analytical chemistry

pubs.acs.org/ac Article

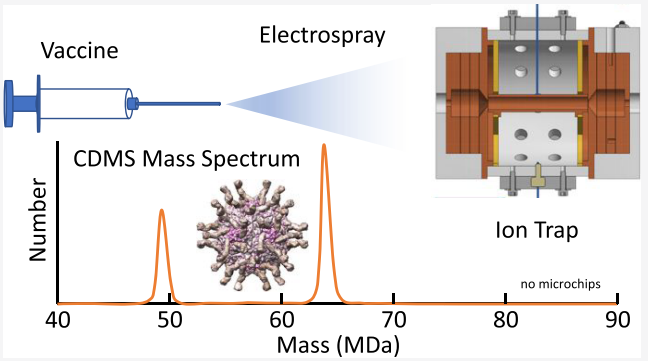
Characterization of Classical Vaccines by Charge Detection Mass Spectrometry

Lohra M. Miller, Kevin M. Bond, Benjamin E. Draper, and Martin F. Jarrold*

Cite This: *Anal. Chem.* 2021, 93, 11965–11972 Read Online

ACCESS | Metrics & More | Article Recommendations

ABSTRACT: Vaccines induce immunity by presenting disease antigens through several platforms ranging from individual protein subunits to whole viruses. Due to the large difference in antigen size, the analytical techniques employed for vaccine characterization are often platform-specific. A single, robust analytical technique capable of widespread, cross-platform use would be of great benefit and allow for comparisons across manufacturing processes. One method that spans the antigen mass range is charge detection mass spectrometry (CDMS). CDMS is a single-ion approach where the mass-to-charge ratio (m/z) and charge are measured simultaneously, allowing accurate mass distributions to be measured for heterogeneous analytes over a broad size range. In this work, CDMS was used to characterize the antigens from three classical multivalent vaccines, inactivated poliomyelitis vaccine (IPOL), RotaTeq, and Gardasil-9, directly from commercial samples. For each vaccine, the antigen purity was inspected, and in the whole virus vaccines, empty virus particles were detected. For IPOL, information on the extent of formaldehyde cross-linking was obtained. RotaTeq shows a narrow peak at 61.06 MDa. This is at a slightly lower mass than expected for the double-layer particle, suggesting that around 10 pentonal trimers are missing. For Gardasil-9, buffer exchange of the vaccine resulted in very broad mass distributions. However, removal of the virus-like particles from the aluminum adjuvant using a displacement reaction generated a spectrum with narrow peaks.

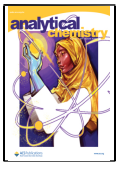


Vaccines have eradicated smallpox, provided a near-global eradication of poliomyelitis, and led to a substantial reduction in many other serious diseases, including measles, rubella, tetanus, and diphtheria.[1] Since the discovery of the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) in late 2019, vaccine development has taken center stage as our best hope to end the coronavirus disease 19 (COVID-19) pandemic. More than 100 vaccine candidates have entered trials. The first two to be granted emergency use authorization by the FDA[2] were based on mRNA, a new vaccination approach. Vaccine formulations are typically mixtures of antigenic particles in a complex matrix. The antigenic particles are often heterogeneous, and the methods used to characterize them are low resolution. Here, we show that charge detection mass spectrometry (CDMS) can characterize the antigenic particles from a variety of vaccine types. Accurate mass information provides insight into the composition of the antigenic particles and can identify misassembled particles.

Classical vaccination approaches can be roughly classified into four main types: whole-inactivated viruses, live-attenuated viruses, virus-like particles, and protein subunits. Each method induces immunity by presenting a disease antigen in different ways. The formulation is critical, often requiring different excipients (nonactive ingredients) such as preservatives, adjuvants, stabilizers, and bulking agents. For a vaccine to become licensed, formulations undergo rigorous safety testing including many years of formulation development and ending with clinical trials. Typical vaccine development proceeds in a linear fashion, and characterization steps can be slow for large, complex molecules that lack fast and robust analytical tools.

For FDA approval of a vaccine, the identity, purity, potency, and stability of the antigen must be characterized, as well as the long-term safety. Physicochemical properties are tested, with techniques such as UV/visible spectroscopy and/or mass spectrometry, amino acid analysis, sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE), and related assays to detect specific proteins.[3] The biological activity can be tested with enzyme-linked immunosorbent assay (ELISA), cytometric analysis, serotyping, neutralization assays, and

Received: May 4, 2021
Accepted: August 16, 2021
Published: August 26, 2021



titrations. Polymerase chain reaction (PCR) is often employed to detect viral contamination.

Orthogonal techniques are essential for complete characterization of commercial vaccines. Examples include electron microscopy (EM), circular dichromism, gas-phase electrophoretic mobility molecular analysis (GEMMA) measurements,[4] and dynamic and multiangle light scattering (MALS) techniques.[5] GEMMA allows the separation of particles based upon their electrophoretic mobility, giving insight into the average size of antigens, and MALS is able to analyze particles in their native buffer. All of these techniques provide valuable information about the antigen but lack the ability to provide detailed compositions for the antigenic particles found in commercial vaccines.

The challenges provided by the development and manufacturing of vaccines provide a significant barrier to entry for companies or governments wishing to develop or produce their own vaccines.[6−10] Some efforts have focused on the improvement of instrumentation and methodologies, which would reduce costs and time associated with the development and production of vaccines.[7−14] However, the diversity of vaccine platforms has made it difficult to identify analytical techniques, with a dynamic range large enough to characterize early in process formulations to final clinical-grade formulations. The development of many vaccines requires a suite of analytical tools to support this production and development process. In this work, we explore whether CDMS can extend the accuracy and precision of mass spectrometry to vaccine characterization. CDMS is a single-particle technique where the mass-to-charge (m/z) and charge (z) are measured for each ion, allowing a direct determination of the mass distribution for even the most heterogeneous analyte.[15−22] CDMS has been used previously to characterize many viral platforms including protein subunits and assembly reactions.[23−32] Herein, CDMS is used to characterize commercially available formulations of whole-inactivated viruses (IPOL, Sanofi Pasteur), live-attenuated viruses (RotaTeq, Merck), and virus-like particles (Gardasil-9, Merck).

