{Net gain} = \prod_{j=1}^N \frac{\zeta_j}{Z_j} \simeq \prod_{j=1}^N \frac{\zeta_j}{\gamma_j}, \quad (6.110)$$

which, with the assumption that other MAPK components behave approximately similarly to that we have here experimentally measured with p38/PerKy-38, can become quite large. That is, phosphate current paths through the cascade that are near resonance (i.e., low net impedance) result in large amplification of the phosphate current along that path. Conversely, with high  $Q$  values, the impedance of pathways composed of elements off-resonance for a given signal will experience high resistance/impedance and low amplification. This is illustrated in Figure 5.5 as the magnitude of the current along a resonant path increasing in magnitude (thickness) along the cascade.



## Chapter 7

# Conclusions

This dissertation has addressed fundamental questions in the biophysics of SARS-CoV-2 assembly through four interconnected studies spanning protein expression, membrane reconstitution, virus-like particle production, and cellular signal transduction. Taken together, the results provide a molecular framework for understanding how the structural proteins of SARS-CoV-2 remodel host membranes, cooperate to form new viral particles, and selectively package RNA cargo. This concluding chapter synthesizes the principal findings, identifies the cross-cutting themes that unify the work, and outlines the most important directions for future investigation.

## 7.1 Summary of Principal Findings

### 7.1.1 High-Yield Expression and Membrane Reconstitution of the M Protein

The membrane protein is the most abundant structural component of the SARS-CoV-2 virion and the central organizer of viral assembly, yet it had remained among the least characterized of the structural proteins due to the extreme difficulty of expressing it in sufficient quantities for biochemical and biophysical study ([Masters 2006](#); [Neuman & Buchmeier 2016](#)). Chapter 2 addressed this bottleneck directly by developing a SUMO-tag-based expression system in *E. coli* that produces 50 to 100 mg of full-length, native M protein per liter of culture — yields orders of magnitude higher than any previously reported method. The availability of large quantities of pure, low-endotoxin M protein enabled a comprehensive biophysical characterization that would not otherwise have been possible.

Insertion of purified M protein dimers into suspended lipid bilayers and imaging by AFM and cryo-EM revealed that the protein adopts exclusively the compact “short” conformation identified in prior structural studies ([Zhang et al. 2022](#); [Dolan et al. 2022](#)). The dimers are oriented perpendicular to the membrane plane with their C-terminal endodomains extending approximately 2.4 nm above the membrane surface, positioning them to interact with other structural proteins during assembly. Atomistic MD simulations confirmed this orientation and revealed a region of localized membrane thinning of approximately 0.5 nm extending over a radius of 12 nm around each dimer, arising from a hydrophobic mismatch between the transmembrane region of the protein and the surrounding

lipid bilayer. This thinning generates a line tension of approximately 0.35 pN that is sufficient to drive the lateral aggregation of M protein dimers even in the absence of specific inter-dimer protein–protein interactions.

Coarse-grained MD simulations incorporating M protein endodomain interactions, spontaneous curvature generation, and line tension demonstrated that budding and spherical particle formation occur only when both endodomain interactions and membrane curvature are present simultaneously. This result establishes that M protein is not merely a passive structural scaffold but an active driver of membrane deformation during viral budding, and it identifies endodomain-mediated inter-dimer interactions and membrane line tension as the two primary physical forces underlying this activity.

### **7.1.2 Extension of the Purification Platform to E Protein and Collaborative Assembly Studies**

The SUMO-tag-based bacterial expression system developed for M protein in Chapter 2 was directly extended to the SARS-CoV-2 envelope (E) protein, the smallest structural protein of the virion. E protein comprises only 75 amino acids yet plays a non-redundant role in particle release and virion morphology: genetic deletion of E results in strongly attenuated virus, and co-expression of M and E is the minimal requirement for VLP formation in cell-based assays (DeDiego et al. 2014; Alvarez et al. 2010; Vennema et al. 1996). The same expression challenges that had impeded M protein research applied equally to E protein, and the high-yield purification method developed here provided, for the first time, quantities of recombinant E protein suitable for membrane reconstitution and structural studies.

Membrane reconstitution experiments on this purified E protein, carried out in collaboration with members of our research group, revealed that E protein forms a heterogeneous population of oligomeric species — monomers, dimers, pentamers, and higher-order assemblies — when reconstituted into supported lipid bilayers composed of ERGIC-mimetic lipids. This finding provides direct structural evidence in a native-like membrane environment and resolves the long-standing controversy over the oligomeric state of this protein ([Pervushin et al. 2009](#); [Surya et al. 2018](#); [Cao et al. 2021](#)). E protein was also found to induce localized membrane thinning and curvature, with the extent of membrane disruption scaling with protein concentration and culminating in bilayer perforations at high density. Confocal microscopy revealed a concentration-dependent morphological transition from spherical multilamellar vesicles to elongated tubular structures in E protein-reconstituted giant unilamellar vesicles (GUVs); chemical modification of the C-terminal domain (CTD) abolished this transition, directly implicating CTD-mediated protein–protein interactions in curvature generation.

When M and E proteins were co-reconstituted into GUVs, tubule diameters increased substantially beyond those produced by either protein alone, and a necking morphology consistent with membrane scission emerged at higher E concentrations. Most strikingly, real-time confocal microscopy by collaborators demonstrated that introduction of pre-assembled N protein–RNA ribonucleoprotein complexes to proteo-GUVs containing both M and E triggered rapid inward vesicle budding, establishing that M, E, and N-RNA together constitute a minimal reconstituted system for SARS-CoV-2 particle assembly *in*

*vitro*. These results, made possible by the protein expression platform of Chapter 2, are being prepared for separate publication.

### 7.1.3 Optimized VLP Production and RNA Packaging

Chapter 4 translated the molecular insights from the reconstitution experiments into a cell-based assembly system by producing SARS-CoV-2 virus-like particles (VLPs) in HEK293T cells. Previous studies had reported conflicting conclusions regarding which combination of structural proteins is sufficient for VLP production and what expression stoichiometries are required (Swann et al. 2020; Plescia et al. 2021). This work resolved part of that ambiguity by demonstrating that expressing all four structural proteins — M, N, E, and S — at ratios reflecting viral mRNA abundance in productively infected cells (Kim et al. 2021; Gourdelier et al. 2022) yields VLPs that are validated by Western blotting, AFM, and dynamic light scattering (DLS). The optimized four-plasmid condition M:N:E:S at 3:12:2:5 produced particles with a size distribution consistent with native SARS-CoV-2 virions.

Analysis of RNA extracted from purified VLP fractions after RNase H treatment to remove surface-associated RNA revealed a size distribution dominated by small RNAs. This finding is consistent with selective RNA packaging rather than nonspecific incorporation of bulk cellular RNA, and it suggests that even in the absence of a complete viral genome, the assembly machinery retains a preference for small structured RNAs. These results provide a biochemical foundation for future sequencing-based characterization of VLP RNA cargo and raise the possibility that the RNA packaging specificity of coronaviruses may be partially

encoded in the intrinsic properties of the assembly machinery rather than requiring specific packaging signals.

