Title: Minimal system for assembly of SARS-CoV-2 virus like particles

Authors:

Heather Swann[1,2]*, Abhimanyu Sharma[2]*, Benjamin Preece[1,2], Abby Peterson[1,2], Crystal Eldridge[1,2], David M. Belnap[3,4], Michael Vershinin[1,2,3,†] and Saveez Saffarian[1,2,3,†]

Affiliations:

[1] Center for Cell and Genome Science, University of Utah

[2] Department of Physics and Astronomy, University of Utah

[3] School of Biological Sciences, University of Utah

[4] Department of Biochemistry, University of Utah

[†] Corresponding authors: Vershinin@Physics.utah.edu and Saffarian@Physics.Utah.edu

*These authors contributed equally to this work

Abstract

SARS-CoV-2 virus is the causative agent of COVID-19. Here we demonstrate that non-infectious SARS-CoV-2 virus like particles (VLPs) can be assembled by co-expressing the viral proteins S, M and E in mammalian cells. The assembled SARS-CoV-2 VLPs display numerous S protein spikes ideal for vaccine development. The particles have a spike to spike size of 103 ± 6 nm and a membrane diameter of 63 ± 5 nm. We further show that SARS-CoV-2 VLPs dried in ambient conditions can retain their structural integrity upon repeated scans with Atomic Force Microscopy up to a peak force of 1 nN.

Main

COVID-19 is a pandemic disease caused by infection of SARS-CoV-2 virus [1]. With more than 4 million cases confirmed and a death toll exceeding several hundred thousand individuals, a search for antiviral therapies as well as vaccine candidates is of utmost urgency. Non-infectious virus like particles (VLPs) displaying essential viral proteins can be used to study the structural properties of the SARS-CoV-2 virions and due to their maximum immunogenicity are also vaccine candidates [2, 3]. VLPs are released from cells with similar mechanisms as fully infectious virions and resemble the shape and composition of fully infectious virions [4]. Most Coronaviruses are pathogens of zoonotic nature with different viruses infecting avian (IBV), bovine (BCoV), porcine (TGEV), feline (FIPV) and murine(MHV) [4] species. There is evidence that Bat-SARS-CoV is the origin of the SARS-CoV virus which first appeared in human hosts in 2003[5]. The genome of SARS-CoV-2 has ~90% similarity to genome of coronaviruses previously identified in bat populations in China [1, 6].

Expression of E and M proteins of TGEV result in release of VLPs with similar sizes to wild type TGEV [7]. Similar results are reported for IBV [8]. Similarly, expression of M, E and S proteins are shown to result in release of morphologically identical particles to wild type SARS-CoV virus [9, 10]. It is known that TGEV VLPs can induce an IFNα response in their hosts [7]. More recently human MERS-CoV VLPs were also shown to elicit an immune response and serve as vaccine candidates [11].

Given the similarities between SARS-CoV-2 and SARS-CoV viruses [1, 12], we set out to create SARS-CoV-2 VLPs by expressing S, M and E proteins of SARS-CoV-2 in mammalian cells. We then tested the structural integrity of the SARS-CoV-2 VLPs attached to dry glass using AFM, since SARS-CoV-2 virions have been reported to survive on solid surfaces in dry conditions for many hours [13].

Methods

SARS-CoV-2 M, S and E protein genes were identified from the full genome sequence of the virus [1] , these genes were then humanized and inserted in CMV driven mammalian expression vectors (see supplement for further details). 24 hours after transfecting a monolayer of 293 cells with a cocktail of S, M and E plasmids, VLPs were harvested from the supernatant (see supplement for further details). The supernatant was filtered using 0.2um filter and VLPs were captured in a 40-10% sucrose step gradient and concentrated using AMICON spin filters (UFC901008, EMD Millipore, Burlington, MA). Figure 1 shows western blot analysis of cell extracts as well as VLPs, the isolated VLPs have both M and S proteins as identified in western blots. The isolated VLPs were further analyzed by running a 10-40% sucrose gradient and sequential fractionation. VLPs were found to be within the density of 20-25% sucrose (Supplement Figure 1).

Negative stain electron microscopy was performed on purified VLPs by applying 3.5 uL of VLPs to a glow-discharged carbon-coated EM grid followed by two de-ionized water washes and staining with 1% uranyl acetate. Imaging was performed in a JOEL JEM1400-Plus microscope with an accelerating voltage of 120 keV. Figure 2 shows a representative image of the electron micrographs with the immobilized VLPs on the EM grids. VLPs can be identified from the uniform size and the presentation of extended S protein spikes on their respective membranes. Using the electron micrographs, we measured a membrane diameter of 63 ± 5 nm and a spike to spike diameter of 103 ± 6 nm (N = 102).

The SARS-CoV-2 VLPs we have characterized are spherical and have a tight size distribution of 103 ± 6 nm measured from spike to spike with a spike density of approximately 100%. Our results are slightly different from general characterization of prior coronaviruses which have a polymorphic shape with size distributions of 100 - 200 nanometers [4]. SARS CoV VLPs have been characterized with cryotomography and range in size distribution of 70 – 110 nm. The shape of such VLPs range from nearly spherical to ellipsoidal with a 2 to 1 aspect ratio and a spike density with 20 – 100 % coverage of the membrane [9]. Our measured VLPs are however consistent with limited electron microscopy data available for the fully infectious SARS-CoV-2 virions observed in patient samples [14].

The structural integrity of VLPs can inform their practical applications as well as serve as an estimate of the stability of fully infections SARS-CoV-2 virions. We therefore further characterized the VLPs on functionalized surfaces with atomic force microscopy (AFM) (Fig. 3a). Briefly, glass coverslips were cleaned and functionalized with biotin-PEG-silane as previously described [15]. The surfaces were then incubated with neutravidin (Thermo Scientific Pierce Protein Biology, Waltham, MA, USA) followed by a buffer exchange and incubation with biotinylated anti-S antibody (11-2001-B, Abeomics, San Diego, CA, USA). Surfaces thus prepared were generally devoid of debris but sometimes had step-edge character suggesting that antibody coating was less than a full monolayer (Figure S2). The VLPs suspended in assay buffer were then incubated with the functionalized surface and finally washed away via brief buffer exchange with dd-H20 followed by drying. All incubations lasted 30 minutes at room temperature. Dimension Icon AFM with an MLCT-BIO-DC probe (Bruker,

Santa Barbara, CA, USA) was then used in PeakForce QNM mode at room temperature. AFM studies (Figure 3B,C and Figure S3) with soft tips and at low peak force allowed the VLP imaging. The VLPs appeared as spherically symmetric particles whose lateral diameter was in excess of 200 nm while heights of the particles varied between 50 nm and 60 nm. The shapes are consistent with VLP dimensions broadened laterally by imaging forces and tip curvature and reduced somewhat in height due to surface adhesion and possibly imaging forces. Although detailed characterization is subject of future work, we found that repeated imaging of VLPs with peak force of 1 nN led to gradual particle deformations: reduced particle heights and non-circular lateral cross-sections – consistent with VLP bursts (Figure 3C and Figure S3). This force is within the range of previously reported values (0.5-5 nN) for bursting various virus capsids[16, 17] although most prior work has been done in liquid. Our observations represent an upper estimate of the rupture force for SARS-CoV-2. This demonstrates that SARS-CoV-2 VLPs can be disrupted by direct application of moderate mechanical perturbations and open the door for future studies of VLP mechanics and integrity.

Expression of just M, S, and E proteins (Figure 1) was sufficient for release of VLPs which were structurally competent for both harvesting and subsequent investigations (Fig 2 and 3). These results open the door for many further studies. They can serve as immunogenic agents in place of the full infectious virus. The VLPs can also serve to study virus interactions with proteins of interest (e.g. receptors) - to date most such studies were only possible at the single protein level or in the context of the infectious virus. In addition, artificial VLPs can now be used to study the mechanical properties of the SARS-CoV-2 virus as well as their dependence on environmental conditions. Finally, VLPs can also be used to develop and validate novel testing strategies for SARS-CoV-2. The expression methodology is robust and should require no modifications to create VLPs for most S, M, and E mutations found in native conditions [18].

Acknowledgements

This study is supported by the NSF RAPID 2026657 award to MV and SS.

Author contributions

HS performed the biochemical characterization AS performed the AFM characterization, BP, AP and CE created reagents, DB performed the electron microscopy characterization, MV and SS designed the research and wrote the manuscript.

Competing financial interests The authors declare no financial conflicts of interest.

