HHS Public Access

Author manuscript

J Am Chem Soc. Author manuscript; available in PMC 2023 December 28.

Published in final edited form as:

J Am Chem Soc. 2022 December 28; 144(51): 23368–23378. doi:10.1021/jacs.2c09160.

Characterizing heterogeneous mixtures of assembled states of the tobacco mosaic virus using charge detection mass spectrometry

Amanda J. Bischoff[1,2], **Conner C. Harper**[1], **Evan R. Williams**[1], **Matthew B. Francis**[1,2]

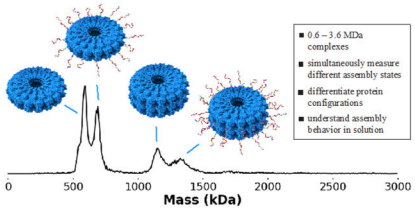
[1]College of Chemistry, University of California, Berkeley, California, 94720, United States

[2]Molecular Biophysics and Integrated Bioimaging Division, Lawrence Berkeley National Laboratories, Berkeley, California, 94720, United States

Abstract

The tobacco mosaic viral capsid protein (TMV) is a frequent target for derivatization for myriad applications including drug delivery, biosensing, and light harvesting. However, solutions of the stacked disk assembly state of TMV are difficult to characterize quantitatively due to their large size and multiple assembled states. Charge detection mass spectrometry (CDMS) addresses the need to characterize heterogeneous populations of large protein complexes in solution quickly and accurately. Using CDMS, previously unobserved assembly states of TMV, including 16-monomer disks and odd-numbered disk stacks, have been characterized. We additionally employed a peptide-protein conjugation reaction in conjunction with CDMS to demonstrate that modified TMV proteins do not redistribute between disks. Finally, this technique was used to discriminate between protein complexes of near-identical mass but different configurations. We have gained a greater understanding of the behavior of TMV, a protein used across a broad variety of fields and applications, in the solution state.

Graphical Abstract



Corresponding Author: **Matthew B. Francis** – Department of Chemistry, University of California, Berkeley, California, 94720, United States; Molecular Biophysics and Integrated Bioimaging Division, Lawrence Berkeley National Laboratories, Berkeley, California 94720; mfrancis@berkeley.edu.

Author Contributions

The manuscript was written through contributions of all authors. / All authors have given approval to the final version of the manuscript.

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website.

Additional experimental details, materials, methods and supporting figures (PDF)

Python script for theoretical mass distribution calculations (txt)

The authors declare that they have no competing interests.

INTRODUCTION

Particles derived from the tobacco mosaic virus (TMV) have broad potential in drug and vaccine development,[1–3] sensory assays,[4,5] filtration,[6] and as mimics of photosynthetic light harvesting complexes,[7–9] among myriad other uses.[10,11] This is partially due to TMV's unique assembly state as a 300 nm helix in its native form or as stacked disks under varied environmental conditions.[12] TMV also has a remarkable capacity for assembling into disk or rod structures when subject to a vast array of mutations to the protein surface, in the pore, and within the cavity between the disks.[8,13–16] The ability of a variety of mutants and conjugates of TMV to maintain ordered assemblies has enabled its functionalization with therapeutics for drug delivery applications,[17,18] or dyes or nanoparticles in synthetic light harvesting systems.[7,8,19] TMV can also be assembled into nanoscale materials with morphologies that do not occur naturally, such as nanosheets exhibiting filtration ability[6] and spherical nanoparticles used for drug delivery.[20, 21]

In the absence of a nucleic acid template, TMV can form multiple assembly states sensitive to environmental conditions. At lower pH and ionic strength TMV monomers assemble into helical rods with 16 $^{2}/_{3}$ monomers per helical turn, similar to the native structure, but assemblies of stacked disks consisting of 17 monomers per disk are formed at neutral pH and higher ionic strengths (PDB ID 1EI7, Figure 1A–C).[22] Mutations made to the TMV capsid can also bias the structure toward forming rods or stacked disks. One example of this is a circular permutant, cpTMV, which favors two-disk stacks under a wider variety of pH and ionic strength conditions than native TMV (Figure 1D–E).[13] In its stacked disk form, solutions of TMV often consist of heterogeneous assembled states of even-numbered disk stacks,[13,15] which can make detailed characterization challenging. Progress has been made recently in low-resolution characterization of TMV particles of various size in the solution state using small angle x-ray scattering.[23] However, probing how single mutations may effect subtle structural changes within a single disk or change the equilibrium of multiple assembly populations present in solution remains a challenge.

TMV conjugates have been studied extensively for therapeutic applications, and we have previously made use of TMV conjugates as a tunable platform for modeling energy transfer processes in photosynthetic light harvesting complexes. These models consist of TMV labeled covalently with synthetic dyes in circular arrays on the assembly surface or within the region between individual disks.[7,8] In measurements of energy transfer in these systems, a detailed understanding of the inter-chromophore distances and relative orientations, as well as protein environment, is crucial to interpreting spectroscopic data.[24] In addition, a concrete picture of the timescales of disassembly and reassembly of the complexes in solution is necessary for experiments involving energy transfer between pairs of donor and acceptor complexes. In the context of TMV conjugate drug and vaccine candidates, a clear analysis of all species present and their assembly kinetics is also vital to understanding the behavior of these conjugates *in vivo*.[25–27] The use of high-resolution structural characterization techniques, including x-ray crystallography, cryo-electron microscopy, and NMR spectroscopy, to characterize protein samples at the unique intersection of large (MDa+) size and sample heterogeneity characteristic of TMV conjugates is typically challenging and labor-intensive.[28] This work seeks to present a detailed and quantitative

analysis of heterogeneous TMV assembly states in solution and their dynamics using native electrospray ionization charge detection mass spectrometry (ESI-CDMS).

Analysis of intact biomolecules by MS, often termed ‘native MS’, has become an important tool in the arsenal of analytical techniques in structural biology. In contrast to many widely used MS-based methods, such as proteomics, where ions are formed from solutions where higher order structure is typically lost and cleavage of noncovalent bonds is used to obtain structure,[29] the goal of native MS is to retain information about higher-order structure and specific non-covalent interactions in solution.[30] However, the inherent heterogeneity of larger biomolecules, such as TMV, due to residual salt adduction and/or incomplete desolvation, can pose significant challenges.[31,32,33] Conventional MS methods work by measuring the mass-to-charge ratio (m/z) of ensembles of analyte ions. The charge and mass are then inferred from either the spacing of isotope peaks of a single charge state or by deconvolving the series of peaks corresponding to the individual charge states of a single analyte. High mass analytes, heterogeneous samples, and adduction of non-volatile solutes, such as Na^+, can lead to broadened peaks and overlapping peaks in charge-state distributions that result in loss of mass information.[32] While methods to extract useful information from overlapping spectra, including Bayesian deconvolution[34] and applying the Fourier transform to the m/z domain,[35] have made analyses of increasingly complex samples possible, they are also ultimately limited by sample heterogeneity. Thus, native MS typically requires increasingly pure samples of the analyte of interest with increasing mass. High concentrations of non-volatile salts used in common biochemical buffers typically lead to interfering salt cluster signals, and so much adduction can occur that no mass information is obtainable.[36] To overcome this adduction problem, most solutions are buffer exchanged into a volatile buffer, such as ammonium acetate, to maintain high ionic strength but reduce peak broadening caused by the adduction of non-volatile salts. Recently, submicron electrospray emitter tips have been demonstrated to reduce adduction in native MS by forming initial droplets that are sufficiently small that they contain a limited number of non-volatile salt molecules.[36] However, the challenge presented by sample heterogeneity grows with molecular size such that even when samples are highly purified with few non-volatile adducts, most conventional native MS experiments are still limited to molecules with masses below a few MDa.[32]

One solution to the analytical challenge presented by sample heterogeneity and interfering ion m/z signals is to remove all possibility of interference by weighing molecules individually using single ion charge detection mass spectrometry (CDMS).[37–39] In CDMS, molecular weight information is obtained from simultaneous measurements of both the m/z and charge of each individual ion. This unconventional approach to molecular mass analysis has several distinct advantages. Because the mass of each molecule is individually measured, there is no interference from other analyte ions, essentially eliminating “chemical noise”, i.e., interferences from other ions. Many individual ion measurements are accumulated to generate a true mass spectrum. These measurements are ideal for highly heterogeneous, high mass samples where charge-state distributions are not resolved in the m/z data due to spectral congestion. CDMS has previously been used to analyze viruses,[40–43] viral capsids,[44,45] and other large, heterogeneous macromolecules.[46–48] In this work, we use unique capabilities of CDMS to probe the assembly state, stoichiometry, and stability of

several TMV mutants under conditions that preclude high resolution analysis by other methods.

RESULTS AND DISCUSSION

Effects of point mutations on cpTMV disk subunit stoichiometry.

The circular permutant of TMV, cpTMV, maintains its ability to fold and assemble when subject to a variety of mutations at sites on the protein surface and within the central cavity.[8] This ability is crucial to its use as a mimic of photosynthetic light harvesting complexes, as it allows for the attachment of chromophores in sites with varying protein, solvent, and adjacent chromophore environments that can then be compared. A primary location for TMV functionalization to light harvesting materials in addition to nanosheets is at residue S3 of recombinant TMV (rTMV) and the analogous S65 position of cpTMV. This site, on the disks' periphery, has been mutated to cysteine for the formation of nanosheet filters and to *p*-aminophenylalanine (pAF) for attachment to nanoparticles (Figure 1F).[6,19] Another widely utilized location for TMV functionalization is position S23 on the face of cpTMV and its analogous position on rTMV, which is readily mutable to cysteine and has been used for attachment of synthetic dyes and small molecule chemotherapeutics (Figure 1F).[7,8,18] However, the effects of point mutations on the assembly state of cpTMV have not previously been studied in detail. Reliably and efficiently measuring small differences in macromolecular assemblies is a strength of CDMS that we sought to apply to understand how point mutations affect both stacking behavior and the intradisk architecture of cpTMV.

