analytical chemistry

pubs.acs.org/ac Article

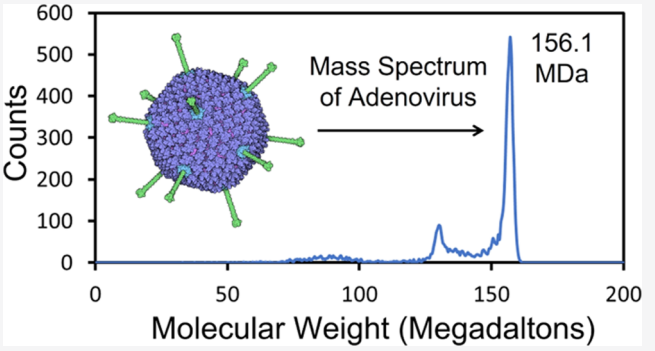
Analysis of Recombinant Adenovirus Vectors by Ion Trap Charge Detection Mass Spectrometry: Accurate Molecular Weight Measurements beyond 150 MDa

Lauren F. Barnes, Benjamin E. Draper, and Martin F. Jarrold*

Cite This: *Anal. Chem.* 2022, 94, 1543−1551 Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Adenovirus is one of the largest nonenveloped, double-stranded DNA viruses. It is widely used as a gene therapy vector and has recently received a lot of attention as a novel vaccine platform for SARS-CoV-2. Human adenovirus 5 (HAdV5) contains over 2500 protein molecules and has a 36 kbp genome. Adenovirus is well beyond the range of conventional mass spectrometry, and it was unclear how well such a large complex could be desolvated. Here, we report molecular weight (MW) distributions measured for HAdV5 and for 11 recombinant AdV vectors with genomes of varying lengths. The MW distributions were recorded using ion trap charge detection mass spectrometry (CDMS), a single-particle technique where m/z and charge are measured for individual ions. The results show that ions as large as 150 MDa can be effectively desolvated and accurate MW distributions obtained. The MW distribution for HAdV5 contains a narrow peak at 156.1 MDa, assigned to the infectious virus. A smaller peak at 129.6 MDa is attributed to incomplete particles that have not packaged a genome. The ions in the 129.6 MDa peak have a much lower average charge than those in the peak at 156.1 MDa. This is attributed to the empty particles missing some or all of the fibers that decorate the surface of the virion. The MW measured for the mature virus (156.1 MDa) is much larger than that predicted from sequence masses and copy numbers of the constituents (142.5 MDa). Measurements performed for recombinant AdV as a function of genome length show that for every 1 MDa increase in the genome MW, the MW of the mature virus increases by around 2.3 MDa. The additional 1.3 MDa is attributed to core proteins that are copackaged with the DNA. This observation suggests that the discrepancy between the measured and expected MWs for mature HAdV5 is due to an underestimate in the copy numbers of the core proteins.



INTRODUCTION

Molecular weight (MW) is a particularly revealing metric, and mass spectrometry is a ubiquitous tool in the characterization of all sorts of molecules. A major limitation with conventional mass spectrometry is its restriction to MWs less than around a megadalton (MDa) unless analytes are extremely homogeneous. There are many samples with much larger MWs, where accurate MW information would be valuable for characterization and quality control, for example, viruses, virus-like particles, and vaccines, to name but a few. In this article, we report the routine measurement of accurate MWs beyond 150 MDa. Specifically, we report MW distributions measured for adenovirus, a widely used gene therapy vector and novel vaccine platform. The detailed MW information provided by these studies opens the door to much more rigorous characterization of recombinant adenovirus vectors and other very high MW pharmaceuticals with MWs in the hundreds of megadaltons.

Adenoviruses are large nonenveloped, double-stranded DNA (dsDNA) viruses. There are over 120 species-specific serotypes. In humans, adenoviruses can cause respiratory, ocular, gastrointestinal, and urinary tract infections.[1] The infections are usually mild but can become life-threatening to the immunocompromised.[2] Adenoviruses have been used as gene therapy and vaccine vectors.[3,4] For example, the Johnson and Johnson COVID-19 vaccine, which was granted an emergency use authorization, uses an adenovirus serotype 26 vector.[5]

Human adenovirus type 5 (HAdV5) is the most widely studied adenovirus. It has an icosahedral pseudo-T = 25 capsid consisting of 240 hexon trimers and 12 penton pentamers (see Figure 1).[6−8] The vertex-to-vertex diameter is around 95 nm. A fiber consisting of trimers of the fiber protein is attached to

Received: June 9, 2021
Accepted: December 30, 2021
Published: January 13, 2022

https://doi.org/10.1021/acs.analchem.1c02439
Anal. Chem. 2022, 94, 1543−1551

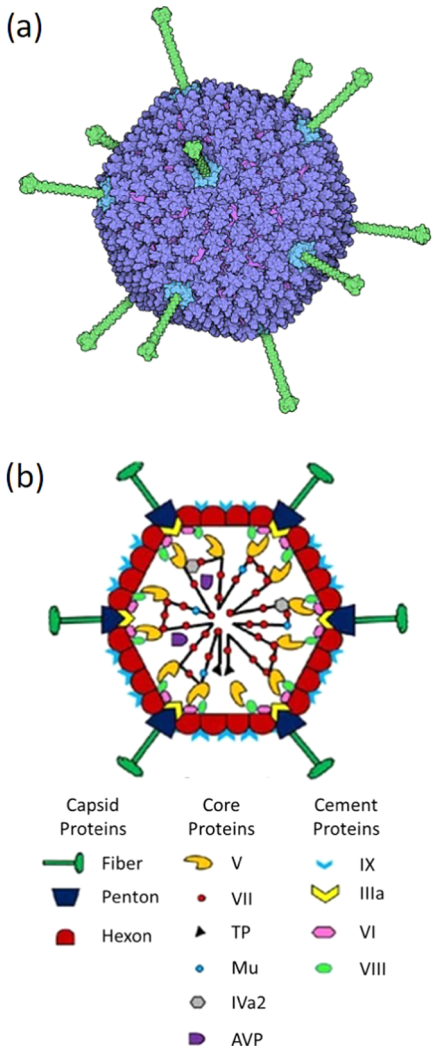


Figure 1. Structure of HAd5. Part (a) shows an exterior view. The hexons are purple, the pentons are blue, and the fibers are green. Part (b) is a schematic cross section of mature HAdV5 showing the locations of the proteins (TP, terminal protein; AVP, adenovirus protease). Part (a) is from ref 13. Part (b) was adapted from ref 12.

each penton, leading to an overall diameter of around 165 nm. In addition to the three structural proteins mentioned above, there are multiple copies of at least 10 more proteins in the mature virion: four cement proteins and six core proteins (see Figure 1b).[9−12] Three of the core proteins (V, VII, and Mu) are highly basic and play an important role in the condensation of the genome. The cement proteins (IIIa, VI, VIII, and IX) help to stabilize the capsid. Altogether, there are over 2500 protein molecules in the HAdV5 virion along with a 35 938 bp linear dsDNA genome. Its MW is expected to be over 100 MDa, well beyond the range of conventional MS.