#### 7.1.4 Bioresonance as a Mechanism of Signaling Specificity

Chapter 5 addressed a fundamental problem in cell biology: how p38 MAPK, a single kinase that exists in either an active phosphorylated state or an inactive dephosphorylated state, achieves the specificity needed to selectively activate the correct subset of downstream targets from among hundreds of possible substrates in response to diverse environmental stresses ([Canovas & Nebreda 2021](#)). To investigate this question, a novel FRET-based direct reporter of p38 activation state (VPC, Venus-p38-Cerulean) was developed and validated, enabling real-time single-cell measurement of p38 phosphorylation dynamics with temporal precision not achievable by existing indirect reporters ([Tomida et al. 2015](#)).

Single-cell FRET measurements revealed that p38 activation state oscillates in a stimulus-specific manner, with the frequency spectrum of these oscillations encoding information about the nature and intensity of the experienced stress. A mathematical framework based on the analogy between bioresonant signaling cascades and alternating current (AC) electronic circuits demonstrated that downstream specificity is achieved when the oscillation frequencies of p38 resonate with the natural frequencies of target substrates, selectively amplifying the response of those substrates whose dynamics are tuned to the stimulus-specific frequencies of p38. This bioresonance mechanism is sufficient to confer selectivity in gene expression responses without requiring allosteric changes to p38 structure or differences in subcellular localization. The quality factor of the p38–PerKy-38 bioresonant interac-

tion compares favorably to that of conventional electronic components, underscoring the quantitative precision of this signaling mechanism.

## 7.2 Cross-Cutting Themes

Several themes emerge from the body of work presented in this dissertation that extend beyond the individual chapters.

**Physical forces drive viral assembly.** A unifying insight from Chapter 2 and the collaborative E protein reconstitution work described above is that the physical properties of membrane proteins — their geometry, their hydrophobic mismatch with surrounding lipids, and the protein–protein interactions mediated by their cytoplasmic domains — are sufficient to drive the membrane deformations necessary for viral budding. M protein line tension, E protein curvature generation, and the synergistic enhancement of tubulation when both are present together constitute a physically coherent assembly pathway that does not require additional viral or host factors beyond the three core structural proteins and the viral ribonucleoprotein.

**Quantitative control of protein stoichiometry is a determinant of assembly efficiency.** The VLP experiments of Chapter 4 demonstrate that the relative amounts of structural proteins, not just their identity, govern assembly outcomes. Expressing M, N, E, and S at ratios calibrated to viral mRNA abundance substantially improves VLP yield and structural fidelity. This principle is further supported by the collaborative E protein reconstitution work, where the M-to-E ratio in reconstituted GUVs controls the diameter and morphology of tubular assemblies, underscoring that stoichiometric control

is a fundamental determinant of coronavirus particle formation independent of the specific experimental system used.

**Dynamics encode biological information.** The bioresonance work of Chapter 5 demonstrates that the temporal dynamics of a signaling molecule — specifically, the frequency spectrum of its activation state oscillations — can encode information that determines the specificity of downstream responses. This principle of dynamic information encoding may be relevant beyond p38 and the MAPK cascade; analogous mechanisms could operate in other contexts where a single molecular species must selectively activate multiple downstream pathways in response to distinct inputs.

## 7.3 Future Directions

The findings presented in this dissertation open several important avenues for future research.

The minimal reconstituted assembly system — in which M, E, and N-RNA together drive inward vesicle budding, as established through the collaborative experiments enabled by the protein expression platform of Chapter 2 — provides a platform for mechanistic dissection of the budding process. Future experiments incorporating purified S protein into this system, together with quantitative imaging of particle size, shape, and protein composition, would establish whether S protein modulates the efficiency or fidelity of assembly and would enable direct comparison between *in vitro* particles and the VLPs produced in Chapter 4.



The RNA packaging results of Chapter 4 establish the size distribution of VLP-associated RNA but leave open the question of sequence specificity. High-throughput RNA sequencing of VLP cargo under different expression conditions would determine whether the small RNA population is enriched for specific sequences or structural motifs, and would test whether the packaging selectivity observed here reflects a biologically conserved packaging signal or an emergent property of the assembly machinery.

The expression system developed for M protein in Chapter 2 could be extended to produce the other structural proteins at the quantities needed for detailed structural studies. In particular, obtaining sufficient quantities of pure M protein for small-angle X-ray scattering (SAXS) or single-particle cryo-EM in detergent-free lipid nanodiscs would enable atomic-resolution structural characterization of M protein in a native membrane environment, resolving remaining questions about the conformational variability of the “short” and “long” forms.

For the bioresonance work, the most pressing future direction is a molecular-level investigation of the mechanisms that set the oscillation frequencies of p38 and its substrates. Identifying the kinases, phosphatases, and feedback interactions that determine these frequencies would enable quantitative prediction of which disease-associated perturbations alter bioresonant specificity, and could open new therapeutic strategies for diseases associated with aberrant p38 signaling ([Canovas & Nebreda 2021](#); [Martínez-Limón et al. 2020](#)).

## 7.4 Closing Remarks

The work presented in this dissertation demonstrates that the physical principles governing membrane remodeling, protein–protein interactions, and molecular dynamics are quantitatively sufficient to account for key aspects of SARS-CoV-2 assembly. By developing new tools for high-yield membrane protein expression, establishing *in situ* AFM as a method for resolving oligomeric heterogeneity in reconstituted membranes, and building a minimal cell-free assembly system from purified components, this dissertation provides both molecular insights and experimental platforms that will accelerate future studies of coronavirus biology. Alongside these virological contributions, the discovery of bioresonance as a mechanism of kinase specificity illustrates a broader principle — that the temporal organization of molecular interactions, and not only their structural complementarity, is a fundamental determinant of biological specificity. Together, these findings reflect the power of quantitative biophysical approaches to reveal the physical logic underlying complex biological processes.

