

Structural characterization of membrane-bound human immunodeficiency virus-1 Gag matrix with neutron reflectometry

Rebecca Eells

Department of Physics, Carnegie Mellon University, Pittsburgh, Pennsylvania 15213

Marilia Barros

Memorial Sloan Kettering Cancer Center, New York, New York 10065

Kerry M. Scott and Ioannis Karageorgos

Biomolecular Measurement Division, National Institute of Standards and Technology, Gaithersburg, Maryland 20899 and Institute for Bioscience and Biotechnology Research, Rockville, Maryland 20850

Frank Heinrich^{a)}

Department of Physics, Carnegie Mellon University, Pittsburgh, Pennsylvania 15213 and NIST Center for Neutron Research, National Institute of Standards and Technology, Gaithersburg, Maryland 20899

Mathias Lösche

Departments of Physics and Biomedical Engineering, Carnegie Mellon University, Pittsburgh, Pennsylvania 15213 and NIST Center for Neutron Research, National Institute of Standards and Technology, Gaithersburg, Maryland 20899

(Received 10 March 2017; accepted 26 April 2017; published 16 May 2017)

The structural characterization of peripheral membrane proteins represents a tremendous challenge in structural biology due to their transient interaction with the membrane and the potential multitude of protein conformations during this interaction. Neutron reflectometry is uniquely suited to address this problem because of its ability to structurally characterize biological model systems nondestructively and under biomimetic conditions that retain full protein functionality. Being sensitive to only the membrane-bound fraction of a water-soluble peripheral protein, neutron reflectometry obtains a low-resolution average structure of the protein-membrane complex that is further refined using integrative modeling strategies. Here, the authors review the current technological state of biological neutron reflectometry exemplified by a detailed report on the structure determination of the myristoylated human immunodeficiency virus-1 (HIV-1) Gag matrix associated with phosphoserine-containing model membranes. The authors found that the HIV-1 Gag matrix is able to adopt different configurations at the membrane in a pH-dependent manner and that the myristate group orients the protein in a way that is conducive to PIP₂-binding. © 2017 Author(s). All article content, except where otherwise noted, is licensed under a Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).

[<http://dx.doi.org/10.1116/1.4983155>]

I. INTRODUCTION

A. Structural biology of peripheral membrane proteins

Water-soluble, peripheral membrane proteins perform a wide range of functions within the cell and are involved in cell signaling: the transfer of chemical information into and out of the cell, enzymatic activities on membrane components, regulation of integral membrane proteins, transport of small molecules or electrons, and structural support for the localization of proteins and protein complexes on the membrane. To achieve their broad range of functionalities, peripheral proteins, in many cases, interact only temporarily with lipid membranes or with receptor sites on integral membrane proteins.^{1,2} Even when interactions between the proteins and the membrane are transient, these interactions can lead to structural rearrangements or conformational changes in the proteins that allow them to perform their function.³ As

such, structures of the solution state of these proteins often provide insufficient information to fully understand the biological processes in which they are involved. Reflectometry techniques, in particular neutron reflectometry (NR), offer distinct advantages over traditional structure determination methods, such as crystallography or NMR, for characterizing the protein in a biomimetic membrane environment even though at a lower intrinsic resolution.^{4–7} The development of in-plane fluid tethered lipid membranes for NR (Refs. 8–10) and the implementation of molecular modeling strategies¹¹ have transformed the technique to allow for routine studies of membrane-associated proteins.^{7,12} NR experiments can be performed over a wide range of environmental conditions in terms of ionic strength, pH value, and temperature including—but not limited to—physiologically relevant ranges. NR can also characterize samples in a changing environment such that biological processes can be simulated *in situ*, for example by introducing cofactors or applying external cues between distinct measurements.¹³ A fully atomistic data

^{a)}Electronic mail: fheinrich@cmu.edu

interpretation can be achieved using integrated modeling strategies involving complementary experimental data and molecular dynamics (MD) simulations.^{14–16}

B. HIV-1 Gag MA

Here, we study the myristoylated membrane-targeting domain of human immunodeficiency virus-1 (HIV-1) Gag polyprotein. Gag is the structural factor essential for capsid formation of nascent daughter virus of HIV-1 and other retroviral species.¹⁷ Expressed in the infected host cell, capsid formation of the daughter virus requires Gag trafficking and binding to the plasma membrane (PM), interaction with neighboring Gag proteins to form the immature protein shell of the capsid, and the binding of viral RNA cosynthesized in the hijacked cell. From its N-terminus to its C-terminus, Gag contains four major structural domains to perform these functions: matrix (MA), capsid (CA), nucleocapsid (NC), and p6, as well as connecting spacer peptides.¹⁸ Membrane targeting of Gag is mediated by the myristoylated N-terminal MA domain. In solution, nonmyristoylated MA adopts a compact, globular fold formed by five α -helices near the N-terminus while the C-terminus is more flexible.^{19–22} Myristoylation does not result in large conformational rearrangements of this structure even though the protein is able to adopt two states in which the myristate is either sequestered or exposed to the cytosol.²³

MA recognizes specific PM components and binds to the membrane surface by a variety of physical interactions. The myristate serves as a hydrophobic anchor,^{24–26} while a conserved patch of basic residues interacts electrostatically with negatively charged lipids in the inner leaflet of the PM.^{25,27,28} In addition, the protein exhibits specificity for phosphoinositol-4,5-diphosphate [PI(4,5)P₂],^{29–31} a characteristic marker of the inner PM.³² The fully assembled Gag lattice that forms the protein backbone of the immature viral capsid with its lipid membrane acquired from the host PM has been studied by electron tomography.³³ However, a dissection of the molecular mechanisms that attract the protein to the lipid surface requires a comparative study of the smaller MA domain under well-controlled synthetic conditions. We previously used NR to study the nonmyristoylated matrix and produced a structural model for the electrostatically driven membrane association.³⁴ The protein was observed in a favorable orientation for Gag lattice formation, but it remained unclear if myristoylation would further modulate the structural organization of the protein at the membrane interface. Recently, we quantified the thermodynamics of MA binding to membranes of varying compositions using surface plasmon resonance (SPR) and studied the contribution of the myristate to membrane interactions.³⁵ While SPR provided the energetics of membrane binding, it does not reveal structural information.

Here, we review recent improvements in neutron reflectometry and the required sample preparation. We then discuss the application of NR to HIV-1 myrMA with a particular focus on the challenges associated with studying a protein with multiple binding motifs and reduced solubility

due to the myristoyl moiety. Finally, we report the structural organization of myrMA on charged membranes and compare it with (–myr)MA. Our results confirm the conformational flexibility of myrMA and show the effect of *pH* on the dynamics of the myristoyl sequestration pocket. The combination of hydrophobic and electrostatic interactions leads to a protein orientation, distinct from that under purely electrostatic association, in which key residues are favorably positioned for PIP₂ binding.

II. BIOLOGICAL NEUTRON REFLECTOMETRY

In a typical application to structural biology, NR provides a one-dimensional compositional profile along the normal vector of a lipid bilayer membrane. Thus, a component volume occupancy (CVO) profile along this direction is obtained that accounts for all molecular components of the interfacial architecture including the lipid bilayer and membrane-associated proteins, peptides, and small molecules. At every position along the bilayer normal, the CVO profile represents an in-plane average over planes that are parallel to the lipid bilayer. Thereby, NR yields a temporal and spatial ensemble average over, for example, membrane-associated protein configurations.

