# Enhanced Reliability of Avian Influenza Virus (AIV) and Newcastle Disease Virus (NDV) Identification Using Matrix-Assisted Laser Desorption/Ionization-Mass Spectrometry (MALDI-MS)

**ARTICLE** *in* ANALYTICAL CHEMISTRY · FEBRUARY 2011

Impact Factor: 5.64 · DOI: 10.1021/ac102846q · Source: PubMed



**7 AUTHORS**, INCLUDING:

Ho Bin Jang
University of Leuven
**27** PUBLICATIONS **223** CITATIONS
SEE PROFILE

Seong Bin Park
Mississippi State University
**35** PUBLICATIONS **251** CITATIONS
SEE PROFILE

Takashi Aoki
Waseda University
**283** PUBLICATIONS **6,257** CITATIONS
SEE PROFILE

Tae Sung Jung
Gyeongsang National University
**104** PUBLICATIONS **960** CITATIONS
SEE PROFILE

analytical chemistry

ARTICLE

pubs.acs.org/ac

# Enhanced Reliability of Avian Influenza Virus (AIV) and Newcastle Disease Virus (NDV) Identification Using Matrix-Assisted Laser Desorption/Ionization-Mass Spectrometry (MALDI-MS)

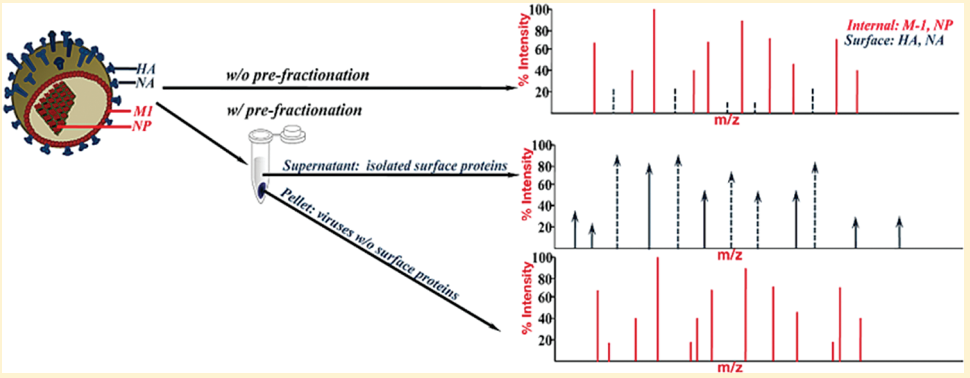
Ho Bin Jang,[†] Haan Woo Sung,[‡] Seong Won Nho,[†] Seong Bin Park,[†] In Seok Cha,[†] Takashi Aoki,[†] and Tae Sung Jung*[,†]

[†]Aquatic Biotechnology Center of WCU project, College of Veterinary Medicine, Gyeongsang National University, Jinju, 660-701, Republic of Korea

[‡]Laboratory of Avian Pathology, College of Veterinary Medicine, Kangwon National University, Chunchon 200-701, Republic of Korea

Supporting Information

**ABSTRACT:**




In-solution enzymatic and nonenzymatic digestion methods have been successfully implemented in matrix-assisted laser desorption/ionization-mass spectrometry (MALDI-MS)-based virus identification, extending to typing/subtyping of deadly influenza viruses. However, these methods are inefficient in obtaining more precise information on surface proteins of myxovirus particles, not only the hemagglutinin and neuraminidase of influenza virus but also the hemagglutinin-neuraminidase of Newcastle disease virus (NDV). Imbalances in viral protein composition cause ion suppression of tryptic fragments from low-abundant target proteins (surface proteins), adversely affecting reproducibility of mass spectra. Additionally, the coexistence of tryptic peptides from several proteins requires sophisticated statistical solutions for precise result interpretations. To circumvent these, we apply detergent-based (gel-free) partitioning of whole viruses into soluble surface proteins and insoluble virus materials, using differential centrifugation. MALDI-TOF or MALDI-TOF/TOF MS was applied to analyze tryptic peptides from separated viral proteins. In this study, we achieved type/subtype of avian influenza virus (AIV) within 5 h, based on 4 major proteins, by significantly reducing ion suppression and signal overlap from various protein sources. Hence, our approach can both yield dependable results and allow Web-based search engines to be directly employed, obviating the need for additional statistical strategy. Additionally, we demonstrate the utility of the method using NDV.