# Bibliography

- Adhikari, P., Li, N., Shin, M., et al. 2020, *Physical Chemistry Chemical Physics*, 22, 18272, doi: [10.1039/D0CP03145C](https://doi.org/10.1039/D0CP03145C)
- Allen, T. M., Romans, A. Y., Kercet, H., & Segrest, J. P. 1980, *Biochimica et Biophysica Acta (BBA) - Biomembranes*, 601, 328, doi: [10.1016/0005-2736\(80\)90537-4](https://doi.org/10.1016/0005-2736(80)90537-4)
- Alvarez, E., et al. 2010, *Virology*, 402, 281, doi: [10.1016/j.virol.2010.03.015](https://doi.org/10.1016/j.virol.2010.03.015)
- Anderson, J. A., Glaser, J., & Glotzer, S. C. 2020, *Computational Materials Science*, 173, 109363, doi: [10.1016/j.commatsci.2019.109363](https://doi.org/10.1016/j.commatsci.2019.109363)
- Armstrong, J., Niemann, H., Smeekens, S., Rottier, P., & Warren, G. 1984, *Nature*, 308, 751, doi: [10.1038/308751a0](https://doi.org/10.1038/308751a0)
- Ashwell, J. D. 2006, *Nature Reviews Immunology*, 6, 532, doi: [10.1038/nri1865](https://doi.org/10.1038/nri1865)
- Atzori, C., Ghetti, B., Piva, R., et al. 2001, *Journal of Neuropathology & Experimental Neurology*, 60, 1190, doi: [10.1093/jnen/60.12.1190](https://doi.org/10.1093/jnen/60.12.1190)
- Baden, L. R., et al. 2021, *New England Journal of Medicine*, 384, 403, doi: [10.1056/NEJMoa2035389](https://doi.org/10.1056/NEJMoa2035389)
- Baró, A. M., & Reifengerger, R. G. 2012, *Atomic Force Microscopy in Liquid: Biological Applications* (Wiley-VCH), doi: [10.1002/9783527649808](https://doi.org/10.1002/9783527649808)
- Bauer, P., Hess, B., & Lindahl, E. 2022, doi: [10.5281/ZENODO.7037337](https://doi.org/10.5281/ZENODO.7037337)
- Bertani, G. 1951, *Journal of Bacteriology*, 62, 293, doi: [10.1128/jb.62.3.293-300.1951](https://doi.org/10.1128/jb.62.3.293-300.1951)
- Bintu, L., Buchler, N. E., Garcia, H. G., et al. 2005a, *Current Opinion in Genetics & Development*, 15, 116, doi: [10.1016/j.gde.2005.02.005](https://doi.org/10.1016/j.gde.2005.02.005)
- . 2005b, *Current Opinion in Genetics & Development*, 15, 125, doi: [10.1016/j.gde.2005.02.006](https://doi.org/10.1016/j.gde.2005.02.006)
- Boson, B., et al. 2021, *Journal of Biological Chemistry*, 296, 100111, doi: [10.1016/j.jbc.2021.100111](https://doi.org/10.1016/j.jbc.2021.100111)
- Bracquemond, D., & Muriaux, D. 2021, *mBio*, 12, e02371, doi: [10.1128/mBio.02371-21](https://doi.org/10.1128/mBio.02371-21)

- Bratsun, D., Volfson, D., Tsimring, L. S., & Hasty, J. 2005, *Proceedings of the National Academy of Sciences*, 102, 14593, doi: [10.1073/pnas.0503858102](https://doi.org/10.1073/pnas.0503858102)
- Brooks, B. R., Brooks, C. L., Mackerell, A. D., et al. 2009, *Journal of Computational Chemistry*, 30, 1545, doi: [10.1002/jcc.21287](https://doi.org/10.1002/jcc.21287)
- Burton, J. C., Antoniadou, W., Okalova, J., Roos, M. M., & Grimsey, N. J. 2021, *International Journal of Molecular Sciences*, 22, 4183, doi: [10.3390/ijms22084183](https://doi.org/10.3390/ijms22084183)
- Canovas, B., & Nebreda, A. R. 2021, *Nature Reviews Molecular Cell Biology*, 22, 346, doi: [10.1038/s41580-020-00322-w](https://doi.org/10.1038/s41580-020-00322-w)
- Cao, Y., et al. 2021, *Nature Nanotechnology*, 16, 1384, doi: [10.1038/s41565-021-00989-2](https://doi.org/10.1038/s41565-021-00989-2)
- Casuso, I., et al. 2012, *Nature Nanotechnology*, 7, 525, doi: [10.1038/nnano.2012.109](https://doi.org/10.1038/nnano.2012.109)
- Chen, B.-K., Huang, C.-C., Chang, W.-C., et al. 2007, *Molecular Biology of the Cell*, 18, 1118, doi: [10.1091/mbc.e06-09-0797](https://doi.org/10.1091/mbc.e06-09-0797)
- Chen, S., Xu, J., Liu, M., et al. 2020, *PLOS ONE*, 15, e0228036, doi: [10.1371/journal.pone.0228036](https://doi.org/10.1371/journal.pone.0228036)
- Cheng, Y. 2015, *Cell*, 161, 450, doi: [10.1016/j.cell.2015.03.049](https://doi.org/10.1016/j.cell.2015.03.049)
- Cohen, J. R., Lin, L.-D., & Machamer, C. E. 2011, *Journal of Virology*, 85, 5794, doi: [10.1128/JVI.00060-11](https://doi.org/10.1128/JVI.00060-11)
- Colié, S., Sarroca, S., Palenzuela, R., et al. 2017, *Scientific Reports*, 7, 45306, doi: [10.1038/srep45306](https://doi.org/10.1038/srep45306)
- Corse, E., & Machamer, C. E. 2000, *Journal of Virology*, 74, 4319, doi: [10.1128/JVI.74.9.4319-4326.2000](https://doi.org/10.1128/JVI.74.9.4319-4326.2000)
- Cross, T. A., et al. 2011, *Trends in Biochemical Sciences*, 36, 117, doi: [10.1016/j.tibs.2010.07.005](https://doi.org/10.1016/j.tibs.2010.07.005)
- Cuenda, A., & Rousseau, S. 2007, *Biochimica et Biophysica Acta*, 1773, 1358, doi: [10.1016/j.bbamcr.2007.03.010](https://doi.org/10.1016/j.bbamcr.2007.03.010)
- Dai, L., & Gao, G. F. 2020, *Nature Reviews Immunology*, 21, 73, doi: [10.1038/s41577-020-00480-0](https://doi.org/10.1038/s41577-020-00480-0)
- Darlyuk-Saadon, I., Heng, C. K. M., Bai, C., et al. 2021, *The FEBS Journal*, 288, 3978, doi: [10.1111/febs.15697](https://doi.org/10.1111/febs.15697)
- DeDiego, M. L., et al. 2007, *Journal of Virology*, 81, 1701, doi: [10.1128/JVI.01467-06](https://doi.org/10.1128/JVI.01467-06)
- . 2014, *Journal of Virology*, 88, 913, doi: [10.1128/JVI.02576-13](https://doi.org/10.1128/JVI.02576-13)
- Dérijard, B., Raingeaud, J., Barrett, T., et al. 1995, *Science*, 267, 682, doi: [10.1126/science.7839144](https://doi.org/10.1126/science.7839144)