A peripheral mutation from serine to the non-canonical amino acid *p*-aminophenylalanine (cpTMV-S65-pAF) of cpTMV was of interest due to its position at the interface of adjacent monomers within a single disk and its use as a handle for nanoparticle attachment.[19] The expected mass for disks consisting of 17 subunits is 302 kDa, resulting in a mass of 604 kDa for the anticipated two disk assembly state. However, the CDMS mass histogram for this mutant displayed a high proportion of 16-monomer disks, with a mass of 570 kDa for a two-disk stack, in addition to the 607 kDa species observed for 17-monomer disks (Figure 2A). Although the helical turn found in wtTMV is made up of 16 $\frac{2}{3}$ monomers per turn, previously determined crystal structures have shown that disks in discrete stacks consist of 17 monomers per disk.[22] However, by electrospray ionization from a 100 mM AA solution and CDMS, two distinct populations of 16- and 17-monomer disks could be measured in tandem, indicating that both are present in solution. No peak consisting of hybridized 16- and 17-mer disks at ~584 kDa was observed, indicating that if a hybridized species of 16 and 17-mer disks exists, its abundance is minimal. Instead, the bias toward forming stacks of disks of like number (e.g. two 16-mers or two 17-mers) suggests that there is a strong pairwise interaction between individual monomers on opposing disks in these cpTMV assemblies.

Several other cpTMV variants with mutations in the S65 position were also assessed for possible assembly into 16-monomer disks. A mutant with 3-nitrotyrosine (3NY) at position S65, but otherwise identical to cpTMV-S65-pAF, is structurally similar but has a decreased charge with respect to the pAF mutant. The cpTMV-S65-3NY mutant still displayed a small population of 16-monomer disks, but at a much lower proportion of 16:17

monomers than for the pAF mutant (Figure 2B). Mass spectra under denaturing conditions of cpTMV-S65-pAF and cpTMV-S65-3NY showed that the monomers of each population have homogeneous masses, indicating that the presence of multiple assembled states is not due to monomer heterogeneity (Figure S1). Other mutants, including cpTMV-S65C and cpTMV without any mutation at the S65 position, also displayed small proportions of 16-monomer populations at lower levels than for the pAF mutant (Figure S2). This fraction of 16-monomer disks is barely distinguishable from noise in several of the mutants; however, in all mass histograms of cpTMV mutants, a small number of 16-monomer disks were detected.

Detecting the subtle difference in size between the 16-monomer and 17-monomer assemblies is difficult using other methods typically used for sizing biomolecules in the size range of these TMV assemblies (>600 kDa). For example, a measurement of cpTMV-S65-pAF using dynamic light scattering (DLS) results in a single, uniform peak with no obvious asymmetries that might be attributable to stacks of both 16-monomer and 17-monomer disks (Figure 2C). A comparison of the cpTMV-S65-pAF mutant to the cpTMV-S65-3NY mutant using size exclusion chromatography-high performance liquid chromatography (SEC-HPLC) yielded essentially identical retention times and only a single resolvable peak in the 600 kDa range (Figure 2D). Assessing and understanding the uniformity of the disk assemblies are crucial to the materials-based applications for which TMV is used, such as components of drug[18] and vaccine candidates[49] and as models for light harvesting systems.

Equilibration of monomeric cpTMV and cpTMV disk assemblies.

In TMV-based artificial light harvesting systems, donor and acceptor chromophores can be embedded within protein complexes in defined locations for the measurement of directional energy transfer.[7] Inter-chromophore energy transfer is likely to be affected by the sensitivity of non-covalent interactions holding the assembled TMV together, and on what scale monomers may be exchanged from one complex to another (a schematic is shown in Figure S3). We investigated a potential equilibrium between intact TMV disk complexes and monomers in solution using CDMS. Often, a small proportion of monomeric TMV is present in purified samples as evidenced by size exclusion chromatography (Figure 2D), but it is difficult to determine whether the disk and monomer states are interchanging or whether the two populations remain unequilibrated. To probe this potential equilibrium, we distinguished one cpTMV population from another through irreversible labeling of each monomer with a 1788 Da peptide, N-acetyl-α-endorphin. We then mixed the α-endorphin-labeled population with unlabeled cpTMV and monitored the mass distributions within the mixture. Convergence to a single mass over time would indicate the exchange of monomers between the unlabeled and labeled species.

Labeling of cpTMV-S23C-S65-pAF with α-endorphin was performed using the enzyme tyrosinase, which activates exposed tyrosine residues toward nucleophilic addition.[50,51] The N-terminal tyrosine of α-endorphin was oxidized to an *o*-quinone with tyrosinase, to which either the engineered cysteine or aniline of pAF in cpTMV-S23C-S65-pAF could add (Figure 3A). This labeling resulted in a highly heterogeneous sample consisting of double disk complexes incorporating 32+ monomers with between 1–3 attached peptides

per monomer. The sample was further complicated by a truncation that occurred during purification, leading to final complexes containing monomers with eight different possible masses (Figure 3B). Additionally, both 32- and 34-monomer double disk assemblies were formed, further increasing the heterogeneity of the sample. These modifications and other molecular variables result in a mixture of 6540 possible unique assembly masses that does not include contributions from stable isotopes and adducts which increases the number substantially. An *m*/*z* spectrum generated from the same data as Figure 3E, which contains a mixture of α-endorphin-labeled and unlabeled cpTMV-S23C-S65-pAF assemblies, is included in the Supporting Information (Figure S4) and shows the broad occupancy of *m*/*z* space that results from such a heterogeneous sample. Statistical analysis based on the relative abundances of each monomer type was used to calculate expected theoretical mass centroids of 590 kDa for the unlabeled 34-mer and 680 kDa for the labeled 34-mer, resulting in an expected mean difference of 89.9 kDa between α-endorphin-labeled and unlabeled samples. The calculated centroid values match relatively well with the observed mass histograms shown in Figure 3C of the unmodified and modified samples, respectively (Figure S5), slightly underestimating the mass of the modified sample. Despite the expected increase in peak width for these samples due to heterogeneity, the two different populations are clearly resolvable, making it possible to monitor any monomer exchange. It is also unlikely that these structures would be differentiated using conventional high resolution mass spectrometry owing to the high heterogeneity of these TMV solutions due to the presence of multiple assembled states, salt adduction, and incomplete desolvation, which results in congested *m*/*z* spectra (Figure S4).

A solution of the mixture of labeled and unlabeled complexes was monitored for changes in mass distribution over time. Directly after mixing the two samples, two distinct populations were seen for both double-layered disks (with centroids for the 34-mer at 590 and 690 kDa, respectively) and for 4-layer disk stacks composed of 68 monomers (Figure 3D). The two cpTMV populations in the mixed sample remained distinct after 3 h and 72 h (Figure 3E, Figure S6). This indicates that, on the timescales used for typical excited state energy transfer experiments, with samples being measured at room temperature for several hours (or kept at 4 °C or frozen for longer periods), there is no significant exchange of monomers between disks. A secondary result of this experiment monitoring the assembly state of cpTMV over time indicates that TMV may have a propensity to stack into larger complexes the longer it is in solution, as evidenced by the appearance of higher MW species after 3 h and 72 h than after initial mixing. While it is possible that variation in the electrospray ionization process may account for some of the observed increase of higher MW species, the label-specific stacking preference is visible even for larger stacks, suggesting that the formation of these structures is favored under the conditions and timescales used in this study. The larger stacks that are formed in solution after 72 h also show a label-specific preference, indicating that their formation in solution is favored over structures consisting of both labeled and unlabeled subunits on this timescale. This intrinsic stability may also contribute to the observation of no exchange for the smaller assemblies even over several days.

Distinguishing between cpTMV conformational states of similar mass.

One goal of the TMV-based light harvesting model is to measure energy transfer between laterally joined cpTMV disks (Figure 4A), imitating adjacent light harvesting complexes within the membranes of photosynthetic organisms. We sought to distinguish between this peripherally joined configuration and four disk stacks of cpTMV protein complexes of nearly identical molecular weight using CDMS. To create an assembly with peripherally connected cpTMV disks, we engineered a cysteine residue onto a surface-exposed position on the periphery of the cpTMV double disk (red residues in Figure 4A–B). When exposed to air, these thiols can form disulfide bonds with neighboring disk assemblies, as has been observed for a cysteine mutation in an analogous position on rTMV.[16] While we have depicted these conjugated assemblies as parallel to one another in Figures 4A, E, and Fto distinguish them from the stacked assembly state, it is likely that there is some flexibility in the disulfide linkage joining the two protein assemblies such that the two bonded complexes are not necessarily parallel to one another. When two double-disk cpTMV assemblies are linked via these peripheral disulfide linkages, they would be expected to form an assembly of ~1.2 MDa with a difference of only a few Da between disulfide-bonded complexes and stacks of four disks as shown in Figure 4B. This small difference relative to the large complex mass makes the two species impossible to differentiate on the basis of mass alone. Because CDMS measures both the m/z and charge of each individual analyte ion independently, species with the same mass can be differentiated based on their charge state. While the two conformations of four disks are nearly identical in molecular weight, their charge states are expected to differ significantly because the apposed assembly of two disulfide-bonded disks (with approximate dimensions of 5 nm × 18 × 36 nm) has a significantly larger surface area available for charging than vertical stacks of four disks (with approximate dimensions of 10 nm × 18 nm × 18 nm). A CDMS histogram for the cpTMV-S65C mutant exhibits a greater proportion of larger disk stoichiometries per assembly than other cpTMV mutants, indicating that the mutation either promotes disk stacking and/or peripheral disulfide bonding (Figure 4C). To support that the species with higher charge in the cpTMV-S65C sample was the expected disulfide-bonded complex, we treated the sample with a reducing agent, TCEP, and monitored the charge vs. mass ratio of the sample over time. After 1.5 h, nearly all of the assemblies containing greater than four disks had disappeared (Figure 4D). Additionally, a sample identical but without the peripheral cysteine mutation, cpTMV-S65, did not have as high a proportion of higher mass species as cpTMV-S65C and did not have an appreciable depletion of higher mass species upon treatment with TCEP (Figure S7A-B).

