Determination of the Complete Amino Acid Sequence for the Coat Protein of Brome Mosaic Virus by Time-of-Flight Mass Spectrometry

EVIDENCE FOR MUTATIONS ASSOCIATED WITH CHANGE OF PROPAGATION HOST

Yi-Min She, Steve Haber, Dallas L. Seifers, Alexander Loboda, Igor Chernushevich, Hélène Perreault and Kenneth G. Standing

From the Department of Physics & Astronomy, University of Manitoba, Winnipeg, MB R3T 2N2, Canada, the Cereal Research Centre, Agriculture & Agrifood Canada, Winnipeg, MB R3T 2M9, Canada, the Agricultural Research Center, Kansas State University, Hays, Kansas 67601-9228, MDS Sciex, Concord, ON L4K 4V8, Canada, and the Department of Chemistry, University of Manitoba, Winnipeg, MB R3T 2N2, Canada

Time-of-flight mass spectrometry (TOFMS) has been applied to determine the complete coat protein amino acid sequences of a number of distinct brome mosaic virus (BMV) isolates. Ionization was carried out by both electrospray ionization and matrix-assisted laser desorption/ionization (MALDI). After determining overall coat protein masses, the proteins were digested with trypsin or Lys-C proteinases, and the digestion products were analyzed in a MALDI QqTOF mass spectrometer. The N terminus of the coat protein was found to be acetylated in each BMV isolate analyzed. In one isolate (BMV-Valverde), the amino acid sequence was identical to that predicted from the cDNA sequence of the "type" isolate, but deviations from the predicted amino acid sequence were observed for all the other isolates analyzed. When isolates were propagated in different host taxa, modified coat protein sequences were observed in some cases, along with the original sequence. Sequencing by TOFMS may therefore provide a basis for monitoring the effects of host passaging on a virus at the molecular level. Such TOFMS-based analyses assess the complete profiles of coat protein sequences actually present in infected tissues. They are therefore not subject to the selection biases inherent in deducing such sequences from reverse-transcribed viral RNA and cloning the resulting cDNA.

Among the parameters characterizing a virus are the mass and the amino acid sequence of the viral coat protein. Molecular masses have traditionally been determined by gel electrophoresis; a method whereby the accuracy and resolution are incapable of distinguishing proteins that differ in mass by less than ~1%. The amino acid sequence provides more definitive information, but most sequences currently in the literature have been deduced from nucleic acid sequences of specific cDNA clones. Thus they fail to take into account changes that are not defined by the nucleic acid, such as post-translational modifications. They also fail to provide a complete profile of the viral coat protein population actually present in infected tissue.

With the development of electrospray ionization (ESI) and matrix-assisted laser desorption/ionization (MALDI), as well as new types of mass analyzers, mass spectrometry (MS) now affords a rapid and efficient approach for obtaining more detailed information about virus coat protein sequences in infected tissue (1–22).

The simplest type of MS observation determines the mass of the coat protein subunit. This can be carried out with an accuracy of a few daltons (in a few tens of kilodaltons), which is usually sufficient to distinguish coat proteins of different viruses or different isolates of the same virus species (15). However, much more information can be deduced by digesting the protein with specific endopeptidases and analyzing the digestion products by MS (i.e. peptide mapping). These analyses can often define the parts of the protein sequences that agree with those predicted from nucleic acid data and thus identify the regions where the amino acid sequence differs from prediction.

Still more detailed information can be obtained from tandem mass spectrometry (MS/MS), in which a given proteolytic fragment ion is selected by one mass analyzer and is forced to collide with the molecules of a target gas. The resulting daughter ions are then characterized by a second mass analyzer. This increases the analytical power of MS and allows the rapid determination of amino acid sequences from multiple protein specimens, even if they are closely related.

We have applied these approaches to investigate a number of plant viruses. Such viruses are of considerable economic importance, but in addition, the use of plant viruses as model systems for basic investigations has several advantages. (a) Propagation and maintenance are easier than for human and animal viruses, allowing faster experiment cycles. This is particularly important in experiments designed to investigate the link between altered virulence achieved by serial passage through alternative hosts and specific changes in viral coat protein amino acid sequence (see the example below). (b) Compared with many human and animal viruses, plant viruses pose little risk to humans. (c) There is little difficulty in meeting requirements for health and care of hosts to be infected, and experiments can accommodate high rates of host destruction if necessary. “Cruelty to plants” has not yet captured the attention of placard-waving demonstrators. (d) Most plant viruses are considerably less complex than most human, animal, and...
bacterial viruses, making it easier to generate well defined test materials for experiments.

We have determined the ratio of $m/z$ (~27,000) for intact virions of brome mosaic virus (BMV) (15, 16), measured the coat protein masses of more than 20 plant virus isolates by time-of-flight mass spectrometry (TOFMS) (17), and subsequently determined the complete amino acid sequences of several groups of ssRNA plant viruses (18–22). These measurements have allowed us to identify variations among individual isolates. Analyses of such variations may point to evolutionary relationships, whether the variations are caused by post-translational modifications, deletions, point mutations, or other mechanisms.

As an illustration of how these methods can be applied to closely related virus isolates, we describe here TOFMS measurements on a group of geographic isolates of brome mosaic virus (BMV). This virus is the type member of the bromoviridae (23) and was one of the first multipartite-genome ssRNA plant viruses to have its genome completely sequenced (24–26). It is found in the Great Plains of North America and has also been reported from Europe and western Asia (27). In nature, BMV predominantly infects grasses and cereals and may cause sporadic crop losses (28). It is unusual in its ability to infect an array of experimental broadleaf hosts (27).