MATERIALS AND METHODS

Sample Preparation. Inactivated poliomyelitis vaccine (IPOL, Sanofi Pasteur SA) and Gardasil-9 (Merck) were purchased from IU Health Center Pharmacy (Bloomington, Indiana). RotaTeq (Merck) was donated to the research group by IU Health Pediatrics (Indianapolis, IN).

Gardasil-9 antigens were removed from the aluminum adjuvant via a displacement reaction. The vaccine was spiked with a 140 mM sodium phosphate buffer and incubated at 37 °C for 4 h. The sample was then centrifuged for 3 min at 8800g to pellet the displaced adjuvant. The supernatant was removed and exchanged into 200 mM ammonium acetate using Micro Bio-Spin P-6 gel columns (Bio-Rad). All other samples were directly exchanged into a 200 mM ammonium acetate solution using the same columns and buffer exchange procedure. The samples showed no signs of degradation during the analysis time.

Charge Detection Mass Spectrometry Measurements. The CDMS instrument used is a home-built instrument that is described in detail elsewhere.[33−40] Briefly, ions are generated by nanoelectrospray and enter the instrument through a metal capillary. Positively charged ions are thermalized and focused through multiple stages of differential pumping and accelerated to an energy of 100 eV/z. A dual hemispherical deflection energy analyzer passes a narrow energy distribution into an electrostatic linear ion trap (ELIT) where ions are trapped for 100 ms. The signal from the oscillating ions is detected by a charge-sensitive amplifier, digitized, and analyzed by fast Fourier transforms. The oscillation frequency and magnitude give the m/z and charge of individual ions allowing for the direct determination of the mass. Measurements are repeated for thousands of ions, and the results are binned to give the mass distribution. For samples with concentrations of around 10^{12} particles/mL, it is usually possible to measure a spectrum containing 5000 ions in less than an hour.

Negative-Stain Electron Microscopy. All electron microscopy (EM) measurements were performed at Indiana University's Electron Microscopy Center on the JEOL JEM 1010 microscope. EM samples were prepared by transferring 10 μL of sample to 300 mesh Formvar/Carbon film copper grids that were glow-discharged with a PELCO easiGlow glow discharger. Samples were dried for 10 min before excess solvent was removed by blotting. Staining was accomplished by the transfer of 10 μL of 2% uranyl acetate stain and subsequent immediate removal. Images were collected with 50 000× magnification.

RESULTS AND DISCUSSION

CDMS Analysis of IPOL Vaccine. IPOL is a trivalent whole-inactivated virus vaccine used to prevent three types of poliomyelitis. Poliomyelitis is caused by the poliovirus.[41−45] The poliovirus has a 7440−7445 nt ssRNA genome (depending on the serotype). The capsid, which is around 30 nm in diameter, contains 60 copies each of the four viral proteins VP1, VP2, VP3, and VP4 arranged in icosahedral symmetry. VP4 is found on the interior surface of the capsid. The vaccine consists of three serotypes Type 1 (Mahoney), Type 2 (MEF-1), and Type 3 (Saukett) to provide more robust immunity. The virions are inactivated by cross-linking with formaldehyde. Poliovirus naturally assembles both with and without the genome. The empty and full particles (called the C-antigen and D-antigen, respectively) can be differentiated by ELISA. ELISA is used for each batch of the vaccine to ensure that the total D-antigen concentration is in the required range and that the ratios of the D-antigens for each serotype (1, 2, and 3) are also within the required range (nominally 5:1:4).[46−48] Table 1 shows a summary of the sequence masses for the C- and D-antigens for the three serotypes and the weighted average values.

A typical CDMS mass histogram for the IPOL vaccine is shown in Figure 1a. The spectrum was collected for 60 min and contains 12 045 ions. It has prominent peaks at 5.94 and 8.58 MDa (according to Gaussian fits). These peaks are assigned to C- and D-antigens (empty and full particles), respectively (see Table 1). Note that there is little intensity

Table 1. Sequence Masses of Poliovirus C- and D-Antigens (Empty and Full Particles) for the Three Serotypes in IPOL

poliovirus strain	sequence mass (in MDa) of C-antigen (empty particle)	sequence mass (in MDa) of D-antigen (full particle)
type 1	5.86	8.25
type 2	5.81	8.21
type 3	5.83	8.22
weighted average	5.84	8.23

https://doi.org/10.1021/acs.analchem.1c01893
Anal. Chem. 2021, 93, 11965−11972

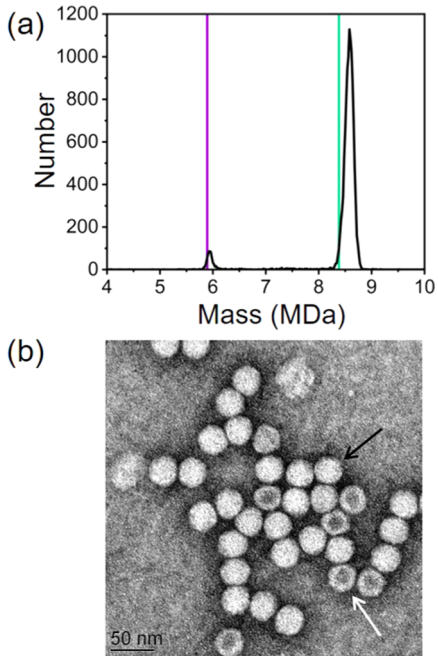


Figure 1. Determination of the relative abundances of C- and D-antigen particles in IPOL vaccine with CDMS and transmission electron microscopy (TEM). Part (a) shows a CDMS mass spectrum of IPOL after buffer exchange into ammonium acetate. A bin size of 20 kDa was used. The expected masses (see the text) of the C- and D-antigen particles are shown by the purple and green lines, respectively. The measured masses are slightly larger than the expected masses because of cross-linking and a small contribution from counterions, residual salt, and trapped solvent (see the text). Part (b) shows a negative-stain EM image of the same IPOL sample. The white arrow shows an empty C-antigen particle and the black arrow shows a full D-antigen particle.

between the two peaks, indicating that the fraction packaging a partial genome is low. The presence of C-antigen can interfere with immunoassays used to determine the D-antigen serotype ratio because both particles can react. Serotype-specific and D-antigen-specific monoclonal antibodies must be used.[49] CDMS shows the relative amounts of C- and D-antigens without relying on highly specific monoclonal antibodies; this is beneficial as formalin cross-linking can alter the binding site of an antigen, further hampering ELISA measurements. Negative-stain EM can also be used to differentiate between the C- and D-antigens. Particles are assigned to empty or full based on whether they have absorbed stain (generally, empty particles absorb stain and full do not). A typical EM image of the IPOL vaccine is shown in Figure 1b. An empty particle is identified by the white arrow and a full by the black. Determining relative abundances in this way is time consuming as thousands of particles must be viewed and assigned. In addition, the results are influenced by how well the particles are stained and there are often some particles that cannot be clearly classified as empty or full.