While the two populations of peripherally bonded versus stacked complexes cannot be distinguished by mass alone, a two-dimensional mass vs. charge histogram of a cpTMV-S65C sample in Figure 4E shows two distinct populations at ~1.2 MDa that are differentiated by ~13 charges (expanded view provided in Figure S8). The stacked disk structure likely corresponds to the population with charge states below the Rayleigh charging limit (red line in Figure 4E), and the peripherally bonded disk structure to the population with charge states above the Rayleigh limit, a characteristic that typically indicates an assembly that deviates from a globular (spherical) structure. The higher-charged species at 1.2 MDa was also depleted after treatment with TCEP, though some of this

species was still visible (Figure 4F). In contrast, cpTMV-S65 did not exhibit multiple charge states at 1.2 MDa and higher masses, either before or after treatment with TCEP (Figure S7C-D). This provides evidence for the identity of the higher charged species as two double-layered disks bonded via disulfides at their periphery and shows how multiple assembly configurations can be distinguished using CDMS. There is also a bimodal charge-state distribution for the ~600 kDa complex. The distribution centered at ~40 charges is too far below the Rayleigh limit for a globular structure formed directly from solution. This indicates that this population is likely formed by gas-phase dissociation of a higher-order structure of the peripherally bonded complexes, consistent with the absence of this population in the cpTMV-S65 data (Figure S7C-D). The observation of some dissociation products despite very soft instrument conditions indicates that the interactions between the two double disks are weak, consistent with a limited contact patch between the assemblies. The dissociation occurs with symmetric charge partitioning between products, consistent with prior results where symmetric dissociation occurs for weakly bound complexes under low energy excitation conditions.[52]

TMV stacking stoichiometries in different buffer solutions.

Due to the sensitivity of TMV's assembly state distribution to buffer conditions, we additionally sought to understand whether the assembly state distribution of TMV particles was affected by buffer identity. The 10 mM NaPhos, pH 7.2 buffer is typically used for the chemical modification of proteins and for the purification and storage of TMV particles due to its buffering capacity and biological compatibility.[8,53] Recombinant TMV particles (rTMV) consisting of monomers with a sequence identical to wtTMV, but lacking N-terminal acylation, were expressed and assembled in *E. coli*. rTMV particles in 10 mM sodium phosphate (NaPhos) were then compared to rTMV exchanged into a volatile buffer of the equivalent ionic strength preferred for mass spectrometry, 18.2 mM ammonium acetate (AA). The versatility of CDMS with respect to sample heterogeneity combined with the use of submicron ESI emitters that reduce the extent of non-volatile adducts[36] made it possible to acquire CDMS mass histograms from each of these solutions (Figure 5A–B). In each of these solutions, rTMV forms even-numbered stacks as expected based on previously determined structures. However, significant populations of stacks of 3 or 5 disks at ~900 and ~1500 kDa were also present, a stacking stoichiometry that has not been observed previously for rTMV. Differences in relative ion count between numbers of disks per stack were observed between the protein solutions in AA and NaPhos, with the solutions in NaPhos appearing to have a greater proportion of higher-order stacks, particularly for the 5-disk stack.

These differences illustrate the utility and importance of CDMS and submicron emitters in analyses from non-volatile buffers and indicate that buffer identity can influence TMV stacking stoichiometries. These data were collected on the same day under identical instrumental conditions to preclude differences in mass distribution due to mass and instrumental biases; however, some variation between the two spectra may be due to environmental factors. Salt adduction and baseline noise are also more pronounced in NaPhos buffer than AA solution, indicating that AA can be preferable for resolving finer differences in mass distributions, such as those shown in Figures 2 and 3. Under denaturing

conditions, a mass spectrum of the rTMV monomer showed that the mass is homogeneous (Figure S1), indicating that the different populations observed using CDMS were due to differing assembly states of identical monomers rather than monomer heterogeneity.

In order to determine whether the odd-numbered stacking behavior occurred as a result of non-specific aggregation during the electrospray ionization (ESI) process or in solution before ESI, we evaluated rTMV in each buffer solution using negative staining and transmission electron microscopy (TEM). TEM images of rTMV were consistent with CDMS observations, with even- and odd-numbered stacks of disks of varying stoichiometry observed in side views of rTMV assemblies (Figure 5C–D; a collection of images can be found in Figure S9). Multiple clear examples of odd-numbered disk stacks were observed in NaPhos buffer (Figure 5D, inlays), illustrating that these structures were present in solution and did not form as a result of non-specific aggregation in the ESI process. A crystal structure of the four-layered aggregate of wtTMV indicates that the four-disk stacks are C_2 symmetric about their center,[22] and solved stacked disk structures of additional TMV mutants also exhibit C_2 symmetry.[13,54] While neither CDMS nor TEM was able to resolve the facial directionality of the individual disks comprising the odd-numbered stacks, they cannot be in fully C_2 symmetric arrangements.

CONCLUSIONS

This work has used CDMS to identify previously unrecognized assembly states of particles derived from TMV, including stacks of disks containing 16 monomers and odd-numbered stacks of disks that necessarily cannot exhibit C_2 symmetry. It has also further emphasized the importance of developing techniques for studying large protein complexes in their native buffers, as buffer exchange can influence the ratio of assembly states present in heterogeneous protein complex solutions. In addition to providing precise mass measurements and proportional estimates of multiple assemblies present in heterogeneous solutions, measuring charge and m/z in tandem allowed for differentiation between species of near identical mass but with different configurations. CDMS also allowed us to probe the possibility of an exchange of monomers between assembly structures in solution and was used to determine that dissociation of assembled TMV complexes into monomers is minimal under the conditions and time periods studied.

The precise morphology, stoichiometry, and arrangement of monomers within quaternary protein structures have broad implications for their function. In the context of TMV, its morphology has been demonstrated to affect its drug delivery capability[18] and has implications for the development of biosensing or light harvesting materials utilizing TMV monomers. Measuring multiple assembled states of TMV present in solution can be accomplished quickly and with minimal protein preparation using CDMS, allowing for the rapid testing of multiple conditions and even time-course experiments. These observations will inform the continued development of light-harvesting models based on mutants of TMV. In addition, the techniques described herein using CDMS are expected to have broad utility for the detailed characterization of heterogeneous populations of large protein complexes.

EXPERIMENTAL SECTION

Synthesis of rTMV and cpTMV mutants.

Synthesis of rTMV and cpTMV mutants was performed as described elsewhere.[8,19,43,53,55] Briefly, for rTMV and cpTMV mutants not containing non-canonical amino acids, BL21-Codonplus (DE3)-RIL cells were transformed with the appropriate pET vector, and colonies were selected for inoculation of Terrific Broth cultures. Cultures were induced with 30 μM isopropyl-β-D-thiogalactopyranoside (IPTG) at an OD_{600} of 0.6–0.8 and allowed to grow 14–18 h at 20 °C before harvesting cell pellets and storing them at −20 °C. For cpTMV mutants containing non-canonical amino acids, DH10B cells were co-transformed with the appropriate pBAD-cpTMV and pDule-pAF or pDule-3NY vectors, and colonies were selected for inoculation in autoinduction media. At an OD_{600} of 0.6–0.8, 1 mM *p*-amino-L-phenylalanine (*p*AF or 3NY) was added and the culture was allowed to grow 18 h at 37 °C. For both rTMV and cpTMV, cell pellets were collected at 8000 rpm for 30 min and stored at −20 °C until purification. For purification, cell pellets were resuspended in 10 mL lysis buffer, 20 mM triethanolamine (TEA) pH 7.2 and lysed by sonication with a 2 s on, 4 s off cycle for a total of 10 min using a standard disruptor horn at 65% amplitude. The resulting lysate was cleared at 14,000 rpm for 30 min. The supernatant was treated with 30–40% volume of saturated ammonium sulfate and allowed to rotate for 10 min at 4 °C. The precipitated protein was collected at 11,000 rpm for 30 min and resuspended in 10 mL lysis buffer, then dialyzed in 1 L lysis buffer over-night with at least one buffer exchange. The resulting protein solution was treated with 5 mL benzonase and 4 mg $MgCl_2$ at room temperature for 30 minutes and purified using a DEAE column with a 0 – 180 mM NaCl gradient elution in 20 mM TEA buffer, pH 7. Further purification was performed using a Sephacryl S-500 column in 10 mM NaPhos pH 7.2 elution buffer. Purity and general assembly state were confirmed by SDS-PAGE, ESI-TOF LC-MS, and HPLC-SEC.

Preparation of rTMV and cpTMV samples for CDMS analysis.

Preparation of rTMV and cpTMV samples was conducted as described previously.[43] TMV assemblies were thawed and then exchanged into the appropriate buffer for subsequent analysis via repeated buffer exchange and concentration using 100 kDa MWCO centrifugal filters. Samples were adjusted to a protein concentration of 0.5 mg/mL protein and filtered through a 0.22 μm filter prior to analysis.

TMV mixing experiments.

For experiments mixing populations of α-endorphin-labeled and unlabeled cpTMV, samples were mixed to a final concentration of 0.5 mg/mL protein (0.25 mg/mL of each individual sample) in 100 mM ammonium acetate. The sample was then incubated at room temperature undisturbed indefinitely. CDMS measurements were recorded directly after mixing and at subsequent time intervals of 3 hours and 3 days.

Procedure for labeling of cpTMV-S23C-S65-pAF with α-endorphin.