**EXPERIMENTAL PROCEDURES**

**Virus Isolates and Sample Preparation**—We first examined the BMV isolates listed in Table I. BMV-P and BMV-N were obtained from the Cereal Research Center collection of virus isolates (28); isolates investigated subsequently, such as BMV-P* and -N*, were derived from these in 1999 by repropagation in alternate hosts. The other isolates listed (BMV1 to BMV4) were supplied to the Manitoba TOFMS laboratory as coded samples from the collection of Kansas State University; their subsequent identification (after the analyses were carried out) is listed in the next column. Care was taken to propagate and purify the BMV isolates separately and to eliminate the possibilities of cross-contamination. Virus purification was carried out according to published methods (28, 29).

After the removal of low molecular weight impurities (~5000 Da) by centrifugal filtration, the purified coat protein was lyophilized to determine its amount. A 1% (w/w, enzyme/protein) trypsin or endoproteinase Lys-C solution was added to the protein solution (1 μg/μl) in 25 mM ammonium bicarbonate, followed by overnight enzymatic digestion at 37 °C. The reaction was terminated by freezing, and the products were stored at −20 °C for subsequent MALDI and ESI analyses.

**Mass Spectrometry**—Measurements of the undigested protein masses were performed on MALDI/TOF (32) and ESI/TOF (33) mass spectrometers. Some of the initial measurements of the proteolytic fragments produced by digestion with trypsin or Lys-C proteinases were carried out on the same instruments. However, mass determinations of proteolytic fragments and MS/MS measurements on the daughter ions in the more difficult cases were made on the Manitoba/Sciex prototype tandem quadrupole/TOF mass spectrometer (QqTOF) (34) coupled to a MALDI source (35, 36), shown in Fig. 1. The MALDI matrix consisted of 2,5-dihydroxybenzoic acid, and argon or nitrogen were used as collision gases; energies of 50–180 eV were applied. For comparison, monoisotopic mass values of the peptide fragments from either enzymatic digestions or collision-induced dissociation (CID) were calculated using the computation program ProMac (Sciex, Concord, ON).

In addition, some ESI measurements, including parent ion scans (37, 38), were made on a Micromass Quattro-LC ESI triple quadrupole instrument. Here the parent ion scans were carried out for arginine, i.e. the third quadrupole was set to detect $m/z$ 175 (diagnostic of Arg), while the first quadrupole was scanned, thus yielding an $m/z$ spectrum of all

---

**TABLE I**

| Sample | Isolate | Ref. | MALDI* | ESI* | Calc. mass# |
|--------|---------|------|--------|------|-------------|
| BMV-P  | Portage (1989) | (28) | 20265.3 | 20265.2 | 20265.4 |
| BMV-T  | Type-pv47 | (29) | 20279.9 | 20279.1 | 20279.4 |
| BMV-1 (T) | Type-pv47 | (29, 31) | 20280.5 | 20280.3 | 20279.4 |
| BMV-2 (W) | Western wheatgrass | (31) | 20163.6 | 20164.8 | 20166.2 |
| BMV-3 (V) | Valverde | (30, 31) | 20296.3 | 20296.8 | 20295.4 |
| BMV-4 (P) | Portage (1989) | (28, 31) | 20265.6 | 20263.9 | 20265.4 |
| BMV-N  | Nebraska (1989) | (28) | 20191.4 | 20191.4 | 20194.3 |

* Measured MALDI masses were obtained from the measured $m/z$ values of the MH+ ions, and measured ESI masses were obtained from the deconvoluted ESI mass spectra.

# Calculated molecular masses were obtained from the measured sequences taking account of acetylation at the N-terminus and the other modifications observed (see text).
parent ions containing arginine. In addition, such scans often give better signal/background ratios than ordinary MS spectra.

RESULTS

Preliminary Characterization of the BMV Isolates

Typical MALDI and ESI spectra of the intact protein sub-units are shown in Figs. 2 and 3. The measured average masses for the coat proteins of the isolates (Table I) all differed from the value of 20,253 Da calculated from the published coat protein sequence of the type isolate (24–26) (derived from the cloned cDNA sequence, assuming deletion of the N-terminal methionine), and the masses of most isolates differed from each other. Clearly a number of modifications in the coat proteins had occurred. To define these modifications, we carried out a series of MS and MS/MS measurements after proteolytic digestion of the proteins, as described above. These digests provided progressively more detailed information that finally made it possible to characterize the differences among the BMV isolates.

In the initial protein sequencing experiments, ESI/TOF MS was used to analyze tryptic digests of BMV-P. Peptide mapping and MS/MS measurements on the tryptic fragments led to confirmation of the type sequence (derived from the published cDNA sequence of BMV (24–26)) for 180 of 189 amino acids (17). The undefined residues were contained within two short regions near the N terminus (residues 2–8 and 20–26), and we deduced from the measurements that these stretches included at least two differences from the type sequence.

To resolve this uncertainty, further digestions of the BMV-P samples (with Lys-C proteinase as well as trypsin) were carried out, and the observed fragments are listed in Table II.

Consistent with the initial analyses, no ions corresponding to the predicted masses of the (2–8) and (20–26) fragments were found. However, a prominent singly charged ion could be observed at m/z = 679.4 in the ESI spectrum, although not in its MALDI counterpart, and we realized that its mass is 42 Da greater than the predicted mass of the fragment encompassing residues 2–8. This is consistent with acetylation at the N terminus, a common post-translational modification. The acetylation was confirmed by the MS/MS spectrum of this parent ion shown in Fig. 6, which contains a complete set of acetylated b ions together with a complete set of y ions, of which only y7 (i.e. the parent ion [M+H]+) is acetylated. Moreover, the diagnostic ion at m/z = 679.4 is present in the ESI spectra from the digests of all samples tested, consistent with N-terminal acetylation of the BMV coat protein (and deletion of the N-terminal methionine) for all the BMV isolates. This observation proved to be the key that allowed us to determine the complete coat protein sequences of all the isolates examined.