Figures:

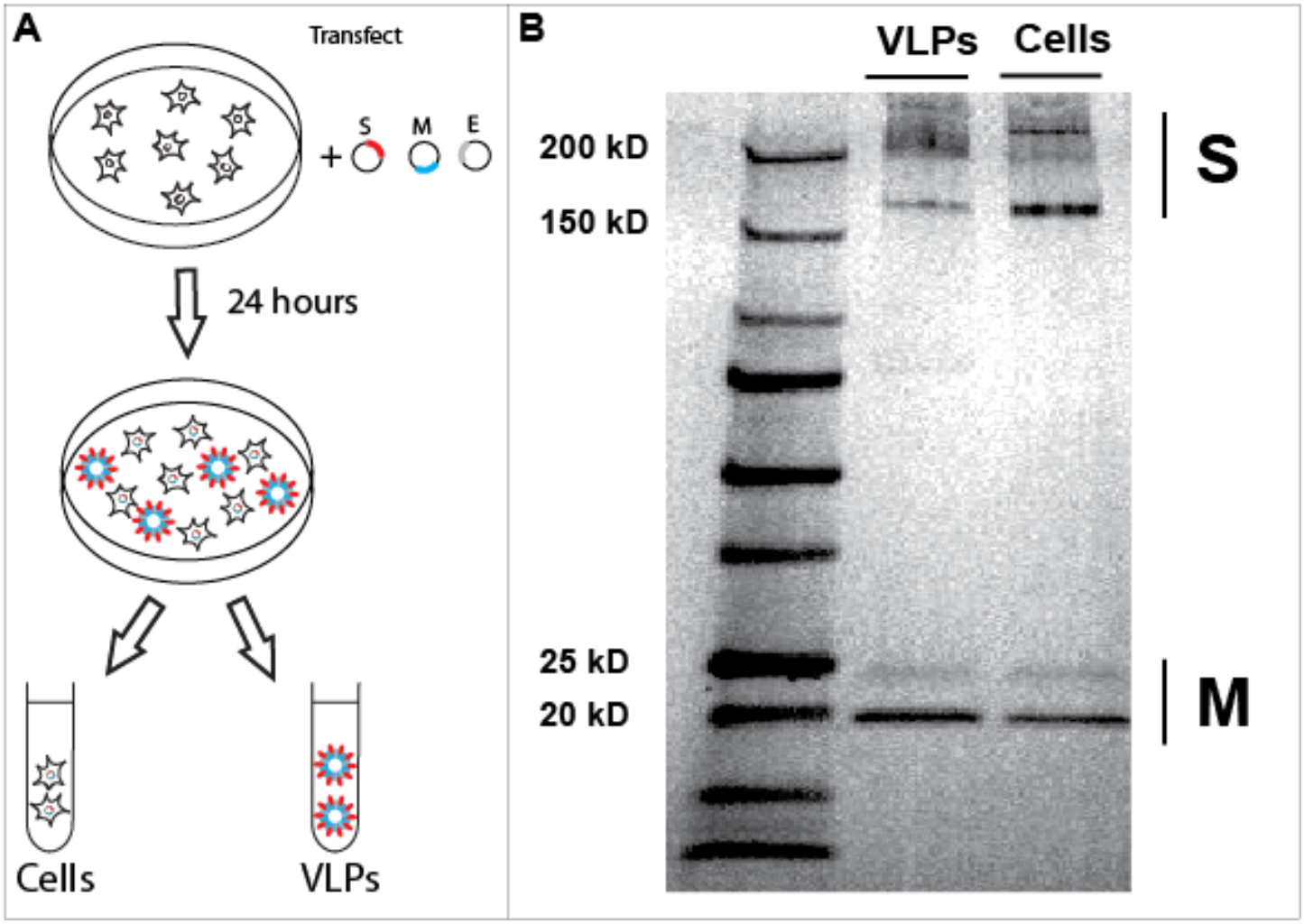

Figure 1. VLP expression and characterization using western blots. (a) Schematic representation of transfection and VLP harvest. (b) Western blot analysis of cells as well as VLPs separated as explained in methods. The blot against S and M proteins has been superimposed for greater clarity. The S protein is clearly identified by two bands running at 160kD and 200kD+. The M protein is running at around 20kD.

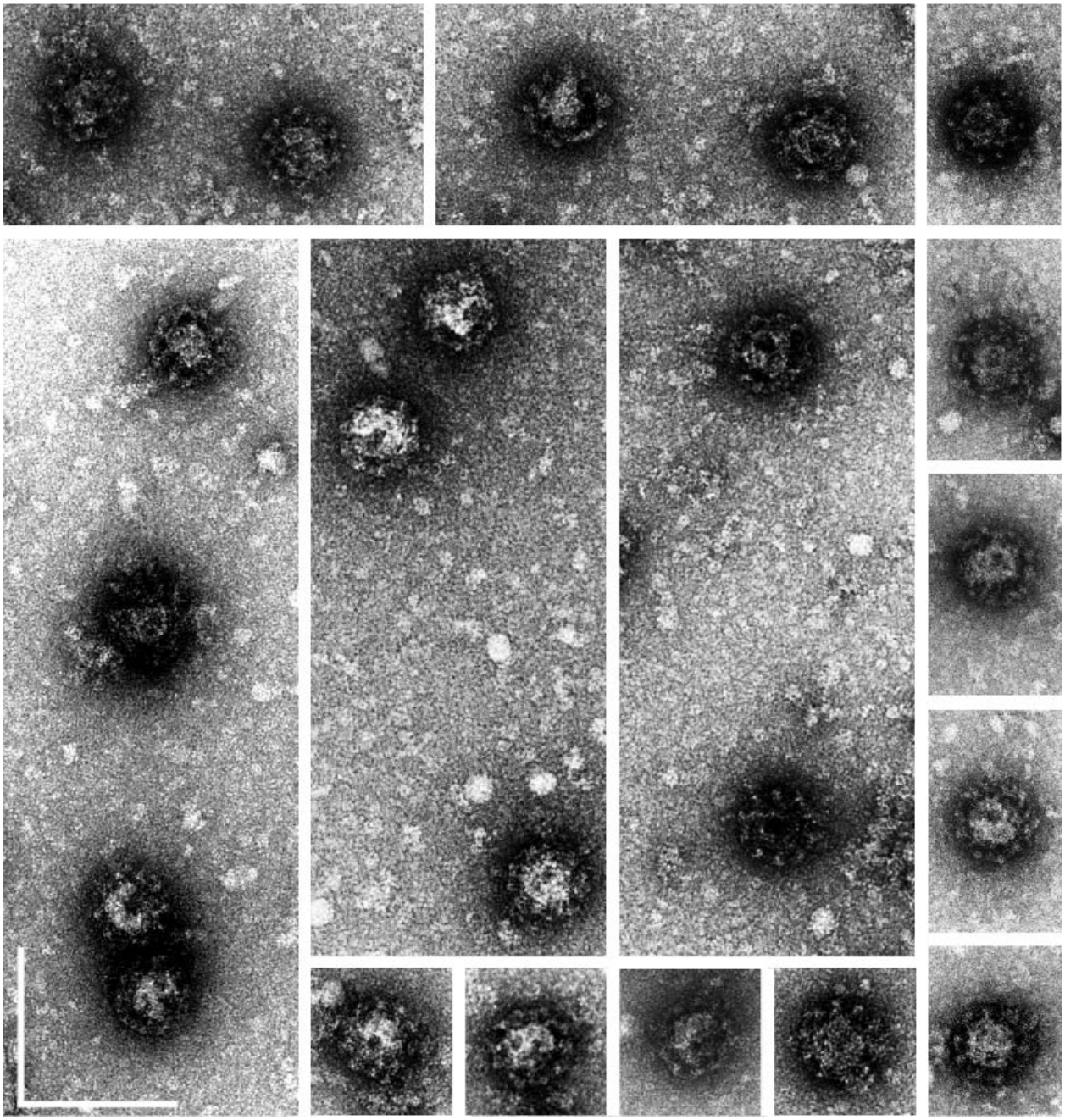

Figure 2. Electron microscopy reveals structurally complete SARS-CoV-2 VLPs. Scale bars (bottom left) : 200 nm . 22 representative VLPs are shown.