In a CDMS spectrum, the heterogeneity of the particles is represented by widths of the peaks. Peak widths were determined by means of a Gaussian fit. The empty peak has a full width at half-maximum (FWHM) of 115 kDa and for the full the FWHM is 187 kDa. The experimental resolution is around 107 kDa at 5.94 MDa and 162 kDa at 8.58 MDa. When we incorporate the heterogeneity from the three serotypes, the expected peak widths increase slightly to 115 and 166 kDa, respectively. Thus, the peak width for the C-antigen (empty particle) agrees with the expected value, and the peak width for the D-antigen is slightly broader than expected. The extra width of the D-antigen peak probably results from heterogeneity derived from the packaged genome, such as counterions and cross-links (see below).

The expected mass of the empty C-antigen particle is 5.89 MDa (purple line in Figure 1a). To calculate the expected mass, we started with the average sequence mass from Table 1 and then incorporated known modifications, such as myristoylation of the N-terminal of VP4 proteins, and additions, such as sphingosine that occupies a hydrophobic pocket of VP1.[50−52] The deviation between the measured mass (5.94 MDa) and the expected mass (5.89 MDa) for the empty C-antigen is around 0.8%. Large protein complexes generated by electrospray of volatile salt solutions often have masses a few tenths of a percent higher than the expected mass due to counterions, residual salt, and trapped solvent. Empty poliovirus capsids are known to have an excess negative charge of around 240 e that is neutralized by counterions.[53] In the present case, an additional contribution comes from formaldehyde cross-linking, which adds 12 Da per cross-link. The three serotypes of poliovirus are incubated in a 1:4000 formalin solution for at least 12 days allowing sufficient cross-linking to stop the release of a genome from the capsid.[54] It is not possible to say with confidence how much of the excess mass is due to cross-linking for the C-antigen. If we attribute a minimum of 0.3% of the 0.8% excess mass to counterions, residual salt, and trapped solvent, then the number of cross-links must be less than 2500.

The expected mass for the D-antigen is 8.43 MDa, which is indicated by the green line in Figure 1a. This value was obtained from the measured mass of the empty capsid (5.94 MDa) and the average sequence mass of the ssRNA genome (2.395 MDa) plus the mass of the counterions associated with the genome. Some of the backbone phosphate groups will be ionized and have counterions.[50−56] In a recent study of DNA packaging in adeno-associated virus, we determined that counterions added around 4% to the genome mass for ions generated by electrospray under similar conditions to those used here.[57] Thus, counterions are expected to contribute around 0.096 MDa to the expected mass of the poliovirus.

The difference between the expected mass of the D-antigen (8.43 MDa) and the measured mass (8.58 MDa) is around 150 kDa. This is probably mainly due to cross-linking. Using 12 Da per cross-link, the D-antigen has around 12 500 cross-links. This is much larger than for the C-antigen where a limit of less than 2500 cross-links was deduced. Formaldehyde reacts with primary and secondary amines to form methylene bridges. It can react with amino acids (e.g., lysine) and with DNA and RNA bases (e.g., adenine). Thus, protein−protein, protein−RNA, and RNA−RNA cross-links are all possible. The number of protein−protein cross-links (from the results for the C-antigen) is much smaller than the number of cross-links involving RNA bases (protein−RNA and RNA−RNA). Note that the increase in the number of cross-links with addition of the RNA (around 10 000) exceeds the number of RNA nucleotides (around 7440). While guanine has two amines that can cross-link, the other bases have only one. The conditions and very long reaction times employed during the manufacture of IPOL (at least 12 days in 1:4000 formalin at 37 °C) may drive the formation of cross-links between mildly nucleophilic sites that are not normally considered reactive. However, we

https://doi.org/10.1021/acs.analchem.1c01893
Anal. Chem. 2021, 93, 11965−11972

also cannot rule out the possibility that some of the additional mass comes from another source. The relatively large shift in the mass for the D-antigen due to cross-linking could be used to track cross-linking during vaccine preparation.

Figure 2a shows a charge vs mass heat map for the IPOL vaccine. Two vertical streaks at around 5.9 and 8.6 MDa correspond to the C- and D-antigens, respectively. It is generally accepted now that large ions generated by electrospray are produced by the charge residue mechanism,[58] where a water droplet deposits its charge on the analyte as it evaporates away. Thus, the charge reflects the physical size of the ion and larger geometries are expected to have a higher charge. As can be seen from Figure 2a, the average charges for the C- and D- antigens are similar, which is consistent with them having similar sizes, the mass difference resulting from the encapsulated genome. The charge distributions for both the empty C-antigen and the full D-antigen are unusually broad, extending from around 180 e (elementary charges) to around 270 e. A closer inspection of Figure 2a reveals what appear to be at least two maxima in the heat map for the full D-antigen. The black line in Figure 2b shows the charge spectrum for IPOL. The IPOL charge distribution appears to contain several components: two prominent peaks centered on around 220 and 250 e, and a broad low charge tail that extends to 170 e. For comparison, the charge spectrum measured for adeno-associated virus 8 (AAV8) reference material (ATCC VR-1816) is shown in red. Like poliovirus, AAV8 is a nonenveloped icosahedral virus, but it has a much narrower charge distribution, centered on around 165 e. According to the Rayleigh model for the maximum charge on a spherical droplet, the charge is proportional to $r^{3/2}$, where r is the droplet radius. AAV8 is slightly smaller than poliovirus (25 vs 30 nm), which explains the lower average charge found for AAV8. However, given a charge of 165 e on 25 nm AAV8, a charge of around 215 e would be expected for a 30 nm poliovirus, assuming an $r^{3/2}$ dependence; 215 e is close to the first peak in the charge spectrum in Figure 2b, which suggests that the second higher charge peak at 250 e is due to a structure that is larger (around 33 nm in diameter if it were roughly spherical and the charge shows an $r^{3/2}$ dependence).