Determination of the Coat Protein Sequences of the Isolates

**BMV-V—**The Valverde isolate (BMV-V) (30) is the simplest case. The N-terminal of its coat protein is acetylated, but its mass and the masses of its peptide digest fragments are otherwise identical with those deduced from the published type sequence (24–26), consistent with the absence of any other modification. MS/MS measurements confirmed this result, as well as the absence of zero-sum mutations.

**BMV-P—**Prominent ions at m/z = 445 are found in both ESI and MALDI spectra (Figs. 4 and 5) and in the ESI parent ion spectrum (Fig. 6), consistent with the predicted mass of residues 20–22, and thus reducing the unknown region to residues 23–26. The only alteration in that region consistent with the
overall mass is a substitution of Arg for Trp at residue 23, and this interpretation is supported by the observation of ions at m/z 5347.0 (sequence 24–26, TAR), 503.2 (sequence 23–26, RTAR) and 601.4 (sequence 20–23, RNRR) (see also BMV-T measurements below). The modification is consistent with a single base change of the nucleic acid sequence (UGG to CGG). Together with acetylation at the N terminus, it then accounts fully for the observed mass differences (Tables I and II) between BMV-P and the published type sequence (24–26).

It is interesting to note that Wang et al. (39, 40) have reported a BMV isolate from Nebraska whose sequence is identical with that of BMV-P, indicating that the same mutation has arisen at different geographic locations.

BMV-T—The ions mentioned above are also observed in the MS spectra of the BMV-T isolate. In addition, they appear in the BMV-T parent ion scan for m/z 5175, an Arg marker (Fig. 7). However, the most convincing evidence for the Trp → Arg modification was obtained when fragments from a Lys-C digestion of BMV-T (pv47) were analyzed in the MALDI-QqTOF instrument (Table III), enabling high mass accuracy to be obtained for both parents and daughters. The ion observed at m/z 53685.106 has a mass 29.973 Da smaller than the value deduced from the published nucleotide sequence for fragment 9–41, consistent with the modification Trp → Arg (calculated mass difference Δm = 29.978 Da). MS/MS measurements confirmed that Trp had indeed mutated to Arg at residue 23 in BMV-T, as well as in BMV-P.

Another anomalous fragment ion was observed at m/z 2886.503 (Table III). This is 14.024 Da larger than the mass of the deduced fragment encompassing residues 166–189, sug-

### Table II

| Peptides | m/z (MALDI) | m/z (ESI) | MH⁺ (calc.) | Peptide sequence |
|----------|-------------|-----------|-------------|-----------------
| 2–8      | 679.4       | 679.7     | 679.7       | STSGTGK (+42 Da)⁹|
| 9–11     | 406.9       | 407.5     | 407.5       | MTR           |
| 9–14     | 762.4       | 762.9     | 762.9       | MTRAQQR       |
| 9–15     | 918.2       | 919.0     | 919.0       | MTRAQQR       |
| 12–14    | 374.0       | 374.4     | 374.4       | AQR           |
| 12–15    | 530.3       | 530.6     | 530.6       | AQR           |
| 16–19    | 388.1       | 388.4     | 388.4       | AAAR          |
| 16–20    | 544.2       | 544.6     | 544.6       | AAAR          |
| 21–22    | 289.9       | 289.3     | 289.3       | NR            |
| 20–22    | 445.2       | 445.5     | 445.5       | RNR           |
| 20–23    | 601.4       | 601.7     | 601.7       | RNRR¹         |
| 23–26    | 503.2       | 503.6     | 503.6       | RTAR¹         |
| 24–26    | 347.0       | 347.4     | 347.4       | TAR           |
| 27–41    | 1506.2      | 1506.8    | 1506.8      | VQP Viveplaqgk |
| 2–41     | 4345.6      | 4348.0    | 4348.0      | STSGTGKMTRARQRRAARRRTARVQ |
| 42–44    | 331.3       | 331.4     | 331.4       | AIK           |
| 45–53    | 909.8       | 910.0     | 910.0       | AIAAGYSISIK |
| 54–64    | 1175.6      | 1179.2    | 1179.2      | WEASSDAITAK |
| 65–81    | 1830.7      | 1830.0    | 1830.0      | ATNAMSSITPHLSSEK |
| 82–89    | 261.7       | 261.3     | 261.3       | NK            |
| 84–86    | 388.0       | 388.5     | 388.5       | ELK           |
| 84–89    | 701.5       | 701.8     | 701.8       | ELKVR         |
| 90–103   | 1494.2      | 1494.9    | 1494.9      | YLLWGLPPSVEGR |
| 87–103   | 1605.9      | 1807.2    | 1807.2      | VGRVLLWGLPPSVEGR |
| 104–105  | 260.1       | 260.4     | 260.4       | IK            |
| 82–105   | 2658.9      | 2661.3    | 2661.3      | NKEKLVGRVLLWGLPPSVEGR |
| 87–105   | 2047.4      | 2048.6    | 2048.6      | VGRVLLWGLPPSVEGR |
| 106–130  | 2506.2      | 2507.8    | 2507.8      | ACVAEKQGAEEAQPVLAVADSSK |
| 131–165  | 3774.7      | 3776.3    | 3776.3      | EVVAAATYDRAFRGATLGDNLNLQY |
| 143–165  | 2420.3      | 2421.8    | 2421.8      | LYASEAVPAK |
| 166–189  | 2872.6      | 2874.3    | 2874.3      | NVVHLHVEHVRPFTDDFPTPVYR |

⁹ The calculated values shown are average masses (Da) taking account of the modifications observed.

b Peptide modified.