Measurement of the MW distribution for adenoviruses would provide important information about their composition and heterogeneity and could provide insight into adenovirus assembly and disassembly, where there remain many unresolved questions.[11,14,15] For another much smaller gene therapy vector, adeno-associated virus (AAV), MW distributions have been used to determine the relative abundances of full and empty particles (AAV particles that have packaged a genome and those that have not).[16] MW distributions can also provide information about particles that have packaged a partial genome or heterogeneous DNA, as well as information on aggregation and multimers.[17] Accurate MW measurements for adenovirus should provide similar information, but for a much larger target, while opening up the possibility of extending MW measurements to even larger targets such as lentivirus.

In conventional MS, the quantity measured is the mass-to-charge ratio (m/z) of the ions. Large ions, particularly those generated by electrospray, are multiply charged, and so the charge must be deduced from the m/z spectrum to determine the mass. This is usually possible if the peaks due to different charge states are resolved in the m/z spectrum. However, heterogeneity broadens the peaks, causing the loss of charge-state resolution, and, consequently, the mass cannot be determined. There are many examples of low-MW samples being heterogeneous,[18] but heterogeneity generally increases with molecular weight. There are a few examples, usually highly purified or manipulated samples, where conventional MS has been applied successfully to molecules with MWs greater than 1 MDa.[19−23]

One solution to this problem is to perform mass measurements for individual ions. This can be accomplished using a charge-stepping approach where two or more m/z measurements are made for the same ion in different charge states.[24−32] For example, Nie and workers used laser-induced acoustic desorption and a cylindrical ion trap to measure the masses of adenovirus, iridovirus, and vaccinia virus.[27,28] The charge-stepping approach is time consuming, and mass measurements are usually only performed for a handful of ions. Another approach to determining the masses of individual ions is to measure each ion's m/z ratio and charge. For example, Chen et al. measured masses for coliphage T4 DNA ions using Fourier transform ion cyclotron resonance (FTICR) MS.[33] The ion's charge was determined from the charge induced by the ion on the detector plates. Since the detector plates do not surround the ion, the induced charge depends on the ion's trajectory in the FTICR trap, and the uncertainty in the charge determination was relatively large (around 10%). In charge detection mass spectrometry (CDMS),[34] the ions pass through a conducting cylinder. If the cylinder is long enough, the induced charge is independent of the ion's trajectory and can be measured accurately.

In single-pass CDMS, the ions travel through the detector once.[34−47] This approach has the advantage of high throughput, but the uncertainty in the charge measurement (which is limited by electrical noise) is high, leading to mass resolving powers in the single digits. The charge uncertainty is substantially reduced in ion trap CDMS, where the detection cylinder is located between the end caps of an electrostatic linear ion trap.[48−56] Trapped ions oscillate back and forth through the detection cylinder, and the resulting signal is processed by fast Fourier transforms (FFTs). The ion's m/z is determined from the oscillation frequency, and the charge is determined from the FFT magnitude. The charge can be determined with an uncertainty of around 0.2 e (elementary charges), which is sufficient to see almost baseline resolution of charge states in the charge spectrum.[57−59] The ions can then be assigned to the nearest integer charge state with a low error rate. Recently, there has been interest in performing single-ion mass measurements on an orbitrap;[60,61] at present, the charge uncertainty is relatively high, around 4%.

Ion trap CDMS has previously been used to investigate a number of issues related to virus composition, assembly, and

https://doi.org/10.1021/acs.analchem.1c02439
Anal. Chem. 2022, 94, 1543−1551

disassembly,[62−75] including accurate MW measurements for 50 MDa bacteriophage P22.[76] In this work, we push the envelope of accurate high-MW measurements even further and show how they are enabling in the characterization of adenovirus. Mass and charge information from ion trap CDMS has been used to identify misassembled capsids and fully assembled capsids while also assessing genome packaging.

MATERIALS AND METHODS

Sample Preparation for HAdV5 and Recombinant Adenoviruses. The HAdV5 sample was purchased from the American Type Culture Collection (ATCC VR-1516). The recombinant adenovirus samples (listed in Table 1) were purchased from Vector Biolabs (Malvern, PA), except for Ad-CMV-GFP/ntLAcZ and Ad-Empty, which were purchased from Viraquest (North Liberty, IA). Sample concentrations are given in Table S2 in the Supporting Information. The proteins in the recombinant adenoviruses have the same sequences as in HAdV5. While the copy numbers for the structural proteins are fixed by symmetry, the copy numbers for the core and cement proteins are not. The genome sequence masses in Table 1 were calculated from sequences provided by the vendors. The samples were manipulated in a class II biosafety cabinet (NuAire LabGard ES type A2). Sample aliquots were stored at −80 °C and thawed at room temperature. Prior to analysis, small volumes (10−20 μL) were desalted and buffer-exchanged into an ammonium acetate (Honeywell 631−31−8) solution by size-exclusion chromatography (Micro Bio-Spin P-6 Gel Columns Bio-Rad).

Table 1. Recombinant Adenovirus Samples Analyzed by Ion Trap CDMS

genome	sequence length (bp)	sequence mass (MDa)
Ad-Empty	32 814	20.28
Ad-CMV-GFP/ntLAcZ	36 545	22.58
Ad-U6-scrmb-shRNA	31 515	19.48
Ad-CMV-Null	32 312	19.97
Ad-RFP	33 019	20.40
Ad-mCherry	33 064	20.43
Ad-CMV-GFP	33 068	20.43
Ad-mCherry-U6-scrmb-shRNA	33 414	20.65
Ad-GFP-U6-scrmb-shRNA	33 418	20.65
Ad-CMV-Luc	33 992	21.01
Ad-CMV-LacZ	35 471	21.92

Charge Detection Mass Spectrometry. In CDMS, the charge and m/z are simultaneously measured for each ion. The m/z and charge are then multiplied to give the mass. This procedure is performed for thousands of ions, and the results are binned into a histogram to give a mass spectrum. The homebuilt prototype CDMS instrument used in this work has been described previously.[49,77,78] Briefly, ions generated by nanoelectrospray (Advion Triversa Nanomate) enter the instrument through a metal capillary and pass through three stages of differential pumping that contain, in turn, a FUNPET,[77] an RF hexapole, and a segmented RF quadrupole. Ions that exit the quadrupole are focused on the dual hemispherical deflection energy analyzer (HDA) that selects a narrow band of ion kinetic energies centered on a nominal value of 100 eV/z. The transmitted ions are focused into an electrostatic linear ion trap (ELIT) based on a cone trap.[79] The ELIT consists of two conical end caps with potentials that can be switched between transmission and reflection modes to trap ions. A detection cylinder is located between the end caps, and trapped ions oscillate back and forth through the detection cylinder. When in the cylinder, the ions induce a charge that is detected by a charge-sensitive amplifier. The resulting signal is amplified, digitized, and then transferred to a computer where