- Deserno, M. 2015, *Chemistry and Physics of Lipids*, 185, 11, doi: [10.1016/j.chemphyslip.2014.05.001](https://doi.org/10.1016/j.chemphyslip.2014.05.001)
- Doktorova, M., Harries, D., & Khelashvili, G. 2017, *Physical Chemistry Chemical Physics*, 19, 16806, doi: [10.1039/C7CP01921A](https://doi.org/10.1039/C7CP01921A)
- Dolan, K. A., Dutta, M., Kern, D. M., et al. 2022, *eLife*, 11, e81702, doi: [10.7554/eLife.81702](https://doi.org/10.7554/eLife.81702)
- Dong, E., Du, H., & Gardner, L. 2020, *The Lancet Infectious Diseases*, 20, 533, doi: [10.1016/S1473-3099\(20\)30120-1](https://doi.org/10.1016/S1473-3099(20)30120-1)
- Drabik, D., Gavutis, M., Valiokas, R., & Ulčinas, A. 2020, *Langmuir*, 36, 13251, doi: [10.1021/acs.langmuir.0c02181](https://doi.org/10.1021/acs.langmuir.0c02181)
- Dubendorf, J. W., & Studier, F. 1991, *Journal of Molecular Biology*, 219, 45, doi: [10.1016/0022-2836\(91\)90856-2](https://doi.org/10.1016/0022-2836(91)90856-2)
- Dykeman, E. C., Stockley, P. G., & Twarock, R. 2014, *Proceedings of the National Academy of Sciences*, 111, 5361, doi: [10.1073/pnas.1319479111](https://doi.org/10.1073/pnas.1319479111)
- García-Hernández, L., García-Ortega, M. B., Ruiz-Alcalá, G., et al. 2022, *International Journal of Molecular Sciences*, 23, doi: [10.3390/ijms23010370](https://doi.org/10.3390/ijms23010370)
- García-Sáez, A. J., Chiantia, S., & Schwill, P. 2007, *Journal of Biological Chemistry*, 282, 33537, doi: [10.1074/jbc.M706162200](https://doi.org/10.1074/jbc.M706162200)
- Garmann, R. F., Comas-Garcia, M., Gopal, A., Knobler, C. M., & Gelbart, W. M. 2014, *Journal of Molecular Biology*, 426, 1050, doi: [10.1016/j.jmb.2013.10.017](https://doi.org/10.1016/j.jmb.2013.10.017)
- Garmann, R. F., Goldfain, A. M., & Manoharan, V. N. 2019, *Proceedings of the National Academy of Sciences*, 116, 22485, doi: [10.1073/pnas.1909223116](https://doi.org/10.1073/pnas.1909223116)
- Gee, M. S., Son, S. H., Jeon, S.-H., et al. 2020, *Alzheimer's Research & Therapy*, 12, 45, doi: [10.1186/s13195-020-00617-2](https://doi.org/10.1186/s13195-020-00617-2)
- Gelbart, W. M., & Knobler, C. M. 2009, *Science*, 323, 1682, doi: [10.1126/science.1170645](https://doi.org/10.1126/science.1170645)
- Geva-Zatorsky, N., Dekel, E., Batchelor, E., Lahav, G., & Alon, U. 2010, *Proceedings of the National Academy of Sciences*, 107, 13550, doi: [10.1073/pnas.1001107107](https://doi.org/10.1073/pnas.1001107107)
- Gong, X., Ming, X., Deng, P., & Jiang, Y. 2010, *Journal of Cellular Biochemistry*, 110, 1420, doi: [10.1002/jcb.22675](https://doi.org/10.1002/jcb.22675)
- Gong, Y., Qin, S., Dai, L., & Tian, Z. 2021, *Signal Transduction and Targeted Therapy*, 6, 396, doi: [10.1038/s41392-021-00809-8](https://doi.org/10.1038/s41392-021-00809-8)
- Gourdelier, M., Swain, J., Arone, C., et al. 2022, *Scientific Reports*, 12, 14651, doi: [10.1038/s41598-022-18681-z](https://doi.org/10.1038/s41598-022-18681-z)

- Gowers, R. J., Linke, M., Barnoud, J., et al. 2016, in Proceedings of the 15th Python in Science Conference, ed. S. Benthall & S. Rostrup, Austin, Texas, 98–105, doi: [10.25080/Majora-629e541a-00e](https://doi.org/10.25080/Majora-629e541a-00e)
- Grodberg, J., & Dunn, J. J. 1988, *Journal of Bacteriology*, 170, 1245, doi: [10.1128/jb.170.3.1245-1253.1988](https://doi.org/10.1128/jb.170.3.1245-1253.1988)
- Gui, D., Gupta, S., Xu, J., et al. 2015, *Journal of Biological Physics*, 41, 135, doi: [10.1007/s10867-014-9370-z](https://doi.org/10.1007/s10867-014-9370-z)
- Gui, J., Zahedi, F., Ortiz, A., et al. 2020, *Nature Cancer*, 1, 603, doi: [10.1038/s43018-020-0064-0](https://doi.org/10.1038/s43018-020-0064-0)
- Guo, S., & Akhremitchev, B. B. 2006, *Biomacromolecules*, 7, 1630, doi: [10.1021/bm0600724](https://doi.org/10.1021/bm0600724)
- Hagan, M. F. 2014, *Advances in Chemical Physics: Volume 155*, 1
- Hale, K. K., Trollinger, D., Rihaneck, M., & Manthey, C. L. 1999, *Journal of Immunology*, 162, 4246. <https://www.jimmunol.org/content/162/7/4246>
- Han, J., Lee, J.-D., Jiang, Y., et al. 1996, *Journal of Biological Chemistry*, 271, 2886, doi: <https://doi.org/10.1074/jbc.271.6.2886>
- Hardenbrook, N. J., & Zhang, P. 2022, *Current Opinion in Virology*, 52, 123, doi: [10.1016/j.coviro.2022.01.007](https://doi.org/10.1016/j.coviro.2022.01.007)
- He, R., et al. 2004, *Biochemical and Biophysical Research Communications*, 316, 476, doi: [10.1016/j.bbrc.2004.02.074](https://doi.org/10.1016/j.bbrc.2004.02.074)
- Hensley, K., Floyd, R. A., Zheng, N.-Y., et al. 1999, *Journal of Neurochemistry*, 72, 2053, doi: [10.1046/j.1471-4159.1999.0722053.x](https://doi.org/10.1046/j.1471-4159.1999.0722053.x)
- Hoffmann, A., Levchenko, A., Scott, M. L., & Baltimore, D. 2002, *Science*, 298, 1241, doi: [10.1126/science.1071914](https://doi.org/10.1126/science.1071914)
- Holasek, S. S., Wengenack, T. M., Kandimalla, K. K., et al. 2005, *Brain Research*, 1045, 185, doi: <https://doi.org/10.1016/j.brainres.2005.03.037>
- Hoover, W. G. 1985, *Physical Review A*, 31, 1695, doi: [10.1103/PhysRevA.31.1695](https://doi.org/10.1103/PhysRevA.31.1695)
- Huang, J., Rauscher, S., Nawrocki, G., et al. 2017, *Nature Methods*, 14, 71, doi: [10.1038/nmeth.4067](https://doi.org/10.1038/nmeth.4067)
- Iserman, C., et al. 2020, *Molecular Cell*, 80, 1078, doi: [10.1016/j.molcel.2020.11.041](https://doi.org/10.1016/j.molcel.2020.11.041)
- Jiménez-Guardenõ, J. M., et al. 2014, *PLOS Pathogens*, 10, e1004320, doi: [10.1371/journal.ppat.1004320](https://doi.org/10.1371/journal.ppat.1004320)
- Jo, S., Kim, T., & Im, W. 2007, *PLoS ONE*, 2, e880, doi: [10.1371/journal.pone.0000880](https://doi.org/10.1371/journal.pone.0000880)