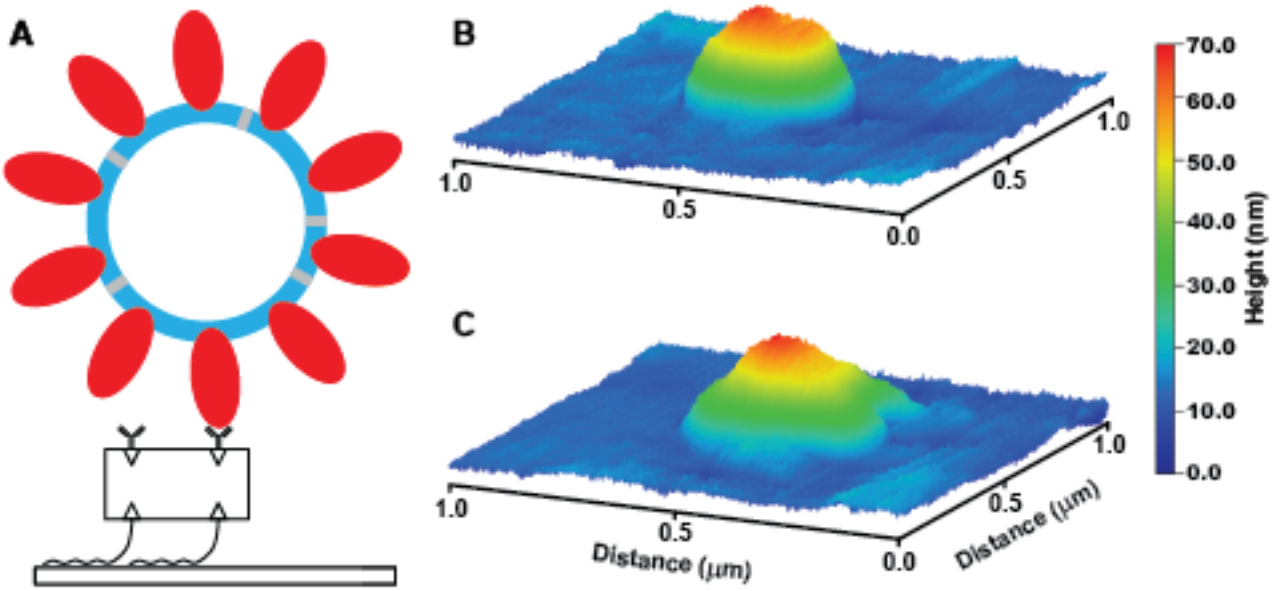

Figure 3. VLP surface immobilization and imaging. (a) Schematic of VLP-surface attachment strategy (bottom to top): glass surfaces were functionalized with Biotin-PEG-Silane, then neutravidin, then biotinylated anti-S antibody which then allowed for the capture of purified VLPs. VLP imaging showed initially symmetric particles (b) which showed prominent deformations likely indicative of bursts upon repeated AFM imaging with 1 nN peak force (c).

References:

1. Wu, F. et al. *Nature* (2020).
2. Prompetchara, E., Ketloy, C. & Palaga, T. *Asian Pac J Allergy Immunol* **38**, 1-9 (2020).
3. Zumla, A., Chan, J.F.W., Azhar, E.I., Hui, D.S.C. & Yuen, K.-Y. *Nat Rev Drug Discov* **15**, 327-347 (2016).
4. Masters, P.S. in Advances in Virus Research, Vol. 66 193-292 (Academic Press, 2006).
5. Lau, S.K.P. et al. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 14040-14045 (2005).
6. Hu, D. et al. *Emerg Microbes Infect* **7**, 154-154 (2018).
7. Baudoux, P., Carrat, C., Besnardeau, L., Charley, B. & Laude, H. *Journal of Virology* **72**, 8636-8643 (1998).
8. Corse, E. & Machamer, C.E. *Journal of Virology* **74**, 4319-4326 (2000).
9. Neuman, B.W. et al. *Journal of Structural Biology* **174**, 11-22 (2011).
10. Mortola, E. & Roy, P. *FEBS Lett* **576**, 174-178 (2004).
11. Wang, C. et al. *Oncotarget* **8**, 12686-12694 (2017).
12. Marra, M.A. et al. *Science* **300**, 1399-1404 (2003).
13. Chin, A.W.H. et al. *The Lancet Microbe* **1**, e10 (2020).
14. Kim, J.-M. et al. *Osong Public Health Res Perspect* **11**, 3-7 (2020).
15. Hodges, J.A. & Saffarian, S. *JoVE*, e51366 (2014).
16. Ivanovska, I.L. et al. *Proc. Natl. Acad. Sci. U. S. A.* **101**, 7600-7605 (2004).
17. Xu, X., Carrasco, C., de Pablo, P.J., Gomez-Herrero, J. & Raman, A. *Biophysical Journal* **95**, 2520-2528 (2008).
18. Korber, B. et al. *bioRxiv*, 2020.2004.2029.069054 (2020).

SUPPLEMENTARY INFORMATION

FOR

MINIMAL SYSTEM FOR ASSEMBLY OF SARS-COV-2 VIRUS LIKE PARTICLES

Authors:

Heather Swann[1,2*], Abhimanyu Sharma[2*], Benjamin Preece[1,2], Abby Peterson[1,2], Crystal Eldridge[1,2], David M. Belnap,[3,4] Michael Vershinin[1,2,3,†] and Saveez Saffarian[1,2,3,†]

Affiliations:

[1] Center for Cell and Genome Science, University of Utah

[2] Department of Physics and Astronomy, University of Utah

[3] School of Biological Sciences, University of Utah

[4] Department of Biochemistry, University of Utah

[†] Corresponding authors: Vershinin@Physics.utah.edu and Saffarian@Physics.Utah.edu

*These authors contributed equally to this work

Cell Culture:

Cells from the human embryonic kidney 293T cell line were maintained in T-25 flasks using TrypLE Express Enzyme (Gibco) and DMEM with L-Glutamine, 4.5g/L glucose and sodium pyruvate (Corning) supplemented with 10% fetal bovine serum (Gibco). The cells were incubated at 37°C. Once over 90% confluent, the 293T cells were plated onto 100 mm tissue culture dishes with 10 mL of medium which would be ready for transfection the next day.

Transfection and VLP harvest:

When the 293T cells were at 50% confluent in the 100 mm dishes, they were co-transfected with 29.5 µg of humanized membrane in pCDNA3.1, 5.9 µg of humanized spike in pCDNA3.1, and 5.9 µg of humanized envelope in pCDNA3.1 (GenScript). This transfection was carried out using 600 mL of Opti-MEM Reduced Serum Medium (Gibco) and 40 mL of Lipofectamine 2000 (Thermo Fisher Scientific). The cells were incubated at 37°C in 9 mL of Full DMEM with FBS supplementation for 24 hours.

Western Blot:

The cell lysates and VLPs were diluted in RIPA Lysis Buffer System (Santa Cruz), prepared with Laemmli sample buffer (BioRad) with 5% BME, and boiled at 95°C for 5 minutes. Standard ladders were prepared with Precision Plus Protein WesternC Standard (BioRad). The proteins were then separated by SDS-PAGE gel electrophoresis in a glycine running buffer. The gel was transferred to a PVDF membrane (Millipore Sigma) via wet transfer at 35 volts and the membrane was probed with anti-SARS-CoV-2 spike monoclonal (GeneTex) and anti-membrane protein 2019 n-CoV polyclonal (eEnzyme) primary antibodies. After overnight incubation at room temperature, the membrane was probed with infrared anti-mouse 680 nm (Invitrogen) and anti-rabbit 800 (BioRad) secondary antibodies. The membrane was scanned using the Odyssey Infrared Imaging System (LI-COR) at 700 and 800 nm. For presentation in Figure 1, the 700 nm imaging was superimposed onto 800 nm imaging and presented as gray scale image.