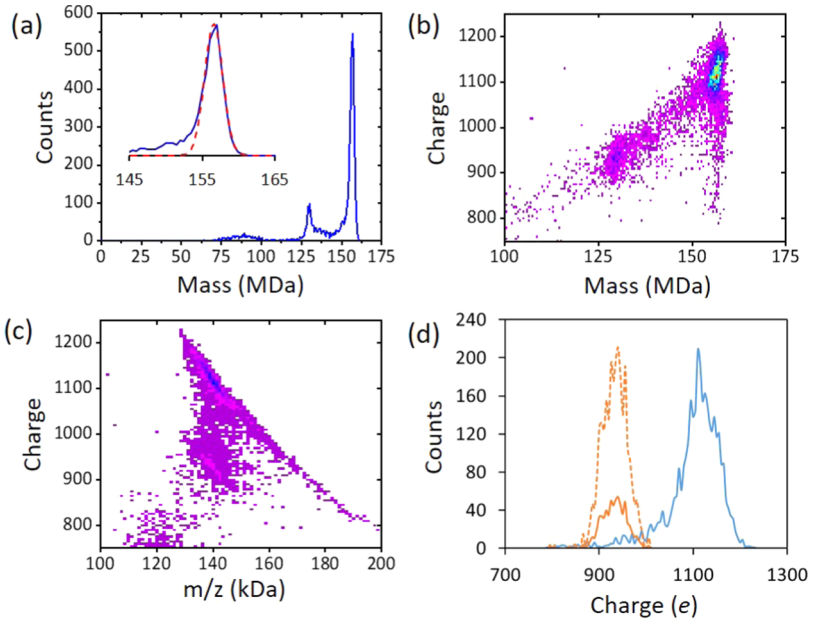


Figure 2. Typical mass and charge distributions measured for HAdV5 by ion trap CDMS. Part (a) shows the mass spectrum. The inset shows an expanded view of the main peak with a Gaussian fit (red dashed line). Part (b) shows a charge versus mass heat map, where warmer-colored pixels represent bins with higher density. Part (c) shows a charge versus m/z heat map. Part (d) shows the charge distributions for ions in the mass peak at around 130 MDa (127−133 MDa) (orange line) and the peak at around 156 MDa (154−160 MDa) (blue line). The dashed line shows the counts for the 130 MDa peak scaled up by a factor of 3.9.

https://doi.org/10.1021/acs.analchem.1c02439
Anal. Chem. 2022, 94, 1543−1551

Table 2. Components of Mature HAdV5[a]

protein	uniprot ID length	location	function	sequence mass (Da)	copy numbers	total mass (MDa)
hexon	P04133 2-952	coat	structure	107 858	720	77.658
penton	P12538 1-571	vertex	structure, entry	63 293	60	3.798
penton fiber	P11818 1-581	vertex	structure, entry	61 585	36	2.217
IIIa	P12537 1-570	cement	cement	63 502	68	4.318
VI	P24937 34-239 240-250	cement	cement, entry, trafficking, early gene expression	23 450	342	8.020
VIII	P24936 2-111 158-227	cement	cement	19 628	128	2.512
IX	P03281 2-140	cement	cement	14 309	247	3.534
V	P24938 2-368	core	core condensation	41 298	157	6.484
VII	P68951 25-198	core	core condensation	19 412	527	10.230
Mu	Q2KS10 33-51	core	core condensation	2 441	290	0.708
TP	P04499 176-671	core	DNA replication	56 303	2	0.113
AVP	P03253 2-204	core	maturation	22 919	15	0.344
IVa2	P03271 1-449	unique vertex	packaging	50 887	7	0.356
dsDNA	35 938 bp	core	genome	22 207 954	1	22.208
total					2599	142.499

[a]Copy numbers were taken from ref 11.

it is analyzed by fast Fourier transforms. The oscillation frequency is related to the *m/z*, and the magnitude is proportional to the charge. Events where the ion was not trapped for the full trapping period (100 ms) were discarded.

RESULTS AND DISCUSSION

A typical CDMS spectrum for HAdV5 (ATCC VR-1516) is shown in Figure 2a. The spectrum contains 6229 individual ion measurements and took around 20 min to record, during which 3−4 μL of a 10 μL sample at around 5.8×10^{11} VP/mL was consumed. There are three prominent populations in the spectrum; the main peak centered on 156.1 MDa, a second smaller peak centered on 129.6 MDa, and a broad, low-intensity distribution centered on around 90 MDa. The peak at 156.1 MDa is attributed to the mature HAdV5 virion. The peak at 129.6 MDa is attributed to incomplete particles that have not packaged a genome (i.e., empty particles).[80,81] Light and heavy particles (where heavy particles are the mature virions) are separable by gradient centrifugation. CDMS measurements for gradient-separated light and heavy recombinant chimpanzee adenovirus C68 have confirmed the assignment of the peaks at around 130 and 156 MDa to empty and full particles, respectively.[82] The broad distribution centered on around 90 MDa is attributed to misassembled or trapped intermediates. Table 2 shows the components of mature HAdV5, along with the masses derived from the sequences (accounting for truncations). The copy numbers for the proteins in Table 2 were taken from ref 11. Adding together the masses of all of the components, including the 35 938 bp linear dsDNA genome, leads to an expected mass of 142.5 MDa. This is significantly less than the measured mass (156.1 MDa).

Ions electrosprayed from volatile salt solutions often have measured masses that are slightly larger than the expected mass.[83,84] The deviation, typically around a percent, is usually attributed to counterions, residual salt, and trapped solvent. However, adenovirus is much larger than the ions usually measured by native MS and it is not known how well such a large species will desolvate. Thus, it is reasonable to ask how much of the difference between the measured and the expected MWs for adenovirus is due to inadequate desolvation? The FUNPET electrospray interface (an ion funnel-ion carpet hybrid) employed in our CDMS instrument was designed to thermalize gigadalton ions after they had been accelerated to supersonic velocities in a capillary interface.[77] The lowest free-energy state of a solvated ion in a low solvent partial pressure environment at room temperature is a desolvated ion and free solvent molecules.[85] The FUNPET is equipped with a long drift region, so that ions are kept in a room-temperature buffer gas for an extended period of time to gently desolvate them. With this interface, the difference between the measured MW and the expected MW is less than 1% for viruses and capsids with MWs in the 5 MDa range.[86] This observation is consistent with molecular dynamics simulations that show virus capsids are porous;[87,88] all of the water molecules inside an empty poliovirus capsid ($\sim 2 \times 10^5$) exchange in solution in around 25 μs.[88] However, the much larger size of adenovirus led us to re-examine the question of desolvation.