---

**Fig. 6.** CID mass spectrum of the singly charged ion at m/z 679.4 measured by ESI MS/MS on the triple quadrupole instrument. An apostrophe denotes an acetylated ion (e.g., b1 in place of b1), i.e., an ion 42 Da heavier than the value predicted by the nucleic acid sequence.

**Fig. 7.** Parent ion scans of m/z 175 on the triple quadrupole instrument for isolates of BMV-T (type pv47), BMV-W (Western wheatgrass), BMV-P (Portage), and BMV-V (Valverde), digested by Lys-C protease.
Beggesting the substitution of Leu or Ile for Val (calculated Δm = 14.016 Da). Comparison of BMV-T and BMV-P spectra in the ESI parent ion scans gave Δ(m/z) = 4.6 (962.8–958.2) for the triply charged ion and Δ(m/z) = 3.7 (722.5–718.8) for the quadruply charged ion (Fig. 7), consistent with this interpretation. It was confirmed and localized by MALDI-MS/MS measurements of the corresponding triply charged ion at m/z = 963 (Table IV). All of the observed b ions from b2 to b24 have masses of 14 Da greater than predicted, localizing the modification at residue 167. Consistent with this observation, the C-terminal ions up to y22 have m/z values that agree with calculations made for the unmodified protein, whereas y23 (i.e. the parent ion [M+H]+) is modified. Substitution of Ile or Leu for Val at residue 167 is therefore the only solution consistent with the MS measurements. It corresponds to a single base change of the nucleic acid sequence (GUC to AUC or CUC; see below). As expected, similar results were obtained from the BMV-1 (pv47) isolate.

**BMV-W**—The overall mass of BMV-W is smaller than that of BMV-P by about 100 Da (Table I), and a prominent doubly charged ion appears in the parent ion spectrum (Fig. 7) of BMV-W at m/z = 868.2 Da, corresponding to a mass of 99 Da less than the calculated mass of fragment 24–41. The mass of this ion was determined accurately by a MALDI QqTOF measurement of the tryptic digest, which yielded the spectrum shown in Fig. 8. In this spectrum, the ion corresponding to fragment 24–41 (expected mass 1834.055 Da) is replaced by another peak at m/z = 1734.981, with a monoisotopic mass 99.074 Da smaller than predicted. Because valine has a mass of

---

**TABLE III**

Peptide mapping of the Lys-C digested BMV-T(pv47) coat protein by MALDI QqTOF MS

| Fragment | m/z (calc.) | m/z (obs.) | Δ (mDa) | Peptide sequence |
|----------|-------------|------------|---------|-----------------|
| 9–11     | 2888.495    | 2888.494   | +9      | (K) AVVVHLEVEHVRPTFDDFTFPPPVR (+) + 14.024 Da (V→I/D) |
| 112–130  | 1505.871    | 1505.869   | +2      | (R) VQPVIVELAAGQK (A) |
| 131–165  | 2417.512    | 2417.513   | +1      | (K) KNEKLGFVRLWLLPGFSVAGPK (A) |
| 166–189  | 2886.503    | 2886.494   | +9      | (K) AVVVHLEVEHVRPTFDDFTFPPPVR (+) + 14.024 Da (V→I/D) |

---

**TABLE IV**

CID fragmentation of the singly charged ion of m/z 2886.50 (MALDI) and the triply charged ion of m/z 963.6 (ESI) from the BMV-T (pv47) peptide (residues 166–189, AVVVHLEVEHVRPTFDDFTFPPPVR)

| b ion | m/z (calc.) | y ion | m/z (calc.) |
|-------|-------------|-------|-------------|
| MALDI | ESI         |       | MALDI       | ESI         |       |
|       | [M+H]+      |       | [M+H]+      | [M+H]+      |       |
|       | [M+H]+      |       | [M+H]+      | [M+H]+      |       |
| b1    | 1127.630    | y1    | 175.6       | 1504.721    | y3    |
| b2    | 1226.689    | y1    | 175.6       | 1504.721    | y3    |
| b3    | 1382.797    | y2    | 175.6       | 1504.721    | y3    |
| b4    | 1478.844    | y3    | 175.6       | 1504.721    | y3    |
| b5    | 1580.891    | y4    | 175.6       | 1504.721    | y3    |
| b6    | 1727.906    | y5    | 175.6       | 1504.721    | y3    |
| b7    | 1842.987    | y6    | 175.6       | 1504.721    | y3    |
| b8    | 1958.014    | y7    | 175.6       | 1504.721    | y3    |
| b9    | 2015.082    | y8    | 175.6       | 1504.721    | y3    |
| b10   | 2125.031    | y9    | 175.6       | 1504.721    | y3    |
| b11   | 2252.150    | y10   | 175.6       | 1504.721    | y3    |
| b12   | 2353.198    | y11   | 175.6       | 1504.721    | y3    |
| b13   | 2450.251    | y12   | 175.6       | 1504.721    | y3    |
| b14   | 2549.319    | y13   | 175.6       | 1504.721    | y3    |
| b15   | 2659.665    | y14   | 175.6       | 1504.721    | y3    |
| b16   | 2868.472    | y15   | 175.6       | 1504.721    | y3    |
| b17   | 3098.185    | y16   | 175.6       | 1504.721    | y3    |
| b18   | 3304.304    | y17   | 175.6       | 1504.721    | y3    |
| b19   | 3537.545    | y18   | 175.6       | 1504.721    | y3    |
| b20   | 3734.381    | y19   | 175.6       | 1504.721    | y3    |
| b21   | 3941.511    | y20   | 175.6       | 1504.721    | y3    |
| b22   | 4143.671    | y21   | 175.6       | 1504.721    | y3    |
| b23   | 4343.831    | y22   | 175.6       | 1504.721    | y3    |
| b24   | 4543.991    | y23   | 175.6       | 1504.721    | y3    |