Enzyme-catalyzed modification of cpTMV-S23C-S65-pAF with α-endorphin was conducted according to literature procedure for with slight modifications.[19,51] The protein was first

exchanged into the reaction buffer (10 mM sodium phosphate, pH 7.2). To 100 μL of α-endorphin (250 μM) was added 400 nM tyrosinase from *Agaricus bisporus* (abTyr), followed by 50 μM protein. The reaction mixture was briefly agitated and then incubated in 1.5 mL Eppendorf tubes at room temperature. After 2 h, the reaction was quenched with 1 mM tropolone and incubated for 5 min. Excess peptide and abTyr were removed via repeated centrifugation through 100 kDa molecular weight cutoff filters. The protein conjugates were analyzed with MS to measure modification levels of monomers.

Reduction of disulfide bonds of cpTMV-S65C.

To a 0.5 mg/mL (22 μM) cpTMV-S65C solution in 100 mM ammonium acetate was added 1 mM tris(2-carboxyethyl)phosphine (TCEP). The reaction was allowed to incubate at least 1 h at room temperature, and no subsequent buffer exchange was performed before CDMS analysis.

Charge detection mass spectrometry.

Individual ion mass measurements are obtained using a home-built charge detection mass spectrometer that is described in detail elsewhere.[46,47] Ions are formed by nanoelectrospray ionization from borosilicate capillaries (1.0 mm outer diameter, 0.78 mm inner diameter) with tips pulled to an inner diameter of 0.5–2.0 μm using a Flaming/Brown P-87 micropipette puller (Sutter Instruments, Novato, CA) with smaller diameter tips (0.5–0.6 μm) used to reduce non-volatile salt adducts from solutions other than ammonium acetate.[36,56] Ions are introduced into the instrument through a modified Waters Z-Spray source (Waters Corporation, Milford, MA) and enter a region containing two rf-only quadrupole ion guides where they are accumulated for up to 1 s. Ions are then pulsed into an electrostatic ion trap containing a charge detection electrode at a pressure of ~3 × 10^{-9} Torr where they are analyzed for 1 s. Typically, several ions are simultaneously trapped and analyzed to reduce the time required to acquire a statistically significant number of individual ions. All experiments in this work required <1 h of data acquisition, with most experiments requiring ~15–20 min to accumulate ion counts. Trapped ions repeatedly induce charge on the detector and the individual oscillation frequencies and amplitudes are used to determine the m/z, charge, and mass of each individual ion. Ion signals are amplified by a CoolFET charge-sensitive preamplifier (Amptek, Bedford, MA) and pass through a custom-built filter stage that removes noise and further amplifies signals. Signals are analyzed using short-time Fourier transform (STFT) methods,[57] with 50 ms segment lengths stepped across the transients in 5 ms increments. Only ions trapped for the entire 1 s trapping period (>75% of all ions) were included in the analysis. The resolution expected for a 1 s data acquisition depends on the inherent sample heterogeneity and instrument performance; a detailed analysis of these factors for the same instrument used in this work is given elsewhere.[43]

Procedure for generating cpTMV homology model.

For visualization purposes, a homology model of cpTMV containing all amino acid residues was generated based on the crystal structure of cpTMV (PDB code 3KML)[13] and the crystal structure of wtTMV (PDB code 1EI7),[22] both obtained from the Protein Data Bank. Initial structural preparation was conducted using Pymol, Version 2.4.2.[58] First, a monomer of

wtTMV was superimposed on a monomer of cpTMV, and the residues resolved in the crystal structure of wtTMV but not cpTMV (residues 92–110 of wtTMV and 1–12, 154–161 of cpTMV) were fused to the unresolved N- and C-termini of cpTMV. The bond between residues 99 and 100 of wtTMV was then cleaved *in silico* and an N-terminal glycine was added to produce the N- and C-termini of cpTMV. Following this, cpTMV was symmetry expanded to create the double disk assembled structure consisting of two C_2-symmetric disks, each containing 17 monomers. The Schrödinger Maestro package (version 2022–1)[59] was used for subsequent structural preparation and molecular dynamics-based side chain minimizations. The Desmond system builder was used to solvate the double disk structure in an orthorhombic box with periodic boundaries at 10 Å from the protein of water molecules described using the TIP3P model,[60,61] neutralized with sodium ions, with the addition of 150 mM NaCl in an OPLS4 force field.[62] A molecular dynamics simulation/minimization was performed with an NPT ensemble of T = 300 K, P = 1 bar, a Coulombic cutoff radius of 9.0 Å, and a 100 ps simulation time with sampling time of 5 ps. This short simulation time was chosen to relax sidechain and solvent interactions for the protein representations shown herein, but not alter the overall quaternary structure of the assembly. The pressure control was applied using the Martyna-Tobias-Klein barostat method[63] with a 2.0 ps relaxation time, and the temperature control was applied using the Nosé-Hoover thermostat method[64] with a 1.0 ps relaxation time. The trajectory was analyzed, and the lowest energy frame was used as the homology model.

Theoretical Mass Distribution Calculations.

To calculate expected mass centroids and distribution for the α-endorphin modified and/or partially truncated TMV double-disks, the percentage of the TMV subunit proteins bearing modifications and/or truncations were measured from conventional mass spectra of the un-assembled subunits (Figure 3B). The unmodified, untruncated subunit has a mass of 17,783 Da. Additions of up to three α-endorphin (net mass added = 1800 Da) were observed, with 7% unmodified, 34% with one addition, 44% with two additions, and 15% with three additions (Figure 3B). Only one truncated subunit mass was observed (16,384 Da, a loss of 1,399 Da) at 26% of the population, with 74% remaining untruncated (Figure 3B). A simple Python script (included in the Supporting Information) was used to simulate the expected net mass added based on these measured values. The generation of 10,000 34-mer double disks were simulated. For each subunit in each 34-mer double disk, the mass added by α-endorphin and/or subtracted by truncation was randomly chosen via weighted probability distributions based on the observed population abundances. The resulting distribution was then fit using a Gaussian function to find the mean net mass added or lost (Figure S5A). Finally, these values were added to the expected mass of an unmodified, untruncated 34-mer double disk (604,622 Da) and the resultant distributions were Gaussian fit to find the expected mass centroids.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGMENT

The authors are grateful for financial support from the National Institutes of Health (5R01GM139338 for E.R.W) and the Director, Office of Science, Chemical Sciences, Geo-sciences, and Biosciences Division, of the U.S. Department of Energy under Contract No. (DEAC02-05CH11231 for M. B. F.) for financial support. A.J.B. also thanks a Chemical Biology Training Grant from the NIH (T32 GM066698) and the NSF Graduate Fellowship Program (DGE 1752814 and 2146752) for financial support. We would also like to thank the staff at the University of California Berkeley Electron Microscope Laboratory for advice and assistance in electron microscopy sample preparation and data collection.

Biographies

Amanda J. Bischoff – Department of Chemistry, University of California, Berkeley, California, 94720, United States; Molecular Biophysics and Integrated Bioimaging Division, Lawrence Berkeley National Laboratories, Berkeley, California 94720;

Conner C. Harper – Department of Chemistry, University of California, Berkeley, California, 94720, United States

Evan R. Williams – Department of Chemistry, University of California, Berkeley, California, 94720, United States;

REFERENCES

(1). Smith ML; Lindbo JA; Dillard-Telm S; Brosio PM; Lasnik AB; McCormick AA; Nguyen LV; Palmer KE Modified Tobacco Mosaic Virus Particles as Scaffolds for Display of Protein Antigens for Vaccine Applications. Virology 2006, 348 (2), 475–488. 10.1016/j.virol.2005.12.039. [PubMed: 16466765]

(2). Venkataraman S; Hefferon K Application of Plant Viruses in Biotechnology, Medicine, and Human Health. Viruses 2021, 13 (9). 10.3390/v13091697.

(3). Chung YH; Cai H; Steinmetz NF Viral Nano-particles for Drug Delivery, Imaging, Immunotherapy, and Theranostic Applications. Advanced Drug Delivery Reviews 2020, 156, 214–235. 10.1016/j.addr.2020.06.024. [PubMed: 32603813]

(4). Koch C; Wabbel K; Eber FJ; Krolla-Sidenstein P; Azucena C; Gliemann H; Eiben S; Geiger F; Wege C Modified TMV Particles as Beneficial Scaffolds to Present Sensor Enzymes. Frontiers in Plant Science 2015, 6, 1137. 10.3389/fpls.2015.01137. [PubMed: 26734040]

(5). Fan XZ; Naves L; Siwak NP; Brown A; Culver J; Ghodssi R Integration of Genetically Modified Virus-like-Particles with an Optical Resonator for Selective Bio-Detection. Nanotechnology 2015, 26 (20), 205501. 10.1088/0957-4484/26/20/205501. [PubMed: 25915182]

(6). Zhang S; Zhang J; Fang W; Zhang Y; Wang Q; Jin J Ultralarge Single-Layer Porous Protein Nanosheet for Precise Nanosize Separation. Nano Lett. 2018, 18 (10), 6563–6569. 10.1021/acs.nanolett.8b03155. [PubMed: 30182720]

(7). Miller RA; Presley AD; Francis MB Self-Assembling Light-Harvesting Systems from Synthetically Modified Tobacco Mosaic Virus Coat Proteins. J. Am. Chem. Soc 2007, 129 (11), 3104–3109. 10.1021/ja063887t. [PubMed: 17319656]

(8). Delor M; Dai J; Roberts TD; Rogers JR; Hamed SM; Neaton JB; Geissler PL; Francis MB; Ginsberg NS Exploiting Chromophore–Protein Interactions through Linker Engineering To Tune Photoinduced Dynamics in a Biomimetic Light-Harvesting Platform. J. Am. Chem. Soc 2018, 140 (20), 6278–6287. 10.1021/jacs.7b13598. [PubMed: 29741876]

(9). Endo M; Fujitsuka M; Majima T Porphyrin Light-Harvesting Arrays Constructed in the Recombinant Tobacco Mosaic Virus Scaffold. Chemistry – A European Journal 2007, 13 (31), 8660–8666. 10.1002/chem.200700895. [PubMed: 17849494]

(10). Koch C; Eber FJ; Azucena C; Förste A; Walheim S; Schimmel T; Bittner AM; Jeske H; Gliemann H; Eiben S; Geiger FC; Wege C Novel Roles for Well-Known Players: From Tobacco

Mosaic Virus Pests to Enzymatically Active Assemblies. Beilstein Journal of Nanotechnology 2016, 7, 613–629. 10.3762/bjnano.7.54. [PubMed: 27335751]

(11). Chu S; Brown AD; Culver JN; Ghodssi R Tobacco Mosaic Virus as a Versatile Platform for Molecular Assembly and Device Fabrication. Biotechnology Journal 2018, 13 (12), 1800147. 10.1002/biot.201800147.