A competitive advantage of NR with respect to traditional techniques in structural biology is that the experiment can be carried out using a fully buffer-immersed sample and by maintaining a high flexibility regarding buffer conditions, such as ionic strength and *pH*. NR is nondestructive with the implication that the sample can be manipulated during a NR measurement series by changing, for example, the environmental conditions or adding other molecular cofactors. Biological processes can thus be mimicked, and the structural evolution of the system under study can be monitored.¹³ Specific isotopic labeling of a subset of proteins that form protein–protein complexes allows for the structural characterization of individual proteins within such a complex.^{36,37} In addition, NR provides particular advantages for the study of disordered proteins and peptides, as well as transiently bound peripheral proteins.^{38,39}

NR is intrinsically a low-resolution technique. With current instrumentation, structural features with a thickness of at least ≈ 10 Å, for example, membrane-bound proteins, can be resolved with a spatial resolution as low as 1 Å. Volume occupancy profiles of components filling only 5%–10% of the volume at any position along the surface normal can be reliably determined using the methods outlined below.

A. Biomimetic model membrane systems

A membrane model system useful for biological NR has to meet several criteria. As a technical requirement, it needs to be planar, of low interfacial roughness, long-time stable, and homogenous over a large sample area (cm²). High interfacial roughness and a curved interface negatively affect the resolution of the measurement.⁴⁰ Long-time stability is required due to the comparatively low flux at current neutron sources and, therefore, long measurement times in NR

amounting to several hours per condition. For the same reason, large sample sizes are advantageous as they make better use of the neutron beam. The homogeneity of the sample is required primarily to ensure a unique analysis of the NR data. Inhomogeneous bilayers, even on length scales below the coherence length of the neutron beam, require a more complex modeling and therefore lower the certainty with which the structural features of interest can be determined. For example, a high density of defects in the lipid bilayer constitutes a significant disadvantage for structural characterization of membrane-associated proteins.

As a biological requirement, a model membrane has to be representative of a lipid membrane *in vivo* and is ideally accessible to buffer and protein exchange during the measurement. A flexible model membrane system supports a wide range of relevant lipid compositions while maintaining lipid diffusion rates that are comparable to biological membranes. It is also structurally inert toward changes in environmental conditions—temperature, *pH*, or ionic strength—which helps to retain a focus on the structural changes induced by membrane-associated proteins.

We devised a family of sparsely-tethered lipid membranes (stBLMs) optimized for biological NR that fulfill these criteria.^{8–10} The tether molecules have a central polyethylene oxide chain of 6–9 repeated units, which are functionalized on one end with two hydrocarbon chains—either saturated or unsaturated—that integrate into the hydrocarbon core of a lipid bilayer membrane. The other end is functionalized with either a thiol or thiol acetate group that binds the tethered membrane to a gold-surface. The polyethylene oxide provides about 15 Å of hydrated space between the lipid bilayer and the gold support. Small thiolated molecules coadsorbed with the tethers compete for binding to the gold-support, thus lowering the density of membrane anchors and passivating exposed areas of the gold surface. The relatively thin sub-membrane space limits the types of proteins that can be studied to those that do not have large extra-membraneous domains on both sides of the lipid bilayer. On the other hand, it keeps the membrane interface flat and of low roughness,

which benefits the resolution of the measurements. We demonstrated that diffusion rates in stBLMs are comparable to those observed in unilamellar vesicles⁴¹ and have not yet encountered any limitations in mimicking relevant *in vivo* lipid compositions with the stBLM system.

We integrated the stBLM platform into NR and complementary surface-sensitive characterization techniques such as SPR (Ref. 35) and electrochemical impedance spectroscopy.^{14,42} For peripheral membrane proteins, such complementary characterization is an indispensable tool to identify optimal conditions for structure determination.

B. Composition-space modeling of neutron reflectometry

Applications of NR in structural biology rely on modeling to yield structural information of membrane-bound proteins. This is best achieved using a molecular-scale description paired with an accurate determination of confidence limits on model parameters. We have consequently retired the traditional slab model used in NR,⁴³ replacing it with a composition-space model that yields CVO profiles of molecular and submolecular groups (see Fig. 1).^{7,11} This model makes full use of auxiliary information such as molecular volumes and chemical connectivity, thereby reducing the number of fit parameters and increasing the confidence on unknown parts of the structure. In the case of a tethered lipid bilayer membrane, the volume occupancies of the headgroups are tied to the respective hydrocarbon chains, as individual volumes for both these parts of a lipid molecule are usually known from auxiliary methods such as x-ray diffraction.⁴⁴ Therefore, aside from parameters for the tether molecules, the only three parameters that are required to describe all submolecular groups of a lipid bilayer are the thicknesses of the substrate-proximal and substrate-distal hydrocarbon chains and the bilayer completeness. Headgroup thicknesses are typically too small to be reliably determined by NR, and fixed values obtained, for example, from MD simulations are used instead.

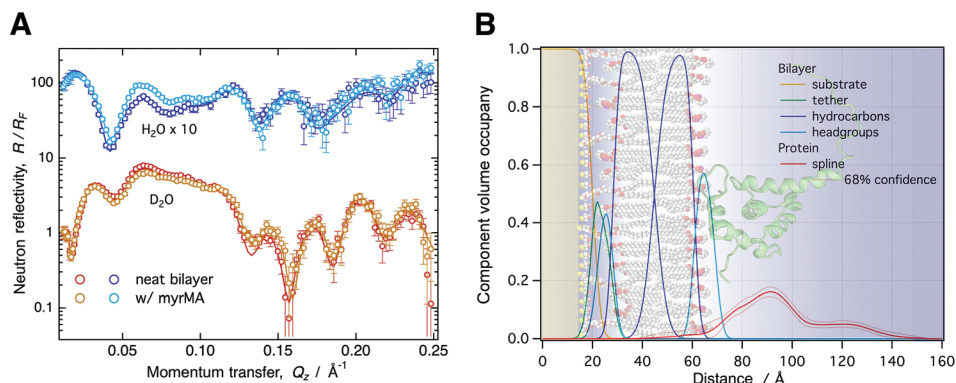


FIG. 1. (a) NR curves normalized by the Fresnel reflectivity (R_F) for a 50:50 DOPC:DOPS stBLM before and after protein addition ($10 \mu\text{M}$ myrMA, *pH* 8.0, 50 mM NaCl). Each condition was characterized using two isotopically distinct bulk solvents (H_2O and D_2O -based buffer) using *in situ* buffer exchange. (b) CVO profiles for the bilayer structure and the membrane-associated protein were obtained by composition-space modeling. The protein orientation was determined by rigid body modeling using the NMR structure in PDB entry 2H3F. The background image visualizes the resulting protein orientation on the stBLM surface.