---

a Values listed are monoisotopic masses (Da).

b Modified peptide.
99.068 Da, this suggests that one of the three valines in the segment (TARVPVIVEPLAAGQGK) has been deleted. MS/MS measurements on the \(m/z = 1735\) ion, shown in Fig. 9a, indicated that both valine residues 30 or 32 were present, but at first inspection a deletion of Val-27 appeared to give good agreement between the MS/MS data and calculations, as shown in Table V. Indeed there would be almost perfect agreement if the mass were measured to only one or two decimal places (typical accuracies for some other methods of MS measurement). However, closer examination showed that all the b-ion mass differences (observed − calculated) are negative, with an average \(\Delta = -11\) mDa, whereas the average \(\Delta\) for the y ions is only \(-0.1\) mDa (note that the latter figure is much less sensitive to residue assignment errors near the N terminus).

We then realized that another possibility lay in replacing Arg with Gly at residue 26, which gives an almost identical calculated mass change \(\Delta m = 99.080\) Da. This stimulated a more careful search of the MS/MS spectrum of Fig. 9a, resulting in the discovery of the small peaks shown in Fig. 9b at \(m/z = 230.118\) and 1605.853 (initially unassigned), corresponding to \(y_2\) and \(y_{15}\) in the breakup of the peptide with the \(R \rightarrow G\) substitution. Furthermore, this substitution yields the daughter ion masses shown in Table VI, where the b-ion mass differences (observed − calculated) are negative, with an average \(\Delta = 0.4\) mDa, a considerable improvement. Thus we conclude that residue 26 in BMV-W has undergone an \(R \rightarrow G\) mutation, corresponding to a single genetic code change of AGG to GGG in the nucleic acid sequence.

**Effect of Change of Propagation Host on BMV Coat Protein Sequences**

A change in virulence of a virus by passaging through a different host is a well established phenomenon (41). It is interesting to see if such passaging is accompanied by a change in the coat protein sequence, a possible explanation of the change in virulence.

In this case, the BMV-P isolate had been propagated in Little Club wheat in 1989 (28), after which the purified lyophilized virus was used for the ESI mass measurement listed in Table I. To provide material for additional analyses, BMV-P was propagated in 1999 from the same source used in 1989 (preserved in the Cereal Research Center collection), except that BW155 wheat was used, as Little Club was no longer readily available. This repropagated virus is denoted as BMV-P*.

Two distinct masses appeared in the deconvoluted ESI spectrum of BMV-P* (Fig. 10). The minor peak (at 20,265 Da) corresponded to the one previously observed for BMV-P, but the predominant peak had a mass indistinguishable from that previously determined for BMV-T (20,279 Da), suggesting a mutation of Val to Ile or Leu at position 167, as deduced above for BMV-T. This was confirmed by mass mapping and by MS/MS measurements (not shown).

A further change in the coat protein sequence of BMV-P was observed when the propagation host was changed from BW155 wheat (BMV-P*) to AC Assiniboia oat (yielding BMV-P*2). In

![FIG. 8. MALDI QqTOF mass spectrum of a tryptic digest of the BMV-W coat protein. The inset shows an expanded view of the abnormal peptide fragment ion with monoisotopic mass 1734.981 Da.](image)

![FIG. 9. CID mass spectra of the parent ion with \(m/z = 1734.98\) measured by MALDI MS/MS on the QqTOF instrument. a, overall spectrum; b, expanded views of the above spectrum near 230 and 1506 Da.](image)

**TABLE V**

| b ion* | \(m/z\) (obs.) | \(MH^+\) (calc.) | \(\Delta\) (mDa) | y ion* | \(m/z\) (obs.) | \(MH^+\) (calc.) | \(\Delta\) (mDa) |
|--------|----------------|-----------------|-----------------|--------|----------------|-----------------|-----------------|
| b1     | —              | 102.056         | —               | y1     | —              | 147.114         | —               |
| b2     | 173.091        | 173.093         | −2              | y2     | 204.133        | 204.135         | −2              |
| b3     | 329.180        | 329.194         | −14             | y3     | 332.191        | 332.194         | −3              |
| b4     | 457.246        | 457.252         | −6              | y4     | 389.219        | 389.215         | +4              |
| b5     | 653.361        | 554.305         | −12             | y5     | 460.249        | 460.252         | −3              |
| b6     | 766.441        | 653.373         | −17             | y6     | 531.283        | 531.289         | +4              |
| b7     | 865.511        | 766.458         | −15             | y7     | —              | 644.373         | —               |
| b8     | 994.566        | 994.569         | −3              | y8     | 741.421        | 741.426         | −5              |
| b9     | 1204.700       | 1091.621        | —               | y9     | 870.470        | 870.469         | +1              |
| b10    | 1346.764       | 1204.705        | −5              | y10    | 969.546        | 969.537         | +9              |
| b11    | 1346.764       | 1204.705        | —               | y11    | 1082.621       | 1082.621        | +8              |
| b12    | 1366.743       | 1275.743        | —               | y12    | —              | 1181.689        | —               |
| b13    | 1346.764       | 1204.705        | −5              | y13    | 1278.747       | 1278.742        | +5              |
| b14    | 1403.801       | 1403.801        | —               | y14    | 1406.821       | 1406.801        | +20             |
| b15    | 1531.843       | 1531.880        | −17             | y15    | 1562.885       | 1562.892        | −17             |
| b16    | 1588.857       | 1588.881        | −24             | y16    | 1633.926       | 1633.939        | −13             |
| b17    | 1716.973       | 1716.976        | −3              | y17    | 1734.977       | 1734.987        | −10             |