In recent decades, matrix-assisted laser desorption/ionization-mass spectrometry (MALDI-MS) has emerged as a powerful tool in virus identification because the time and accuracy requirements of clinical diagnosis are fulfilled.[1−4] Using this method, analysis is complete within seconds. Currently, advances in MS instrumentation and methodological techniques[1−3] permit direct virus identification using various methods.[5−8] Recent health threats such as pandemic flu outbreak caused by H1N1 have stimulated the development of confirmative MS-based proteomic tools for analysis of potentially deadly viruses such as the influenza virus.[9−13] As a part of these efforts over the past few years, various processing procedures to generate viral proteins and proteolytic peptides have been continuously suggested as alternative approaches to each other. These commonly consist of suspending the whole virus in a solution for detergent-based solubilization,[7] enzymatic digestion in organic solvent[6,10−12] and chemical digestion under microwave irradiation.[8]

**Received:** November 5, 2010
**Accepted:** January 11, 2011
**Published:** February 04, 2011

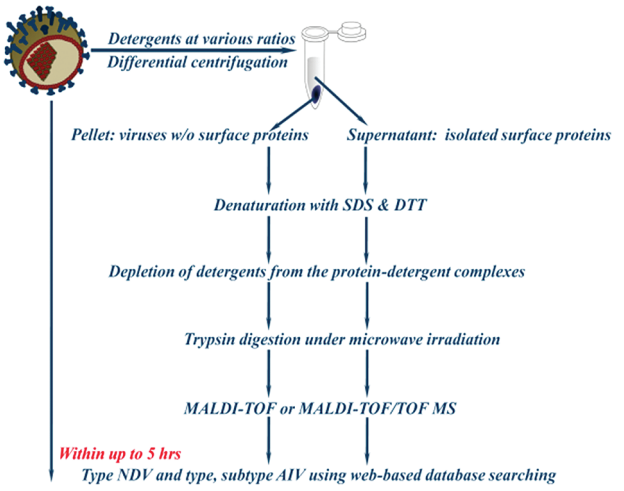



**Figure 1.** Schematic work-flow.

Currently, fingerprint matching of proteolytic peptides based on these enzymatic and nonenzymatic digestion (in-solution digestion) have been successfully used for MALDI-MS analysis, since it can provide a rapid (in a few minutes) analysis by minimizing requirements for sample handling and reduce the possibility of false identification by direct analysis of a number of viral protein extracts.[8]

However, such methods may be inadequate if used to detect viruses with relatively large number of proteins, unless additional statistical analysis is employed. The disruption of whole viruses such as Sindbis virus (SIN) and AIV may induce mixtures of structural and nonstructural proteins with complicated MALDI−MS spectra,[6,10−13] making viral identification using Web-based protein search engines[14−16] difficult. Consequently, new strategies using statistical algorithms combined with construction of sequence information on digested peptides must be provided for unequivocal virus identification.[6,10−13]

When in-solution digestion is applied to whole AIV, the relative disparities in the composition of AIV proteins, such as matrix protein 1 (M1), nucleoprotein (NP), hemeagglutinin (HA), and neuraminidase (NA),[17,39] can affect the numbers and intensities of MALDI-MS mass spectra[20] by virtue of an ion suppression effect.[18] Together with several undesirable features of sample preparation,[18] such imbalances in viral protein composition can also significantly reduce ion signals from enzymatically derived peptide fragments of HA, eventually leading to unexpected reproducibility of mass spectra of such peptides.[10,11] Problems in MS reproducibility can seriously affect the identification specificity.[19,38] In addition, given the fact that the influenza A virus is studded with glycoprotein spikes of HA and NA in a ratio of approximately 4:1 to 5:1,[17] detection of proteolytic fragments of NA might be affected by the presence of M1 and NP that are both well-ionized and present in high amounts. Hence, NA-derived proteolytic peptides, which are also of great importance in subtyping AIV,[17] might be barely detectable in MALDI mass spectra of whole viral digests.[20]

Hence, to unambiguously type and subtype AIV in MALDI-MS and alleviate a painstaking process of statistical analysis, we introduce a new strategy in which the whole virus is separated into its soluble surface proteins and insoluble cellular materials by applying prefraction step of detergent-based isolation followed by differential centrifugation (Figure 1).

In this manner, both major structural proteins of not only HA but also NA and nonstructural proteins of M1 and NP were comprehensively identified using Web-based searching programs, or in-house version of Mascot, which eventually enables us to type and subtype AIV in a limited time (maximally 5 h). For comparison, the results after in-solution digestion of the whole virus lysate, using detergent-based disruption without prefraction, are also shown. In addition, the wider applicability of the method was tested using Newcastle disease virus (NDV).