In distinction to slab models, the composition-space approach allows for spatially overlapping molecular distributions. Molecular profiles of unknown shape, such as those of membrane-associated proteins, require free-form models. We implemented approaches based upon Hermite splines that accurately join protein profiles with CVO profiles of the molecular constituents of the lipid bilayer.⁷ An unbiased determination of modeling uncertainties—a necessity for free-form modeling due to the inherent risk of overparameterization—is carried out using a Monte Carlo Markov Chain optimizer.⁴⁵

C. Integrative modeling of interfacial structures

The intrinsically low resolution of NR in 1D calls for integrative modeling strategies to obtain high-resolution 3D structures of membrane-bound proteins. We routinely integrate available crystallographic or NMR structures into the refinement of NR data from surface-associated proteins (see Fig. 1).^{7,14,15} A comparison of a free-form protein profile and profiles obtained using an atomistically resolved protein structure reveals whether the protein undergoes reorganization upon binding to the membrane. When varying the orientation of the high-resolution protein structure with respect to the membrane (rigid body modeling), such a comparison can determine its orientation if the protein does not undergo significant structural changes, within the resolution of NR, upon binding. However, if no agreement with the free-form profile can be achieved, significant structural changes are evident that require additional modeling. Such an alternative strategy generally involves MD or Monte Carlo simulations of the membrane-bound protein, which can be conducted with steering by the experimental data. Additional experimental data, for example local structural information from NMR, can also be integrated in the form of constraining potentials in the simulation. While we have started utilizing simulation-based integrative modeling strategies,^{15,16,38,46} a complete integrative modeling framework for NR has not yet been established.

III. APPLICATION: THE MEMBRANE BOUND STRUCTURE OF THE MYRISTOYLATED HIV-1 GAG MATRIX

A. Materials and methods

1. Protein expression and purification

Standard laboratory chemicals, culture media, myristic acid, isopropyl- β -D-thio-galactopyranoside (IPTG), and phenylmethylsulfonyl fluoride (PMSF) were purchased from Sigma-Aldrich (St. Louis, MO), unless otherwise noted. Tris(2-carboxyethyl)phosphine hydrochloride (TCEP-HCl) was purchased from Thermo Fisher Scientific (Waltham, MA). The plasmid used for the myrMA preparation via co-expression of MA protein and *N*-myristoyltransferase was kindly provided by Michael Summers (University of Maryland Baltimore County). Transformed *Escherichia coli* BL21 (DE3) cells containing the expression vector were

grown while shaking (250 Hz) at 37 °C to OD₆₀₀ = 0.4. Cells were supplemented with 1 ml of myristic acid (10 mg/ml) per liter of culture and kept growing to OD₆₀₀ = 0.8. At this point, protein expression was induced by adding IPTG to a concentration of 1 mmol/l (1 mM), and cells were kept at 30 °C overnight. The cells were harvested by centrifugation at 6000g for 15 min at 4 °C, washed with phosphate-buffered saline, and held frozen at -80 °C. Five grams (wet-weight) of cells were resuspended in 30 ml of lysis buffer [20 mM Tris, 300 mM NaCl, 10% glycerol, 1 mM PMSF, 1× protease inhibitor mixture set I (Calbiochem; EMD Millipore, Billerica, MA), and 1 mM TCEP, pH 7.4] and disrupted by sonication on ice. The cell lysate was centrifuged at 10 000g for 30 min at 4 °C, and the protein was purified by immobilized metal affinity chromatography. Monomeric MA was separated by size exclusion chromatography on a Superdex-75 10/30 GL column using an AKTA purifier system (Amersham Biosciences, Little Chalfont, UK).

2. Mass spectrometry

Protein identification and purity were established using an electrospray ionization interface on an Agilent 6550 quadrupole time of flight mass spectrometer coupled with an Agilent 1200 high performance liquid chromatography column (Santa Clara, CA). Protein was eluted from a C18 column (3 μ m, 3 × 150 mm; Waters, Milford, MA) over a 30 min gradient from 3% to 60% acetonitrile containing 0.1% formic acid at a flow rate of 9 μ l/min. Data were acquired in the positive ion mode with the following settings: capillary temperature, 290 °C; capillary voltage, 3500 V; fragmentor, 300 V; and a *m/z* 300–3200 mass range. Mass deconvolution was performed using the Agilent MassHunter (version B.06) software.

Peptic peptides of myrMA were generated by passing 17 pmol of protein through an Enzymate pepsin column (Waters, Milford, MA) and identified using tandem mass spectrometry (MS/MS) on a Thermo LTQ Orbitrap Elite unit (Thermo Fisher Scientific, Waltham, MA). One full mass spectral acquisition triggered six scans of MS/MS with activation by collision-induced dissociation on the most abundant precursor ions. Peptides were identified using the MASCOT (Matrix Science, Oxford, UK) database search engine with the following parameters: enzyme, none; oxidation (M) as a variable modification; MS tolerance, 20 ppm; MS/MS tolerance, 0.6 Da; peptide charge of +2, +3, and +4.

3. HDX-MS and HDX data processing

D₂O was obtained from Cambridge Isotope Laboratories Inc. (Andover, MA). For hydrogen-deuterium exchange MS (HDX-MS) analyses, the myrMA protein stock was diluted in H₂O buffer (20 mmol/l Tris, 150 mmol/l sodium chloride, 2 mmol/l TCEP at pH 7.4 and 8.0) to prepare a 5 μ mol/l final concentration, and equilibrated at 1 °C. HDX was conducted on a HDX PAL robot (LEAP Technologies, Carrboro, NC). Protein solutions (5 μ l) were diluted in 25 μ l of D₂O buffer (20 mmol/l Tris, 150 mmol/l sodium chloride, 2 mmol/L TCEP at pH 7.4 and 8.0) at 25 °C. At the selected times

(0 s, 30 s, 5 min, 15 min, 1 h, and 4 h), the HDX sample was quenched by mixing with 35 μ l of quench buffer (3 mol/l urea, 0.1 mol/l sodium phosphate at pH 2.5) at 1 °C. The quenched solution was injected into an on-line immobilized pepsin column for 3 min. The digested protein solution was trapped on a C18 guard column (1.0 mm diameter \times 10 cm length, 5 μ m; Grace Discovery Sciences, Deerfield, IL) and separated with a C18 analytical column (1.0 mm diameter \times 5 cm length, 1.9 μ m, Hypersil GOLD; Thermo Fisher Scientific, Waltham, MA) via a Dionex Ultimate 3000 UPLC with a 9.5 min gradient operated with a binary mixture of solvents, A (water containing 0.1% formic acid) and B (80% acetonitrile and 20% water containing 0.1% formic acid), at a flow rate of 50 μ l/min. The gradient settings were: 5%–35% solvent B for 3 min, 35%–60% solvent B for 5 min, 60%–100% solvent B for 0.5 min, isocratic flow at 100% solvent B for 0.5 min, and a return in 5% solvent B for 0.5 min. LC connection lines and valves were housed in a refrigerated compartment at 2 °C. Peptides were analyzed on a Thermo Orbitrap Elite unit (Thermo Fisher Scientific, Waltham, MA). The instrument settings were: spray voltage, 3.7 kV; sheath gas flow rate, 25 (arbitrary units); capillary temperature, 275 °C. In the Orbitrap stage, MS spectra were acquired with the resolution set at 60 000. Three replicates for each ion-exchange time point were obtained.