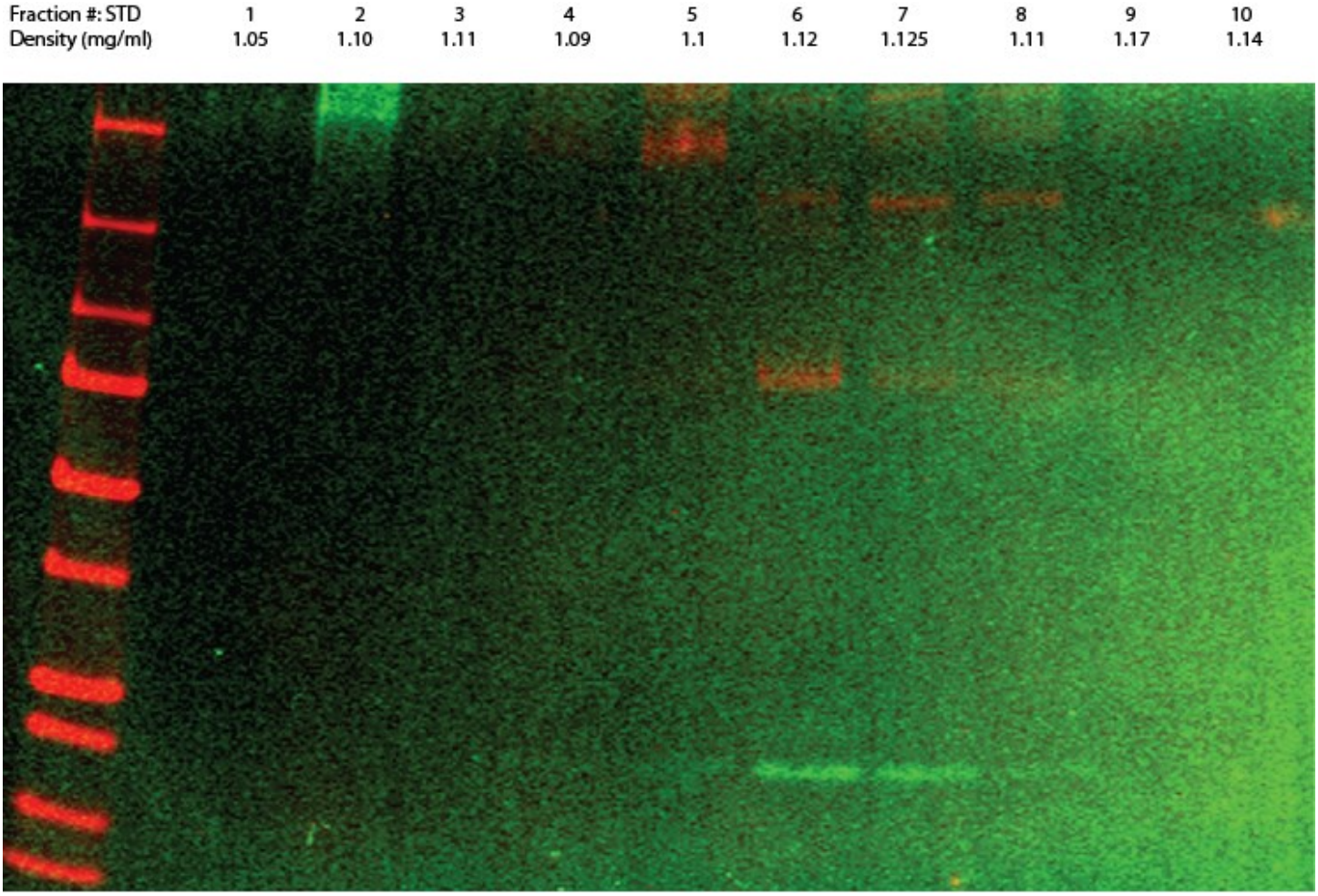

Figure S1. Sucrose fractionation of VLPs.

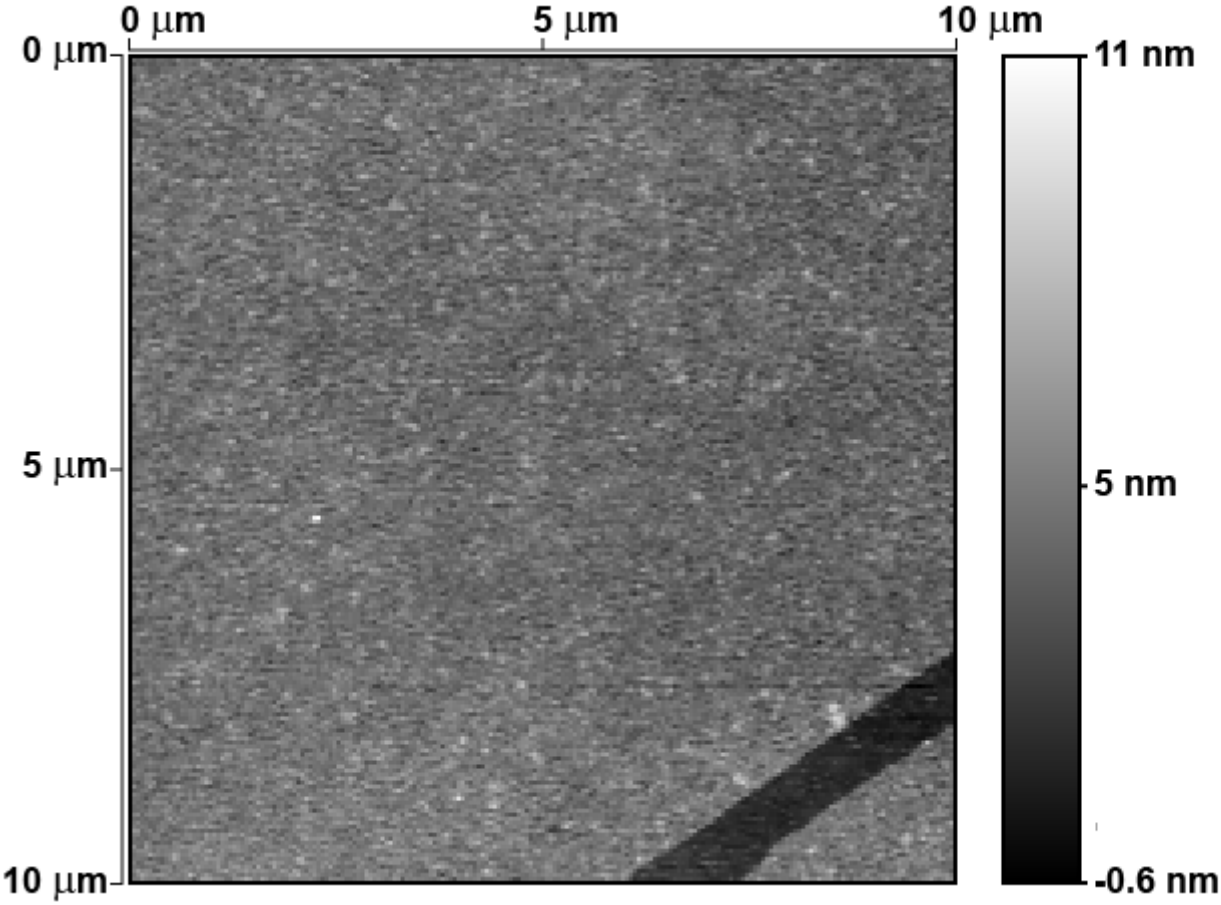

Figure S2. AFM imaging of glass surface functionalized with anti-S antibody. Glass surface functionalized as described in main text typically has large flat debris-free areas with occasional shallow step edges. These latter features are only seen after antibody adsorption and therefore correspond to sub-monolayer antibody coverage.

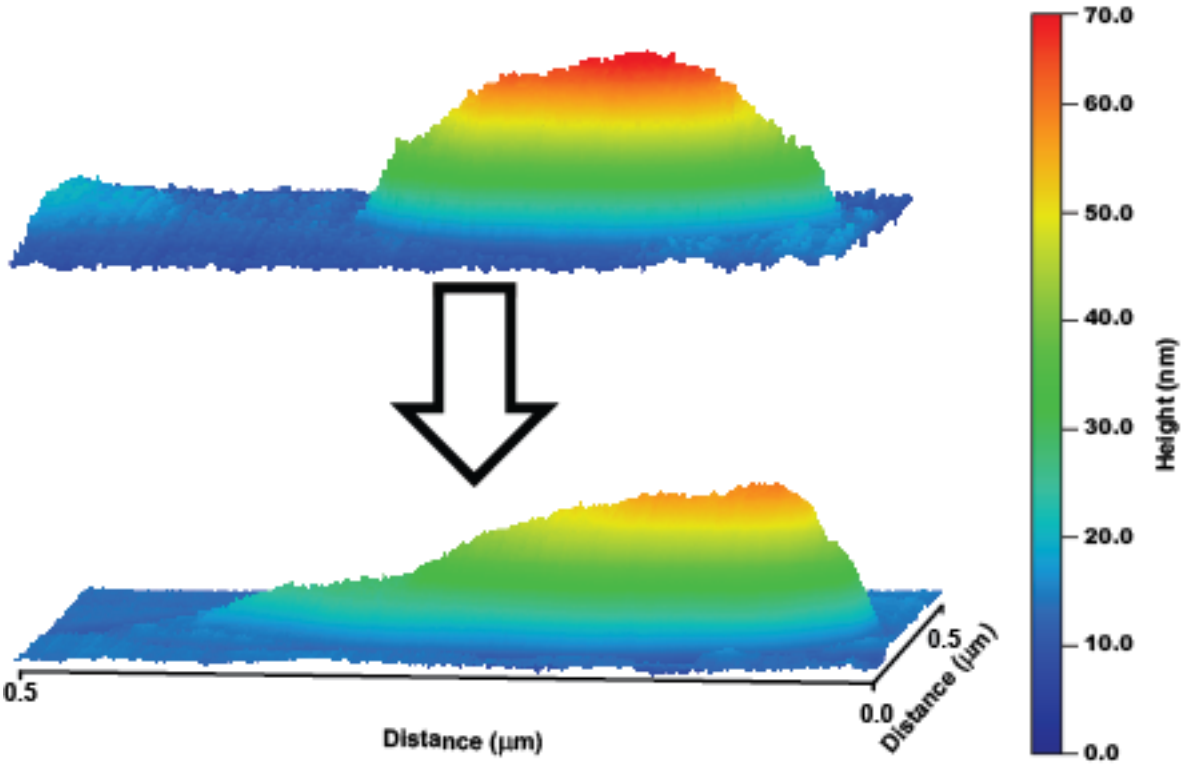

Figure S3. AFM imaging and bursting of a VLP. Several repeated scans at 1 nN peak force lead intact VLP (top) to burst (bottom).