## EXPERIMENTAL SECTION

**Materials.** Acetonitrile (ACN), ammonium bicarbonate, dithiothreitol (DTT), iodoacetamide (IAA), Nonidet P-40 substitute (NP-40S) Tergitol solution, octyl-$\beta$-D-glucopyranoside (OG), sodium deoxysulfate (SDS), trifluoroacetic acid (TFA; 99%), urea, and α-cyano-4-hydroxycinnamic acid (CHCA) were purchased from Sigma-Aldrich (St. Louis, MO). Sequencing-grade modified trypsin was the product of Promega (Madison, WI, USA). The Bicinchoninic acid (BCA) protein assay reagent kit was from Thermo Scientific (Bonn, Germany) and used to determine the concentration of protein. Ultrapure 18.2-MΩ water from Millipore Milli-Q-system (Bedford, MA) was employed to make sample buffers. Allantoic fluid of 10 day-old embryonated chicken eggs containing Influenza A virus (A/Chicken/Korea/MS96/96(H9N2)) or Newcastle disease virus (NDV B1) were kindly provided by Kangwon National University, purified as previously described[40,41] and stored at −70 °C prior to use. Briefly, at 3 days post infection, the supernatant of each virus was harvested and clarified (2000 × *g*, 20 min, at 4 °C). The clarified supernatant was concentrated by ultracentrifugation (28 000 rpm, 90 min, at 4 °C in a SW 70 Ti rotor [Beckman Coulter, Fullerton, CA, USA]). The concentrated virus of NDV was purified through a 20−50% sucrose gradient (27 000 rpm, 90 min, at 4 °C in a SW 40Ti rotor) and for AIV, 30−60% sucrose gradient (26 000 rpm, 90 min, at 4 °C in a SW 40 Ti rotor). The virus bands were collected, diluted with TNE buffer (100 mM NaCl, 10 mM Tris-Cl (pH7.4), 1 mM EDTA), pelleted (28 000 rpm, 90 min, at 4 °C in a SW 70 Ti rotor), and resuspended in 2 mL of TNE buffer. Typical protein yields at 1.81 $\mu$g/$\mu$L of H9N2 and 1.79 $\mu$g/$\mu$L of NDV were obtained. Both viruses were carefully handled, as appropriate for "Biohazard Level Two" microorganisms.

**Selective Whole Virus Solubilization.** Different concentrations of nonionic detergent mixtures were prepared in ultrapure 18.2 MΩ water to a total concentration of 200, 150, 100, or 50 parts viral protein to 1 part OG (w/w), each mixture containing 800:1, viral proteins: NP-40S (w/w). The virus and detergent mixtures were prepared to a final volume of at least 60 $\mu$L. One part virus (1.81 $\mu$g/$\mu$L of H9N2 or 1.79 $\mu$g/$\mu$L of NDV) was mixed with 2 parts detergent mixture, held at room temperature for 30 min with occasional mixing, and centrifuged at 18 000 × *g* for 40 min at 4 °C in an Allegra 64R centrifuge (Beckman Coulter, Fullerton, CA, USA). The separation of structural and nonstructural proteins was confirmed by SDS-PAGE (12.5% separation gel) and gels were silver-stained.

Subsequently, supernatants were aspirated with a fine needle to ensure that pellets were not disturbed. Pellets containing insoluble protein aggregates were resuspended in TNE buffer (10 mM Tris, 150 mM NaCl, 1 mM EDTA, pH 7.4) equivalent to 2/3rd of the original volume of the mixture. Finally, SDT lysis buffer (2% [w/v] SDS, 100 mM Tris/HCl, 100 mM DTT

pH 7.6) equivalent to 1/3rd of the original volume was added into both the supernatants and the resuspended pellets prior to final incubation at 95 °C for 3 min to denature and release virus internal proteins. For comparison, the same amount of untreated H9N2 and NDV were disrupted in an equal volume of SDT lysis buffer (2% [w/v] SDS, 100 mM Tris/HCl, 100 mM DTT, pH 7.6), followed by denaturation under the same conditions for both fractions.