From the mass spectra obtained during HDX-MS experiments, the centroid of each deuterated peptide envelope and the relative deuterium uptake by each peptide were calculated using HDX WorkBench (Scripps Research Institute, Jupiter, FL). Corrections for back exchange were made by considering the values of 80% deuterium content of the exchange buffer and an estimated 70% deuterium recovery. Paired t-tests were used to verify deuterium uptake differences.

4. Lipid and liposome preparation

1,2-dioleoyl-*sn*-glycero-3-phosphocholine (DOPC), 1,2-dioleoyl-*sn*-glycero-3-phospho-L-serine (DOPS), 1-palmitoyl-2-oleoyl-*sn*-glycero-3-phosphocholine (POPC), 1-palmitoyl- d_{31} -2-oleoyl-*sn*-glycero-3-phosphocholine (d_{31} -POPC), and 1-palmitoyl-2-oleoyl-*sn*-glycero-3-phospho-L-serine (POPS) were all purchased from Avanti Polar Lipids, Inc. The tether compound HC18 [Z20-(Z-octadec-9-enyloxy)-3,6,9,12,15,18,22-heptaaxatetracont-31-ene-1-thiolacetate] was synthesized and characterized as described.¹⁰ Lipids from stock solutions in chloroform were mixed at desired molar ratios. The organic solvent was removed from the lipid mixtures via evaporation under vacuum for 12 h. The lipid films were hydrated in high salt aqueous buffer (1 M NaCl, 10 mM NaPO₄ at pH 7.4) to a lipid concentration of 5 mg/ml and then sonicated until clear.

5. Preparation of sparsely tethered bilayer lipid membranes

1 \times 3 in. microscopy glass slides (Thermo Fisher Scientific, Waltham, MA; for SPR) and 3 in. diameter, 5 mm

thick n-type Si:P[100] wafers (El-Cat Inc., Ridgefield Park, NJ; for NR) were cleaned with 5 vol. % Hellmanex solution (Hellma Analytics, Müllheim, Germany) and then sulphuric acid plus Nochromix (Godax Laboratories, Cabin John, MD) followed by extensive rinsing with ultrapure water (EMD Millipore, Billerica, MA) and pure ethanol (EtOH, Pharmo-Aaper, Shelbyville, KY) and dried in a N₂ gas stream. The substrates were coated with Cr (\sim 20 Å) and Au (\sim 150 and \sim 450 Å for NR and SPR, respectively) by magnetron sputtering (ATC Orion; AJA International, Scituate, MA). Coated substrates were immediately soaked in a 7:3 (mol/mol) ethanol solution of HC18 and β -mercaptoethanol (BME) at a total concentration of 0.2 mM to form a self-assembled monolayer (SAM). A 5:5 (mol/mol) HC18:BME ratio was used for the preparation of the 70:30 d_{31} -POPC:POPS stBLM for NR to aid complete bilayer formation. Vesicle solutions were allowed to incubate the dry SAM for \sim 2 h, and then the system was flushed with low ionic strength buffer (50 mM NaCl, 10 mM NaPO₄, pH 7.4) to complete stBLM formation.

6. Neutron reflectometry

NR measurements were performed at the NG7 horizontal and CGD-Magik⁴⁷ reflectometers at the NIST Center for Neutron Research (NCNR). Reflectivity curves were recorded for momentum transfer values $0.01 \leq q_z \leq 0.25 \text{ \AA}^{-1}$. For each measurement, adequate counting statistics were obtained after 5–7 h. The NCNR flow cell⁴⁵ allows for *in situ* buffer exchange; thereby, subsequent measurements were performed on the same sample area. The entire flow cell was maintained at room temperature (RT). After *in situ* completion of the stBLM, NR data were sequentially collected with H₂O buffer and D₂O buffer in the measurement cell. Buffer exchange was accomplished by flushing \sim 10 ml of buffer through the cell (volume \sim 1.3 ml) using a syringe.

Two methods were used to prepare protein for the measurement. If the protein stock buffer did not match the desired working buffer, spin columns (Micro Bio-Spin 6 Columns; Bio-Rad, Hercules, CA) were used for buffer exchange followed by a measurement of the protein concentration post exchange. Protein at the desired concentration (typically 10 μ M) was then prepared by diluting an aliquot from the stock with the working buffer (H₂O or D₂O) to a final volume of \sim 1.5 ml. If the protein stock buffer was the same as the working buffer, an aliquot of protein from the main stock was diluted to achieve the desired concentration. Protein was introduced to the NR cell via a syringe, and its incubation with the bilayer was measured in both H₂O and D₂O buffer contrasts. The system was then rinsed and measured in both contrasts again to detect any remaining, tightly bound protein at the membrane interface.

1D-structural profiles along the lipid bilayer normal were parameterized using a stratified slab model for the solid substrate,⁴³ a continuous distribution model for the stBLM,¹¹ and a monotonic Hermite spline for the model-free protein distribution.⁷ Individual slabs were implemented for the bulk

silicon, silicon oxide, chromium, and gold layers. Fit parameters are thickness and neutron scattering length density (nSLD) for each layer, except for the bulk silicon for which the nSLD is known. One global roughness fit parameter was applied to all the substrate interfaces. Individual submolecular groups implemented in the continuous distribution model were: β ME, tether polyethylene glycol chains, tether glycerol groups, substrate-proximal and substrate-distal phosphatidylcholine and phosphatidylserine (PS) headgroups, substrate-proximal and substrate-distal methylene chains and methyls of lipid and tether molecules. Fit parameters were the bilayer hydrocarbon thickness for each bilayer leaflet, bilayer completeness, tether surface density, tether thickness, and β ME surface density. One roughness fit parameter was applied to all distributions.

The Hermite spline that described the CVO profiles of the protein was defined by control points which were on average 15 Å apart. The spatial extension of the protein along the bilayer normal determined the number of control points which were iteratively refined during model optimization. Fit parameters for each control point were the volume occupancy of the envelope and deviations from the equidistant separation of control points.

To determine the orientation of myrMA on the membrane, a high-resolution NMR structure [protein data bank (PDB) entry 2H3F]³⁰ was used within a rigid body modeling approach after hydrogens were added using MolProbity.⁴⁸ While this structure describes the unmyristoylated form of the protein, it is identical within the NR resolution to the NMR structure for myrMA (PDB entry 1UPH²³) for the compact core-region of MA, residues Val-7–Ile-104 (0.5 ± 0.1 Å root mean square deviation).²³ The conformation of the N-terminus depends on the state of the myristoyl group (sequestered, free, membrane-inserted), and we expect that the structure of the N-terminal in PDB entry 2H3F would best describe that of the membrane-bound state.