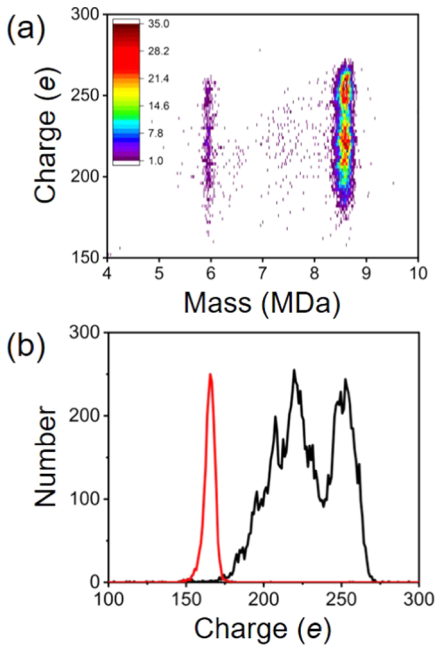


Figure 2. CDMS analysis of IPOL vaccine. Part (a) shows a charge versus mass heat map for IPOL. Part (b) shows a charge spectrum of IPOL in black. The red line shows the charge spectrum for the AAV8 reference standard (see the text). A bin size of 20 kDa and 2 e were used for part (a), and a bin size of 1 e was used for part (b). The AAV8 charge spectrum has been scaled down by 0.67 to match the maximum intensity in the IPOL spectrum.

As previously mentioned, IPOL vaccine is composed of three serotypes in a 5:1:4 ratio of type 1 (Mahoney), type 2 (MEF-1), and type 3 (Saukett), respectively. The structural differences between the three serotypes arise at the antigen binding sites on the capsid,[59,60] and it is unlikely that these differences are responsible for the broad charge distributions. However, when the poliovirus binds to its receptor, it undergoes a conformational change that leads to externalization of VP4 and the N-terminus of VP1 with a segment of VP1 being cleaved.[45,61,62] Reversible externalization can also be induced by heating to 37 °C.[63] During manufacture, the vaccine is exposed to 1:4000 formalin at 37 °C for at least 12 days where cross-linking could trap externalized VP1 or VP4. The CDMS mass measurements rule out significant loss of VP1 or VP4 in the IPOL vaccine. However, externalization of some of the VP1 and VP4 N-termini could account for the high charge population detected in the CDMS measurements.

CDMS Analysis of RotaTeq. Rotavirus, known to cause severe gastroenteritis in young children, has been greatly diminished by the RotaTeq vaccine. As a pentavalent, live-attenuated vaccine, it offers a broad immunity by using a reassortant combination of five human-bovine rotavirus strains.[64] Most of the analysis of RotaTeq centers on electropherotyping of the dsRNA to make sure the correct genetic composition was achieved.[42] Genotyping with PCR to detect the five strains is also done. PCR is a robust platform that identifies specific genomic sequences of the viral genome but requires specified primers to the sequences of interest. In addition, PCR gives no information about the integrity of the reassortment assembly.

In nature, infectious rotavirus forms 80 nm triple-layered particles (TLPs) packaging an 18 550 bp double-stranded RNA (dsRNA)-segmented genome. Shedding of the outer-layer proteins, VP4 and VP7, creates a noninfectious double-layer particle (DLP). The DLP is commonly found when the viral particles are attenuated for vaccine purposes.[65] The core of the rotavirus consists of a VP2 capsid that surrounds the genome and 11 or 12 copies of VP1 (viral polymerase) and VP3 (viral capping enzyme). To obtain an expected mass for the rotavirus, the sequence masses of the structural proteins and the dsRNA were determined for each strain and then averaged. The results along with the expected copy numbers are shown in Table 2.[66,67] In the case of the dsRNA genome, the mass has been scaled up by 4% in anticipation of copackaged counterions.[55] The core has an expected mass of 27.17 MDa. The expected masses of the double-layer and triple-layer particles are 62.16 and 106.77 MDa, respectively.

As an oral vaccine, the reassortants are dispensed in a buffered solution that contains sugars, salts, and the surfactant polysorbate 80. The vaccine was buffer-exchanged into 200 mM ammonium acetate before electrospray. A typical CDMS mass spectrum measured for RotaTeq is shown in Figure 3a. It contains 3013 ions collected in under 20 min. The spectrum covers a 30−80 MDa range as core particles and TLPs were not detected at lower or higher masses, respectively. The main

https://doi.org/10.1021/acs.analchem.1c01893
Anal. Chem. 2021, 93, 11965−11972

Table 2. Rotavirus Protein Copy Number and Average Molecular Weights

viral protein	location	sequence mass (kDa)	copy number	cumulative mass (MDa)
VP1	core	124.92	11−12	
VP2	core	102.36	120	
VP3	core	97.91	11−12	
dsRNA	core	12 320[a]	1	27.17
VP6	middle layer	44.87	780	62.16
VP4	outer layer	86.57	180	
VP7	outer layer	32.21	780	106.77

[a]Average mass of the reassortant genomes. The mass difference between the different genomes is 5 kDa. The mass in the table includes counterions assuming that they contribute 4% to the mass. The mass without counterions is 11 840 kDa.

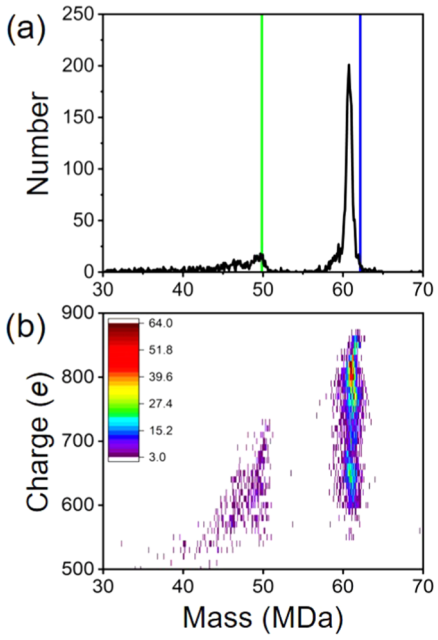


Figure 3. CDMS spectra of RotaTeq. Part (a) shows a typical mass spectrum of RotaTeq after buffer exchange into ammonium acetate (black line). Colored vertical lines show expected masses of the DLP with all pentonal trimers (blue) and the DLP with all pentonal trimers but missing the genome (green). The bin size is 100 kDa. Part (b) shows a charge versus mass heat map of RotaTeq. The bin sizes are 100 kDa and 10 e.

species detected by CDMS, the DLP, is thought to have a small antigenic response. The main antigen in the RotaTeq vaccine is the TLP that has a concentration of between 5×10^6 and 80×10^6 infectious units/mL in the vaccine. This is below the current limit of detection for CDMS (which is around 330×10^6 particle/mL for hepatitis B virus).[68] The concentration of the DLP is not specified, but it is clearly much larger than for the TLP.