To investigate how well adenovirus is desolvated, measurements were performed as a function of the dc voltage across the FUNPET. As the voltage is reduced, ions spend longer in the drift region and have more time to desolvate. When the voltage is lowered to extend the time spent in the drift region, the signal decreases, dropping precipitously for voltages below 80 V. Figure 3 shows the percent deviation in the mass of the main peak in the spectrum plotted against the FUNPET dc

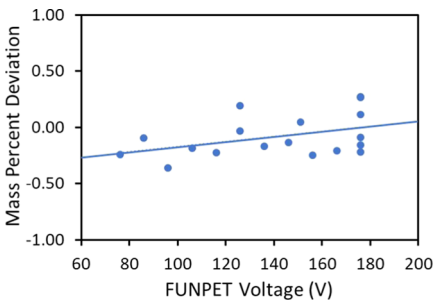


Figure 3. Desolvation of HAdV5. Plot of the mass percent deviation of the main peak in the HAdV5 spectrum against the FUNPET DC voltage. The points are the measurements, and the line is a linear least-squares fit.

https://doi.org/10.1021/acs.analchem.1c02439
Anal. Chem. 2022, 94, 1543−1551

voltage. The points are the measurements, and the line is a linear least-squares fit. The change in the mass as the FUNPET voltage is lowered is small, less than 0.25%. This is in line with the deviations we have observed in smaller systems where the mass is known more precisely.

The approach adopted here may appear counterintuitive to those familiar with traditional electrospray interfaces where an increase in the DC voltage across the interface leads to increased collisional activation and better desolvation. In these interfaces, the time available for desolvation is limited and collisional activation is used to compensate. However, collisional activation is not necessary to achieve desolvation, as our prior results for particles in the 5 MDa range,[86] and the results in Figure 3, demonstrate. In Figure 3, the measured mass decreases as the voltage across the interface is lowered. This result suggests that poor desolvation is not the main factor responsible for the 10% deviation between the measured and expected masses of HAdV5.

The inset in Figure 1a shows an expanded view of the main peak in the spectrum for HAdV5. The dashed red line is a Gaussian fit with a full width at half-maximum of 3.08 MDa. The peak width includes contributions from the instrumental resolution and the underlying peak width due to sample heterogeneity. Recall that the mass is determined from the product of each ion's m/z and charge, so uncertainties in both quantities contribute to the mass resolution. Because the charge on the 156 MDa ions is so large (see below), the relative uncertainty in the charge is much smaller than that in the m/z. Thus, the mass resolution under the conditions used to record the spectrum in Figure 2a is dominated by the uncertainty in the m/z determination. Assuming that the mass resolution and the underlying peak width due to sample heterogeneity are uncorrelated, the underlying peak width is around 2.3 MDa. Since the genome has a fixed length, and the numbers of hexons, pentons, and fiber proteins are fixed by symmetry, the peak width probably results mainly from variability in the numbers of cement and core proteins. In addition, there is probably a distribution in the number of counterions, residual salt, and trapped solvent molecules that contribute to the excess mass. In particular, the dsDNA genome probably has counterions associated with it; the number of counterions probably depends on the number of highly basic core proteins (V, VII, and Mu) associated with the DNA.

In addition to the 13 proteins listed in Table 2, which are known to be present in the mature virion, there are at least 6 proteins that do not appear in the mature virion but play a role in its assembly, packaging, and maturation.[11] For most of these proteins, there is no information available on the number of copies present in the empty HAdV5 particle. Removing the genome and the core condensation proteins from the list in Table 2 leads to a mass of 102.9 MDa for the remaining capsid and cement proteins. This value is significantly less than the mass of the lower-intensity peak at 129.6 MDa in Figure 2a. Most of the deficit probably result from scaffolding proteins involved in capsid assembly.

Since the charge and mass are independent variables, it is instructive to see how they are correlated. Large ions generated by electrospray are thought to be produced by the charge residue mechanism where the evaporating droplet deposits its charge on the analyte.[89,90] The charge on the analyte is then given approximately by the charge on a water droplet of the same size, which is, in turn, determined by the Rayleigh limit.[91] Thus, higher charges are expected to correspond to larger or more elongated structures, while lower charges correspond to more compact structures.[65] In previous studies of AAV, we found that empty particles and particles that had packaged a full genome had similar charges,[16,17] a result consistent with the genome being packaged inside the capsid. Thus, the empty and full AAV capsids have different masses but similar sizes and hence similar charges. Figure 2b,c shows heat maps of charge versus mass and charge versus m/z for HAdV5. In these plots, the warmer-colored pixels represent bins with higher density. It is evident from Figure 2b that the charge on the lighter 129.6 MDa particles is significantly lower than on those that make up the heavier 156.1 MDa particles. Figure 2d shows charge distributions for the lighter particles (orange line) and the heavier particles (blue line). The charge distributions are centered on around 930 e for the lighter particles and 1120 e for the heavier. There is a significantly low charge tail that extends to below 1000 e for the heavier particles. The lower charges observed for the lighter particles suggest that they have more compact structures than the heavier particles. One plausible explanation for this difference is that the lighter particles are missing some or all of the fibers that decorate the surface of the mature HAdV5. This is consistent with the previous work showing that the abundance of the fiber proteins is substantially diminished for empty particles.[80] The relatively narrow charge distribution indicates that the distribution of missing fibers is narrow.

The loss of fibers cannot be responsible for the low charge tail observed for the heavier particles because there is no corresponding decrease in mass as the charge decreases. The fibers are flexible and are thought to bend after initial attachment to the host cell to allow the penton base to bind and initiate internalization of the capsid.[14,92] Thus, we postulate that the low charge tail is due to fibers folding down on the surface of the capsid. The absence of a low charge tail for the lighter particles is consistent with this explanation. Finally, the inspection of Figure 2b reveals that the most intense region of the charge versus mass heat map (the blue and light blue areas) is not symmetric, but as the charge decreases, the mass also decreases slightly. This may indicate that some of the particles in this region of the heat map are missing a few fibers, causing the charge to decrease.

In the charge versus m/z heat map in Figure 2c, the two main features in the mass distribution (the light and heavy particles at 129.6 and 156.1 MDa, respectively) have similar m/z values. Thus, the m/z spectra for these two mass peaks overlap. This leads to a very complex m/z spectrum, but with the charge measurement, it is easy to resolve these components.