**Removal of the Detergents and in-Solution Digestion of Viral Proteins.** To deplete detergents from virus lysates, filter-aided sample preparation (FASP)[21] was employed, with slight modifications, as follows: Amicon Ultra YM-10 (Millipore) filter tubes containing 60 $\mu$L supernatant (corresponding to 24 $\mu$g of H9N2 or 25.8 $\mu$g of NDV-B1) or 60 $\mu$L of resuspended pellet (21 $\mu$g H9N2 or 12 $\mu$g NDV-B1) mixed with 400 $\mu$L of 8 M urea (100 mM Tris/HCl, pH8.5) were centrifuged twice at 14 000 × $g$ for 15 min at 25 °C. 200 $\mu$L of 50 mM IAA in 8 M urea/100 mM Tris/HCl (pH 8.5) was added and, after mixing briefly in a thermo-mixer, tubes were incubated for 10 min, and centrifuged. After three times centrifugation in 200 $\mu$L of 8 M urea and 400 $\mu$L of 50 mM ammonium bicarbonate for the supernatant and the pellet, respectively, the concentrated solute was recovered by placing the filter device upside down in a clean micro centrifuge tube and centrifuged at 1000 × $g$ for 2 min at 25 °C. In-solution digestion was achieved by adding trypsin (10 ng $\mu$L$^{-1}$ in 50 mM ammonium bicarbonate) at a ratio of 1:50 (enzyme to protein, w/w), and facilitated by application of microwave irradiation (700 W at 2450 MHz), for 150 s, in a domestic microwave oven (model RE-442B, Samsung, Suweon, Korea).[22] Particularly with the AIV pellet fraction, tryptic digests were obtained by centrifugation at 14 000 × $g$ for 20 min after digestion of the concentrate with trypsin at a ratio of 1:50 (enzyme to protein, w/w), under microwave irradiation for 150 s. The remaining digests were further collected by centrifugation in 50 $\mu$L amounts of 0.5 M NaCl. Reactions were terminated by addition of the same volume of 0.5% (v/v) TFA and, finally, tryptic digests were desalted using $C_{18}$ ZipTips (Millipore).

**Mass Spectrometry Analysis and Database Searching.** After desalting, peptide solutions from NDV supernatant and pellet fraction and AIV pellet fraction were sequentially analyzed in a Voyager-DE STR MALDI-TOF mass spectrometer (Applied Biosystems, Framingham, MA) equipped with a 337-nm nitrogen laser, in positive ion mode, with delayed extraction using reflectron mode. CHCA, as the MALDI matrix, was prepared by dissolving the matrix compounds in 70% ACN/ultrapure 18.2-MΩ water (v/v) containing 0.1% (v/v) TFA to a concentration of 50 mM using the "dried droplet method".[23]

The internal standard (IS) used was 1 pmol each of ACTH fragment 18−39 ($[M + H]^+$ of 2,465.1989) and Angiotensin II ($[M + H]^+$ of 1046.5423). Matrix/IS (4parts matrix +1part IS) was prepared immediately, mixed with the peptide mixture (2 part peptide mixture +1part matrix/IS), spotted on a target plate, and air-dried. Acquired data were processed using peak deisotoping, with the aid of Data Explorer 4.0 (Applied Biosystems).

Mass data on all digest fragments were delivered to the Mascot search program[15] and the UCSF Protein Prospector,[14] interrogating both the NCBInr and MSDS databases. The matrix signal (568.15 [3M + H]) was excluded.[35] The mass tolerance was kept at 20 ppm for all searches with no taxonomical restriction on species. Carbamidomethylation of cysteine residues and acetylation of protein N-termini were permitted. The Mascot scoring system was used to determine protein expectation values, which correspond to the number of matches that are expected to occur by chance alone.[15] Each protein was analyzed with the probability-based MOWSE threshold [$p < 0.05$] and the match with the lowest expectation value (the highest Mascot score) was considered as the best match.

In separate experiments, proteins were identified by mass spectrometry of the AIV supernatant fraction. The mass spectrometer used was an ABI 4800 Plus MALDI-TOF/TOF instrument (Applied Biosystems) equipped with a 200 Hz ND:YAG laser operating at 355 nm. CHCA, which formed the MALDI matrix, was prepared in the manner described above for MALDI-TOF. External calibration was performed using a mixture of peptides with known masses (Calibration Mixture 1 of the Peptide Mass Standard Kit for the 4700 Proteomics Analyzer, Applied Biosystems). MS spectra were searched against the NCBInr database by ProteinPilot v3.0 (with Mascot as the database search engine) with mass tolerances of 20 ppm and no taxonomical restriction on species. All mass-assigned ion signals matching HA and NA peptides within 20 ppm, ranging in size from 500 to 4000 Da, were manually selected for subsequent MS/MS analysis in the 2 kV mode, and 800−1000 consecutive laser shots were fired. Air was used as the collision gas. The MS/MS tolerance was set to 50 ppm with a taxonomical restriction on viruses. Carbamidomethylation of cysteine residues and acetylation of protein N-termini were permitted. The minimal acceptance criterion was set to >95% confidence that the identified protein or highly homologous protein was found (typically, Mascot probability-based score >70).