The Rotateq spectrum in Figure 3a has a narrow prominent peak centered on a mass of 61.06 MDa with a smaller peak at 49.64 MDa. The lower mass peak has a low mass tail that extends to below 30 MDa. The blue line in Figure 3a shows the expected mass of the DLP (62.16 MDa; see Table 2). The line is at a higher mass than the main peak in the spectrum at 61.06 MDa, which indicates that some of the components counted in Table 2 must be missing in the sample. Previous studies have found that the pentonal trimers dissociate from the transcriptionally active DLP in Wa strains.[64,69] There are five pentonal trimers around each 5-fold axis, so altogether, there are 60. In the Wa strain, >90% are lost. The loss of a small number of pentonal trimers is a plausible explanation for the lower-than-expected mass found here. If 10 pentonal trimers are lost, the expected mass would be 60.81 MDa, which is 0.4% less than the measured value. This discrepancy is within the range normally found for counterions, salt adducts, and trapped solvents. While loss of around 10 pentonal trimers can account for the measured mass and seems the most likely explanation, we only measure the total mass and so we cannot exclude other possible explanations.

The lower abundance peak in Figure 3a at around 50 MDa probably results from DLPs that do not contain a genome. The green line shows the expected mass of an otherwise intact DLP that does not have a genome (49.85 MDa). The measured peak is at a slightly lower mass (49.64 MDa), indicating that some additional components are missing, possibly a few copies of VP1, VP2, or a small number of pentonal trimers.

The peak width can be used to assess sample heterogeneity. The width of the measured peak at 61 MDa is 1.02 MDa (FWHM). Considering the expected instrumental mass resolution for this peak (0.84 MDa), the underlying peak width is 0.58 MDa (FWHM). While at first glance, this seems broad, bear in mind that this is for a 61 MDa peak. In fact, the relative peak width is less than 1%. Factors that contribute to this peak width include possible variations in the copy numbers of VP1 and VP3 and in the number of missing pentonal trimers. The distribution of excess mass due to counterions, salt adducts, and trapped solvents also contributes.

Figure 3b shows a charge versus mass heat map for the RotaTeq vaccine. The charge distributions are elongated, like the results for IPOL. For the main peak at 61 MDa, the charge population extends from 570 to 860 e, spanning almost 300 e. There appear to be several subpopulations at around 650, 770, and 810 e. Since RotaTeq is pentavalent, it is possible that the broad charge population is related to the different strains. The charge distribution for the smaller 50 MDa peak in Figure 3b is not as broad as for the 61 MDa peak and only extends up to around 750 e. In this regard, the charge distributions for RotaTeq and IPOL differ. For IPOL, the peak due to the empty and full capsids had similar charge distributions. The higher charge population for the DLP could be due to the partially exposed genome, a suggestion that is supported by the complete absence of the higher charge population in the empty DLP.

CDMS Analysis of Gardasil-9. Gardasil-9 is a recombinant vaccine made from nine strains of the human papillomavirus's (HPV) major capsid protein (L1) formed into virus-like particles (VLPs). These VLPs provide immunity to multiple cancers caused by nine separate forms of HPV.[70] The VLPs are assembled from 72 pentameric capsomers arranged in a $T = 7$ icosahedral lattice that closely resembles that of the native human virions.[71,72] Table 3 provides molecular weight and relative abundance information for the HPV strains in Gardasil-9.

For a better immunological response, the L1 VLPs are adsorbed onto an aluminum adjuvant and prepared as a liquid suspension containing additional salts, amino acids, and polysorbate 80. Figure 4a shows the CDMS spectrum measured following buffer exchange of the Gardasil-9 into 200 mM ammonium acetate before electrospray. There are no discernible peaks, and the mass distribution is very broad with

https://doi.org/10.1021/acs.analchem.1c01893
Anal. Chem. 2021, 93, 11965−11972

Table 3. Strains Present in Gardasil-9 and the VLP Molecular Weights

HPV strain	L1 molecular weight (Da)	VLP molecular weight (MDa)	relative abundance
6	55 566.94	20.00	0.111
11	55 835.27	20.10	0.148
33	55 902.73	20.12	0.074
16	56 278.19	20.26	0.222
31	56 352.03	20.29	0.074
58	59 038.44	21.25	0.074
52	59 468.93	21.41	0.074
45	60 310.74	21.71	0.074
18	63 623.72	22.90	0.148

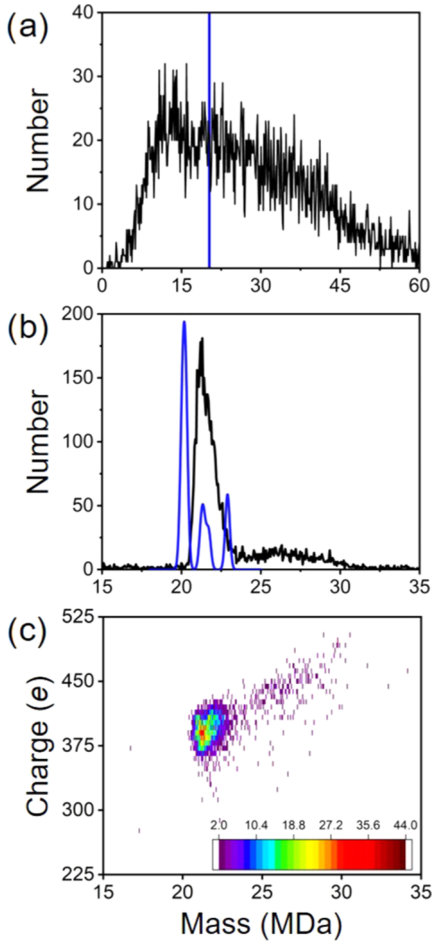


Figure 4. CDMS spectra of Gardasil-9. Part (a) shows a mass spectrum of Gardasil-9 buffer-exchanged into 200 mM ammonium acetate. The bins are 110 kDa. The blue line is the weighted average mass of the 9 VLPs in Gardasil-9. Part (b) shows a mass spectrum of Gardasil-9 after removal of aluminum adjuvant and buffer exchange into 200 mM ammonium acetate (black line). The bins are 55 kDa. The blue line shows the expected mass distribution for Gardasil-9 considering the experimental resolution (see the text). Part (c) shows a charge versus mass heat map for Gardasil-9. The bins are 55 kDa and 6 e.