Two continuous fit parameters, the two Euler angles β and γ , describe the protein orientation at the interface. The reference orientation $(\beta, \gamma) = (0^\circ, 0^\circ)$ is defined by the orientation of the protein given in PDB entry 2H3F and in the reference orientation; by definition, the intrinsic protein coordinate system and the extrinsic bilayer coordinate system are aligned. The z-axis in the PDB file is aligned parallel to the bilayer normal and points toward the substrate-distal side of the membrane. The x and y-axes are parallel to the plane of the bilayer. Orientations (β, γ) are then obtained by extrinsic rotations of the protein around the axes of the bilayer coordinate system. First, the protein is rotated by $0^\circ \leq \gamma < 360^\circ$ counter-clockwise about the membrane normal, z. Then, the protein is rotated by $0^\circ \leq \beta \leq 90^\circ$ about the x-axis of the bilayer coordinate system (extrinsic rotation) that generally differs from a rotation around the (intrinsic) protein coordinate system at this point.

Optimization of model parameters was performed using the *ga_refl* and *Refl1D* software packages developed at the NCCR.⁴⁵ All the reflectivity curves of one dataset were fit simultaneously with the same model, sharing fit parameters, for example, for the solid substrate. A Monte Carlo Markov

Chain-based global optimizer⁴⁵ was used to determine fit parameter confidence limits.

7. Surface plasmon resonance

SPR measurements were conducted in a single-batch mode at (25.00 ± 0.01) °C using a custom-built instrument (SPR Biosystems, Germantown, MD). Gold-coated glass slides with a SAM layer were assembled in the Kretschmann configuration by index matching to a prism and stBLMs prepared via vesicle fusion *in situ*, as described above. A 2D-CCD detector records the intensity of light reflected from the glass/buffer interface that carries the membrane, and the position of the intensity minimum is recorded as a function of time. SPARia (SPR Biosystems) was used for real-time acquisition of the response, R , measured as the displacement of the reflection minimum on the CCD (in pixels). To determine a baseline, the neat bilayer was measured before adding protein in increasing concentrations. Time courses of R were recorded for each protein concentration, c_p , until equilibrated at R_{eq} . The change in R_{eq} as a function of c_p was fitted with a Langmuir isotherm,

$$R_{eq} = \frac{R_\infty \cdot c_p}{c_p + K_d},$$

where K_d is the equilibrium dissociation constant and R_∞ is the saturation of the SPR response as $c \rightarrow \infty$.

B. Results and discussion

1. MA membrane binding depends on pH

Biophysical studies of protein membrane binding are ideally carried out under physiological conditions to allow for an extrapolation of the experimental results to the *in vivo* situation. In practice, however, this is often impossible. As a truncation product of full-length Gag, the isolated MA domain studied here lacks self-interaction mechanisms that promote Gag multimerization at the membrane via CA dimerization and RNA-binding by NC.⁴⁹ Therefore, a reduction in the ionic strength was necessary to shift the binding equilibrium of MA for a characterization of the membrane-bound state.^{34,35} Other techniques, such as NMR, required a lower pH to permit measurements of protein–lipid interactions.³¹

Based on a comprehensive set of myrMA binding data to lipid membranes,³⁵ optimal buffer conditions for a structural characterization using NR were found to be pH 7.4 and 50 mM NaCl where the membrane dissociation constant is $K_d \sim 5 \mu\text{M}$. While SPR indicated Langmuir binding behavior, NR revealed membrane remodeling and proteinaceous multilayers at the interface under these conditions, which prohibited a conclusive structural characterization of membrane-bound myrMA, in contrast to (–myr)MA which forms well-defined protein monolayers at the membrane.³⁴ NMR studies of myrMA showed that the myristate group can adopt sequestered and exposed states with only minor conformational changes.²³ MA oligomerization has been associated with myristate exposure and also depends on

protein concentration and pH .^{23,50} At pH 7.0, the monomer/trimer dissociation constant is $(1.4 \pm 0.2) 10^{-8} M^2$, whereas at pH 8, myrMA was found to be purely monomeric.⁵⁰ At a protein concentration of $10 \mu M$ used in NR, the trimer concentration is $<50 nM$, insignificant for membrane binding. However, if the trimer assembles in a lattice-competent structure, binding interfaces from more than one monomer would lead to a significant enhancement in binding and could explain the formation of protein multilayers.

An increase in the buffer pH to 8.0 shifted the monomer/trimer equilibrium entirely toward the monomer. SPR experiments on a 70:30 POPC:POPS stBLM showed a reduction in the binding affinity by a factor of ~ 4 (Fig. 2), consistent with a shift toward the myristoyl-sequestered state^{31,50} and a change in protein charge from $+3.7e$ to $+2.9e$.⁵¹ The saturation surface coverage was not significantly affected by the pH change, and the binding curve is well described by the Langmuir model.

2. pH Dependence of MA conformational flexibility

myrMA was analyzed by HDX-MS to probe the effect of pH on the conformational flexibility in solution and to correlate dynamic changes with myristic acid exposure. The experiments detect the exchange of backbone amide hydrogen atoms of individual amino acids and reveal the extent of amide hydrogen bonding and solvent accessibility of a protein. By comparing measurements for different conditions, such as varying pH , changes in protein conformation and/or dynamics can be determined. The data (see supplementary material, Figs. S1–S4, for more detail)⁵⁵ reveal that 34% of reporting amides were affected by pH changes from 7.4 to 8.0 and exhibit higher deuterium uptake rates over time

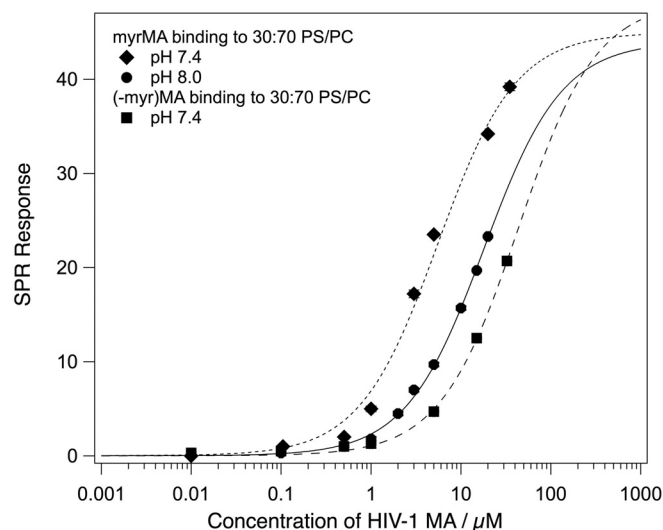


FIG. 2. MA binding to stBLMs containing 30% PS at 50 mM NaCl. The equilibrium response is plotted as a function of MA monomer concentration and fitted with the Langmuir isotherm, and the binding curve for (–myr)MA at pH 7.4 is included for comparison. K_d of myrMA increases from ~ 5 to $\sim 18 \mu M$ between pH 7.4 and 8 but remains significantly lower than for (–myr)MA at either pH . This shows that the myristate still contributes to membrane binding at pH 8.

between 6% and 30% (Fig. S5). The remaining 66% of the protein was not affected by pH changes. The overall deuterium uptake rate increase suggests that myrMA becomes more dynamic in a number of distinct regions at pH 8.0, and the majority of those regions belong to the helices that form the myristate pocket (see Fig. 3). We conclude that the higher pH increases the flexibility of the pocket, which promotes myristate sequestration at pH 8.0 in accordance with SPR binding data and literature.^{23,50}

3. MA adopts different configurations at the membrane depending on myristoylation, pH , and surface charge

To directly probe if the optimized conditions prevent the formation of overlayers, $10 \mu M$ myrMA was studied with NR at a 70:30 d_{31} -POPC:POPS stBLM at pH 8 and 50 mM NaCl (see Fig. S6). The thus obtained free-form protein CVO profile indicated a single layer of protein at the membrane interface (Fig. 4). However, a detailed analysis revealed that this profile is inconsistent with a single protein conformation at the membrane as it exhibits a broad maximum that exceeds the dimensions of a single MA molecule by $\approx 50\%$. Rinsing the sample with buffer did not significantly alter the profile (data not shown), indicating stable and irreversible protein binding. This result agrees with coarse-grained MD simulations of myrMA (Ref. 52) and electrostatic modeling⁵³ that suggested dynamic binding of myrMA in multiple conformations to the membrane. We concluded that a slight tendency for protein multimerization at the interface remains a possibility. In addition, the low volume occupancy of the protein (5%–10%) leaves the analysis vulnerable to small systematic errors, which may be as large as 3%, in the CVO profile.