It is likely that the significant discrepancy between the measured mass (156.1 MDa) and the expected mass (142.5 MDa) of mature HAdV5 results mainly from underestimation of the copy numbers of some of the proteins in Table 2. The copy numbers for the capsid proteins (hexon, penton, and fiber) are defined by symmetry. Copy numbers for the cement and core proteins are not defined by symmetry, and there is some uncertainty associated with them. The copy numbers in Table 2 are derived from the 2016 review by Ahi and Mittal.[11] There is reason to believe that some of the copy numbers in Table 2 may be underestimated. In a 2014 compilation, Benevento et al. reported a significantly higher copy number for protein VII (833 versus 527 in Table 2).[10] Furthermore, in measurements reported in ref 10, the copy number for protein

https://doi.org/10.1021/acs.analchem.1c02439
Anal. Chem. 2022, 94, 1543−1551

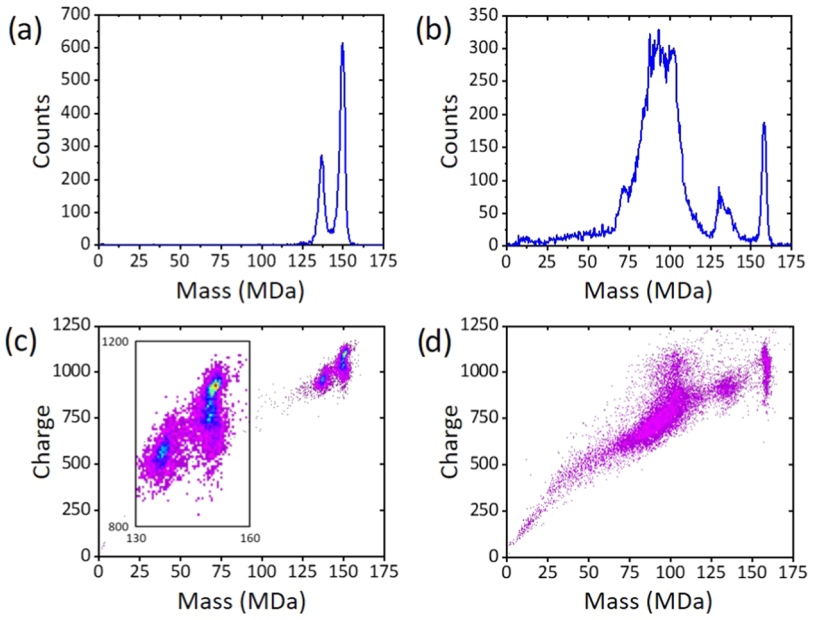


Figure 4. Mass histograms of (a) Ad-empty and (b) Ad-CMV-GFP/ntLacZ. Parts (c) and (d) show charge versus mass heat maps of Ad-Empty and Ad-CMV-GFP/ntLacZ. Warmer-colored pixels represent bins with higher density. The inset in part (c) shows an expanded view of the high mass and charge region.

IX is larger than in Table 2 (676 vs 247). To investigate this issue further and get insight into whether the issue is mainly with core or cement proteins, we made a series of measurements with recombinant adenovirus with genomes of varying lengths. Table 1 provides a list of the recombinant adenovirus samples that were investigated.

Figure 4 shows the mass spectra and charge versus mass heat maps for Ad-Empty and Ad-CMV-GFP/ntLAcZ. Ad-Empty has a genome where E1 and E3 are deleted but there is no promoter or transgene. The Ad-Empty mass spectrum shown in Figure 4a has a peak at 149.6 MDa that is attributed to the mature particle. A second peak around 136.9 MDa is likely an immature capsid or trapped intermediates. Figure 4c shows charge versus mass heat maps for Ad-Empty. The charge on the higher mass peak (150 MDa) is centered on around 1040 e, and the charge for the lower mass peak (137 MDa) is centered on around 900 e. The substantial decrease in the charge of the lower mass peak is similar to the results for HAdV5 described above (930 e for the 130 MDa particles and 1120 e for the 156 MDa) and consistent with the lower mass peak for Ad-Empty missing some or all of the fibers on the capsid exterior.

The mass spectrum for Ad-CMV-GFP/ntLAcZ (Figure 4b) shows a sharp peak at 157.7 MDa, which is attributed to the mature particle. There are two broad peaks at lower masses: the smaller peak centered on around 133 MDa and the larger centered on 95 MDa. Most of the intensity is in the 95 MDa peak. The charge versus mass heat map is shown in Figure 4d. The charges for the peak attributed to mature particles (158 MDa) are similar to those for the mature particles for HAdV5 and Ad-Empty.

Mass distributions recorded for the other nine recombinant AdV listed in Table 1 were dominated by a single peak that was assigned to the mature particle, and the abundances of lower mass ions were considerably less than for Ad-Empty and Ad-CMV-GFP/ntLAcZ in Figure 4. Typical mass distributions and charge versus mass heat maps are given in the Supporting Information.

Figure 5 shows the measured masses of the mature particle for HAdV5 and the recombinant AdV samples listed in Table 1 plotted against the genome sequence masses. The points are the average of three independent measurements. A table of the average measured masses and the standard deviations is given in the Supporting Information (Table S1). Run-to-run differences in the measured masses appear to be associated with the number of freeze−thaw cycles the samples experienced. An increase in the number of freeze−thaw cycles also reduced the signal intensity. The dashed line in Figure 5 shows a linear least-squares fit to the points, where the slope is 2.315 and the intercept is 105.4 MDa. The slope indicates that for every additional 1 MDa of genome, the particle mass increases by an additional 1.3 MDa. This additional mass is most likely due to the proteins that are closely associated with the packaged DNA: proteins V, VII, and Mu (those listed with a function of core condensation in Table 2). The intercept should correspond to the mass of the empty particle (i.e., the mass of the virion minus the mass of the genome and core condensation proteins, which is 102.9 MDa, according to the values in Table 2). The intercept (105.4 MDa) is close to the expected mass of the empty particle.

According to the preceding discussion, the total mass of the core condensation proteins should be around 1.3 times the mass of the genome. The genome mass for HAdV5 is 22.2 MDa, so the core condensation proteins should weigh in at around 28.9 MDa. In Table 2, the total mass of the core condensation proteins is 17.4 MDa. We hypothesize that the copy numbers for the core proteins in Table 2 are underestimated. If we use a value of 28.9 MDa for the mass of the core proteins, the expected mass of the mature capsid increases from 142.5 to 154.0 MDa, which is much closer to the measured value of 156.1 MDa.

https://doi.org/10.1021/acs.analchem.1c02439
Anal. Chem. 2022, 94, 1543−1551

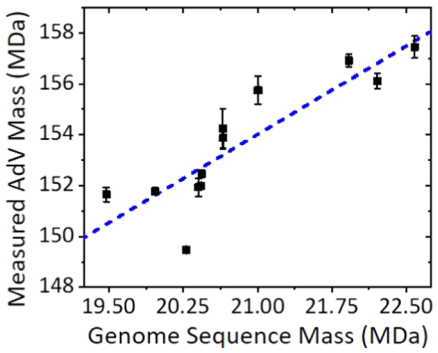


Figure 5. Plot of measured masses of the mature particles for HAdV5 and the recombinant AdV samples listed in Table 1 versus the genome sequence masses. The points show average values for triplicate measurements, and error bars are ±1 standard deviation. Some error bars fall within the points. The dashed blue line shows a linear least-squares fit with a slope of 2.315 and an intercept of 105.4 MDa.

CONCLUSIONS

In this work, we have extended the ability to measure accurate MW distributions by mass spectrometry to over 150 MDa. Higher mass ions have been measured by single-pass CDMS[38−41] and ion trap CDMS.[53,77] However, these measurements were performed on heterogeneous samples and the peak widths were very broad. Here, we have shown that analytes with MWs over 150 MDa can be effectively desolvated, giving rise to relatively narrow MW distributions. While these enormous ions are not completely desolvated, the excess mass due to counterions, salt adducts, and trapped solvent appears to be relatively small and not much larger than found with smaller ions.