**Database Construction and Query.** The information on amino acid sequences of HN and NC of NDV B1, and HA, NA, NP, and M-1 of AIV (A/Chicken/Korea/MS96/96(H9N2)) were obtained from the NCBInr protein database, and each in silico digested tryptic peptide was generated using MS-digested software (http://prospector.ucsf.edu/prospector/cgi-bin/msform.cgi?form=msdigest) with the following parameters: database of NCBInr 2010. 3. 30, trypsin digest with maximum missed cleavage of 1, constant/variable modification of carbamidomethyl (C)/acetyl (N), and peptide mass range of 500−4000 Da. Experimentally observed masses in whole virus lysates unidentified by Mascot search were further analyzed by comparing with respective hypothetical peptide masses using Mascot search.[15] A match was counted if a selected peptide from the whole virus lysate corresponded within 20 ppm to an identical protein in Mascot.

## RESULTS AND DISCUSSION

**Selective Whole Virus Solubilization.** When the first attempt to separate hemagglutinating component from influenza virus was conducted in 1950, various methods that are still in vogue were developed for study of myxoviruses.[24,25] In most such approaches, ionic/nonionic detergent-based solubilization of surface proteins was employed to characterize biological and immunological activity of various viral components.[24−26] Recently, a detergent mixture was used to characterize the AIV envelope and its association with surface proteins.[27]

In the present study, which sought to improve the dynamic range of detection of tryptic fragments derived from viral proteins by MALDI-MS, by minimizing ion suppression effect of complex mixture, we applied selective solubilization of whole viruses to study not only AIV but also NDV.

To determine the appropriate ratios of detergents required to effectively remove surface proteins from whole NDV, we performed SDS-PAGE on soluble (supernatants) and insoluble fractions (pellets). The isolation efficiencies of surface proteins from whole NDV using an OG/NP40S mixture was found to depend upon the amount of OG/NP40S employed. At a ratio of 50:1 (OG) and 800:1 (NP40S), most NDV surface proteins, including HN and F, were in the supernatant, whereas the pellet contained relatively high amounts of nonstructural NC and M proteins (Supporting Information Figure S-1).

When samples of AIV solubilized in OG/NP40S were analyzed by SDS-PAGE, release of HA and NA proteins from the pellet fractions was apparent when the amount of OG increased. NP was reliably present only in the insoluble fractions after use of various nonionic detergents, whereas M1 was present in both fractions, being relatively more abundant in the insoluble fraction (Figure S-2). This is in line with the suggested structural association of M1 with surface proteins of HA and NA.[28]

Thus, partitioning of whole virus into soluble and insoluble fractions can increase the sensitivity of subsequent MALDI-MS analysis, which may be compromised if complicated viral protein complexes are present.

**Removal of the Detergents from Viral Protein-Detergent Complex.** When filter-aided sample preparation (FASP) was used to deplete detergent, most hydrophobic proteins could not be recovered. Binding of hydrophobic proteins to filtration membrane is a serious problem encountered during purification/buffer exchange step.[29,30] Indeed, despite recent modifications to the Amicon Ultra device that are designed to enhance protein recovery,[31] we experienced severe loss of hydrophobic proteins of AIV pellet fraction, attributable to adsorption. Thus, to maximize sample recovery, we employed two different elution methods for the supernatant and pellet fractions, as described in Experimental Procedures.

**Mass Spectrometry.** *Detection of NDV Proteins with MALDI-TOF.* Since first identified in 1926, Newcastle disease virus (NDV), the infectious agent of Newcastle disease, has caused great economic losses in the poultry industry.[32,33] Among a total of six NDV structural and nonstructural proteins, the nucleocapsid protein (NC) is the most abundant, comprising approximately 45% of total viral protein.[32] The hemagglutinin-neuraminidase (HN) of an NDV surface protein, along with the fusion (F) protein, has been known as a key factor in determining virulence,[33,34] which account for 20% of NDV proteins.[32]

The influence of NDV protein structural characteristic on MALDI-mass spectra is well demonstrated in Figure 2. Interestingly, similar mass spectra were obtained by direct analysis of whole NDV lysates (A) and of an insoluble fraction containing nonstructural proteins (B). However, the soluble NDV fraction (C) exhibited significant increases in both signal intensities and the number of distinguishable peaks associated with HN-derived tryptic peptides, compared to what was seen in A and B.