*Values listed are monoisotopic masses (Da).
same three peaks were observed, but there was a higher pro-
P*4), the purified virus no longer contained the protein of
inoculated on TR241 barley and propagated (generating BMV-
20,225 Da) whose mass did not correspond to that of any of the
previously measured for BMV-N, but there was a new peak (at
20,295 Da). The sequence of the 20,295 Da peak was
changed instead to Seneca maize (generating BMV-P* 3), the

addition to peaks corresponding to sequences identical to
BMV-P (20,265 Da) and BMV-T (20,279 Da), BMV-P*2 had a
new peak at 20,315 Da (Table VII). Analysis by tandem QqTOF
mass spectrometry showed that this new sequence had a mu-

caused by a substitution of Ala → Val at residue 12 (corre-
sponding to a single base change GCC to GUG). Second, the ion
at m/z = 1904.984 corresponding to residues 112–130 in
BMV-W (and BMV-T, see Table III) was present, but there was
an additional peak at m/z = 1934.993. MS/MS measurements
on this ion (not shown) indicated that it also corresponded to
residues 112–130, but with an Ala → Thr substitution at posi-
tion 122 (corresponding to the single base change GCA to ACA).
Thus the 30.009 Da mass difference observed in the 112–130
doublet (calculated separation 30.011 Da) is accounted for, as
well as the difference in the overall masses between BMV-N
and BMV-N*.

BMV-N appeared less responsive than BMV-P to further
changes in propagation host. After the initial set of changes
induced by the use of BW155 rather than Little Club wheat
(Table VII), which generated BMV-N*, no further changes were
observed when the propagation host was subsequently shifted
from BW155 wheat to AC Assiniboia oat, then from BW155
wheat to Seneca maize, and finally re-propagated in Seneca
maize.

**DISCUSSION**

**Mass Spectrometric Methods**—The measurements reported
here were carried out over a period of several years (15–20),
during which techniques and available instrumentation im-
proved greatly. In particular, the mass measurement accuracy
for both proteolytically generated fragments and daughter ions
improved by about two orders of magnitude from the original
observations in 1996 to the recent measurements with the
QqTOF instrument. This led to a marked increase in the level
of confidence in the sequence assignments, notably in distin-
guishing between Val deletion and Arg → Gly substitution in
BMV-W, as described above. In that case, the ability to obtain
a mass accuracy better than 10 mDa for the daughter ions
was clearly important.

In addition, it is useful to note that the peak that defined the
N-terminal acetylation for all the BMV samples was observed
here were carried out over a period of several years (15–20),
during which techniques and available instrumentation im-
proved greatly. In particular, the mass measurement accuracy
for both proteolytically generated fragments and daughter ions
improved by about two orders of magnitude from the original
observations in 1996 to the recent measurements with the
QqTOF instrument. This led to a marked increase in the level
of confidence in the sequence assignments, notably in distin-

**Significance of Polymorphism in BMV Coat Protein (CP)
Sequences**—Coat protein molecular weights inferred from the
sequence analysis of cDNA clones, while accurate, do not show
the complete profile of the viral coat protein population actu-
ally present in infected plant tissue. Comparisons of the CP
sequences of a set of distinct BMV isolates not only show

**TABLE VI**

Comparison of CID fragmentation of the singly charged ion (m/z 1734.98) from the BMV-W peptide (residues 24–41) with mass values calculated assuming a mutation of Arg to Gly at position 26, i.e. from TAGVQPVIVEPLAAGQGK

| b ion a | n/z (obs.) | MH+ (calc.) | Δ (mDa) | y ion b | m/z (obs.) | MH+ (calc.) | Δ (mDa) |
|---------|------------|-------------|---------|---------|------------|-------------|---------|
| b1      | 173.091    | 173.093     | -2      | y1      | 120.652    | 120.656     | +4      |
| b2      | 230.118    | 230.114     | +4      | y2      | 173.091    | 173.093     | -2      |
| b3      | 239.151    | 239.152     | -2      | y3      | 328.192    | 328.194     | -3      |
| b4      | 457.246    | 457.241     | +1      | y4      | 489.249    | 489.252     | +4      |
| b5      | 489.249    | 489.246     | -3      | y5      | 531.293    | 531.289     | +4      |
| b6      | 653.361    | 653.362     | -1      | y6      | 741.421    | 741.428     | -5      |
| b7      | 766.441    | 766.446     | -5      | y7      | 870.470    | 870.469     | +1      |
| b8      | 865.515    | 865.517     | -9      | y8      | 971.546    | 969.537     | +9      |
| b9      | 994.566    | 994.557     | +9      | y9      | 1082.629   | 1082.621    | +8      |
| b10     | 1204.700   | 1204.694    | +6      | y10     | 1278.747   | 1278.742    | +5      |
| b11     | 1275.731   | 1275.731    | -4      | y11     | 1406.821   | 1406.801    | +20     |
| b12     | 1346.764   | 1346.768    | -4      | y12     | 1505.853   | 1505.869    | -16     |
| b13     | 1403.790   | 1531.849    | -6      | y13     | 1562.885   | 1562.891    | -6      |
| b14     | 1588.857   | 1588.870    | -13     | y14     | 1633.926   | 1633.928    | -2      |
| b15     | 1716.973   | 1716.965    | +8      | y15     | 1734.977   | 1734.975    | +2      |

a Values listed are monoisotopic masses (Da).

b m/z = 1734.98 (corresponding to residues 24–41 in BMV-W) as described above. In that case, the ability to obtain a mass accuracy better than 10 mDa for the daughter ions was clearly important.