a high mass tail extending to beyond 60 MDa. The blue vertical line in Figure 4a shows the weighted average for the nine VLPs in Gardasil-9 (see Table 3). We attributed the broad mass distribution in Figure 4a to the problem of separating the VLPs from the aluminum adjuvant. Figure 4b shows a typical CDMS spectrum measured after removal of the aluminum adjuvant via a displacement reaction (see the Materials and Methods section) followed by exchange into 200 mM ammonium acetate. The spectrum contains 8845 ions and took 276 min to collect due to the low signal intensity from this sample. In addition to the main peak centered on 21.44 MDa, there is a broad distribution with a maximum at around 26 MDa that extends from the main peak to beyond 30 MDa. The blue line in Figure 4b shows the spectrum expected for the VLPs listed in Table 3. The spectrum was calculated by representing each VLP by a Gaussian located at the VLP mass, with an intensity given by its relative abundance, and an FWHM given by the experimental resolution (0.34 MDa). The nine Gaussians were then summed to give the blue line in Figure 4b. According to the calculated spectrum in Figure 4b, it should be possible to resolve three subpopulations for the Gardasil-9 vaccine. The most intense peak at 20.2 MDa (0.45 MDa FWHM) is due to strains 6, 11, 16, 31, and 33. The broader peak at around 21.5 MDa (0.66 MDa FWHM) is due to strains 45, 52, and 58. Finally, the peak at 22.9 MDa (0.34 MDa FWHM) is due to strain 18.

The measured spectrum does not show the three predicted subpopulations, and the measured peak appears to be at a significantly higher mass than the main peak in the predicted spectrum. To overlap with the main peak in the measured spectrum with the predicted spectrum, it is necessary to shift the measured spectrum to lower mass by around 1 MDa. An excess mass of around 5% is much larger than usually associated with counterions, salt adducts, and trapped solvents. The large excess mass is probably due to an aluminum adjuvant that is not completely removed, though we cannot rule out the presence of misassembled particles.

If the measured spectrum is shifted down by 1 MDa, the peak at 22.9 MDa in the simulated spectrum (the peak due to strain 18) falls outside the range of the main peak in the measured spectrum and lacks an obvious corresponding feature in the measured spectrum. It is possible that the broad higher mass feature in the measured spectrum centered on around 26 MDa is due to strain 18 (the 22.9 MDa peak in the simulated spectrum) that has retained significantly more residual aluminum adjuvant than the other strains. Thus, adjuvant release may depend on the strain.

Figure 4c shows the charge versus mass heat map for Gardasil-9. For the main peak in the mass distribution, there is a single charge group at around 390 e. This differs from the elongated charge distributions found for IPOL and RotaTeq. Note that the charges for the broad higher mass distribution are systematically larger than that for the main peak. The higher charge indicates a larger object, which would be consistent with the aluminum adjuvant bound to the exterior of the VLPs.

CONCLUSIONS

CDMS provides a new way to characterize vaccines, offering substantial benefits in both their development and mass production. Representative examples of three classical vaccine strategies were investigated: IPOL, an inactivated whole virus vaccine; RotaTeq, a live-attenuated virus; and Gardasil-9, a recombinant vaccine using VLP's. Protein subunits are the other major types of classical vaccine strategy, and we recently reported CDMS measurements for the SARS-CoV-2 spike protein.[32] In each case, the ability to analyze the viral antigen goes beyond the realm of a typical mass spectrometer because

https://doi.org/10.1021/acs.analchem.1c01893
Anal. Chem. 2021, 93, 11965−11972

of both the size of the antigen and the sample heterogeneity. CDMS can provide the mass distribution for the antigen and masses and relative abundances of impurities (i.e., empty and defective particles). CDMS is a robust technique that does not require sample-specific standards. The depth of information available from a CDMS assay that is quick (usually <1 h analysis time), efficient (sample size 10−20 μL), and sensitive enough to detect low abundance intermediates and misassembled particles has the potential to increase the speed and reduce the cost of vaccine development.

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

Lohra M. Miller − *Chemistry Department, Indiana University, Bloomington, Indiana 47405, United States*

Kevin M. Bond − *Chemistry Department, Indiana University, Bloomington, Indiana 47405, United States*

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

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.analchem.1c01893

Notes

The authors declare the following competing financial interest(s): Authors except BED and MFJ declare no financial interest. BED is an employee of Megadalton Solutions, a company engaged in commercializing CDMS. MFJ and BED are shareholders in Megadalton Solutions.

ACKNOWLEDGMENTS

The authors are grateful to the NSF IUCRC Center for Bioanalytic Metrology (CBM) for financial support provided under Grant No. IIP-1916645 and for valuable discussions with CBM industry partners and staff. BED is funded through an NIH SBIR grant (5R44GM136095) awarded to Megadalton Solutions.

REFERENCES

(1) Centers for Disease Control and Prevention (CDC). *Ten Great Public Health Achievements-United States, 1900-1999*; Morb. Mortal. Wkly. Rep, National Library of Medicine, 1999; pp 241−243.

(2) U.S. Food and Drug, FDA News Release. *Takes Additional Action in Fight Against COVID-19 By Issuing Emergency Use Authorization for Second COVID-19 Vaccine*. Last revised December 2020. https://www.fda.gov/news-events/press-announcements/fda-takes-additional-action-fight-against-covid-19-issuing-emergency-use-authorization-second-covid (accessed 2021-01-24).

(3) U.S. Food and Drug, Center for Biologics Evaluation and Research. Docket number 98D-0401. Content and Format of Chemistry, Manufacturing and Controls Information and Establishment Description Information for a Vaccine or Related Product. Last revised August 2019. https://www.fda.gov/media/73614/download (accessed 2021-1-24).

(4) Weiss, V. U.; Pogan, R.; Zoratto, S.; Bond, K.; Boulanger, P.; Jarrold, M. F.; N Lyktey, N.; Pahl, D.; Puffler, N.; Schelhaas, M.; Selivanovitch, E.; Uetrecht, C.; Allmaier, G. *Anal. Bioanal. Chem.* **2019**, *411*, 5951−5962.