The mass measured for HAdV5 by ion trap CDMS is considerably higher than that predicted from sequence masses and copy numbers. Measurements for recombinant AdV, as a function of the genome length, suggest that the copy numbers for the core proteins may have been significantly underestimated. When this is considered, the measured mass and expected mass are close. In addition to the mass information obtained from CDMS in this study, the correlation of the mass and charge revealed additional information. The much lower charges found for empty HAdV5 particles than for the mature virion were attributed to the empty particles lacking the fibers that decorate the surface of the mature particles.

Finally, the results presented here show that CDMS can be used to characterize recombinant adenovirus samples. In addition to a single narrow peak at around 150 MDa attributed to the mature particle with an encapsidated genome, empty particles and other trapped intermediates were resolved. Detection of these particles may be important for therapeutic applications, as formulation impurities could impact therapeutic efficacy and have side effects. CDMS allows a much more rigorous characterization of formulation stabilities, lot-to-lot variation, and the effectiveness of separation and purification technologies. The use of mass spectrometry to characterize such large biopharmaceuticals is groundbreaking. Furthermore, the analysis can be accomplished with a small sample volume (10 μL) and in a relatively short time (typically less than an hour).

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.analchem.1c02439.

Three figures showing mass distributions and charge versus mass heat maps for recombinant AdV listed in Table 1 (except for Ad-Empty and Ad-CMV-GFP/ntLacZ); and tables of measured masses and recombinant AdV sample concentrations (PDF)

AUTHOR INFORMATION

Corresponding Author

Martin F. Jarrold − *Chemistry Department, Indiana University, Bloomington, Indiana 47405, United States;* orcid.org/0000-0001-7084-176X; Email: mfj@iu.edu

Authors

Lauren F. Barnes − *Chemistry Department, Indiana University, Bloomington, Indiana 47405, United States*

Benjamin E. Draper − *Megadalton Solutions, Inc., Bloomington, Indiana 47401, United States*

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.analchem.1c02439

Notes

The authors declare the following competing financial interest(s): MFJ and BED are shareholders in Megadalton Solutions, a company that is engaged in the commercialization of CDMS. BED is an employee of Megadalton Solutions.

ACKNOWLEDGMENTS

This work is funded, in part, by a grant from the NSF (IIP-2031083). BED is funded through an NIH SBIR grant (5R44GM136095) awarded to Megadalton Solutions, Inc.

REFERENCES

(1) Ghebremedhin, B. *Eur. J. Microbiol. Immunol.* **2014**, *4*, 26−33.
(2) Leen, A. M.; Rooney, C. M. *Br. J. Haematol.* **2005**, *128*, 135−144.
(3) Lee, C. S.; Bishop, E. S.; Zhang, R.; Yu, X.; Farina, E. M.; Yan, S.; Zhao, C.; Zheng, Z.; Shu, Y.; Wu, X.; Lei, J.; Li, Y.; Zhang, W.; Yang, C.; Ke Wu, K.; Wu, Y.; Ho, S.; Athiviraham, A.; Lee, M. J.; Wolf, J. M.; Reid, R. R.; He, T.-C. *Genes Dis.* **2017**, *4*, 43−63.
(4) Kremer, E. *Mol. Ther.* **2020**, *28*, 2303−2304.
(5) FDA issues emergency use authorization for third COVID-19 vaccine. *FDA News Release*, 2021. https://www.fda.gov/news-events/press-announcements/fda-issues-emergency-use-authorization-third-covid-19-vaccine (accessed 03-01-2021).
(6) Russell, W. C. *J. Gen. Virol.* **2009**, *90*, 1−20.
(7) Liu, H.; Jin, L.; Koh, S. B. S.; Atanasov, I.; Schein, S.; Wu, L.; Zhou, Z. H. *Science* **2010**, *329*, 1038−1043.
(8) Reddy, V. S.; Natchiar, S. K.; Stewart, P. L.; Nemerow, G. R. *Science* **2010**, *329*, 1071−1075.
(9) Reddy, V. S.; Nemerow, G. R. *Proc. Natl. Acad. Sci. U.S.A.* **2014**, *111*, 11715−11720.
(10) Benevento, M.; Di Palma, S.; Snijder, J.; Moyer, C. L.; Reddy, V. S.; Nenerow, G. R.; Heck, A. J. R. *J. Biol. Chem.* **2014**, *289*, 11421−11430.
(11) Ahi, Y. S.; Mittal, S. K. *Front. Microbiol.* **2016**, *7*, No. 1503.
(12) Kulanayake, S.; Tikoo, S. K. *Viruses* **2021**, *13*, 388.
(13) https://pdb101.rcsb.org/motm/132 (accessed 5/11/2021).
(14) San Martín, C. *Viruses* **2012**, *4*, 847−877.
(15) Condezo, G. N.; San Martín, C. *PLoS Pathog.* **2017**, *13*, No. e1006320.