- Jo, S., Kim, T., Iyer, V. G., & Im, W. 2008, *Journal of Computational Chemistry*, 29, 1859, doi: [10.1002/jcc.20945](https://doi.org/10.1002/jcc.20945)
- Jo, S., Lim, J. B., Klauda, J. B., & Im, W. 2009, *Biophysical Journal*, 97, 50, doi: [10.1016/j.bpj.2009.04.013](https://doi.org/10.1016/j.bpj.2009.04.013)
- K.E. Medders, M. K. 2011, *J Neuroimmune Pharmacol* 6, 202-215
- Khan, H. M., MacKerell, A. D., & Reuter, N. 2019, *Journal of Chemical Theory and Computation*, 15, 7, doi: [10.1021/acs.jctc.8b00839](https://doi.org/10.1021/acs.jctc.8b00839)
- Kheiri, G., Dolatshahi, M., Rahmani, F., & Rezaei, N. 2019, *Reviews in the Neurosciences*, 30, 9, doi: [doi:10.1515/revneuro-2018-0008](https://doi.org/10.1515/revneuro-2018-0008)
- Khoo, S. H., et al. 2021, *Nature Medicine*, 27, 1404, doi: [10.1038/s41591-021-01445-3](https://doi.org/10.1038/s41591-021-01445-3)
- Kim, D., et al. 2021, *Nature Communications*, 12, 5120, doi: [10.1038/s41467-021-25387-5](https://doi.org/10.1038/s41467-021-25387-5)
- Klumperman, J., et al. 1994, *Journal of Virology*, 68, 6523
- Kolberg, L., Raudvere, U., Kuzmin, I., Vilo, J., & Peterson, H. 2023, *Nucleic Acids Research*, 51, W207, doi: [10.1093/nar/gkad347](https://doi.org/10.1093/nar/gkad347)
- Krammer, F. 2020, *Nature*, 586, 516, doi: [10.1038/s41586-020-2798-3](https://doi.org/10.1038/s41586-020-2798-3)
- Kuo, L., & Masters, P. S. 2002, *Journal of Virology*, 76, 4987, doi: [10.1128/JVI.76.10.4987-4999.2002](https://doi.org/10.1128/JVI.76.10.4987-4999.2002)
- Kuzmin, P. I., Akimov, S. A., Chizmadzhev, Y. A., Zimmerberg, J., & Cohen, F. S. 2005, *Biophysical Journal*, 88, 1120, doi: [10.1529/biophysj.104.048223](https://doi.org/10.1529/biophysj.104.048223)
- Lahav, G., Rosenfeld, N., Sigal, A., et al. 2004, *Nature Genetics*, 36, 147, doi: [10.1038/ng1293](https://doi.org/10.1038/ng1293)
- Laude, H., Gelfi, J., Lavenant, L., & Charley, B. 1992, *Journal of Virology*, 66, 743, doi: [10.1128/jvi.66.2.743-749.1992](https://doi.org/10.1128/jvi.66.2.743-749.1992)
- Lawson, H. S., Holló, G., Horvath, R., Kitahata, H., & Lagzi, I. 2020, *The Journal of Physical Chemistry Letters*, 11, 3014, doi: [10.1021/acs.jpcllett.0c00586](https://doi.org/10.1021/acs.jpcllett.0c00586)
- Lee, J., Hitzenberger, M., Rieger, M., et al. 2020, *The Journal of Chemical Physics*, 153, 035103, doi: [10.1063/5.0012280](https://doi.org/10.1063/5.0012280)
- Lee, J., Cheng, X., Swails, J. M., et al. 2016, *Journal of Chemical Theory and Computation*, 12, 405, doi: [10.1021/acs.jctc.5b00935](https://doi.org/10.1021/acs.jctc.5b00935)
- Lee, J., Patel, D. S., Stähle, J., et al. 2019, *Journal of Chemical Theory and Computation*, 15, 775, doi: [10.1021/acs.jctc.8b01066](https://doi.org/10.1021/acs.jctc.8b01066)
- Lee, J. K., & Kim, N.-J. 2017, *Molecules*, 22, doi: [10.3390/molecules22081287](https://doi.org/10.3390/molecules22081287)

- Levine, J. H., Lin, Y., & Elowitz, M. B. 2013, *Science*, 342, 1193, doi: [10.1126/science.1239999](https://doi.org/10.1126/science.1239999)
- Li, S., Erdemci-Tandogan, G., Wagner, J., Van Der Schoot, P., & Zandi, R. 2017, *Physical Review E*, 96, 022401, doi: [10.1103/PhysRevE.96.022401](https://doi.org/10.1103/PhysRevE.96.022401)
- Li, S., & Zandi, R. 2022, *Viruses*, 14, 2089, doi: [10.3390/v14102089](https://doi.org/10.3390/v14102089)
- Lindahl, E., Hess, B., & van der Spoel, D. 2001, *Journal of Molecular Modeling*, 7, 306, doi: [10.1007/s008940100045](https://doi.org/10.1007/s008940100045)
- Liu, H., Fu, H., Chipot, C., Shao, X., & Cai, W. 2021, *Journal of Chemical Theory and Computation*, 17, 3908, doi: [10.1021/acs.jctc.1c00219](https://doi.org/10.1021/acs.jctc.1c00219)
- Lloret, A., Fuchsberger, T., Giraldo, E., & Viña, J. 2015, *Free Radical Biology and Medicine*, 83, 186, doi: [10.1016/j.freeradbiomed.2015.02.028](https://doi.org/10.1016/j.freeradbiomed.2015.02.028)
- Love, M. I., Huber, W., & Anders, S. 2014, *Genome Biology*, 15, 550, doi: [10.1186/s13059-014-0550-8](https://doi.org/10.1186/s13059-014-0550-8)
- Low, K. B. 1974, *The Quarterly Review of Biology*, 49, 151, doi: [10.1086/408025](https://doi.org/10.1086/408025)
- Lu, S., et al. 2021, *Nature Communications*, 12
- Lua, L. H. L., et al. 2014, *Biotechnology and Bioengineering*, 111, 425, doi: [10.1002/bit.25159](https://doi.org/10.1002/bit.25159)
- Ludwig, C., & Wagner, R. 2007, *Current Opinion in Biotechnology*, 18, 537, doi: [10.1016/j.copbio.2007.10.013](https://doi.org/10.1016/j.copbio.2007.10.013)
- Lyonnais, S., et al. 2021, *Scientific Reports*, 11, 11885, doi: [10.1038/s41598-021-91287-8](https://doi.org/10.1038/s41598-021-91287-8)
- Ma Y, N. J. 2023, *Epub* 2023
- Mahmood, T., & Yang, P.-C. 2012, *North American Journal of Medical Sciences*, 4, 429, doi: [10.4103/1947-2714.100998](https://doi.org/10.4103/1947-2714.100998)
- Mahtarin, R., Islam, S., Islam, M. J., et al. 2022, *Journal of Biomolecular Structure and Dynamics*, 40, 4725, doi: [10.1080/07391102.2020.1861983](https://doi.org/10.1080/07391102.2020.1861983)
- Markwardt, M. L., Kremers, G.-J., Kraft, C. A., et al. 2011, *PLoS ONE*, 6, e17896, doi: [10.1371/journal.pone.0017896](https://doi.org/10.1371/journal.pone.0017896)
- Marrink, S. J., & Tieleman, D. P. 2013, *Chemical Society Reviews*, 42, 6801, doi: [10.1039/c3cs60093a](https://doi.org/10.1039/c3cs60093a)
- Martin, V., Zhuang, F., Zhang, Y., Pinheiro, K., & Gordân, R. 2023, *Nucleic Acids Research*, 51, 11600, doi: [10.1093/nar/gkad947](https://doi.org/10.1093/nar/gkad947)
- Martínez-Limón, A., Joaquin, M., Caballero, M., Posas, F., & De Nadal, E. 2020, *International Journal of Molecular Sciences*, 21, 1913, doi: [10.3390/ijms21061913](https://doi.org/10.3390/ijms21061913)