HN (gi|11545724) was unambiguously identified, after prefractionation, with an increased number of matched peptide (20 peptides), sequence coverage (35%), and an expectation value of $2.4 \times 10^{-12}$ (Table 1). Compared to the mass spectra of whole virus lysates, of 20 peptides, 15 reproducible peptide peaks with high signal intensities were surprisingly observed only in soluble fraction with an average mass error of 5.9 ppm. Relative ion intensities of the 5 additional spectra were also increased in intensity ratio from 4.48 to 10.72 ($I_S/I_W$; $I_S$ indicates intensity of HN in soluble fraction (C) and $I_W$, intensity of HN in whole virus lyates (A)), and S/N ratios also increased by a factor of from 6.14 to 21.65. In addition, Table 1 shows that the observed mass spectra of NC peptides in insoluble fraction increased from 13 to 18 and matched hypothetical NC (gi|14190063) tryptic peptides of NDV B1 with 47% sequence coverage and statistical significance above 95% identity threshold (Table 1).

This result strongly supports our assumption that MALDI signals are suppressed in a complex viral protein mixture. Hence, given the difficulty to obtain sufficient HN tryptic fragments by directly applying in-solution digestion, our approach more effectively yields conclusive information on not only the identification but the virulence of NDV.

**Detection of AIV Proteins with MALDI-TOF or MALDI-TOF/TOF.** As shown in the experiment using NDV B1, direct analysis of whole virus lysates with MALDI-MS might be impractical when information on low-abundance target proteins critical for virulence or subtype determination is sought. The low relative intensity of a peptide at $m/z$ 2,096.1108 leads to an unexpected presence of a tryptic peptide in the repeat experiments (upper insets of Figure 3A). This can give an inconsistent reproducibility in detecting residues 238−254 with sequence R.IDYYWSVLK-PGQTLRVR.S of HA (gi|269826347). In addition, the intricate mass spectra, shown in Figure 3A, arising from the presence of many other peptides in whole AIV lysates could also lead to serious consistency in AIV identification using MALDI-MS. For example, overlapping mass peptide ions in the region, $m/z$ value of 2,032.2272 to 2,037.0752 attributable to the coexistence of $[M + H^+]$ 2,029.9430 associated with residues 423−441 of NP (gi|7861789) and $[M + H^+]$ 2,031.9413 corresponding to resdues 78−95 of M-1 (gi|7861792) (lower insets of Figure 3A), cause a lack of recognition of monoisotopic mass. The identification of AIV was compounded by inaccurate identity of ion signals within poorly resolved ion cluster. Four different subtype of Influenza A virus (A/duck/Yangzhou/02/2005(H8N4)), (A/swine/Bakum/5/95(H1N1)), (A/Chicken/Taiwan/2383 V/00(H6N1)), and (A/aquatic bird/Korea/w216/2007(H5N2)) corresponding to each M-1 and NP, were identified in repeat experiments (Table 2A). Hence, to correctly determine type/subtype of AIV in whole virus lysates, it is important to clearly detect a large number of peptides from respective AIV protein with high quality in terms of signal intensity, S/N ratio and resolution, which can exploit the increased sequence coverage as well as mass spectra reproducibility thus afforded.

Table 2 shows the number of matched tryptic peptides seen in unseparated AIV and in fractionated samples, indicating that the total number of HA- and NA-derived tryptic peptides in supernatant fraction is much higher than that in unfractionated AIV. HA digests of whole virus lysates (yielding 11 tryptic fragments) contained approximately 2-fold fewer peptide fragments that seen in supernatant fraction (18 fragments). More interestingly, the number of tryptic fragments from NA increased from 7 in whole virus lysates to 11 in supernatant fraction. Considering that detergent-based solubilization of whole AIV lysates, used here as a control procedure, was more effective than was enzymatic digestion with organic solvent,[21] our method offers a considerable improvement in generation of reproducible HA and NA-associated mass spectra. In addition, the reproducibility of MALDI-mass spectra is generally regarded as a crucial factor in both protein profiling[19] and reliable microbial identification[38] using MALDI-MS. With the addition of a prefractionation step, the quality of MALDI mass spectra relevant to HA and NA was also significantly more reproducible when compared to those of whole AIV lysates (Figure 3A and C). Although no definitive criteria on the appropriate level of sequence coverage for correct