Humanized S protein expression plasmid:

GACGGATCGGGAGATCTCCCGATCCCCTATGGTGCACTCTCAGTACAATCTGCTCTGATGCCGCATAGTTAAGCCAGTATCTGCTCCC
TGCTTGTGTGTTGGAGGTCGCTGAGTAGTGCGCGAGCAAAATTTAAGCTACAACAAGGCAAGGCTTGACCGACAATTGCATGAAGA
ATCTGCTTAGGGTTAGGCGTTTTGCGCTGCTTCGCGAGTACATTTATATTGGCTCATGTCCAATATGACCGCCATGTTGACATTGATTA
TTGACTAGTTATTAATAGTAATCAATTACGGGGTCATTAGTTCATAGCCCATATATGGAGTTCCGCGTTACATAACTTACGGTAAATGG
CCCGCCTGGCTGACCGCCCAACGACCCCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCAT
TGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTCCGCCCCCTATTGACG
TCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTACGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGT
CATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCAC
CCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAA
TGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTCAGATCCTCACTCTCTTCCGCATCGCTG
TCTGCGAGGGCCAGCTGTTGGGCTCGCGGTTGAGGACAAACTCTTCGCGGTCTTTCCAGTACTCTTGGATCGGAAACCCGTCGGCCT
CCGAACGGTACTCCGCCACCGAGGGACCTGAGCGAGTCCGCATCGACCGGATCGGAAAACCTCTCGAGAAAGGCGTCTAACCAGTC
ACAGTCGCAAGGTAGGCTGAGCACCGTGGCGGGCGGCAGCGGGTGGCGGTCGGGGTTGTTTCTGGCGGAGGTGCTGCTGATGATG
TAATTAAAGTAGGCGGTCTTGAGACGGCGGATGGTCGAGGTGAGGTGTGGCAGGCTTGAGATCCAGCTGTTGGGGTGAGTACTCCC
TCTCAAAAGCGGGCATTACTTCTGCGCTAAGATTGTCAGTTTCCAAAAACGAGGAGGATTTGATATTCACCTGGCCCGATCTGGCCAT
ACACTTGAGTGACAATGACATCCACTTTGCCTTTCTCTCCACAGGTGTCCACTCCCAGGTCCAAGTTTAAACTTTAATACGACTCACTAT
AGGGGCCGCCACCAAGCTTNNNNNGGTACCNNNNNATGTTTGTGTTCCTGGTGCTGCTGCCACTGGTGTCCAGCCAGTGTGTGAAC
CTGACCACCAGGACCCAACTTCCTCCTGCCTACACCAACTCCTTCACCAGGGGAGTCTACTACCCTGACAAGGTGTTCAGGTCCTCTGT
GCTGCACAGCACCCAGGACCTGTTCCTGCCATTCTTCAGCAATGTGACCTGGTTCCATGCCATCCATGTGTCTGGCACCAATGGCACC
AAGAGGTTTGACAACCCTGTGCTGCCATTCAATGATGGAGTCTACTTTGCCAGCACAGAGAAGAGCAACATCATCAGGGGCTGGATT
TTTGGCACCACCCTGGACAGCAAGACCCAGTCCCTGCTGATTGTGAACAATGCCACCAATGTGGTGATTAAGGTGTGTGAGTTCCAG
TTCTGTAATGACCCATTCCTGGGAGTCTACTACCACAAGAACAACAAGTCCTGGATGGAGTCTGAGTTCAGGGTCTACTCCTCTGCCA
ACAACTGTACCTTTGAATATGTGAGCCAACCATTCCTGATGGACTTGGAGGGCAAGCAGGGCAACTTCAAGAACCTGAGGGAGTTTG
TGTTCAAGAACATTGATGGCTACTTCAAGATTTACAGCAAACACACACCAATCAACCTGGTGAGGGACCTGCCACAGGGCTTCTCTGC
CTTGGAACCACTGGTGGACCTGCCAATTGGCATCAACATCACCAGGTTCCAGACCCTGCTGGCTCTGCACAGGTCCTACCTGACACCT
GGAGACTCCTCCTCTGGCTGGACAGCAGGAGCAGCAGCCTACTATGTGGGCTACCTCCAACCAAGGACCTTCCTGCTGAAATACAAT
GAGAATGGCACCATCACAGATGCTGTGGACTGTGCCCTGGACCCACTGTCTGAGACCAAGTGTACCCTGAAATCCTTCACAGTGGAG
AAGGGCATCTACCAGACCAGCAACTTCAGGGTCCAACCAACAGAGAGCATTGTGAGGTTTCCAAACATCACCAACCTGTGTCCATTTG
GAGAGGTGTTCAATGCCACCAGGTTTGCCTCTGTCTATGCCTGGAACAGGAAGAGGATTAGCAACTGTGTGGCTGACTACTCTGTGC
TCTACAACTCTGCCTCCTTCAGCACCTTCAAGTGTTATGGAGTGAGCCCAACCAAACTGAATGACCTGTGTTTCACCAATGTCTATGCT
GACTCCTTTGTGATTAGGGGAGATGAGGTGAGACAGATTGCCCCTGGACAAACAGGCAAGATTGCTGACTACAACTACAAACTGCCT
GATGACTTCACAGGCTGTGTGATTGCCTGGAACAGCAACAACCTGGACAGCAAGGTGGGAGGCAACTACAACTACCTCTACAGACTG
TTCAGGAAGAGCAACCTGAAACCATTTGAGAGGGACATCAGCACAGAGATTTACCAGGCTGGCAGCACACCATGTAATGGAGTGGA