**Discussion**

Mass Spectrometric Methods—The measurements reported
here were carried out over a period of several years (15–20),
during which techniques and available instrumentation im-
proved greatly. In particular, the mass measurement accuracy
for both proteolytically generated fragments and daughter ions
improved by about two orders of magnitude from the original
observations in 1996 to the recent measurements with the
QqTOF instrument. This led to a marked increase in the level
of confidence in the sequence assignments, notably in distin-
...
TABLE VII
Coat protein average molecular masses (Da) of BMV specimens propagated in different hosts

| Sample     | Host change | MALDI | ESI  | Calc. mass | Intensity |
|------------|-------------|-------|------|------------|-----------|
| BMV-P      | wheat, 89   | 20265.3 | 20265.2 | 20265.4 | 100%      |
| BMV-P*     | wheat, 99   | 20278.2 | 20265.5 | (20265.4 | 7%        |
| BMV-Pa2    | wheat→oat   | 20265.4 | 20265.4 | 20265.4 | 22%       |
| BMV-Pb3    | wheat→corn  | 20265.5 | 20265.5 | 20265.5 | 20%       |
| BMV-Pc+4   | corn→barley | 20278.8 | 20279.4 | 20279.4 | 44%       |
| BMV-N      | wheat, 89   | 20191.4 | 20194.3 | 20194.3 | 100%      |
| BMV-Na*    | wheat, 99   | 20215.8 | 20194.3 | 20194.3 | 66%       |
| BMV-Nb2    | wheat→oat   | 20194.6 | 20194.3 | 20194.3 | 40%       |
| BMV-Nc3    | wheat→corn  | 20194.8 | 20194.3 | 20194.3 | 38%       |
| BMV-Nd4    | corn→corn   | 20225.4 | 20224.3 | 20224.3 | 62%       |
|            |             |        |       |            |           |

a Calculated molecular mass was obtained from the measured sequence taking account of the observed acetylation at the N-terminus, as also shown in Table I.

b The intensity was determined by integrating the area under the peaks.

Fig. 11. Measured amino acid sequences of the BMV coat protein. The top line shows the type amino acid sequence deduced from a cDNA clone (24–26). The mutations observed are shown; dashes indicate agreement with the type sequence. N-terminal acetylation is indicated by the symbol @. I at residue 167 in BMV-T indicates Ile or Leu, but most likely Ile (see “Discussion”). The modifications induced by passing through various hosts are not shown but are described in the text.

consistent discrete polymorphisms, but also identify a specific host-attenuation effect on the profile of the CP population of the BMV-N and BMV-P isolates. If CP sequences of defined BMV isolates remained invariant, we would expect the type isolate BMV-T (pv47) CP sequence to be identical with that reported in the literature (27, 29). Instead, the BMV-T (pv47) sequence differs from the reported one by a Trp change from GUC to AUC (transition), whereas substitution of Leu requires a change from GUC to CUC (i.e. G→C, a purine–pyrimidine transversion). In general, transitions are significantly more probable than transversions (44), consistent with examinations of viral coat proteins; 20 of 28 nucleic acid substitutions that we have observed in these compounds have been transitions. We conclude that the substituent at residue 167 in BMV-T is probably Ile.

Role of Mass Spectrometry in Plant Virus Research—Analysis by MALDI or ESI mass spectrometry of the viral coat protein is an approach that can distinguish virus isolates or closely related virus strains rapidly and definitively. Moreover, this method reveals directly any post-translational modifications, which of course could not be deduced from nucleic acid sequencing. Here, for example, the N-terminal residue in all the BMV isolates studied was modified by acetylation, a common phenomenon in ssRNA plant virus coat proteins, reported elsewhere for members of the tobamo-, potex-, and poty-virus groups. Structural information of this kind may contribute to understanding how the virus functions, and what factors account for variations in virulence among different isolates.

Moreover, reverse transcription of viral RNA and amplification of the reverse-transcribed cDNA will only yield the protein that is coded by the particular nucleic acid sequence selected. Mass spectrometry, by contrast, gives a reasonably accurate picture of the distribution of sequences in a mixed population of virus coat proteins, because the relatively small changes in sequence observed here are unlikely to produce much change in MS sensitivity. The data shown in Table VII and in Figs. 10 and 11 show that it is now possible to measure by TOFMS the change in profile of the population of coat proteins of a given virus isolate when the propagation host is changed. The changes we observe in the coat protein sequence provide one possible explanation for the change in virulence produced by passing through a different host.

Acknowledgments—We thank A. N. Krutchinsky, R. J. McNabb, V. Spicer, and M. Bromirski for skillful assistance in the TOF laboratory, and J. Ackerman and B. Gillis for excellent work in the propagation and purification of virus preparations.

REFERENCES
1. Gorman, J. J., Corino, G. L., and Shiell, B. (1990) J. Biomed. Environ. Mass Spectrom. 19, 646–654
2. Henderson, L. E., Bowers, M. A., Sewder II, R. C., Serabyn, S. A., Johnson, D. G., Bess, J. W., Jr., Arthur, L. O., Bryant, D. K., and Fenselau, C. (1992) J. Virol. 66, 1856–1865
3. Siuzdak, G., Bother, B., Teager, M., Brugidou, C., Fauquet, C. M., Hoey, K., and Chang, C. M. (1996) Chem. Biol. 3, 45–48
4. Despeyroux, D., Philipps, R., and Watts, P. (1996) Rapid Commun. Mass Spectrom. 10, 937–941
5. Jonscher, K. R., and Yates, J. R., III (1997) J. Biol. Chem. 272, 1735–1741
Gorman, J. J., Ferguson, B. L., Speelman, D., and Mills, J. (1997) *Protein Sci.* **6**, 1308–1315