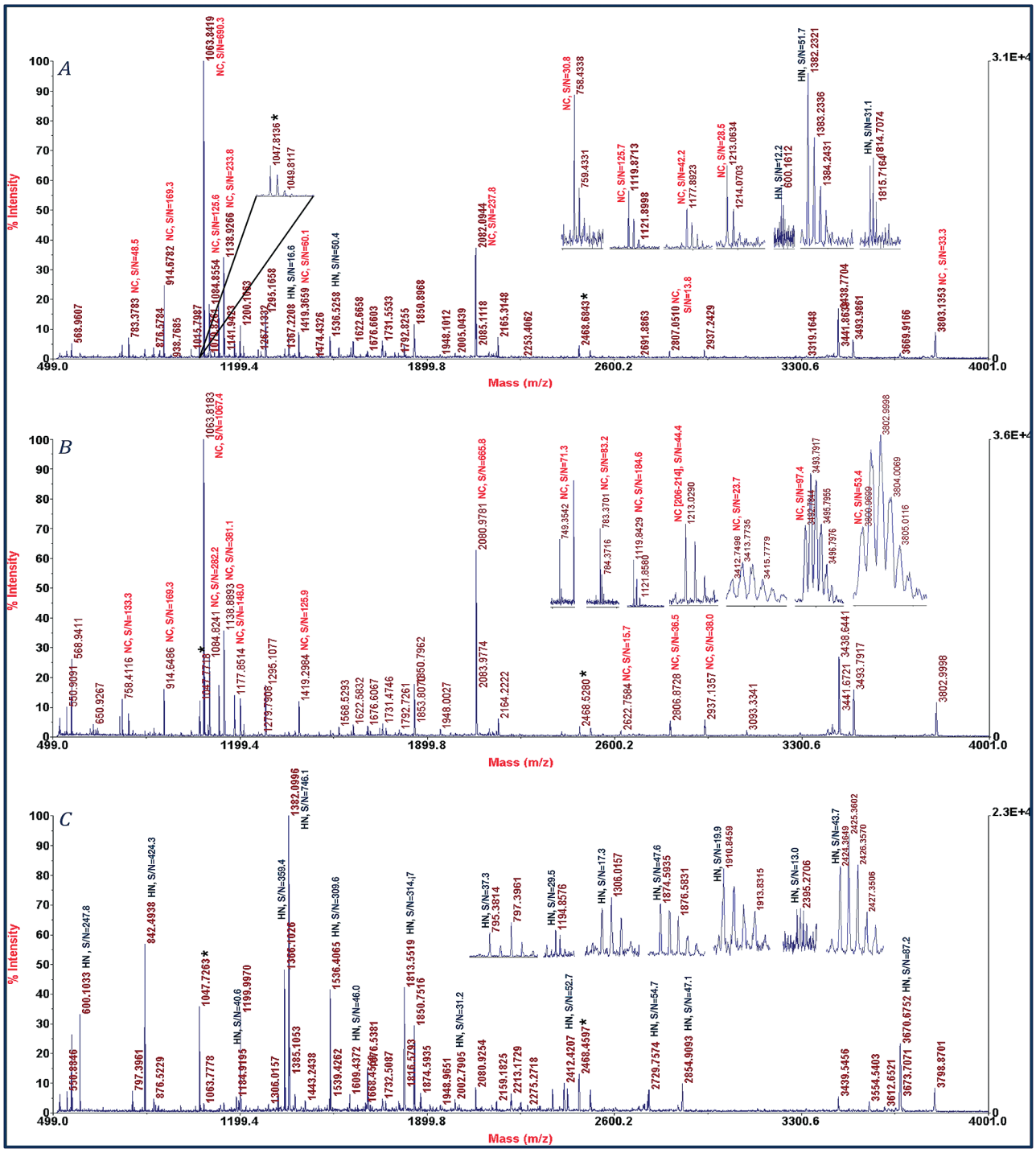



**Figure 2.** MALDI-TOF mass spectra of tryptic peptides from NDV obtained without prefractionation (A) and with prefractionation (B, C). (A) A typical spectrum of tryptic digest of the whole NDV lysate; (B) mass spectra of insoluble NDV material; (C) soluble surface proteins. Peaks are labeled with signal-to-noise ratio (S/N). The S/N of relevant peaks is calculated by root-mean-square (rms) calculation method. The observed internal calibrants are represented with asterisks (*).

identification of biological entities have appeared,[36] an increase in the number and reproducibility of tryptic fragments from relevant surface proteins both yields dependable results and allows Web-based search engines to be directly employed, obviating the need for any additional statistical strategy.