(12). Butler PJG The Current Picture of the Structure and Assembly of Tobacco Mosaic Virus. Journal of General Virology, 1984, 65, 253–279. [PubMed: 6363621]

(13). Dedeo MT; Duderstadt KE; Berger JM; Francis MB Nanoscale Protein Assemblies from a Circular Permutant of the Tobacco Mosaic Virus. Nano Lett. 2010, 10 (1), 181–186. 10.1021/nl9032395. [PubMed: 19924865]

(14). Kadri A; Maiß E; Amsharov N; Bittner AM; Balci S; Kern K; Jeske H; Wege C Engineered Tobacco Mosaic Virus Mutants with Distinct Physical Characteristics in Planta and Enhanced Metallization Properties. Virus Research 2011, 157 (1), 35–46. 10.1016/j.virusres.2011.01.014. [PubMed: 21310199]

(15). Zhang J; Zhou K; Wang Q Tailoring the Self-Assembly Behaviors of Recombinant Tobacco Mosaic Virus by Rationally Introducing Covalent Bonding at the Protein–Protein Interface. Small 2016, 12 (36), 4955–4959. 10.1002/smll.201503487. [PubMed: 27061916]

(16). Zhang J; Wang X; Zhou K; Chen G; Wang Q Self-Assembly of Protein Crystals with Different Crystal Structures Using Tobacco Mosaic Virus Coat Protein as a Building Block. ACS Nano 2018, 12 (2), 1673–1679. 10.1021/acsnano.7b08316. [PubMed: 29350903]

(17). Lin RD; Steinmetz NF Tobacco Mosaic Virus Delivery of Mitoxantrone for Cancer Therapy. Nanoscale 2018, 10 (34), 16307–16313. 10.1039/C8NR04142C. [PubMed: 30129956]

(18). Finbloom JA; Aanei IL; Bernard JM; Klass SH; Elledge SK; Han K; Ozawa T; Nicolaides TP; Berger MS; Francis MB Evaluation of Three Morphologically Distinct Virus-Like Particles as Nanocarriers for Convection-Enhanced Drug Delivery to Glioblastoma. Nanomaterials 2018, 8 (12). 10.3390/nano8121007.

(19). Ramsey AV; Bischoff AJ; Francis MB Enzyme Activated Gold Nanoparticles for Versatile Site-Selective Bioconjugation. J. Am. Chem. Soc 2021, 143 (19), 7342–7350. 10.1021/jacs.0c11678. [PubMed: 33939917]

(20). Bruckman MA; VanMeter A; Steinmetz NF Nanomanufacturing of Tobacco Mosaic Virus-Based Spherical Biomaterials Using a Continuous Flow Method. ACS Biomater. Sci. Eng 2015, 1 (1), 13–18. 10.1021/ab500059s. [PubMed: 25984569]

(21). Bruckman MA; Czapar AE; VanMeter A; Randolph LN; Steinmetz NF Tobacco Mosaic Virus-Based Protein Nanoparticles and Nanorods for Chemotherapy Delivery Targeting Breast Cancer. Journal of Controlled Release 2016, 231, 103–113. 10.1016/j.jconrel.2016.02.045. [PubMed: 26941034]

(22). Bhyravbhatla B; Watowich SJ; Caspar DLD Refined Atomic Model of the Four-Layer Aggregate of the Tobacco Mosaic Virus Coat Protein at 2.4-Å Resolution. Biophysical Journal 1998, 74 (1), 604–615. 10.1016/S0006-3495(98)77819-1. [PubMed: 9449361]

(23). Ksenofontov AL; Petoukhov MV; Prusov AN; Fedorova NV; Shtykova EV Characterization of Tobacco Mosaic Virus Virions and Repolymerized Coat Protein Aggregates in Solution by Small-Angle X-Ray Scattering. Biochemistry (Moscow) 2020, 85 (3), 310–317. 10.1134/S0006297920030062. [PubMed: 32564735]

(24). Miller RA; Stephanopoulos N; McFarland JM; Rosko AS; Geissler PL; Francis MB Impact of Assembly State on the Defect Tolerance of TMV-Based Light Harvesting Arrays. J. Am. Chem. Soc 2010, 132 (17), 6068–6074. 10.1021/ja909566z. [PubMed: 20392093]

(25). Vugmeyster Y; Xu X; Theil F-P; Khawli LA; Leach MW Pharmacokinetics and Toxicology of Therapeutic Proteins: Advances and Challenges. World J Biol Chem 2012, 3 (4), 73–92. 10.4331/wjbc.v3.i4.73. [PubMed: 22558487]

(26). Ducry L Challenges in the Development and Manufacturing of Antibody–Drug Conjugates. In Therapeutic Proteins: Methods and Protocols; Voynov V., Caravella JA, Eds.; Humana Press: Totowa, NJ, 2012; pp 489–497. 10.1007/978-1-61779-921-1_29.

(27). Vhora I; Patil S; Bhatt P; Misra A Chapter One - Protein- and Peptide-Drug Conjugates: An Emerging Drug Delivery Technology. In Advances in Protein Chemistry and Structural

Biology; Donev R, Ed.; Academic Press, 2015; Vol. 98, pp 1–55. 10.1016/bs.apcsb.2014.11.001. [PubMed: 25819275]

(28). Heck AJR Native Mass Spectrometry: A Bridge between Interactomics and Structural Biology. Nature Methods 2008, 5 (11), 927–933. 10.1038/nmeth.1265. [PubMed: 18974734]

(29). Catherman AD; Skinner OS; Kelleher NL Top Down Proteomics: Facts and Perspectives. Biochemical and Biophysical Research Communications 2014, 445 (4), 683–693. 10.1016/j.bbrc.2014.02.041. [PubMed: 24556311]

(30). Leney AC; Heck AJR Native Mass Spectrometry: What Is in the Name? J. Am. Soc. Mass Spectrom 2017, 28 (1), 5–13. 10.1007/s13361-016-1545-3.

(31). Weiss VU; Bereszczak JZ; Havlik M; Kallinger P; Gösler I; Kumar M; Blaas D; Marchetti-Deschmann M; Heck AJR; Szymanski WW; Allmaier G Analysis of a Common Cold Virus and Its Subviral Particles by Gas-Phase Electrophoretic Mobility Molecular Analysis and Native Mass Spectrometry. Anal. Chem 2015, 87 (17), 8709–8717. 10.1021/acs.analchem.5b01450. [PubMed: 26221912]

(32). Lössl P; Snijder J; Heck AJR Boundaries of Mass Resolution in Native Mass Spectrometry. J. Am. Soc. Mass Spectrom 2014, 25 (6), 906–917. 10.1007/s13361-014-0874-3. [PubMed: 24700121]

(33). McKay AR; Ruotolo BT; Ilag LL; Robinson CV Mass Measurements of Increased Accuracy Resolve Heterogeneous Populations of Intact Ribosomes. J. Am. Chem. Soc 2006, 128 (35), 11433–11442. 10.1021/ja061468q. [PubMed: 16939266]

(34). Marty MT; Baldwin AJ; Marklund EG; Hochberg GKA; Benesch JLP; Robinson CV Bayesian Deconvolution of Mass and Ion Mobility Spectra: From Binary Interactions to Polydisperse Ensembles. Anal. Chem 2015, 87 (8), 4370–4376. 10.1021/acs.analchem.5b00140. [PubMed: 25799115]

(35). Cleary SP; Prell JS Distinct Classes of Multi-Subunit Heterogeneity: Analysis Using Fourier Transform Methods and Native Mass Spectrometry. Analyst 2020, 145 (13), 4688–4697. 10.1039/D0AN00726A. [PubMed: 32459233]

(36). Susa AC; Xia Z; Williams ER Small Emitter Tips for Native Mass Spectrometry of Proteins and Protein Complexes from Nonvolatile Buffers That Mimic the Intracellular Environment. Anal. Chem 2017, 89 (5), 3116–3122. 10.1021/acs.analchem.6b04897. [PubMed: 28192954]

(37). Keifer DZ; Pierson EE; Jarrold MF Charge Detection Mass Spectrometry: Weighing Heavier Things. Analyst 2017, 142 (10), 1654–1671. 10.1039/C7AN00277G. [PubMed: 28443838]

(38). Kafader JO; Melani RD; Senko MW; Makarov AA; Kelleher NL; Compton PD Measurement of Individual Ions Sharply Increases the Resolution of Orbitrap Mass Spectra of Proteins. Anal. Chem 2019, 91 (4), 2776–2783. 10.1021/acs.analchem.8b04519. [PubMed: 30609364]

(39). Harper CC; Elliott AG; Oltrogge LM; Savage DF; Williams ER Multiplexed Charge Detection Mass Spectrometry for High-Throughput Single Ion Analysis of Large Molecules. Anal. Chem 2019, 91 (11), 7458–7465. 10.1021/acs.analchem.9b01669. [PubMed: 31082222]

(40). Keifer DZ; Motwani T; Teschke CM; Jarrold MF Acquiring Structural Information on Virus Particles with Charge Detection Mass Spectrometry. J. Am. Soc. Mass Spectrom 2016, 27 (6), 1028–1036. 10.1007/s13361-016-1362-8. [PubMed: 27020925]

(41). Harper CC; Brauer DD; Francis MB; Williams ER Direct Observation of Ion Emission from Charged Aqueous Nanodrops: Effects on Gaseous Macromolecular Charging. Chem. Sci 2021, 12 (14), 5185–5195. 10.1039/D0SC05707J. [PubMed: 34168773]