(16) Pierson, E. E.; Keifer, D. Z.; Asokan, A.; Jarrold, M. F. *Anal. Chem.* **2016**, *88*, 6718−6725.
(17) Barnes, L. F.; Draper, B. E.; Chen, Y.-T.; Powers, T. W.; Jarrold, M. F. *Mol. Ther.–Methods Clin. Dev.* **2021**, *23*, 87−97.
(18) Miller, L. M.; Barnes, L. F.; Raab, S. A.; Draper, B. E.; El-Baba, T. J.; Lutomski, C. A.; Robinson, C. V.; Clemmer, D. E.; Jarrold, M. F. *J. Am. Chem. Soc.* **2021**, *143*, 3959−3966.
(19) Tito, M. A.; Tars, K.; Valegard, K.; Hajdu, J.; Robinson, C. V. *J. Am. Chem. Soc.* **2000**, *122*, 3550−3551.
(20) Uetrecht, C.; Versluis, C.; Watts, N. R.; Roos, W. H.; Wuite, G. J. L.; Wingfield, P. T.; Steven, A. C.; Heck, A. J. R. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 9216−9220.
(21) Snijder, J.; Rose, R. J.; Veesler, D.; Johnson, J. E.; Heck, A. J. R. *Angew. Chem., Int. Ed.* **2013**, *52*, 4020−4023.
(22) Snijder, J.; van de Waterbeemd, M.; Damoc, E.; Denisov, E.; Grinfeld, D.; Bennett, A.; Agbandje-McKenna, M.; Makarov, A.; Heck, A. J. R. *J. Am. Chem. Soc.* **2014**, *136*, 7295−7299.
(23) van de Waterbeemd, M.; Snijder, J.; Tsvetkova, I. B.; Dragnea, B. G.; Cornelissen, J. J.; Heck, A. J. R. *J. Am. Soc. Mass Spectrom.* **2016**, *27*, 1000−1009.
(24) Philip, M. A.; Gelbard, F.; Arnold, S. *J. Colloid Interface Sci.* **1983**, *91*, 507−515.
(25) Hars, G.; Tass, Z. *J. Appl. Phys.* **1995**, *77*, 4245−4250.
(26) Peng, W.-P.; Yang, Y.-C.; Kang, M.-W.; Lee, Y. T.; Chang, H.-C. *J. Am. Chem. Soc.* **2004**, *126*, 11766−11767.
(27) Nie, Z.; Tzeng, Y.-K.; Chang, H.-C.; Chiu, C.-C.; Chang, C.-Y.; Chang, C.-M.; Tao, M.-H. *Angew. Chem., Int. Ed.* **2006**, *45*, 8131−8134.
(28) Zhang, N.; Zhan, L.; Xiong, C.; Nie, Z. *Sci. Sininica Chim.* **2014**, *44*, 801−806.
(29) Schlemmer, S.; Illemann, J.; Wellert, S.; Gerlich, D. *J. Appl. Phys.* **2001**, *90*, 5410−5418.
(30) Smith, R. D.; Cheng, X.; Bruce, J. E.; Hofstadler, S. A.; Anderson, G. A. *Nature* **1994**, *369*, 137−139.
(31) Bruce, J. E.; Cheng, X.; Bakhtiar, R.; Wu, Q.; Hofstadler, S. A.; Anderson, G. A.; Smith, R. D. *J. Am. Chem. Soc.* **1994**, *116*, 7839−7847.
(32) Cheng, X.; Camp, D. G.; Wu, Q.; Bakhtiar, R.; Springer, D. L.; Morris, B. J.; Bruce, J. E.; Anderson, G. A.; Edmonds, C. G.; Smith, R. D. *Nucleic Acids Res.* **1996**, *24*, 2183−2189.
(33) Chen, R.; Cheng, X.; Mitchell, D.; Hofstadler, S.; Wu, Q.; Rockwood, A.; Sherman, Ml.; Smith, R. *Anal. Chem.* **1995**, *67*, 1159−1163.
(34) Fuerstenau, S. D.; Benner, W. H. *Rapid Commun. Mass Spectrom.* **1995**, *9*, 1528−1538.
(35) Schultz, J. C.; Hack, C. A.; Benner, W. H. *J. Am. Soc. Mass Spectrom.* **1998**, *9*, 305−313.
(36) Schultz, J. C.; Hack, C. A.; Benner, W. H. *Rapid Commun. Mass Spectrom.* **1999**, *13*, 15−20.
(37) Doussineau, T.; Kerleroux, M.; Dagany, X.; Clavier, C.; Barbaire, M.; Maurelli, J.; Antoine, R.; Dugourd, P. *Rapid Commun. Mass Spectrom.* **2011**, *25*, 617−623.
(38) Doussineau, T.; Bao, C. Y.; Antoine, R.; Dugourd, P.; Zhang, W.; D'Agosto, F.; Charleux, B. *ACS Macro Lett.* **2012**, *1*, 414−417.
(39) Doussineau, T.; Désert, A.; Lambert, O.; Taveau, J.-C.; Lansalot, M.; Dugourd, P.; Bourgeat-Lami, E.; Ravaine, S.; Duguet, E.; Antoine, R. *J. Phys. Chem. C* **2015**, *119*, 10844−10849.
(40) Warren, N. J.; Mykhaylyk, O. O.; Ryan, A. J.; Williams, M.; Doussineau, T.; Dugourd, P.; Antoine, R.; Portale, G.; Armes, S. P. *J. Am. Chem. Soc.* **2015**, *137*, 1929−1937.
(41) Antoine, R. *Rapid Commun. Mass Spectrom.* **2020**, *34*, No. e8539.
(42) Hendricks, C. D. *J. Colloid Sci.* **1962**, *17*, 249−259.
(43) Maze, J. T.; Jones, T. C.; Jarrold, M. F. *J. Phys. Chem. A* **2006**, *110*, 12607−12612.
(44) Mabbett, S. R.; Zilch, L. W.; Maze, J. T.; Smith, J. W.; Jarrold, M. F. *Anal. Chem.* **2007**, *79*, 8431−8439.
(45) Zilch, L. W.; Maze, J. T.; Smith, J. W.; Ewing, G. E.; Jarrold, M. F. *J. Phys. Chem. A* **2008**, *112*, 13352−13363.
(46) Zilch, L. W.; Maze, J. T.; Smith, J. W.; Jarrold, M. F. *Int. J. Mass Spectrom.* **2009**, *283*, 191−199.
(47) Gao, J.; Austin, D. E. *J. Am. Soc. Mass Spectrom.* **2020**, *31*, 2044−2052.
(48) Benner, W. *Anal. Chem.* **1997**, *69*, 4162−4168.
(49) Contino, N.; Jarrold, M. *Int. J. Mass Spectrom.* **2013**, *345-347*, 153−159.
(50) Contino, N.; Pierson, E.; Keifer, D.; Jarrold, M. *J. Am. Soc. Mass Spectrom.* **2013**, *24*, 101−108.
(51) Pierson, E.; Keifer, D.; Contino, N.; Jarrold, M. *Int. J. Mass Spectrom.* **2013**, *337*, 50−56.
(52) Pierson, E.; Contino, N.; Keifer, D.; Jarrold, M. *J. Am. Soc. Mass Spectrom.* **2015**, *26*, 1213−1220.
(53) Elliott, A. G.; Merenbloom, S. I.; Chakrabarty, S.; Williams, E. R. *Int. J. Mass Spectrom.* **2017**, *414*, 45−56.
(54) Elliott, A. G.; Harper, C. C.; Lin, H.-W.; Susa, A. C.; Xia, Z.; Williams, E. R. *Anal. Chem.* **2017**, *89*, 7701−7708.
(55) Harper, C. C.; Elliot, A. G.; Oltrogge, L. M.; Savage, D. F.; Williams, E. R. *Anal. Chem.* **2019**, *91*, 7458−7465.
(56) Harper, C. C.; Brauer, D. D.; Francis, M. B.; Williams, E. R. *Chem. Sci.* **2021**, *12*, 5185.
(57) Keifer, D.; Shinholt, D.; Jarrold, M. *Anal. Chem.* **2015**, *87*, 10330−10337.
(58) Todd, A. R.; Alexander, A. W.; Jarrold, M. F. *J. Am. Soc. Mass Spectrom.* **2020**, *31*, 146−154.
(59) Todd, A. R.; Jarrold, M. F. *J. Am. Soc. Mass Spectrom.* **2020**, *31*, 1241−1248.
(60) Kafader, J. O.; Melani, R. D.; Durbin, K. R.; Ikwuagwu, B.; Early, B. P.; Fellers, R. T.; Beu, S. C.; Zabrouskov, V.; Makarov, A. A.; Maze, J. T.; Shinholt, D. L.; Yip, P. F.; Tullman-Ercek, D.; Senko, M. W.; Compton, P. D.; Kelleher, N. L. *Nat. Methods* **2020**, *17*, 391−394.
(61) Wörner, T. P.; Snijder, J.; Bennett, A.; Agbandje-McKenna, M.; Makarov, A. A.; Heck, A. J. R. *Nat. Methods* **2020**, *17*, 395−398.
(62) Pierson, E. E.; Keifer, D. Z.; Selzer, L.; Lee, L. S.; Contino, N. C.; Wang, J. C.-Y.; Zlotnick, A.; Jarrold, M. F. *J. Am. Chem. Soc.* **2014**, *136*, 3536−3541.
(63) Kukreja, A. A.; Wang, J. C.-Y.; Pierson, E.; Keifer, D. Z.; Selzer, L.; Tan, Z.; Dragnea, B.; Jarrold, M. F.; Zlotnick, A. *J. Virol.* **2014**, *88*, 14105−14115.
(64) Pierson, E. E.; Keifer, D. Z.; Kukreja, A. A.; Wang, J. C.-Y.; Zlotnick, A.; Jarrold, M. F. *J. Mol. Biol.* **2016**, *428*, 292−300.
(65) Keifer, D. Z.; Motwani, T.; Teschke, C. M.; Jarrold, M. F. *J. Am. Soc. Mass Spectrom.* **2016**, *27*, 1028−1036.
(66) Chen, C.; Wang, J. C.-Y.; Pierson, E. E.; Kiefer, D. Z.; Delaleau, M.; Gallucci, L.; Cazenave, C.; M Kann, M.; Jarrold, M. F.; Zlotnick, A. *PLoS Pathog.* **2016**, *12*, No. e1005802.
(67) Motwani, T.; Lokareddy, R. K.; Dunbar, C. A.; Cortines, J. R.; Jarrold, M. F.; Cingolani, G.; Teschke, C. M. *Sci. Adv.* **2017**, *3*, No. e1700423.
(68) Lee, L. S.; Brunk, N.; Haywood, D. G.; Keifer, D.; Pierson, E.; Kondylis, P.; Wang, J. C.-Y.; Jacobson, S. C.; Jarrold, M. F.; Zlotnick, A. *Protein Sci.* **2017**, *26*, 2170−2180.
(69) Lutomski, C.; Lyktey, N.; Zhao, Z.; Pierson, E.; Zlotnick, A.; Jarrold, M. *J. Am. Chem. Soc.* **2017**, *139*, 16932−16938.
(70) Lutomski, C. A.; Lyktey, N. A.; Pierson, E. E.; Zhao, Z.; Zlotnick, A.; Jarrold, M. F. *J. Am. Chem. Soc.* **2018**, *140*, 5784−5790.
(71) Dunbar, C. A.; Callaway, H. M.; Parrish, C. R.; Jarrold, M. F. *J. Am. Chem. Soc.* **2018**, *140*, 15701−15711.
(72) Dunbar, C. A.; Rayaprolu, V.; Wang, C.-Y.; Brown, C. J.; Leishman, E.; Jones-Burrage, S.; Trinidad, J. C.; Bradshaw, H. B.; Clemmer, D. E.; Mukhopadhyay, S.; Jarrold, M. F. *ACS Infect. Dis.* **2019**, *5*, 892−902.
(73) Bond, K. M.; Lyktey, N. A.; Tsvetkova, I. B.; Dragnea, B.; Jarrold, M. F. *J. Phys. Chem B* **2020**, *124*, 2124−2131.
(74) Bond, K.; Tsvetkova, I.; Wang, J. C.-Y.; Jarrold, M. F.; Dragnea, B. *Small* **2020**, *16*, No. 2004475.
(75) Zhao, Z.; Wang, J. C.-Y.; Zhang, M.; Lyktey, N. A.; Jarrold, M. F.; Jacobson, S. C.; Zlotnick, A. *Nat. Commun.* **2021**, *12*, No. 589.