- Masters, P. S. 2006, *Advances in Virus Research*, 66, 193, doi: [10.1016/S0065-3527\(06\)66005-3](https://doi.org/10.1016/S0065-3527(06)66005-3)
- McBride, R., van Zyl, M., & Fielding, B. C. 2014, *Viruses*, 6, 2991, doi: [10.3390/v6082991](https://doi.org/10.3390/v6082991)
- Medders, K. E., Sejbuk, N. E., Maung, R., Desai, M. K., & Kaul, M. 2010, *The Journal of Immunology*, 185, 4883, doi: [10.4049/jimmunol.0902535](https://doi.org/10.4049/jimmunol.0902535)
- Michaud-Agrawal, N., Denning, E. J., Woolf, T. B., & Beckstein, O. 2011, *Journal of Computational Chemistry*, 32, 2319, doi: [10.1002/jcc.21787](https://doi.org/10.1002/jcc.21787)
- Miroux, B., & Walker, J. E. 1996, *Journal of Molecular Biology*, 260, 289, doi: [10.1006/jmbi.1996.0399](https://doi.org/10.1006/jmbi.1996.0399)
- Müller, D. J., & Dufrêne, Y. F. 2008, *Nature Nanotechnology*, 3, 261, doi: [10.1038/nnano.2008.100](https://doi.org/10.1038/nnano.2008.100)
- Munoz, L., & Ammit, A. J. 2010, *Neuropharmacology*, 58, 561, doi: <https://doi.org/10.1016/j.neuropharm.2009.11.010>
- Nagai, T., Ibata, K., Park, E. S., et al. 2002, *Nature Biotechnology*, 20, 87, doi: [10.1038/nbt0102-87](https://doi.org/10.1038/nbt0102-87)
- Nal, B., et al. 2005, *Journal of General Virology*, 86, 1423, doi: [10.1099/vir.0.80671-0](https://doi.org/10.1099/vir.0.80671-0)
- Narayanan, K., et al. 2000, *Journal of Virology*, 74, 8127, doi: [10.1128/JVI.74.17.8127-8134.2000](https://doi.org/10.1128/JVI.74.17.8127-8134.2000)
- Neuman, B., & Buchmeier, M. 2016, in *Advances in Virus Research*, Vol. 96 (Elsevier), 1–27, doi: [10.1016/bs.aivir.2016.08.005](https://doi.org/10.1016/bs.aivir.2016.08.005)
- Neuman, B. W., et al. 2011a, *Journal of Structural Biology*, 174, 11, doi: [10.1016/j.jsb.2010.11.021](https://doi.org/10.1016/j.jsb.2010.11.021)
- Neuman, B. W., Kiss, G., Kunding, A. H., et al. 2011b, *Journal of Structural Biology*, 174, 11, doi: [10.1016/j.jsb.2010.11.021](https://doi.org/10.1016/j.jsb.2010.11.021)
- Noad, R., & Roy, P. 2003, *Trends in Microbiology*, 11, 438, doi: [10.1016/S0966-842X\(03\)00208-7](https://doi.org/10.1016/S0966-842X(03)00208-7)
- Nogales, E. 2016, *Nature Methods*, 13, 24, doi: [10.1038/nmeth.3694](https://doi.org/10.1038/nmeth.3694)
- Nosé, S. 1984, *Molecular Physics*, 52, 255, doi: [10.1080/00268978400101201](https://doi.org/10.1080/00268978400101201)
- Nosé, S., & Klein, M. 1983, *Molecular Physics*, 50, 1055, doi: [10.1080/00268978300102851](https://doi.org/10.1080/00268978300102851)
- Oostra, M., De Haan, C. A. M., De Groot, R. J., & Rottier, P. J. M. 2006, *Journal of Virology*, 80, 2326, doi: [10.1128/JVI.80.5.2326-2336.2006](https://doi.org/10.1128/JVI.80.5.2326-2336.2006)
- Panahandeh, S., Li, S., Dragnea, B., & Zandi, R. 2022, *ACS Nano*, 16, 317, doi: [10.1021/acsnano.1c06335](https://doi.org/10.1021/acsnano.1c06335)