(42). Pierson EE; Keifer David. Z.; Asokan A; Jarrold MF Resolving Adeno-Associated Viral Particle Diversity With Charge Detection Mass Spectrometry. Anal. Chem 2016, 88 (13), 6718–6725. 10.1021/acs.analchem.6b00883. [PubMed: 27310298]

(43). Harper CC; Miller ZM; Lee H; Bischoff AJ; Francis MB; Schaffer DV; Williams ER Effects of Molecular Size on Resolution in Charge Detection Mass Spectrometry. Anal. Chem 2022, 94 (33), 11703–11712. 10.1021/acs.analchem.2c02572. [PubMed: 35961005]

(44). Lutomski CA; Lyktey NA; Zhao Z; Pierson EE; Zlotnick A; Jarrold MF Hepatitis B Virus Capsid Completion Occurs through Error Correction. J. Am. Chem. Soc 2017, 139 (46), 16932–16938. 10.1021/jacs.7b09932. [PubMed: 29125756]

(45). Lutomski CA; Lyktey NA; Pierson EE; Zhao Z; Zlotnick A; Jarrold MF Multiple Pathways in Capsid Assembly. J. Am. Chem. Soc 2018, 140 (17), 5784–5790. 10.1021/jacs.8b01804. [PubMed: 29672035]

(46). Elliott AG; Harper CC; Lin H-W; Williams ER Mass, Mobility and MSn Measurements of Single Ions Using Charge Detection Mass Spectrometry. Analyst 2017, 142 (15), 2760–2769. 10.1039/C7AN00618G. [PubMed: 28636005]

(47). Elliott AG; Merenbloom SI; Chakrabarty S; Williams ER Single Particle Analyzer of Mass: A Charge Detection Mass Spectrometer with a Multi-Detector Electrostatic Ion Trap. International Journal of Mass Spectrometry 2017, 414, 45–55. 10.1016/j.ijms.2017.01.007. [PubMed: 29129967]

(48). Lutomski CA; Gordon SM; Remaley AT; Jarrold MF Resolution of Lipoprotein Subclasses by Charge Detection Mass Spectrometry. Anal. Chem 2018, 90 (11), 6353–6356. 10.1021/acs.analchem.8b01127. [PubMed: 29756771]

(49). Royal JM; Simpson CA; McCormick AA; Phillips A; Hume S; Morton J; Shepherd J; Oh Y; Swope K; DeBeauchamp JL; Webby RJ; Cross RW; Borisevich V; Geisbert TW; Demarco JK; Bratcher B; Haydon H; Pogue GP Development of a SARS-CoV-2 Vaccine Candidate Using Plant-Based Manufacturing and a Tobacco Mosaic Virus-like Nano-Particle. Vaccines 2021, 9 (11). 10.3390/vaccines9111347.

(50). Maza JC; Bader DLV; Xiao L; Marmelstein AM; Brauer DD; ElSohly AM; Smith MJ; Krska SW; Parish CA; Francis MB Enzymatic Modification of N-Terminal Proline Residues Using Phenol Derivatives. J. Am. Chem. Soc 2019, 141 (9), 3885–3892. 10.1021/jacs.8b10845. [PubMed: 30726077]

(51). Marmelstein AM; Lobba MJ; Mogilevsky CS; Maza JC; Brauer DD; Francis MB Tyrosinase-Mediated Oxidative Coupling of Tyrosine Tags on Peptides and Proteins. J. Am. Chem. Soc 2020, 142 (11), 5078–5086. 10.1021/jacs.9b12002. [PubMed: 32093466]

(52). Jurchen JC; Williams ER Origin of Asymmetric Charge Partitioning in the Dissociation of Gas-Phase Protein Homodimers. J. Am. Chem. Soc 2003, 125 (9), 2817–2826. 10.1021/ja0211508. [PubMed: 12603172]

(53). Hamerlynck LM; Bischoff AJ; Rogers JR; Roberts TD; Dai J; Geissler PL; Francis MB; Ginsberg NS Static Disorder Has Dynamic Impact on Energy Transport in Biomimetic Light-Harvesting Complexes. J. Phys. Chem. B 2022, 126 (40), 7981–7991. 10.1021/acs.jpcb.2c06614. [PubMed: 36191182]

(54). Li X; Song B; Chen X; Wang Z; Zeng M; Yu D; Hu D; Chen Z; Jin L; Yang S; Yang C; Chen B Crystal Structure of a Four-Layer Aggregate of Engineered TMV CP Implies the Importance of Terminal Residues for Oligomer Assembly. PLOS ONE 2013, 8 (11), e77717. 10.1371/journal.pone.0077717. [PubMed: 24223721]

(55). Miller ZM; Harper CC; Lee H; Bischoff AJ; Francis MB; Schaffer DV; Williams ER Apodization Specific Fitting for Improved Resolution, Charge Measurement, and Data Analysis Speed in Charge Detection Mass Spectrometry. J. Am. Soc. Mass Spectrom 2022. 10.1021/jasms.2c00213.

(56). Jordan JS; Xia Z; Williams ER Tips on Making Tiny Tips: Secrets to Submicron Nanoelectrospray Emitters. J. Am. Soc. Mass Spectrom 2022, 33 (3), 607–611. 10.1021/jasms.1c00372. [PubMed: 35157433]

(57). Harper CC; Williams ER Enhanced Multiplexing in Fourier Transform Charge Detection Mass Spectrometry by Decoupling Ion Frequency from Mass to Charge Ratio. J. Am. Soc. Mass Spectrom 2019, 30 (12), 2637–2645. 10.1007/s13361-019-02330-3. [PubMed: 31720975]

(58). The PyMOL Molecular Graphics System.

(59). Schrödinger Release 2021–4: Desmond Molecular Dynamics System, 2022.

(60). Mahoney MW; Jorgensen WL A Five-Site Model for Liquid Water and the Reproduction of the Density Anomaly by Rigid, Nonpolarizable Potential Functions. J. Chem. Phys 2000, 112 (20), 8910–8922. 10.1063/1.481505.

(61). Jorgensen WL; Chandrasekhar J; Madura JD; Impey RW; Klein ML Comparison of Simple Potential Functions for Simulating Liquid Water. J. Chem. Phys 1983, 79 (2), 926–935. 10.1063/1.445869.

(62). Lu C; Wu C; Ghoreishi D; Chen W; Wang L; Damm W; Ross GA; Dahlgren MK; Russell E; Von Bargen CD; Abel R; Friesner RA; Harder ED OPLS4: Improving Force Field Accuracy on Challenging Regimes of Chemical Space. J. Chem. Theory Comput 2021, 17 (7), 4291–4300. 10.1021/acs.jctc.1c00302. [PubMed: 34096718]

(63). Martyna GJ; Tobias DJ; Klein ML Constant Pressure Molecular Dynamics Algorithms. J. Chem. Phys 1994, 101 (5), 4177–4189. 10.1063/1.467468.

(64). Martyna GJ; Klein ML; Tuckerman M Nosé–Hoover Chains: The Canonical Ensemble via Continuous Dynamics. J. Chem. Phys 1992, 97 (4), 2635–2643. 10.1063/1.463940.

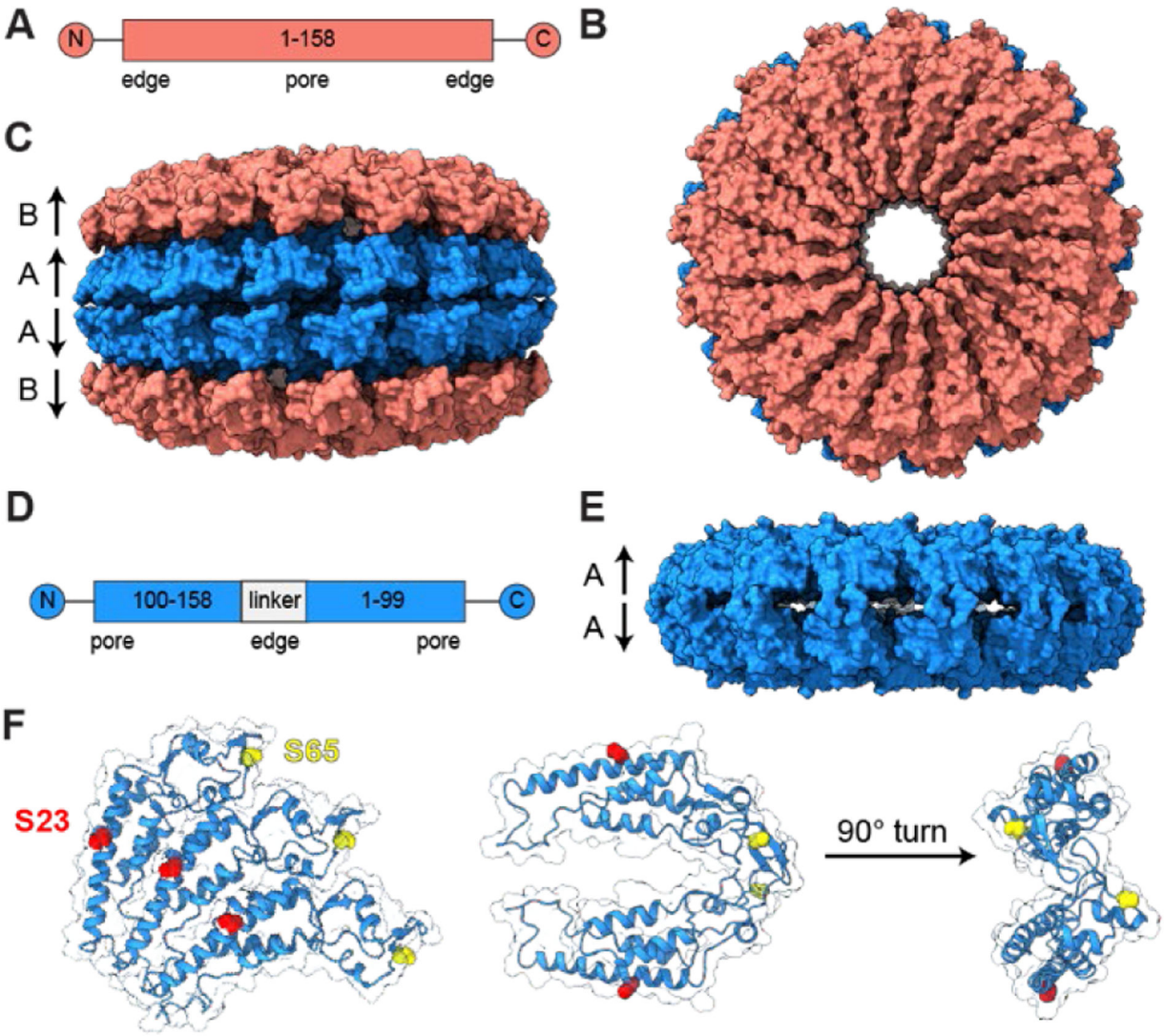


Figure 1. Structures of TMV used as light harvesting complex mimics.