GGGCTTCAACTGTTACTTTCCACTCCAATCCTATGGCTTCCAACCAACCAATGGAGTGGGCTACCAACCATACAGGGTGGTGGTGCTG
TCCTTTGAACTGCTCCATGCCCCTGCCACAGTGTGTGGACCAAAGAAGAGCACCAACCTGGTGAAGAACAAGTGTGTGAACTTCAAC
TTCAATGGACTGACAGGCACAGGAGTGCTGACAGAGAGCAACAAGAAGTTCCTGCCATTCCAACAGTTTGGCAGGGACATTGCTGA
CACCACAGATGCTGTGAGGGACCCACAGACCTTGGAGATTCTGGACATCACACCATGTTCCTTTGGAGGAGTGTCTGTGATTACACCT
GGCACCAACACCAGCAACCAGGTGGCTGTGCTCTACCAGGATGTGAACTGTACTGAGGTGCCTGTGGCTATCCATGCTGACCAACTT
ACACCAACCTGGAGGGTCTACAGCACAGGCAGCAATGTGTTCCAGACCAGGGCTGGCTGTCTGATTGGAGCAGAGCATGTGAACAA
CTCCTATGAGTGTGACATCCCAATTGGAGCAGGCATCTGTGCCTCCTACCAGACCCAGACCAACAGCCCAAGGAGGGCAAGGTCTGT
GGCAAGCCAGAGCATCATTGCCTACACAATGAGTCTGGGAGCAGAGAACTCTGTGGCTTACAGCAACAACAGCATTGCCATCCCAAC
CAACTTCACCATCTCTGTGACCACAGAGATTCTGCCTGTGAGTATGACCAAGACCTCTGTGGACTGTACAATGTATATCTGTGGAGAC
AGCACAGAGTGTAGCAACCTGCTGCTCCAATATGGCTCCTTCTGTACCCAACTTAACAGGGCTCTGACAGGCATTGCTGTGGAACAG
GACAAGAACACCCAGGAGGTGTTTGCCCAGGTGAAGCAGATTTACAAGACACCTCCAATCAAGGACTTTGGAGGCTTCAACTTCAGC
CAGATTCTGCCTGACCCAAGCAAGCCAAGCAAGAGGTCCTTCATTGAGGACCTGCTGTTCAACAAGGTGACCCTGGCTGATGCTGGC
TTCATCAAGCAATATGGAGACTGTCTGGGAGACATTGCTGCCAGGGACCTGATTTGTGCCCAGAAGTTCAATGGACTGACAGTGCTG
CCTCCACTGCTGACAGATGAGATGATTGCCCAATACACCTCTGCCCTGCTGGCTGGCACCATCACCTCTGGCTGGACCTTTGGAGCAG
GAGCAGCCCTCCAAATCCCATTTGCTATGCAGATGGCTTACAGGTTCAATGGCATTGGAGTGACCCAGAATGTGCTCTATGAGAACC
AGAAACTGATTGCCAACCAGTTCAACTCTGCCATTGGCAAGATTCAGGACTCCCTGTCCAGCACAGCCTCTGCCCTGGGCAAACTCCA
AGATGTGGTGAACCAGAATGCCCAGGCTCTGAACACCCTGGTGAAGCAACTTTCCAGCAACTTTGGAGCCATCTCCTCTGTGCTGAAT
GACATCCTGAGCAGACTGGACAAGGTGGAGGCTGAGGTCCAGATTGACAGACTGATTACAGGCAGACTCCAATCCCTCCAAACCTAT
GTGACCCAACAACTTATCAGGGCTGCTGAGATTAGGGCATCTGCCAACCTGGCTGCCACCAAGATGAGTGAGTGTGTGCTGGGACA
AAGCAAGAGGGTGGACTTCTGTGGCAAGGGCTACCACCTGATGAGTTTTCCACAGTCTGCCCCTCATGGAGTGGTGTTCCTGCATGT
GACCTATGTGCCTGCCCAGGAGAAGAACTTCACCACAGCCCCTGCCATCTGCCATGATGGCAAGGCTCACTTTCCAAGGGAGGGAGT
GTTTGTGAGCAATGGCACCCACTGGTTTGTGACCCAGAGGAACTTCTATGAACCACAGATTATCACCACAGACAACACCTTTGTGTCT
GGCAACTGTGATGTGGTGATTGGCATTGTGAACAACACAGTCTATGACCCACTCCAACCTGAACTGGACTCCTTCAAGGAGGAACTG
GACAAATACTTCAAGAACCACACCAGCCCTGATGTGGACCTGGGAGACATCTCTGGCATCAATGCCTCTGTGGTGAACATCCAGAAG
GAGATTGACAGACTGAATGAGGTGGCTAAGAACCTGAATGAGTCCCTGATTGACCTCCAAGAACTGGGCAAATATGAACAATACATC
AAGTGGCCATGGTACATCTGGCTGGGCTTCATTGCTGGACTGATTGCCATTGTGATGGTGACCATAATGCTGTGTTGTATGACCTCCT
GTTGTTCCTGTCTGAAAGGCTGTTGTTCCTGTGGCTCCTGTTGTAAGTTTGATGAGGATGACTCTGAACCTGTGCTGAAAGGAGTGAA
ACTGCACTACACCTGANNNNNTCTAGANNNNNGCGGCCGCCGAATTCGGGCCCGTTTAAACCCGCTGATCAGCCTCGACTGTGCCT
TCTAGTTGCCAGCCATCTGTTGTTTGCCCCTCCCCCGTGCCTTCCTTGACCCTGGAAGGTGCCACTCCCACTGTCCTTTCCTAATAAAAT
GAGGAAATTGCATCGCATTGTCTGAGTAGGTGTCATTCTATTCTGGGGGGTGGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGG
AAGACAATAGCAGGCATGCTGGGGATGCGGTGGGCTCTATGGCTTCTGAGGCGGAAAGAACCAGCTGGGGCTCTAGGGGGTATCC
CCACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGGTGTGGTGGTTACGCGCAGCGTGACCGCTACACTTGCCAGCGCCCTAGCGC
CCGCTCCTTTCGCTTTCTTCCCTTCCTTTCTCGCCACGTTCGCCGGCTTTCCCCGTCAAGCTCTAAATCGGGGGCTCCCTTTAGGGTTCC
GATTTAGTGCTTTACGGCACCTCGACCCCAAAAAACTTGATTAGGGTGATGGTTCACGTAGTGGGCCATCGCCCTGATAGACGGTTTT

TCGCCCTTTGACGTTGGAGTCCACGTTCTTTAATAGTGGACTCTTGTTCCAAACTGGAACAACACTCAACCCTATCTCGGTCTATTCTTT
TGATTTATAAGGGATTTTGCCGATTTCGGCCTATTGGTTAAAAAATGAGCTGATTTAACAAAAATTTAACGCGAATTAATTCTGTGGA
ATGTGTGTCAGTTAGGGTGTGGAAAGTCCCCAGGCTCCCCAGCAGGCAGAAGTATGCAAAGCATGCATCTCAATTAGTCAGCAACCA
GGTGTGGAAAGTCCCCAGGCTCCCCAGCAGGCAGAAGTATGCAAAGCATGCATCTCAATTAGTCAGCAACCATAGTCCCGCCCCTAA
CTCCGCCCATCCCGCCCCTAACTCCGCCCAGTTCCGCCCATTCTCCGCCCCATGGCTGACTAATTTTTTTTATTTATGCAGAGGCCGAGG
CCGCCTCTGCCTCTGAGCTATTCCAGAAGTAGTGAGGAGGCTTTTTTGGAGGCCTAGGCTTTTGCAAAAAGCTCCCGGGAGCTTGTAT
ATCCATTTTCGGATCTGATCAGCACGTGATGAAAAAGCCTGAACTCACCGCGACGTCTGTCGAGAAGTTTCTGATCGAAAAGTTCGAC
AGCGTCTCCGACCTGATGCAGCTCTCGGAGGGCGAAGAATCTCGTGCTTTCAGCTTCGATGTAGGAGGGCGTGGATATGTCCTGCGG
GTAAATAGCTGCGCCGATGGTTTCTACAAAGATCGTTATGTTTATCGGCACTTTGCATCGGCCGCGCTCCCGATTCCGGAAGTGCTTG
ACATTGGGGAATTCAGCGAGAGCCTGACCTATTGCATCTCCCGCCGTGCACAGGGTGTCACGTTGCAAGACCTGCCTGAAACCGAAC
TGCCCGCTGTTCTGCAGCCGGTCGCGGAGGCCATGGATGCGATCGCTGCGGCCGATCTTAGCCAGACGAGCGGGTTCGGCCCATTC
GGACCGCAAGGAATCGGTCAATACACTACATGGCGTGATTTCATATGCGCGATTGCTGATCCCCATGTGTATCACTGGCAAACTGTG
ATGGACGACACCGTCAGTGCGTCCGTCGCGCAGGCTCTCGATGAGCTGATGCTTTGGGCCGAGGACTGCCCCGAAGTCCGGCACCTC
GTGCACGCGGATTTCGGCTCCAACAATGTCCTGACGGACAATGGCCGCATAACAGCGGTCATTGACTGGAGCGAGGCGATGTTCGG
GGATTCCCAATACGAGGTCGCCAACATCTTCTTCTGGAGGCCGTGGTTGGCTTGTATGGAGCAGCAGACGCGCTACTTCGAGCGGAG
GCATCCGGAGCTTGCAGGATCGCCGCGGCTCCGGGCGTATATGCTCCGCATTGGTCTTGACCAACTCTATCAGAGCTTGGTTGACGG
CAATTTCGATGATGCAGCTTGGGCGCAGGGTCGATGCGACGCAATCGTCCGATCCGGAGCCGGGACTGTCGGGCGTACACAAATCG
CCCGCAGAAGCGCGGCCGTCTGGACCGATGGCTGTGTAGAAGTACTCGCCGATAGTGGAAACCGACGCCCCAGCACTCGTCCGAGG
GCAAAGGAATAGCACGTGCTACGAGATTTCGATTCCACCGCCGCCTTCTATGAAAGGTTGGGCTTCGGAATCGTTTTCCGGGACGCC
GGCTGGATGATCCTCCAGCGCGGGGATCTCATGCTGGAGTTCTTCGCCCACCCCAACTTGTTTATTGCAGCTTATAATGGTTACAAAT
AAAGCAATAGCATCACAAATTTCACAAATAAAGCATTTTTTTCACTGCATTCTAGTTGTGGTTTGTCCAAACTCATCAATGTATCTTATC
ATGTCTGTATACCGTCGACCTCTAGCTAGAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATT
CCACACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCA
CTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGG
GCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATA
CGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCG
CGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAG
GACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCC
TTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCT
GTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATC
GCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACT
ACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCG
GCAAACAAACCACCGCTGGTAGCGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGA
TCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTA

GATCCTTTTAAATTAAAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTG
AGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTA
CCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGG
GCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAG
TTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCC
CAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAAGT
TGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGT
GAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAATACCGCGCCA
CATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTT
CGATGTAACCCACTCGTGCACCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAA
TGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAGCATTTATCAGGGT
TATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCCAC
CTGACGTC

Humanized M protein expression plasmid:

GACGGATCGGGAGATCTCCCGATCCCCTATGGTGCACTCTCAGTACAATCTGCTCTGATGCCGCATAGTTAAGCCAGTATCTGCTCCC
TGCTTGTGTGTTGGAGGTCGCTGAGTAGTGCGCGAGCAAAATTTAAGCTACAACAAGGCAAGGCTTGACCGACAATTGCATGAAGA
ATCTGCTTAGGGTTAGGCGTTTTGCGCTGCTTCGCGATGTACGGGCCAGATATACGCGTTGACATTGATTATTGACTAGTTATTAATA
GTAATCAATTACGGGGTCATTAGTTCATAGCCCATATATGGAGTTCCGCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCG
CCCAACGACCCCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGG
AGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTACGCCCCCTATTGACGTCAATGACGGTAAATG
GCCCGCCTGGCATTATGCCCAGTACATGACCTTATGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGG
TGATGCGGTTTTGGCAGTACATCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATG
GGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAACAACTCCGCCCCATTGACGCAAATGGGCGGTAGGCGTG
TACGGTGGGAGGTCTATATAAGCAGAGCTCTCTGGCTAACTAGAGAACCCACTGCTTACTGGCTTATCGAAATTAATACGACTCACTA
TAGGGAGACCCAAGCTGGCTAGCGCAGCATCTTCAGCTCTTGGGTTTAGCAAAAGTTCATCTCCTTCTGCATCCTTAACTGAGAATGA
GCTTTTGTGGGAGCCCACACCAGTCAAGTTGGATTTGAACCCAGCTGCTCTGTACAAGCTTGGTACCGCCACCATGGCCGACAGCAA
CGGCACCATCACCGTGGAGGAGCTGAAGAAGCTGCTGGAGCAGTGGAACCTGGTGATCGGCTTCCTGTTCCTGACCTGGATCTGCCT
GCTGCAGTTCGCCTACGCCAACAGGAACAGGTTCCTGTACATCATCAAGCTGATCTTCCTGTGGCTGCTGTGGCCCGTGACCCTGGCC
TGCTTCGTGCTGGCCGCCGTGTACAGGATCAACTGGATCACCGGCGGCATCGCCATCGCCATGGCCTGCCTGGTGGGCCTGATGTGG
CTGAGCTACTTCATCGCCAGCTTCAGGCTGTTCGCCAGGACCAGGAGCATGTGGAGCTTCAACCCCGAGACCAACATCCTGCTGAAC
GTGCCCCTGCACGGCACCATCCTGACCAGGCCCCTGCTGGAGAGCGAGCTGGTGATCGGCGCCGTGATCCTGAGGGGCCACCTGAG
GATCGCCGGCCACCACCTGGGCAGGTGCGACATCAAGGACCTGCCCAAGGAGATCACCGTGGCCACCAGCAGGACCCTGAGCTACT
ACAAGCTGGGCGCCAGCCAGAGGGTGGCCGGCGACAGCGGCTTCGCCGCCTACAGCAGGTACAGGATCGGCAACTACAAGCTGAA

CACCGACCACAGCAGCAGCAGCGACAACATCGCCCTGCTGGTGCAGTAATCTAGAGGGCCCGTTTAAACCCGCTGATCAGCCTCGAC
TGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGCCCCTCCCCCGTGCCTTCCTTGACCCTGGAAGGTGCCACTCCCACTGTCCTTTCCTA
ATAAAATGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTCATTCTATTCTGGGGGGTGGGGTGGGGCAGGACAGCAAGGGGGAG
GATTGGGAAGACAATAGCAGGCATGCTGGGGATGCGGTGGGCTCTATGGCTTCTGAGGCGGAAAGAACCAGCTGGGGCTCTAGGG
GGTATCCCCACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGGTGTGGTGGTTACGCGCAGCGTGACCGCTACACTTGCCAGCGCC
CTAGCGCCCGCTCCTTTCGCTTTCTTCCCTTCCTTTCTCGCCACGTTCGCCGGCTTTCCCCGTCAAGCTCTAAATCGGGGGCTCCCTTTA
GGGTTCCGATTTAGTGCTTTACGGCACCTCGACCCCAAAAAACTTGATTAGGGTGATGGTTCACGTAGTGGGCCATCGCCCTGATAG
ACGGTTTTTCGCCCTTTGACGTTGGAGTCCACGTTCTTTAATAGTGGACTCTTGTTCCAAACTGGAACAACACTCAACCCTATCTCGGT
CTATTCTTTTGATTTATAAGGGATTTTGCCGATTTCGGCCTATTGGTTAAAAAATGAGCTGATTTAACAAAAATTTAACGCGAATTAAT
TCTGTGGAATGTGTGTCAGTTAGGGTGTGGAAAGTCCCCAGGCTCCCCAGCAGGCAGAAGTATGCAAAGCATGCATCTCAATTAGTC
AGCAACCAGGTGTGGAAAGTCCCCAGGCTCCCCAGCAGGCAGAAGTATGCAAAGCATGCATCTCAATTAGTCAGCAACCATAGTCCC
GCCCCTAACTCCGCCCATCCCGCCCCTAACTCCGCCCAGTTCCGCCCATTCTCCGCCCCATGGCTGACTAATTTTTTTTATTTATGCAGA
GGCCGAGGCCGCCTCTGCCTCTGAGCTATTCCAGAAGTAGTGAGGAGGCTTTTTTGGAGGCCTAGGCTTTTGCAAAAAGCTCCCGGG
AGCTTGTATATCCATTTTCGGATCTGATCAAGAGACAGGATGAGGATCGTTTCGCATGATTGAACAAGATGGATTGCACGCAGGTTCT
CCGGCCGCTTGGGTGGAGAGGCTATTCGGCTATGACTGGGCACAACAGACAATCGGCTGCTCTGATGCCGCCGTGTTCCGGCTGTCA
GCGCAGGGGCGCCCGGTTCTTTTTGTCAAGACCGACCTGTCCGGTGCCCTGAATGAACTGCAGGACGAGGCAGCGCGGCTATCGTG
GCTGGCCACGACGGGCGTTCCTTGCGCAGCTGTGCTCGACGTTGTCACTGAAGCGGGAAGGGACTGGCTGCTATTGGGCGAAGTGC
CGGGGCAGGATCTCCTGTCATCTCACCTTGCTCCTGCCGAGAAAGTATCCATCATGGCTGATGCAATGCGGCGGCTGCATACGCTTGA
TCCGGCTACCTGCCCATTCGACCACCAAGCGAAACATCGCATCGAGCGAGCACGTACTCGGATGGAAGCCGGTCTTGTCGATCAGGA
TGATCTGGACGAAGAGCATCAGGGGCTCGCGCCAGCCGAACTGTTCGCCAGGCTCAAGGCGCGCATGCCCGACGGCGAGGATCTCG
TCGTGACCCATGGCGATGCCTGCTTGCCGAATATCATGGTGGAAAATGGCCGCTTTTCTGGATTCATCGACTGTGGCCGGCTGGGTG
TGGCGGACCGCTATCAGGACATAGCGTTGGCTACCCGTGATATTGCTGAAGAGCTTGGCGGCGAATGGGCTGACCGCTTCCTCGTGC
TTTACGGTATCGCCGCTCCCGATTCGCAGCGCATCGCCTTCTATCGCCTTCTTGACGAGTTCTTCTGAGCGGGACTCTGGGGTTCGAA
ATGACCGACCAAGCGACGCCCAACCTGCCATCACGAGATTTCGATTCCACCGCCGCCTTCTATGAAAGGTTGGGCTTCGGAATCGTTT
TCCGGGACGCCGGCTGGATGATCCTCCAGCGCGGGGATCTCATGCTGGAGTTCTTCGCCCACCCCAACTTGTTTATTGCAGCTTATAA
TGGTTACAAATAAAGCAATAGCATCACAAATTTCACAAATAAAGCATTTTTTTCACTGCATTCTAGTTGTGGTTTGTCCAAACTCATCA
ATGTATCTTATCATGTCTGTATACCGTCGACCTCTAGCTAGAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTAT
CCGCTCACAATTCCACACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCACATTAATT
GCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGG
TTTGCGTATTGGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCA
AAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACC
GTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGC
GAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGG
ATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCT

CCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAG
ACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGT
GGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTA
GCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTC
AAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAA
AAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACC
AATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGATAACTACG
ATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAAC
CAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGA
GTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTATGGCTT
CATTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGAT
CGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTCATGCCATCCGTAAGAT
GCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACG
GGATAATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAACTCTCAAGGATCTTACC
GCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAA
AAACAGGAAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTTCCTTTTTCAATATTATT
GAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATT
TCCCCGAAAAGTGCCACCTGACGTC

Humanized E protein expression plasmid sequence:

GACGGATCGGGAGATCTCCCGATCCCCTATGGTGCACTCTCAGTACAATCTGCTCTGATGCCGCATAGTTAAGCCAGTATCTGCTCCC
TGCTTGTGTGTTGGAGGTCGCTGAGTAGTGCGCGAGCAAAATTTAAGCTACAACAAGGCAAGGCTTGACCGACAATTGCATGAAGA
ATCTGCTTAGGGTTAGGCGTTTTGCGCTGCTTCGCGATGTACGGGCCAGATATACGCGTTGACATTGATTATTGACTAGTTATTAATA
GTAATCAATTACGGGGTCATTAGTTCATAGCCCATATATGGAGTTCCGCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCG
CCCAACGACCCCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGG
AGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTACGCCCCCTATTGACGTCAATGACGGTAAATG
GCCCGCCTGGCATTATGCCCAGTACATGACCTTATGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGG
TGATGCGGTTTTGGCAGTACATCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATG
GGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAACAACTCCGCCCCATTGACGCAAATGGGCGGTAGGCGTG
TACGGTGGGAGGTCTATATAAGCAGAGCTCTCTGGCTAACTAGAGAACCCACTGCTTACTGGCTTATCGAAATTAATACGACTCACTA
TAGGGAGACCCAAGCTGGCTAGCGCAGCATCTTCAGCTCTTGGGTTTAGCAAAAGTTCATCTCCTTCTGCATCCTTAACTGAGAATGA
GCTTTTGTGGGAGCCCACACCAGTCAAGTTGGATTTGAACCCAGCTGCTCTGTACAAGCTTGGTACCGCCACCATGTACAGCTTCGTG
AGCGAGGAGACCGGCACCCTGATCGTGAACAGCGTGCTGCTGTTCCTGGCCTTCGTGGTGTTCCTGCTGGTGACCCTGGCCATCCTG

ACCGCCCTGAGGCTGTGCGCCTACTGCTGCAACATCGTGAACGTGAGCCTGGTGAAGCCCAGCTTCTACGTGTACAGCAGGGTGAAG
AACCTGAACAGCAGCAGGGTGCCCGACCTGCTGGTGTAATCTAGAGGGCCCGTTTAAACCCGCTGATCAGCCTCGACTGTGCCTTCT
AGTTGCCAGCCATCTGTTGTTTGCCCCTCCCCCGTGCCTTCCTTGACCCTGGAAGGTGCCACTCCCACTGTCCTTTCCTAATAAAATGA
GGAAATTGCATCGCATTGTCTGAGTAGGTGTCATTCTATTCTGGGGGGTGGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAA
GACAATAGCAGGCATGCTGGGGATGCGGTGGGCTCTATGGCTTCTGAGGCGGAAAGAACCAGCTGGGGCTCTAGGGGGTATCCCC
ACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGGTGTGGTGGTTACGCGCAGCGTGACCGCTACACTTGCCAGCGCCCTAGCGCCC
GCTCCTTTCGCTTTCTTCCCTTCCTTTCTCGCCACGTTCGCCGGCTTTCCCCGTCAAGCTCTAAATCGGGGGCTCCCTTTAGGGTTCCGA
TTTAGTGCTTTACGGCACCTCGACCCCAAAAAACTTGATTAGGGTGATGGTTCACGTAGTGGGCCATCGCCCTGATAGACGGTTTTTC
GCCCTTTGACGTTGGAGTCCACGTTCTTTAATAGTGGACTCTTGTTCCAAACTGGAACAACACTCAACCCTATCTCGGTCTATTCTTTTG
ATTTATAAGGGATTTTGCCGATTTCGGCCTATTGGTTAAAAAATGAGCTGATTTAACAAAAATTTAACGCGAATTAATTCTGTGGAAT
GTGTGTCAGTTAGGGTGTGGAAAGTCCCCAGGCTCCCCAGCAGGCAGAAGTATGCAAAGCATGCATCTCAATTAGTCAGCAACCAG
GTGTGGAAAGTCCCCAGGCTCCCCAGCAGGCAGAAGTATGCAAAGCATGCATCTCAATTAGTCAGCAACCATAGTCCCGCCCCTAAC
TCCGCCCATCCCGCCCCTAACTCCGCCCAGTTCCGCCCATTCTCCGCCCCATGGCTGACTAATTTTTTTTATTTATGCAGAGGCCGAGGC
CGCCTCTGCCTCTGAGCTATTCCAGAAGTAGTGAGGAGGCTTTTTTGGAGGCCTAGGCTTTTGCAAAAAGCTCCCGGGAGCTTGTATA
TCCATTTTCGGATCTGATCAAGAGACAGGATGAGGATCGTTTCGCATGATTGAACAAGATGGATTGCACGCAGGTTCTCCGGCCGCT
TGGGTGGAGAGGCTATTCGGCTATGACTGGGCACAACAGACAATCGGCTGCTCTGATGCCGCCGTGTTCCGGCTGTCAGCGCAGGG
GCGCCCGGTTCTTTTTGTCAAGACCGACCTGTCCGGTGCCCTGAATGAACTGCAGGACGAGGCAGCGCGGCTATCGTGGCTGGCCAC
GACGGGCGTTCCTTGCGCAGCTGTGCTCGACGTTGTCACTGAAGCGGGAAGGGACTGGCTGCTATTGGGCGAAGTGCCGGGGCAG
GATCTCCTGTCATCTCACCTTGCTCCTGCCGAGAAAGTATCCATCATGGCTGATGCAATGCGGCGGCTGCATACGCTTGATCCGGCTA
CCTGCCCATTCGACCACCAAGCGAAACATCGCATCGAGCGAGCACGTACTCGGATGGAAGCCGGTCTTGTCGATCAGGATGATCTGG
ACGAAGAGCATCAGGGGCTCGCGCCAGCCGAACTGTTCGCCAGGCTCAAGGCGCGCATGCCCGACGGCGAGGATCTCGTCGTGACC
CATGGCGATGCCTGCTTGCCGAATATCATGGTGGAAAATGGCCGCTTTTCTGGATTCATCGACTGTGGCCGGCTGGGTGTGGCGGAC
CGCTATCAGGACATAGCGTTGGCTACCCGTGATATTGCTGAAGAGCTTGGCGGCGAATGGGCTGACCGCTTCCTCGTGCTTTACGGT
ATCGCCGCTCCCGATTCGCAGCGCATCGCCTTCTATCGCCTTCTTGACGAGTTCTTCTGAGCGGGACTCTGGGGTTCGAAATGACCGA
CCAAGCGACGCCCAACCTGCCATCACGAGATTTCGATTCCACCGCCGCCTTCTATGAAAGGTTGGGCTTCGGAATCGTTTTCCGGGAC
GCCGGCTGGATGATCCTCCAGCGCGGGGATCTCATGCTGGAGTTCTTCGCCCACCCCAACTTGTTTATTGCAGCTTATAATGGTTACA
AATAAAGCAATAGCATCACAAATTTCACAAATAAAGCATTTTTTTCACTGCATTCTAGTTGTGGTTTGTCCAAACTCATCAATGTATCTT
ATCATGTCTGTATACCGTCGACCTCTAGCTAGAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACA
ATTCCACACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGC
TCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATT
GGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTA
ATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGG
CCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGA
CAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCC

GCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGG
GCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTT
ATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTA
ACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATC
CGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCC
TTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTC
ACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAAT
CAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGATAACTACGATACGGGAG
GGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCG
GAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTT
CGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCC
GGTTCCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCAGAA
GTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTG
ACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAATACC
GCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGAT
CCAGTTCGATGTAACCCACTCGTGCACCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAG
GCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAGCATTTAT
CAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAG
TGCCACCTGACGTC