(5) Torisu, T.; Shikama, S.; Nakamura, K.; Enomoto, K.; Maruno, T.; Mori, A.; Uchiyama, S.; Satou, T. *J. Pharm. Sci.* **2021**, *110*, 2121−2129.

(6) Smith, J.; Lipsitch, M.; Almond, J. R. W. *Lancet* **2011**, *378*, 428−438.

(7) Plotkin, S.; Robinson, J. M.; Cunningham, G.; Iqbal, R.; Larsen, S. *Vaccine* **2017**, *35*, 4064−4071.

(8) Milstien, J.; Batson, A.; Meaney, W. *Vaccine* **1997**, *15*, 1358−1363.

(9) Kaufmann, S. H. E.; McElrath, M. J.; Lewis, D. J. M.; Giudice, G. D. *Curr. Opin. Immunol.* **2014**, *28*, 18−26.

(10) Milstien, J. *Procedia Vaccinol.* **2009**, *1*, 183−188.

(11) Hebert, C. G.; DiNardo, N.; Evans, Z. L.; Hart, S. J.; Hachmann, A.-B. *Vaccine* **2018**, *36*, 6061−6069.

(12) Pisano, G. P. *The Development Factory: Unlocking the Potential of Process Innovation*; Harvard Business School Press, Cambridge, MA, 1996.

(13) Bolton, G. R.; Violand, B. N.; Wright, R. S.; Sun, S.; Sunasara, K. M.; Waltson, K.; Coffman, J. L.; Gallo, C.; Godavarti, R. *BioPharm. Int.* **2011**, *24*, s8−s15.

(14) Soema, P. C.; Kompier, R.; Amorij, J.-P.; Kersten, G. F. A. *Eur J. Pharm. Biopharm.* **2015**, *94*, 251−263.

(15) Fuerstenau, S. D.; Benner, W. H. *Rapid Commun. Mass Spectrom.* **1995**, *9*, 1528−1538.

(16) Benner, W. H. *Anal. Chem.* **1997**, *69*, 4162−4168.

(17) Mabbett, S. R.; Zilch, L. W.; Maze, J. T.; Smith, J. W.; Jarrold, M. F. *Anal. Chem.* **2007**, *79*, 8431−8439.

(18) Contino, N. C.; Jarrold, M. F. *Int. J. Mass Spectrom.* **2013**, *345-347*, 153−159.

(19) Doussineau, T.; Kerleroux, M.; Dagany, X.; Clavier, C.; Barbaire, M.; Maurelli, J.; Antoine, R.; Dugourd, P. *Rapid Commun. Mass Spectrom.* **2011**, *25*, 617−623.

(20) Doussineau, T.; Santacreu, M.; Antoine, R.; Dugourd, P.; Zhang, W.; Chaduc, I.; Lansalot, M.; D'Agosto, F.; Charleux, B. *ChemPhysChem* **2013**, *14*, 603−609.

(21) Harper, C. C.; Williams, E. R. *J. Am. Soc. Mass Spectrom.* **2019**, *30*, 2637−2645.

(22) Harper, C. C.; Elliott, A. G.; Oltrogge, L. M.; Savage, D. F.; Williams, E. R. *Anal. Chem.* **2019**, *91*, 7458−7465.

(23) 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.

(24) 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.

(25) Keifer, D. Z.; Motwani, T.; Teschke, C. M.; Jarrold, M. F. *Rapid Commun. Mass Spectrom.* **2016**, *30*, 1957−1962.

(26) 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.

(27) 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.

(28) Lutomski, C. A.; Lyktey, N. A.; Zhao, Z.; Pierson, E. E.; Zlotnick, A.; Jarrold, M. F. *J. Am. Chem. Soc.* **2017**, *139*, 16932−16938.

(29) Lutomski, C. A.; Lyktey, N. A.; Pierson, E. E.; Zhao, Z.; Zlotnick, A.; Jarrold, M. F. *J. Am. Chem. Soc.* **2018**, *140*, 5784−5790.

(30) Bond, K. M.; Lyktey, N. A.; Tsvetkova, I. B.; Dragnea, B.; Jarrold, M. F. *J. Phys. Chem. B* **2020**, *124*, 2124−2131.

(31) Bond, K.; Tsvetkova, I.; Wang, J. C.-Y.; Jarrold, M. F.; Dragnea, B. *Small* **2020**, *16*, No. 2004475.

(32) 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.

(33) Contino, N. C.; Jarrold, M. F. *Int. J. Mass Spectrom.* **2013**, *345-347*, 153−159.

(34) Keifer, D. Z.; Shinholt, D. L.; Jarrold, M. F. *Anal. Chem.* **2015**, *87*, 10330−10337.