- Park, S., Choi, Y. K., Kim, S., Lee, J., & Im, W. 2021, *Journal of Chemical Information and Modeling*, 61, 5192, doi: [10.1021/acs.jcim.1c00770](https://doi.org/10.1021/acs.jcim.1c00770)
- Parrinello, M., & Rahman, A. 1981, *Journal of Applied Physics*, 52, 7182, doi: [10.1063/1.328693](https://doi.org/10.1063/1.328693)
- Paternostre, M. T., Roux, M., & Rigaud, J. L. 1988, *Biochemistry*, 27, 2668, doi: [10.1021/bi00408a006](https://doi.org/10.1021/bi00408a006)
- Perdikari, T. M., et al. 2020, *EMBO Journal*, 39, e106478, doi: [10.15252/embj.2020106478](https://doi.org/10.15252/embj.2020106478)
- Perlmutter, J. D., Qiao, C., & Hagan, M. F. 2013, *eLife*, 2, e00632, doi: [10.7554/eLife.00632](https://doi.org/10.7554/eLife.00632)
- Pervushin, K., et al. 2009, *PLOS Pathogens*, 5, e1000511, doi: [10.1371/journal.ppat.1000511](https://doi.org/10.1371/journal.ppat.1000511)
- Peterson, A. F., Ingram, K., Huang, E., et al. 2022, *Cell Systems*, 13, 885, doi: <https://doi.org/10.1016/j.cels.2022.10.003>
- Pettersen, E. F., Goddard, T. D., Huang, C. C., et al. 2004, *Journal of Computational Chemistry*, 25, 1605, doi: [10.1002/jcc.20084](https://doi.org/10.1002/jcc.20084)
- Plescia, C. B., et al. 2021, *Journal of Biological Chemistry*, 296, 100103, doi: [10.1016/j.jbc.2021.100103](https://doi.org/10.1016/j.jbc.2021.100103)
- Plotnikov, A., Zehorai, E., Procaccia, S., & Seger, R. 2011, *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research*, 1813, 1619, doi: <https://doi.org/10.1016/j.bbamcr.2010.12.012>
- Polack, F. P., et al. 2020, *New England Journal of Medicine*, 383, 2603, doi: [10.1056/NEJMoa2034577](https://doi.org/10.1056/NEJMoa2034577)
- Purvis, J. E., & Lahav, G. 2013, *Cell*, 152, 945, doi: [10.1016/j.cell.2013.02.005](https://doi.org/10.1016/j.cell.2013.02.005)
- Raingeaud, J., Whitmarsh, A. J., Barrett, T., Dérijard, B., & Davis, R. J. 1996, *Molecular and Cellular Biology*, 16, 1247, doi: [10.1128/MCB.16.3.1247](https://doi.org/10.1128/MCB.16.3.1247)
- Rigaud, J.-L., & Lévy, D. 2003, in *Methods in Enzymology*, Vol. 372 (Elsevier), 65–86, doi: [10.1016/S0076-6879\(03\)72004-7](https://doi.org/10.1016/S0076-6879(03)72004-7)
- Rigaud, J. L., Paternostre, M. T., & Bluzat, A. 1988, *Biochemistry*, 27, 2677, doi: [10.1021/bi00408a007](https://doi.org/10.1021/bi00408a007)
- Roden, C. A., et al. 2022, *Nucleic Acids Research*, 50
- Rohm, S., Schröder, M., Dwyer, J. E., et al. 2020, *European Journal of Medicinal Chemistry*, 208, 112721, doi: [10.1016/j.ejmech.2020.112721](https://doi.org/10.1016/j.ejmech.2020.112721)
- Rottier, P., et al. 1984, *Journal of Virology*, 50, 222

- Savastano, A., et al. 2020, Nature Communications, 11, 6041, doi: [10.1038/s41467-020-19843-1](https://doi.org/10.1038/s41467-020-19843-1)
- Schneider, C. A., Rasband, W. S., & Eliceiri, K. W. 2012, Nature Methods, 9, 671, doi: [10.1038/nmeth.2089](https://doi.org/10.1038/nmeth.2089)
- Schoeman, D., & Fielding, B. C. 2019, Virology Journal, 16, 69, doi: [10.1186/s12985-019-1182-0](https://doi.org/10.1186/s12985-019-1182-0)
- Shang, J., et al. 2020a, Proceedings of the National Academy of Sciences, 117, 11727, doi: [10.1073/pnas.2003138117](https://doi.org/10.1073/pnas.2003138117)
- Shang, J., Ye, G., Shi, K., et al. 2020b, Nature, 581, 221, doi: [10.1038/s41586-020-2179-y](https://doi.org/10.1038/s41586-020-2179-y)
- Shankaran, H., Ippolito, D. L., Chrisler, W. B., et al. 2009, Molecular Systems Biology, 5, 332, doi: [10.1038/msb.2009.90](https://doi.org/10.1038/msb.2009.90)
- Sheng, J. G., Jones, R. A., Zhou, X. Q., et al. 2001, Neurochemistry International, 39, 341, doi: [https://doi.org/10.1016/S0197-0186\(01\)00041-9](https://doi.org/10.1016/S0197-0186(01)00041-9)
- Siu, Y. L., et al. 2008, Journal of Virology, 82, 11318, doi: [10.1128/JVI.01052-08](https://doi.org/10.1128/JVI.01052-08)
- Sorkin, L., Boyle, D., Hammaker, D., et al. 2009, Neuroscience, 162, 462, doi: <https://doi.org/10.1016/j.neuroscience.2009.05.008>
- Stertz, S., et al. 2007, Virology, 361, 304, doi: [10.1016/j.virol.2006.11.027](https://doi.org/10.1016/j.virol.2006.11.027)
- Stetefeld, J., McKenna, S. A., & Patel, T. R. 2016, Biophysical Reviews, 8, 409, doi: [10.1007/s12551-016-0218-6](https://doi.org/10.1007/s12551-016-0218-6)
- Stirling, D. R., Swain-Bowden, M. J., Lucas, A. M., et al. 2021, BMC Bioinformatics, 22, 433, doi: [10.1186/s12859-021-04344-9](https://doi.org/10.1186/s12859-021-04344-9)
- Studier, F. 1991, Journal of Molecular Biology, 219, 37, doi: [10.1016/0022-2836\(91\)90855-Z](https://doi.org/10.1016/0022-2836(91)90855-Z)
- Studier, F., & Moffatt, B. A. 1986, Journal of Molecular Biology, 189, 113, doi: [10.1016/0022-2836\(86\)90385-2](https://doi.org/10.1016/0022-2836(86)90385-2)
- Stukowski, A. 2010, Modelling and Simulation in Materials Science and Engineering, 18, 015012, doi: [10.1088/0965-0393/18/1/015012](https://doi.org/10.1088/0965-0393/18/1/015012)
- Sun, A., Liu, M., Nguyen, X. V., & Bing, G. 2003, Experimental Neurology, 183, 394, doi: [https://doi.org/10.1016/S0014-4886\(03\)00180-8](https://doi.org/10.1016/S0014-4886(03)00180-8)
- Sun, L., Yang, G., Zaidi, M., & Iqbal, J. 2008a, Biochemical and Biophysical Research Communications, 371, 900, doi: [10.1016/j.bbrc.2008.03.114](https://doi.org/10.1016/j.bbrc.2008.03.114)
- . 2008b, Biochemical and Biophysical Research Communications, 371, 912, doi: [10.1016/j.bbrc.2008.03.112](https://doi.org/10.1016/j.bbrc.2008.03.112)