In a subsequent NR experiment, we increased the fraction of PS to 50% to yield a higher protein surface coverage and provide a more homogeneous binding interface, potentially reducing the propensity of the protein to adopt multiple configurations at the membrane [Fig. 1(a)]. These conditions, indeed, yielded a CVO profile compatible with a single myrMA configuration at the membrane [see Figs. 1(b) and 5(a)]. The orientation of myrMA was determined using the

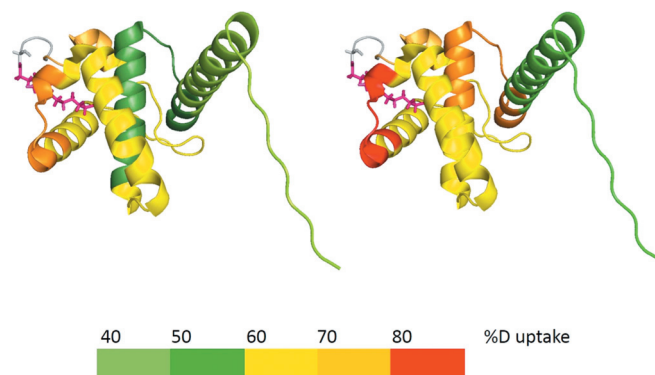


FIG. 3. Deuteration profile of myrMA at pH 7.4 (left) and pH 8.0 (right) after 30 min of incubation in D_2O .

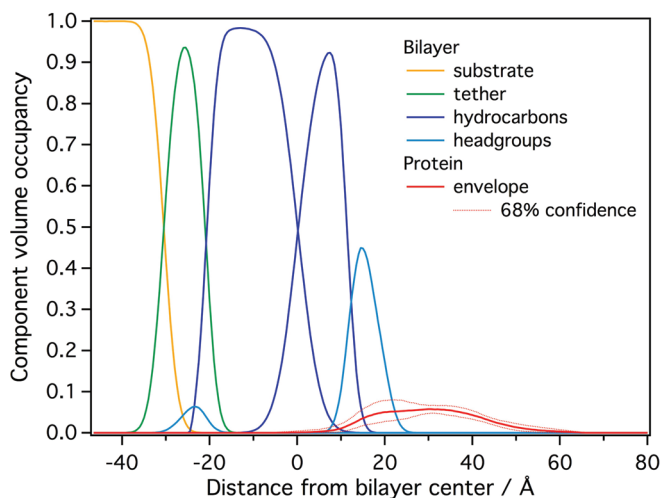


FIG. 4. NR CVO profile of myrMA on a 70:30 d_{31} -POPC:POPS stBLM at pH 8.0 after incubation with $10 \mu\text{M}$ protein. The median protein envelope, shown with 68% confidence intervals (red traces), is inconsistent with a single MA conformation at the membrane interface.

ensemble average of the NMR structures in PDB entry 2H3F [see Fig. 5(b)]. A comparison of the free-form CVO spline profile with the result of the rigid body modeling shows a very good agreement for the folded core but an imperfect agreement for the flexible N and C termini. The observed

extra density of the spline CVO in the membrane region is consistent with myristate insertion and the insertion of N-terminal residues. This interpretation is supported by recent NR results that indicated a deep penetration of a peptide that represents myrMA truncated to its residues 2–32 into the hydrocarbon region of the lipid bilayer.⁵⁴ The observed discrepancy between the free-form spline CVO profile and the profile based upon the NMR structure in the flexible C-terminal region shows that the ensemble of solution structures is too compact. While the orientation of the myrMA at the membrane is largely constrained by the asymmetric rigidly folded core, the discrepancies in both the termini introduce some uncertainty to the final result. Ultimately, an integrative modeling approach will be required to find a fully satisfying interpretation of the free-form profile of the membrane-associated myrMA.

In the most likely orientation of myrMA at the membrane [$\beta \sim 20^\circ$, $\gamma \sim 335^\circ$, shown in Fig. 5(b)], helix I and residues 31–35 of helix II are in close contact with the lipid membrane surface. While the overall membrane penetration is shallow, residues in the highly basic region [HBR; residues 15–31 (Ref. 25)] penetrate more deeply into the bilayer. In particular, lysine residues K26, K27, K30, and K32 penetrate deeply into the headgroup layer. Based on simulations and the previous NR measurements of (–myr)MA, these residues mediate