Suizdak, G. (1998) *J. Mass Spectrom.* **33**, 203–211

Lopatnicki, S., Morrow, C. J., and Gorman, J. J. (1998) *J. Mass Spectrom.* **33**, 950–960

Booth, B., Dong, X. F., Bibbs, L., Johnson, J. E., and Siuzdak, G. (1998) *J. Biol. Chem.* **273**, 673–676

Jeyarajah, S., Parker, C. E., Sumner, M. T., and Tomer, K. B. (1998) *J. Am. Soc. Mass Spectrom.* **9**, 157–165

Bothner, B., Dong, X. F., Bibbs, L., Johnson, J. E., and Siuzdak, G. (1998) *J. Biol. Chem.* **273**, 673–676

Lewis, J. K., Bendahan, M., Smith, T. J., Beachy, R. N., and Siuzdak, G. (1998) *Proc. Natl. Acad. Sci. U. S. A.* **95**, 8596–8601

Lopaticki, S., Morrow, C. J., and Gorman, J. J. (1998) *J. Mass Spectrom.* **33**, 950–960

Loboda, A., Ens, W., She, Y. M., and Standing, K. G. (1999) *Rapid Commun. Mass Spectrom.* **13**, 508–518

She, Y. M., Haber, S., Seifers, D., Loboda, A., Bromirski, M., Ens, W., and Standing, K. G. (1999) *Rapid Commun. Mass Spectrom.* **13**, 508–518

Seifers, D. L., Salomon, R., Marie-Jeanne, V., Allist, B., Signoret, P., Haber, S., Loboda, A., Ens, W., She, Y. M., and Standing, K. G. (2000) *Phytopathology* **90**, 565–575

Seifers, D. L., Salomon, R., Marie-Jeanne, V., Allist, B., Signoret, P., Haber, S., Loboda, A., Ens, W., She, Y. M., and Standing, K. G. (2000) *Phytopathology* **90**, 565–575

Seifers, D. L., Salomon, R., Marie-Jeanne, V., Allist, B., Signoret, P., Haber, S., Loboda, A., Ens, W., She, Y. M., and Standing, K. G. (2000) *Phytopathology* **90**, 565–575

Seifers, D. L., Salomon, R., Marie-Jeanne, V., Allist, B., Signoret, P., Haber, S., Loboda, A., Ens, W., She, Y. M., and Standing, K. G. (2000) *Phytopathology* **90**, 565–575

Seifers, D. L., Salomon, R., Marie-Jeanne, V., Allist, B., Signoret, P., Haber, S., Loboda, A., Ens, W., She, Y. M., and Standing, K. G. (2000) *Phytopathology* **90**, 565–575

Seifers, D. L., Salomon, R., Marie-Jeanne, V., Allist, B., Signoret, P., Haber, S., Loboda, A., Ens, W., She, Y. M., and Standing, K. G. (2000) *Phytopathology* **90**, 565–575

Seifers, D. L., Salomon, R., Marie-Jeanne, V., Allist, B., Signoret, P., Haber, S., Loboda, A., Ens, W., She, Y. M., and Standing, K. G. (2000) *Phytopathology* **90**, 565–575

Seifers, D. L., Salomon, R., Marie-Jeanne, V., Allist, B., Signoret, P., Haber, S., Loboda, A., Ens, W., She, Y. M., and Standing, K. G. (2000) *Phytopathology* **90**, 565–575

Seifers, D. L., Salomon, R., Marie-Jeanne, V., Allist, B., Signoret, P., Haber, S., Loboda, A., Ens, W., She, Y. M., and Standing, K. G. (2000) *Phytopathology* **90**, 565–575

Seifers, D. L., Salomon, R., Marie-Jeanne, V., Allist, B., Signoret, P., Haber, S., Loboda, A., Ens, W., She, Y. M., and Standing, K. G. (2000) *Phytopathology* **90**, 565–575

Seifers, D. L., Salomon, R., Marie-Jeanne, V., Allist, B., Signoret, P., Haber, S., Loboda, A., Ens, W., She, Y. M., and Standing, K. G. (2000) *Phytopathology* **90**, 565–575

Seifers, D. L., Salomon, R., Marie-Jeanne, V., Allist, B., Signoret, P., Haber, S., Loboda, A., Ens, W., She, Y. M., and Standing, K. G. (2000) *Phytopathology* **90**, 565–575

Seifers, D. L., Salomon, R., Marie-Jeanne, V., Allist, B., Signoret, P., Haber, S., Loboda, A., Ens, W., She, Y. M., and Standing, K. G. (2000) *Phytopathology* **90**, 565–575

Seifers, D. L., Salomon, R., Marie-Jeanne, V., Allist, B., Signoret, P., Haber, S., Loboda, A., Ens, W., She, Y. M., and Standing, K. G. (2000) *Phytopathology* **90**, 565–575

Seifers, D. L., Salomon, R., Marie-Jeanne, V., Allist, B., Signoret, P., Haber, S., Loboda, A., Ens, W., She, Y. M., and Standing, K. G. (2000) *Phytopathology* **90**, 565–575

Seifers, D. L., Salomon, R., Marie-Jeanne, V., Allist, B., Signoret, P., Haber, S., Loboda, A., Ens, W., She, Y. M., and Standing, K. G. (2000) *Phytopathology* **90**, 565–575

Seifers, D. L., Salomon, R., Marie-Jeanne, V., Allist, B., Signoret, P., Haber, S., Loboda, A., Ens, W., She, Y. M., and Standing, K. G. (2000) *Phytopathology* **90**, 565–575

Seifers, D. L., Salomon, R., Marie-Jeanne, V., Allist, B., Signoret, P., Haber, S., Loboda, A., Ens, W., She, Y. M., and Standing, K. G. (2000) *Phytopathology* **90**, 565–575