(A) A gene map, (B) face view, and (C) side view of a four-layer stack of wtTMV disks displays C_2 symmetry about the horizontal axis, with A disks being identical to one another and B disks identical to one another but with slight structural variation from A disks (PDB ID: 1EI7). (D) A gene map and (E) side view of a homology model of cpTMV show the favored structure consisting of two disks, containing 17 monomers each, with C_2 symmetry (PDB ID: 3KML). (F) Engineered residues used for attachment to chromophores or surfaces in light harvesting mimics are shown for three monomers on one cpTMV disk, with the S23 site shown in red and S65 site in yellow. The S23 and S65 sites are also shown for two monomers on opposing disks from a cut-away and edge-on view.

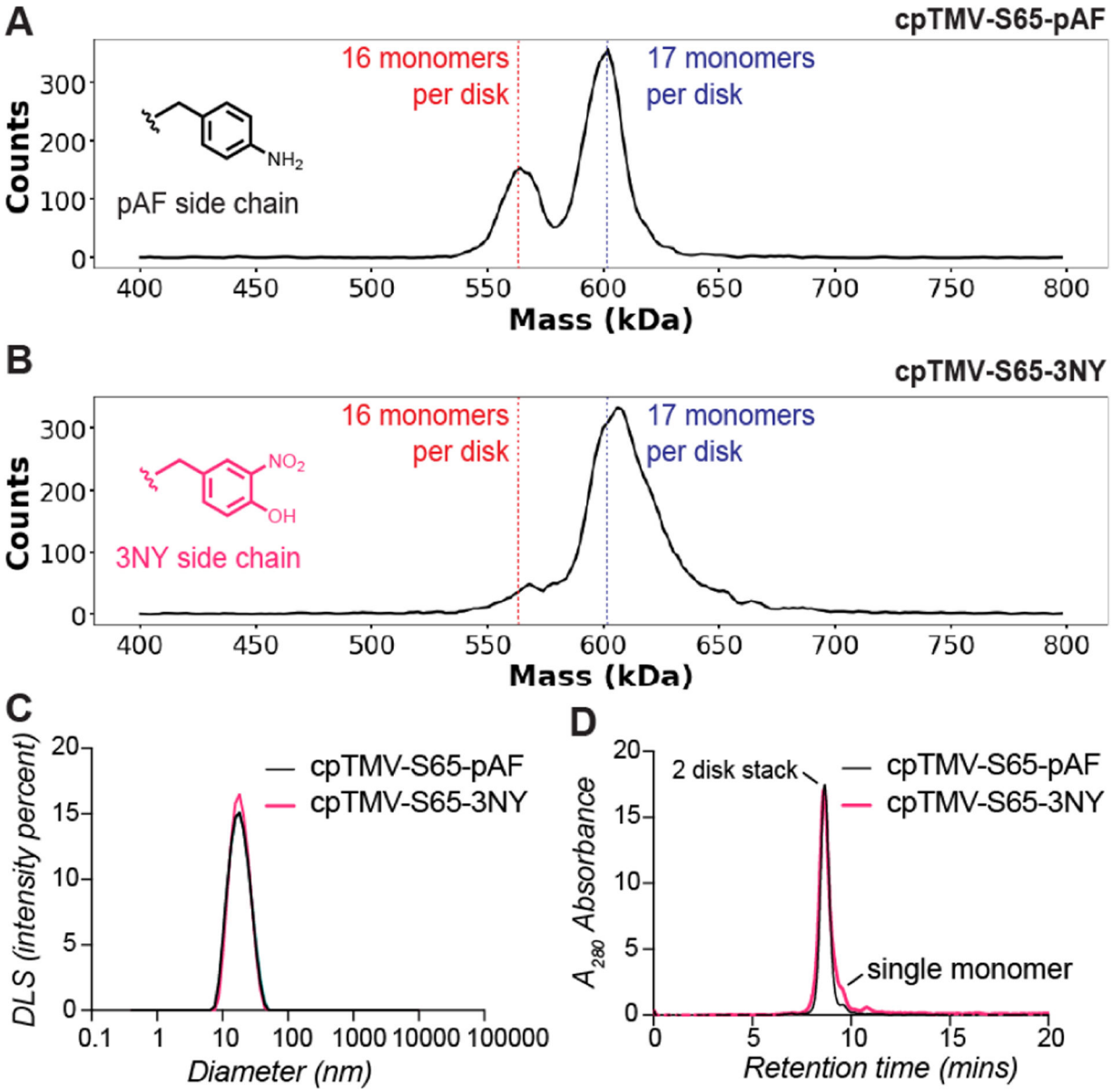


Figure 2.

Mutants of TMV displaying disk assemblies with different symmetries. (A) A mass histogram of a circular permutant of TMV with a peripheral p-aminophenylalanine (pAF) mutation (cpTMV-S65-pAF) in 100 mM ammonium acetate solution clearly differentiates between 16-fold symmetric and 17-fold symmetric stacks of two disks. (B) A mutant of cpTMV with 3-nitrotyrosine (3NY) at the same site, cpTMV-S65-3NY, in 100 mM ammonium acetate solution contains a much smaller ratio of 16:17-mers. Distinguishing between 16- and 17-mers is challenging by other common characterization methods, with no differentiation observed using (C) dynamic light scattering and (D) size exclusion chromatography.

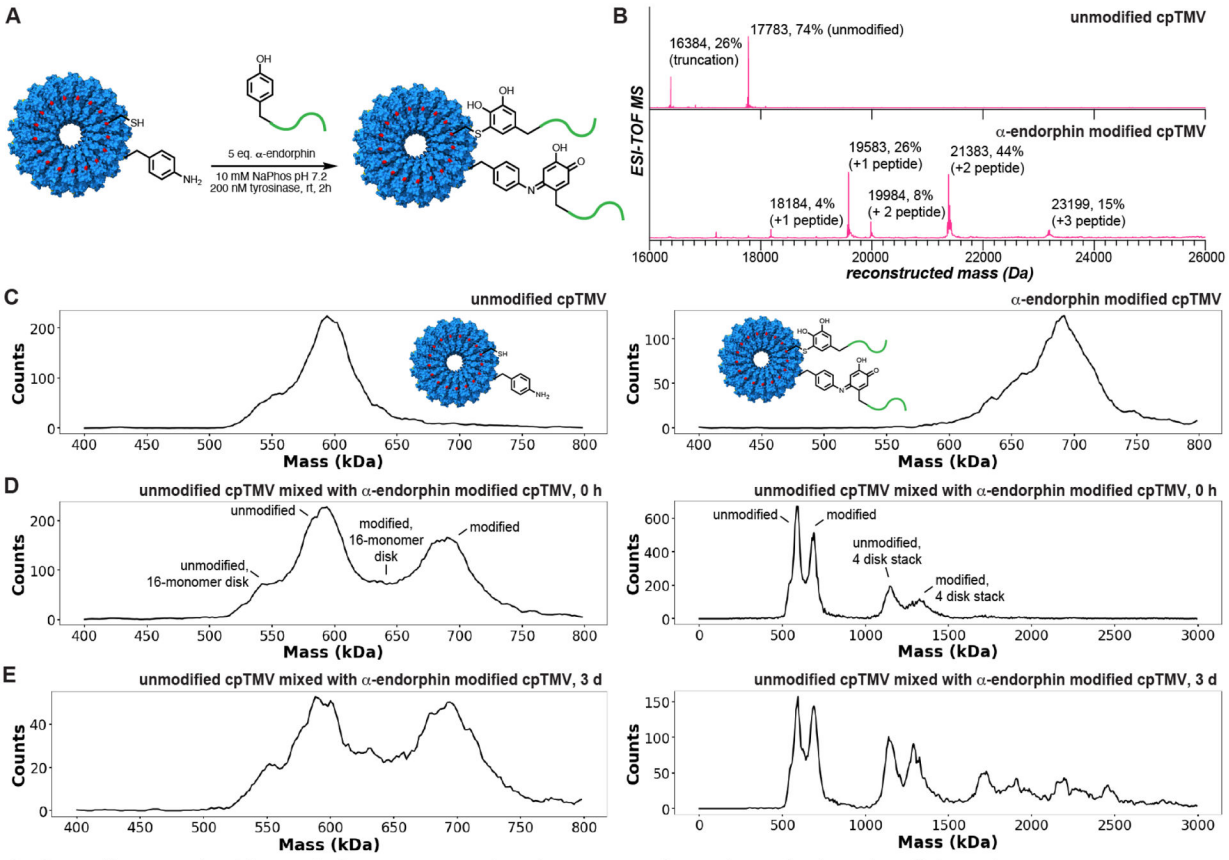


Figure 3.