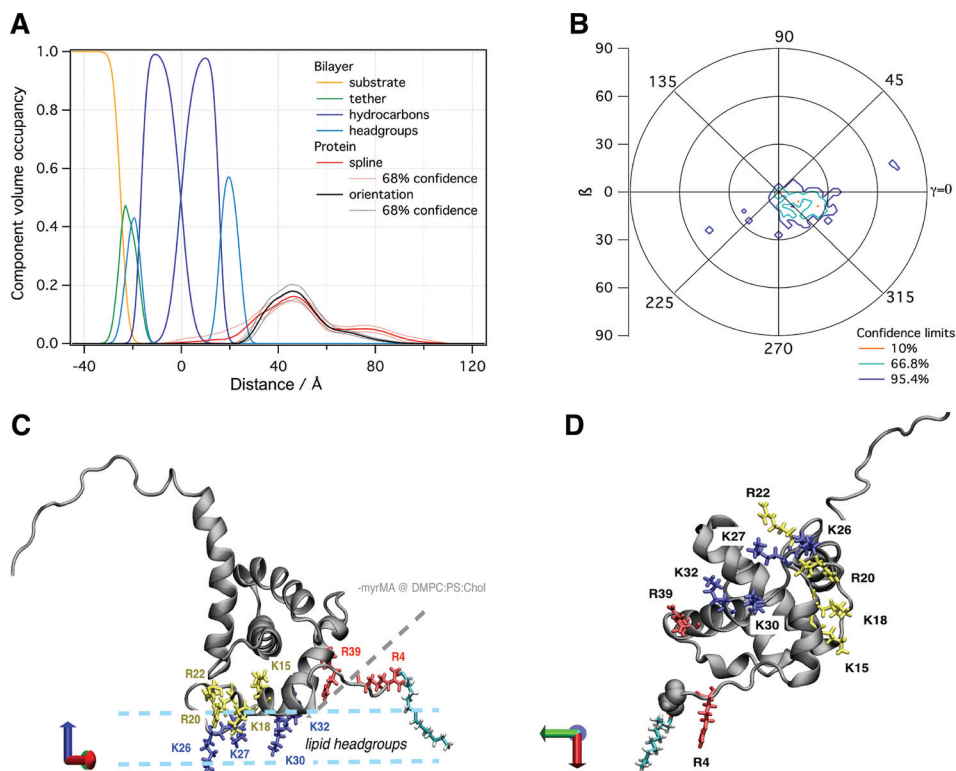


FIG. 5. (a) NR CVO profile of myrMA on a 50:50 DOPS/DOPC stBLM at pH 8, 50 mM NaCl after incubation with $10 \mu\text{M}$ protein. The median protein envelope is shown with 68% confidence intervals (red traces). The median orientation fit using the MA NMR structure (PDB entry 2H3F) is shown for comparison (black trace). (b) Probability distribution of myrMA orientations with respect to the 50:50 DOPS/DOPC bilayer normal. (c) The most likely orientation of the protein ($\beta \sim 20^\circ$, $\gamma \sim 335^\circ$) on the membrane. Lysine residues (K26, K27, K30, and K32) that penetrate deeply into the lipid headgroup region are highlighted in blue. Basic residues with a peripheral interaction (K15 and R22) or slight penetration (K18 and R20) are shown in yellow. Arginine residues (R4 and R39) that were previously shown in (–myr)MA to interact closely with the membrane but are more peripheral in myrMA are shown in red. A myristate group, shown in cyan, was added to the protein structure to highlight its location. (d) Bottom view of myrMA in its membrane bound orientation.

membrane interactions via electrostatic interactions when PIP₂ cannot be engaged.^{34,52} When PIP₂ is present, these residues are still involved in membrane interactions. K30 and K32, specifically, were shown to be important for PIP₂ binding,³¹ and mutation of these residues to glutamate retargeted Gag to intracellular compartments.²⁸ In addition, R4 was also identified to interact with the membrane,^{31,34} but rigid body modeling of the NR data does not place this residue in direct contact with the headgroup region. On the other hand, with the proposed flexibility of the N terminus, R4 may penetrate the lipid headgroups as well.

Other HBR residues also interact closely with the membrane in our NR model: K15 and K18 from helix I and R20 and R22 from the loop connecting helix I and II. Based on the rigid body modeling, K18 and R20 penetrate slightly into the headgroup layer, and K15 and R22 may also insert into the headgroup region depending on their side-chain conformations. The NR characterization of (–myr)MA also identified these residues as interacting with the membrane, albeit more peripherally. Notably, R39, which showed a significant overlap with the headgroup region for (–myr)MA, is further away from the membrane surface for myrMA. The change in the orientation between the (–myr)MA and myrMA structures brings helix I and the HBR into more direct contact with the membrane surface while positioning R39 on helix II further away [Fig. 5(c)]. In simulations, a similar change in orientation was observed between the myristate exposed and myristate sequestered states, and myristate–membrane interactions resulted in a similar shift in membrane contacts from helix II to helix I.⁵² Although the membranes used in this study do not contain PIP₂, our NR results show that the myristate interaction with the membrane positions the HBR such that it is conducive to PIP₂ binding.

IV. CONCLUSION

Structural characterization of peripheral membrane proteins poses unique challenges due to their interaction with the membrane, which is often transient, and the variety of protein conformations that can occur during this interaction. Detailed knowledge of membrane binding characteristics and aggregation behavior of a protein is required before a structural characterization can be pursued. A deviation from *in vivo* conditions, for example, by decreasing the ionic strength or altering the pH of the buffer, is often necessary to stabilize one particular conformation.

The current study of myrMA is an important step toward a full structural characterization of HIV-1 Gag membrane binding and viral assembly. We identified experimental conditions that overcome the challenges presented by this particular peripheral membrane protein. Our results for membrane-bound myrMA showed deviations of the membrane-bound protein structure from the high-resolution NMR structure in solution. By comparing the myrMA structure with the previously determined structure of (–myr)MA, we observed that their membrane complexes are similar to each other. However, the presence of the myristate results in a reorientation of MA

that positions the HBR to allow for PIP₂ engagement. This orientation agrees with simulations that suggested strong interactions between the membrane and helix I for the myristate-exposed form of the protein. In addition, the key basic residues K26, K30, and K32 were found to interact with PS when PIP₂ lipids could not be engaged, and we observe penetration of these residues into the lipid headgroup region of PS-containing membranes in the absence of PIP₂.

ACKNOWLEDGMENTS

The authors thank D. J. Vanderah for providing HC18 and M. Summers for the plasmid used for myrMA expression. R. Eells thanks the Molecular Biophysics and Structural Biology (MBSB) program for support through a NIH Training Grant (T32 GM88119). Research was performed in part at the National Institute of Standards and Technology (NIST) Center for Nanoscale Science and Technology. This work was supported by the NIH (R01GM101647) and the NIST IMS program “Precision Measurements for Integral Membrane Proteins.” Certain commercial materials, equipment, and instruments are identified in this work to describe the experimental procedure as completely as possible. In no case does such an identification imply a recommendation or endorsement by NIST nor does it imply that the materials, equipment, or instrument identified are necessarily the best available for the purpose.

¹S. Oellerich, S. Lecomte, M. Paternostre, T. Heimburg, and P. Hildebrandt, *J. Phys. Chem. B* **108**, 3871 (2004).

²A. M. Whited and A. Johs, *Chem. Phys. Lipids* **192**, 51 (2015).

³S. A. Tatulian, *Methods Enzymol.* **583**, 197 (2017).

⁴H. P. Wacklin, *Curr. Opin. Colloid Interface Sci.* **15**, 445 (2010).

⁵G. Fragneto, *Eur. Phys. J.-Spec. Top.* **213**, 327 (2012).

⁶L. A. Clifton, C. Neylon, and J. H. Lakey, *Methods Mol. Biol.* **974**, 119 (2013).

⁷F. Heinrich and M. Lösche, *Biochim. Biophys. Acta* **1838**, 2341 (2014).

⁸D. J. McGillivray, G. Valincius, D. J. Vanderah, W. Febo-Ayala, J. T. Woodward, F. Heinrich, J. J. Kasianowicz, and M. Lösche, *Biointerphases* **2**, 21 (2007).

⁹F. Heinrich, T. Ng, D. J. Vanderah, P. Shekhar, M. Mihailescu, H. Nanda, and M. Lösche, *Langmuir* **25**, 4219 (2009).

¹⁰R. Budvytyte *et al.*, *Langmuir* **29**, 8645 (2013).

¹¹P. Shekhar, H. Nanda, M. Lösche, and F. Heinrich, *J. Appl. Phys.* **110**, 102216 (2011).

¹²M. Belicka, Y. Gerelli, N. Kučerka, and G. Fragneto, *Soft Matter* **11**, 6275 (2015).

¹³S. A. K. Datta, F. Heinrich, S. Raghunandan, S. Krueger, J. E. Curtis, A. Rein, and H. Nanda, *J. Mol. Biol.* **406**, 205 (2011).