https://doi.org/10.1021/acs.analchem.1c02439
Anal. Chem. 2022, 94, 1543−1551

(76) Keifer, D. Z.; Motwani, T.; Teschke, C. M.; Jarrold, M. F. *Rapid Commun. Mass Spectrom.* **2016**, *30*, 1957−1962.
(77) Draper, B.; Anthony, S.; Jarrold, M. *J. Am. Soc. Mass Spectrom.* **2018**, *29*, 2160−2172.
(78) Draper, B.; Jarrold, M. *J. Am. Soc. Mass Spectrom.* **2019**, *30*, 898−904.
(79) Schmidt, H. T.; Cederquist, H.; Jensen, J.; Fardi, A. *Nucl. Instrum. Methods Phys. Res., Sect. B* **2001**, *173*, 523−527.
(80) Takahashi, E.; Cohen, S. L.; Tsai, P. K.; Sweeney, J. A. *Anal. Biochem.* **2006**, *349*, 208−217.
(81) Condezo, G. N.; Marabini, R.; Ayora, S.; Carazo, J. M.; Alba, R.; Chillón, M.; San Martín, C. *J. Virol.* **2015**, *89*, 9653−9664.
(82) Mullins, E.; Powers, T.; Zobel, J.; Clawson, K.; Barnes, L.; Draper, B.; Zou, Q.; Binder, J.; Dai, S.; Zhang, K.; Friese, O.; Runnels, H.; Jarrold, M.; Thompson, L. *Front. Bioeng. Biotechnol.* **2021**, *9*, No. 753480.
(83) McKay, A. R.; Ruotolo, B. T.; Ilag, L. L.; Robinson, C. V. *J. Am. Chem. Soc.* **2006**, *128*, 11433−11442.
(84) Lössl, P.; Snijder, J.; Heck, A. J. R. *J. Am. Soc. Mass Spectrom.* **2014**, *25*, 906−917.
(85) Jarrold, M. F. *Annu. Rev. Phys. Chem.* **2000**, *51*, 179−207.
(86) Todd, A. R.; Barnes, L. F.; Young, K.; Zlotnick, A.; Jarrold, M. F. *Anal. Chem.* **2020**, *92*, 11357−11364.
(87) Larsson, D. S. D.; Liljas, L.; van der Spoel, D. *PLoS Comput. Biol.* **2012**, *8*, No. e1002502.
(88) Andoh, Y.; Yoshii, N.; Yamada, A.; Fujimoto, K.; Kojima, H.; Mizutani, K.; Nakagawa, A.; Nomoto, A.; Okazaki, S. *J. Chem. Phys.* **2014**, *141*, No. 165101.
(89) Fernandez de la Mora, J. *Anal. Chim. Acta* **2000**, *406*, 93−104.
(90) Konermann, L.; Ahadi, E.; Rodriguez, A.; Vahidi, S. *Anal. Chem.* **2013**, *85*, 2−9.
(91) Rayleigh, L. *London, Edinburgh Dublin Philos. Mag. J. Sci.* **1882**, *14*, 184−186.
(92) Wu, E.; Pache, L.; Von Seggern, D.; Mullen, T.; Mikyas, Y.; Stewart, P.; Nemerow, G. *J. Virol.* **2003**, *77*, 7225−7235.

https://doi.org/10.1021/acs.analchem.1c02439
Anal. Chem. 2022, 94, 1543−1551