Monitoring protein assembly dynamics using charge detection mass spectrometry. (A) cpTMV-S23C-S65-pAF was modified with α-endorphin, a 2 kDa peptide, using the enzyme tyrosinase from *Agaricus bisporus* (abTyr). (B) One to two copies of α-endorphin were attached per cpTMV-S23C-S65-pAF monomer as assessed by ESI-TOF mass spectrometry. (C) A comparison of the mass histograms of unmodified and α-endorphin-modified cpTMV-S23C-S65-pAF disks in 100 mM ammonium acetate solution shows a clear difference in size of approximately 90 kDa. A mixture of the unmodified and α-endorphin-modified disks in 100 mM ammonium acetate solution (D) immediately after mixing and (E) incubated at rt 3 days after mixing displays two distinct populations that do not equilibrate over time. For (D) and (E), both a window showing only the two-disk stack and a larger window showing higher stacking stoichiometries are displayed. A statistical analysis of the intact disk distributions based on the monomer % modification is shown in Figure S5.

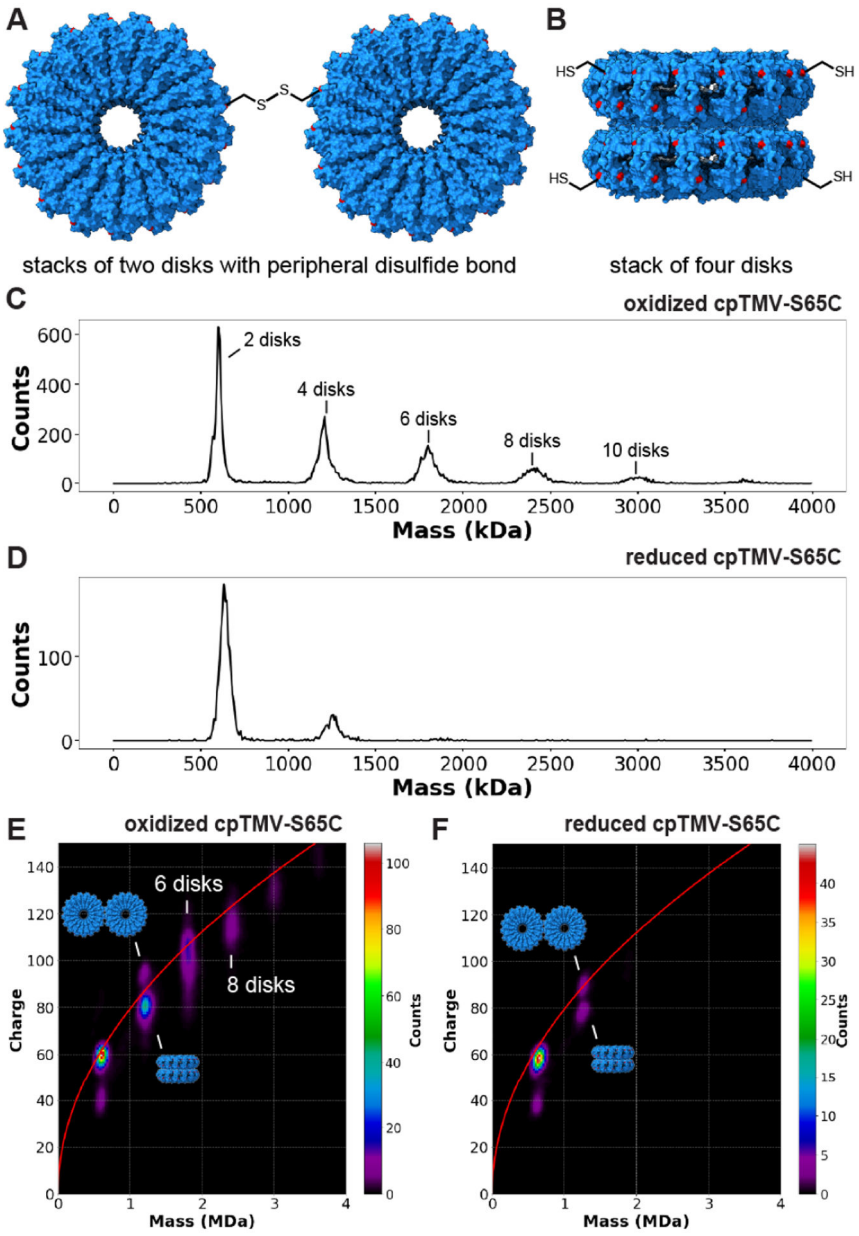


Figure 4.

Distinguishing between assembly states of nearly identical mass using CDMS. (A) The cpTMV-S65C mutant can both form disulfide bonds at the disk's periphery and (B) assemble into stacks of four disks, resulting in two structures with a nearly equal mass at 1.2 MDa. The location of the S65C mutation is shown in red. (C) A mass histogram of cpTMV-S65C in 100 mM ammonium acetate solution oxidized in air displays assemblies with even-numbered disk stoichiometries varying from two to twelve disks per assembly. (D) After reduction with 1 mM tris(2-carboxyethyl)phosphine (TCEP) for 1 h, there is a decrease in the proportion of larger-sized assemblies of cpTMV-S65C. (E) A two-dimensional mass vs. charge histogram of air-oxidized cpTMV-S65C in 100 mM ammonium acetate solution displays two distinct populations at 1.2 MDa above and below the Rayleigh limit (red line).

It also shows broad charge distributions for species with greater than four disks which indicate a mixture of different stacked and edge-on conformers. (F) After treatment with TCEP for 1 h, populations larger than four disks were no longer visible, and only a small population of the higher charged species at 1.2 MDa remained, indicating that this is the peripherally disulfide bonded species.

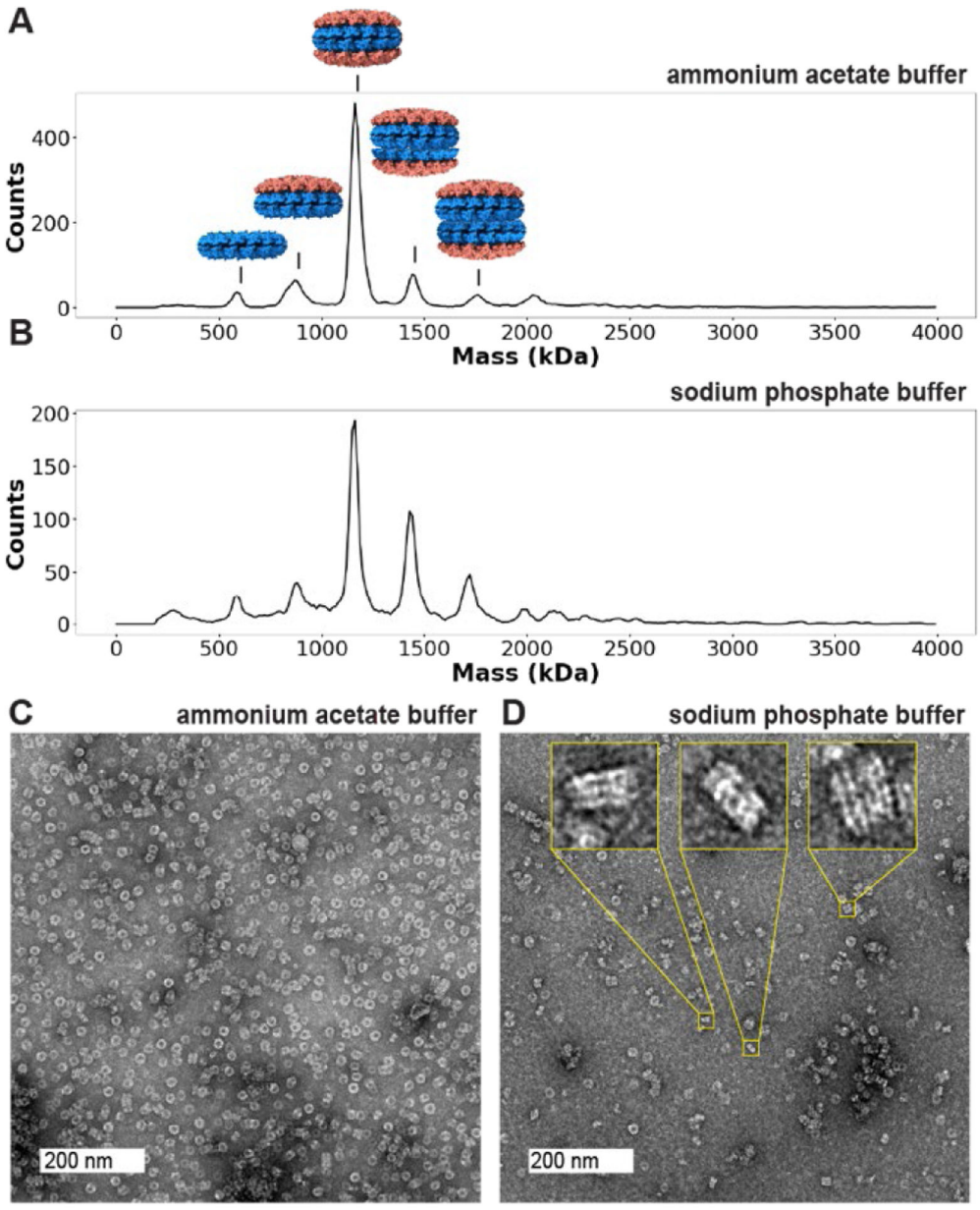


Figure 5.

Odd-numbered stacks of disks in recombinant tobacco mosaic viral capsids (rTMV). Mass histograms of rTMV capsid ions formed by electrospray from (A) 18.2 mM ammonium acetate pH 7 and (B) 10 mM sodium phosphate pH 7.2 solutions display different ratios of number of disks per stack, with a single disk having a molecular weight of 300 kDa. (C) A representative TEM image of rTMV in 18.2 mM ammonium acetate pH 7 shows many disks on their faces with a few larger stacks. (D) A representative TEM image of rTMV in 10 mM sodium phosphate pH 7.2 shows several examples of odd-numbered stacks of rTMV disks. Inlays show stacks with odd numbers of layers (3 and 5 disks per stack). A collection of images can be found in SI Figure S9.