¹⁴D. J. McGillivray, G. Valincius, F. Heinrich, J. W. F. Robertson, D. J. Vanderah, W. Febo-Ayala, I. Ignatjev, M. Lösche, and J. J. Kasianowicz, *Biophys. J.* **96**, 1547 (2009).

¹⁵F. Heinrich, H. Nanda, H. Z. Goh, C. Bachert, M. Lösche, and A. D. Linstedt, *J. Biol. Chem.* **289**, 9683 (2014).

¹⁶H. Nanda, F. Heinrich, and M. Lösche, *Methods* **77–78**, 136 (2014).

¹⁷I. M. Jones and Y. Morikawa, *Rev. Med. Virol.* **8**, 87 (1998).

¹⁸A. Ono, *Future Virol.* **4**, 241 (2009).

¹⁹M. A. Massiah, M. R. Starich, C. Paschall, M. F. Summers, A. M. Christensen, and W. I. Sundquist, *J. Mol. Biol.* **244**, 198 (1994).

²⁰S. Matthews, P. Barlow, N. Clark, S. Kingsman, A. Kingsman, and I. Campbell, *Biochem. Soc. Trans.* **23**, 725 (1995).

²¹S. Matthews *et al.*, *Nature* **370**, 666 (1994).

²²C. P. Hill, D. Worthylake, D. P. Bancroft, A. M. Christensen, and W. I. Sundquist, *Proc. Natl. Acad. Sci. U. S. A.* **93**, 3099 (1996).

- ²³C. Tang, E. Loeliger, P. Luncsford, I. Kinde, D. Beckett, and M. F. Summers, *Proc. Natl. Acad. Sci. U. S. A.* **101**, 517 (2004).
- ²⁴M. Bryant and L. Ratner, *Proc. Natl. Acad. Sci. U. S. A.* **87**, 523 (1990).
- ²⁵W. J. Zhou, L. J. Parent, J. W. Wills, and M. D. Resh, *J. Virol.* **68**, 2556 (1994).
- ²⁶P. Spearman, R. Horton, L. Ratner, and I. Kuli-Zade, *J. Virol.* **71**, 6582 (1997).
- ²⁷E. O. Freed, G. Englund, and M. A. Martin, *J. Virol.* **69**, 3949 (1995).
- ²⁸A. Ono, J. M. Orenstein, and E. O. Freed, *J. Virol.* **74**, 2855 (2000).
- ²⁹W. J. Zhou and M. D. Resh, *J. Virol.* **70**, 8540 (1996).
- ³⁰J. S. Saad, J. Miller, J. Tai, A. Kim, R. H. Ghanam, and M. F. Summers, *Proc. Natl. Acad. Sci. U. S. A.* **103**, 11364 (2006).
- ³¹P. Y. Mercredi *et al.*, *J. Mol. Biol.* **428**, 1637 (2016).
- ³²G. van Meer, D. R. Voelker, and G. W. Feigenson, *Nat. Rev. Mol. Cell Biol.* **9**, 112 (2008).
- ³³G. Zhao *et al.*, *Nature* **497**, 643 (2013).
- ³⁴H. Nanda, S. A. K. Datta, F. Heinrich, M. Lösche, A. Rein, S. Krueger, and J. E. Curtis, *Biophys. J.* **99**, 2516 (2010).
- ³⁵M. Barros, F. Heinrich, S. A. K. Datta, A. Rein, I. Karageorgos, H. Nanda, and M. Lösche, *J. Virol.* **90**, 4544 (2016).
- ³⁶F. Heinrich, *Methods Enzymol.* **566**, 211 (2016).
- ³⁷T. L. Yap, Z. Jiang, F. Heinrich, J. M. Gruschus, C. M. Pfefferkorn, M. Barros, J. E. Curtis, E. Sidransky, and J. C. Lee, *J. Biol. Chem.* **290**, 744 (2015).
- ³⁸C. M. Pfefferkorn, F. Heinrich, A. J. Sodt, A. S. Maltsev, R. W. Pastor, and J. C. Lee, *Biophys. J.* **102**, 613 (2012).
- ³⁹Z. Jiang, S. K. Hess, F. Heinrich, and J. C. Lee, *J. Phys. Chem. B* **119**, 4812 (2015).
- ⁴⁰C. F. Majkrzak, E. Carpenter, F. Heinrich, and N. F. Berk, *J. Appl. Phys.* **110**, 102212 (2011).
- ⁴¹S. Shenoy, R. Moldovan, J. Fitzpatrick, D. J. Vanderah, M. Deserno, and M. Lösche, *Soft Matter* **2010**, 1263.
- ⁴²G. Valincius, F. Heinrich, R. Budvytyte, D. J. Vanderah, D. J. McGillivray, Y. Sokolov, J. E. Hall, and M. Lösche, *Biophys. J.* **95**, 4845 (2008).
- ⁴³J. F. Ankner and C. F. Majkrzak, *Proc. SPIE* **1738**, 260 (1992).
- ⁴⁴J. F. Nagle and S. A. Tristram-Nagle, *Biochim. Biophys. Acta* **1469**, 159 (2000).
- ⁴⁵B. J. Kirby, P. A. Kienzle, B. B. Maranville, N. F. Berk, J. Krycka, F. Heinrich, and C. F. Majkrzak, *Curr. Opin. Colloid Interface Sci.* **17**, 44 (2012).
- ⁴⁶J. E. Curtis, H. Zhang, and H. Nanda, *Comput. Phys. Commun.* **185**, 3010 (2014).
- ⁴⁷J. A. Dura *et al.*, *Rev. Sci. Instrum.* **77**, 74301 (2006).
- ⁴⁸V. B. Chen, W. B. I. Arendall, J. J. Headd, D. A. Keedy, R. M. Immormino, G. J. Kapral, L. W. Murray, J. S. Richardson, and D. C. Richardson, *Acta Crystallogr., D* **66**, 12 (2010).
- ⁴⁹V. Chukkappalli and A. Ono, *J. Mol. Biol.* **410**, 512 (2011).
- ⁵⁰E. L. Fledderman, K. Fujii, R. H. Ghanam, K. Waki, P. E. Prevelige, E. O. Freed, and J. S. Saad, *Biochemistry* **49**, 9551 (2010).
- ⁵¹B. Bjellqvist, G. J. Hughes, C. Pasquali, N. Paquet, F. Ravier, J. C. Sanchez, S. Frutiger, and D. Hochstrasser, *Electrophoresis* **14**, 1023 (1993).
- ⁵²L. Charlier, M. Louet, L. Chaloin, P. Fuchs, J. Martinez, D. Muriaux, C. Favard, and N. Floquet, *Biophys. J.* **106**, 577 (2014).
- ⁵³P. S. Murray, Z. H. Li, J. Y. Wang, C. L. Tang, B. Honig, and D. Murray, *Structure* **13**, 1521 (2005).
- ⁵⁴L. O'Neil *et al.*, *Biochim. Biophys. Acta* **1858**, 3071 (2016).
- ⁵⁵See supplementary material at <http://dx.doi.org/10.1116/1.4983155> for information on the protein expression and characterization, and additional HDX-MS and NR results.