- Sung, M.-H., Salvatore, L., De Lorenzi, R., et al. 2009, PLoS ONE, 4, e7163, doi: [10.1371/journal.pone.0007163](https://doi.org/10.1371/journal.pone.0007163)
- Surya, W., et al. 2018, Biochimica et Biophysica Acta – Biomembranes, 1860, 1309, doi: [10.1016/j.bbamem.2018.02.017](https://doi.org/10.1016/j.bbamem.2018.02.017)
- Swain, J., et al. 2022, bioRxiv, doi: [10.1101/2022.03.08.483451](https://doi.org/10.1101/2022.03.08.483451)
- Swann, H., et al. 2020, Scientific Reports, 10, 21877, doi: [10.1038/s41598-020-78762-5](https://doi.org/10.1038/s41598-020-78762-5)
- Tartoff, K. D., & Hobbs, C. A. 1987, Bethesda Research Laboratories Focus, 9
- Thomas, S. 2020, Pathogens and Immunity, 5, 342, doi: [10.20411/pai.v5i1.377](https://doi.org/10.20411/pai.v5i1.377)
- Tomida, T., Takekawa, M., & Saito, H. 2015, Nature Communications, 6, 8350, doi: [10.1038/ncomms9350](https://doi.org/10.1038/ncomms9350)
- Tortarolo, M., Veglianesi, P., Calvaresi, N., et al. 2003, Molecular and Cellular Neuroscience, 23, 180, doi: [https://doi.org/10.1016/S1044-7431\(03\)00022-8](https://doi.org/10.1016/S1044-7431(03)00022-8)
- Trepolec, N., Dave-Coll, N., & Nebreda, A. R. 2013, Cell, 152, 924, doi: <https://doi.org/10.1016/j.cell.2013.01.047>
- Tresset, G., Chen, J., Chevreuil, M., et al. 2017, Physical Review Applied, 7, 014005, doi: [10.1103/PhysRevApplied.7.014005](https://doi.org/10.1103/PhysRevApplied.7.014005)
- Tseng, Y.-T., et al. 2010, Journal of Biomedical Science, 17, 88, doi: [10.1186/1423-0127-17-88](https://doi.org/10.1186/1423-0127-17-88)
- Uhlik, M., Abell, A., Johnson, N., et al. 2003, Nature Cell Biology, 5, 1104, doi: [10.1038/ncb1068](https://doi.org/10.1038/ncb1068)
- Ujike, M., & Taguchi, F. 2015, Viruses, 7, 1700, doi: [10.3390/v7041700](https://doi.org/10.3390/v7041700)
- Usery, R. D., Enoki, T. A., Wickramasinghe, S. P., et al. 2017, Biophysical Journal, 112, 1431, doi: [10.1016/j.bpj.2017.02.033](https://doi.org/10.1016/j.bpj.2017.02.033)
- Vennema, H., et al. 1996, EMBO Journal, 15, 2020, doi: [10.1002/j.1460-2075.1996.tb00553.x](https://doi.org/10.1002/j.1460-2075.1996.tb00553.x)
- Viorel, V. I., Pastorello, Y., Bajwa, N., & Slevin, M. 2024, Neural Regeneration Research, 19, 1649, doi: [10.4103/1673-5374.389645](https://doi.org/10.4103/1673-5374.389645)
- Walls, A. C., et al. 2020, Cell, 181, 281, doi: [10.1016/j.cell.2020.02.058](https://doi.org/10.1016/j.cell.2020.02.058)
- Wang, Q., et al. 2020, Cell, 181, 894, doi: [10.1016/j.cell.2020.03.045](https://doi.org/10.1016/j.cell.2020.03.045)
- Waters, K. M., Cummings, B. S., Shankaran, H., Scholpa, N. E., & Weber, T. J. 2014, Chemical Research in Toxicology, 27, 1496, doi: [10.1021/tx500085u](https://doi.org/10.1021/tx500085u)
- Werner, S. L., Barken, D., & Hoffmann, A. 2005, Science, 309, 1857, doi: [10.1126/science.1113319](https://doi.org/10.1126/science.1113319)

- White, S. H., et al. 2001, Journal of Biological Chemistry, 276, 32395, doi: [10.1074/jbc.R100008200](https://doi.org/10.1074/jbc.R100008200)
- Wilkins, M. R., Gasteiger, E., Bairoch, A., et al. 1999, Springer
- World Health Organization. 2023, World Health Organization. <https://covid19.who.int>
- Wrapp, D., et al. 2020, Science, 367, 1260, doi: [10.1126/science.abb2507](https://doi.org/10.1126/science.abb2507)
- Wu, E. L., Cheng, X., Jo, S., et al. 2014, Journal of Computational Chemistry, 35, 1997, doi: [10.1002/jcc.23702](https://doi.org/10.1002/jcc.23702)
- Wu, F., et al. 2020, Nature, 579, 265, doi: [10.1038/s41586-020-2008-3](https://doi.org/10.1038/s41586-020-2008-3)
- Wu, H., Rothmel, B., Kanatous, S., et al. 2001, The EMBO Journal, 20, 6414, doi: [10.1093/emboj/20.22.6414](https://doi.org/10.1093/emboj/20.22.6414)
- Yao, H., et al. 2020, Cell, 183, 730, doi: [10.1016/j.cell.2020.09.018](https://doi.org/10.1016/j.cell.2020.09.018)
- Young, P. R., McLaughlin, M. M., Kumar, S., et al. 1997, Journal of Biological Chemistry, 272, 12116, doi: <https://doi.org/10.1074/jbc.272.18.12116>
- Yurkovetskiy, L., Wang, X., Pascal, K. E., et al. 2020, Cell, 183, 739, doi: [10.1016/j.cell.2020.09.032](https://doi.org/10.1016/j.cell.2020.09.032)
- Zandi, R., Dragnea, B., Travesset, A., & Podgornik, R. 2020, Physics Reports, 847, 1, doi: [10.1016/j.physrep.2019.12.005](https://doi.org/10.1016/j.physrep.2019.12.005)
- Zarubin, T., & Han, J. 2005, Cell Research, 15, 11, doi: [10.1038/sj.cr.7290257](https://doi.org/10.1038/sj.cr.7290257)
- Zeltins, A. 2013, Molecular Biotechnology, 53, 92, doi: [10.1007/s12033-012-9598-4](https://doi.org/10.1007/s12033-012-9598-4)
- Zhang, Y., Liu, H., Yan, F., & Zhou, J. 2017, Scientific Reports, 7, 11495, doi: [10.1038/s41598-017-11891-w](https://doi.org/10.1038/s41598-017-11891-w)
- Zhang, Y., et al. 2024, Science Advances, 10
- Zhang, Z., Nomura, N., Muramoto, Y., et al. 2022, Nature Communications, 13, 4399, doi: [10.1038/s41467-022-32019-3](https://doi.org/10.1038/s41467-022-32019-3)
- Zhou, P., et al. 2020, Nature, 579, 270, doi: [10.1038/s41586-020-2012-7](https://doi.org/10.1038/s41586-020-2012-7)
- Zhu, X., Rotkamp, C. A., Boux, H., et al. 2000, Journal of Neuropathology & Experimental Neurology, 59, 880, doi: [10.1093/jnen/59.10.880](https://doi.org/10.1093/jnen/59.10.880)
- Zuo, X., Li, S., Hall, J., et al. 2005, Journal of Structural and Functional Genomics, 6, 103, doi: [10.1007/s10969-005-2664-4](https://doi.org/10.1007/s10969-005-2664-4)