(35) Hogan, J. A.; Jarrold, M. F. *J. Am. Soc. Mass Spectrom.* **2018**, *29*, 2086−2095.
(36) Draper, B. E.; Anthony, S. N.; Jarrold, M. F. *J. Am. Soc. Mass Spectrom.* **2018**, *29*, 2160−2172.
(37) Draper, B. E.; Jarrold, M. F. *J. Am. Soc. Mass Spectrom.* **2019**, *30*, 898−904.
(38) Todd, A. R.; Alexander, A. W.; Jarrold, M. F. *J. Am. Soc. Mass Spectrom.* **2020**, *31*, 146−154.
(39) Todd, A. R.; Jarrold, M. F. *J. Am. Soc. Mass Spectrom.* **2020**, *31*, 1241−1248.
(40) Todd, A. R.; Barnes, L. F.; Young, K.; Zlotnick, A.; Jarrold, M. F. *Anal. Chem.* **2020**, *92*, 11357−11364.
(41) Tuthill, T. J.; Groppelli, E.; Hogle, J. M.; Rowlands, D. J. Picornaviruses. In *Cell Entry by Non-Enveloped Viruses*; Johnson, J. E., Ed.; Current Topics in Microbiology and Immunology; Springer, 2010; Vol. *343*, pp 43−89.
(42) Jiang, P.; Liu, Y.; Ma, H.-C.; Paul, A. V.; Wimmer, E. *Microbiol. Mol. Biol. Rev.* **2014**, *78*, 418−437.
(43) Kitamura, N.; Semler, B. L.; Rothberg, P. G.; Larsen, G. R.; Adler, C. J.; Dorner, A. J.; Emini, E. A.; Hanecak, R.; Lee, J. J.; van der Werf, S.; Anderson, C. W.; Wimmer, E. *Nature* **1981**, *291*, 547−553.
(44) Nomoto, A.; Omata, T.; Toyoda, H.; Kuge, S.; Horie, H.; Kataoka, Y.; Genba, Y.; Nakano, Y.; Imura, N. *Proc. Natl. Acad. Sci. U.S.A.* **1982**, *79*, 5793−5797.
(45) Bubeck, D.; Filman, D. J.; Cheng, N.; Steven, A. C.; Hogle, J. M.; Belnap, D. M. *J. Virol.* **2005**, *79*, 7745−7755.
(46) Hummeler, K.; Anderson, T. F.; Brown, R. A. *Virology* **1962**, *16*, 84−90.
(47) Beale, A. J.; Mason, P. J. *J. Hyg.* **1962**, *60*, 113−121.
(48) Sawyer, L. A.; McInnis, J.; Albrecht, P. *Biologicals* **1993**, *21*, 169−177.
(49) Singer, C.; Knauert, F.; Bushar, G.; Klutch, M.; Lundquist, R.; Quinnan, G. V., Jr. *J. Biol. Stand.* **1989**, *17*, 137−150.
(50) Moscufo, N.; Chow, M. *J. Virol.* **1992**, *66*, 6849−6857.
(51) Basavappa, R.; Filman, D. J.; Syed, R.; Flore, O.; Icenogle, J. P.; Hogle, J. M. *Protein Sci.* **1994**, *3*, 1651−1669.
(52) Moscufo, N.; Simons, J.; Chow, M. *J. Virol.* **1991**, *65*, 2372−2380.
(53) 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.
(54) Sanofi Pasteur, S. A. Poliovirus Vaccine Inactivated (Monkey Kidney Cell), IPOL [package insert]. Food and Drug Administration. 2019, https://www.fda.gov/files/vaccines%2C%20blood%20%26%20biologics/published/Package-Insert---IPOL.pdf (accessed Jan 24, 2021).
(55) Ansardi, D. C.; Porter, D. C.; Anderson, M. J.; Morrow, C. D. Poliovirus Assembly and Encapsidation of Genomic RNA. In *Advances in Virus Research*; Maramorosch, K.; Murphy, F. A.; Shatkin, A. J., Eds.; Elsevier, San Diego, CA, 1996; Vol. *46*, pp 1−68.
(56) Strauss, M.; Filman, D. J.; Belnap, D. M.; Cheng, N.; Noel, R. T.; Hogle, J. M. *J. Virol.* **2015**, *89*, 4143−4157.
(57) Barnes, L. F.; Draper, B. E.; Chen, Y.-T.; Powers, T. W.; Jarrold, M. F. Quantitative analysis of genome packaging in recombinant AAV vectors by charge detection mass spectrometry. *Mol. Ther. Methods Clin. Dev.*
(58) Fernandez de la Mora, J. *Anal. Chim. Acta* **2000**, *406*, 93−104.
(59) Wilton, T. Molecular Characterisation of Poliovirus Inactivation with Formaldehyde or Other Alternative Chemical Compounds. PhD Thesis, Imperial College London, Department of Medicine, Division of Infectious Diseases, 2012.
(60) Lentz, K. N.; Smith, A. D.; Geisler, S. C.; Cox, S.; Buontempo, P.; Skelton, A.; DeMartino, J.; Rozhon, E.; Schwartz, J.; Girijavallabhan, V.; O'Connell, J.; Arnold, E. *Structure* **1997**, *5*, 961−978.
(61) Curry, S.; Chow, M.; Hogle, J. M. *J. Virol.* **1996**, *70*, 7125−7131.
(62) Wilton, T.; Dunn, G.; Eastwood, D.; Minor, P. D.; Martin, J. *J. Virol.* **2014**, *88*, 11955−11964.
(63) Lin, J.; Lee, L. Y.; Roivainen, M.; Filman, D. J.; Hogle, J. M.; Belnap, D. M. *J. Virol.* **2012**, *86*, 5959−5962.
(64) Merck Sharp & Dohme Corp. Rotavirus Vaccine, Live, Oral, Pentavalent, RotaTeq [package insert]. U.S. Food and Drug, Center for Biologics Evaluation and Research. https://www.fda.gov/vaccines-blood-biologics/vaccines/rotateq.2020. (accessed 2021-1-24).
(65) RotaTeq. INN-Rotavirus Vaccine, Live, Oral.https://www.ema.europa.eu/en/documents/scientific-discussion/rotateq-epar-scientificdiscussion_en.pdf (accessed Jan 24, 2021).
(66) McClain, B.; Settembre, E.; Temple, B. R. S.; Bellamy, A. R.; Harrison, S. C. *J. Mol. Biol.* **2010**, *397*, 587−599.
(67) Matthijnssens, J.; Joelsson, D. B.; Warakomski, D. J.; Zhou, T.; Mathis, P. K.; van Maanen, M.-H.; Ranheim, T. S.; Ciarlet, M. *Virology* **2010**, *403*, 111−127.
(68) Todd, A. R.; Jarrold, M. F. *Anal. Chem.* **2019**, *91*, 14002−14008.
(69) Greig, S. L.; Berriman, J. A.; O'Brien, J. A.; Taylor, J. A.; Bellamy, A. R.; Yeager, M. J.; Mitra, A. K. *J. Mol. Biol.* **2006**, *356*, 209−221.
(70) Merck Sharp & Dohme Corp. Human Papillomavirus 9-valent Vaccine, Recombinant, Gardasil 9. https://www.fda.gov/vaccines-blood-biologics/vaccines/gardasil-9. (accessed 2021-1-24).
(71) Zhao, Q.; Potter, C. S.; Carragher, B.; Lander, G.; Sworen, J.; Towne, V.; Abraham, D.; Duncan, P.; Washabaugh, M. W.; Sitrin, R. D. *Hum. Vaccines Immunother.* **2014**, *10*, 734−739.
(72) Mcbride, A. A.; Porter, S. S. Human Papillomaviruses.In *Reference Module in Life Sciences*; Academic Press, 2019; Vol. *2*, pp 493−501.

https://doi.org/10.1021/acs.analchem.1c01893
Anal. Chem. 2021, 93, 11965−11972