Both Mascot[15] and Protein Prospector[14] protein search engines correctly identified the mixture of hemagglutinin [Influenza A virus (A/Chicken/Korea/MS96-CE6/1996-(H9N2))] and neuraminidase subtype 2 [Influenza A virus (A/Chicken/Korea/MS96/96(H9N2))] with Mascot score of 184 and 100, and MOWSE score of 2.2e+7 and 14,446 (data not shown), respectively. Of which sequence coverages were 33% in HA (gi|269826347) and 25%, NA (gi|7861787) (Table 2C).


dx.doi.org/10.1021/ac102846q |*Anal. Chem.* 2011, 83, 1717–1725

**Table 1. Masses of the Matched Tryptic Peptides Observed in MALDI Mass Spectra without and with Prefraction of NDV Bl[a]**

| *m/z* (monoH M + HT1 observed | | | mass error(ppm) | | relative intensity(%) | | *m/z* (monoH M + HT1 observed | | | mass error(ppm) | | relative intensity(%) | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| w/o prefraction | w/ prefraction | amino acid residue | | | | | w/o prefraction | w/ prefraction | amino acid residue | | | | |
| HN | HN | | | | | | NC | NC | | | | | |
| M.c.:1 78 s.c: 10% | M.c.:l87 s.c.: 35% | | (a) | (b) | (a) | (b) | M.c.:1 78 s.c: 28% | M.c.:l87 s.c.: 47% | | (a) | (b) | (a) | (b) |
| 599.3045 | 599.3045 | $^{270}$MVHCR$^{274}$ | −6 | −6 | 2.12 | 12.73 | | 748.3820 | $^{453}$AVANSMR$^{459}$ | | 7 | | 2.71 |
| | 794.4091 | 357SSYKPCR363 | | −8 | | 3.31 | 757.42S6 | 757.4313 | $^{391}$GLAAAAQR$^{398}$ | −8 | 0 | 3.95 | 4.88 |
| | 841.4810 | 334YVIYKR339 | | −14 | | 33.88 | 782.3467 | 782.3676 | $^{197}$YMQQGR$^{202}$ | −19 | 8 | 6.17 | 3.62 |
| | 1183.6247 | $^{357}$SSYK-PGRGGGK$^{284}$ | | 2 | | 5.9 | 913.5259 | 913.5317 | $^{390}$RGLAAAAQR$^{398}$ | −7 | −1 | 19.65 | 6.75 |
| | 1193.5548 | $^{275}$LGFDGQYHEK$^{236}$ | | −3 | | 3.15 | 1062.5565 | 1062.5754 | $^{128}$FAMIAGSLPR$^{214}$ | −19 | .-1 | 100 | 63.31 |
| | 1303.6577 | $^{226}$SINLDDTQNRK$^{236}$ | | −2 | | 2.98 | 1083.5515 | 1083.5638 | 207YILYPVCR214 | −13 | −2 | 20.19 | 14.96 |
| 1364.7030 | 1364.6649 | $^{449}$LNPASAVFDSTSR$^{511}$ | 17 | −11 | 4.19 | 41.34 | 1118.5369 | 1118.5537 | SGVSEERAQR$^{127}$ | −16 | −1 | 20.22 | 8.88 |
| 1380.6730 | 1380.6494 | $^{486}$GVFGTMLDGEQAR$^{498}$ | 11 | −6 | 9.32 | 100 | 1137.5756 | 1137.5845 | $^{197}$YMQQGRVQK$^{205}$ | −7 | 1 | 43.71 | 20.85 |
| 1534.8390 | 1534.8375 | $^{404}$ILTVGTSHFLYQR$^{416}$ | 1 | −0 | 9.16 | 43.1 | 1176.5075 | 1176.5147 | $^{353}$DFMSTSSFWR$^{361}$ | −6 | 0 | 9.09 | 7.9 |
| | 1607.8127 | $^{499}$LNPASAVFDSTSR$^{513}$ | | −1 | | 6.67 | 1211.6484 | 1211.6637 | $^{206}$KYILYPVCR$^{214}$ | −10 | 3 | 3.93 | 2.5 |
| 1811.8040 | 1811.7745 | $^{254}$ATETEEEDYNSAVPTR$^{269}$ | 6 | −10 | 9.22 | 52.18 | 1417.7761 | 1417.7686 | $^{125}$AQRFAMIAGSLPR$^{137}$ | 2 | −3 | 13.03 | 8.9 |
| | 1872.7709 | $^{237}$SCSVSATPLGCDMLCSK$^{253}$ | | −13 | | 8.95 | 2078.9685 | 2078.9465 | $^{434}$SQGQPEAGDGETQFLDLMR$^{452}$ | 11 | 1 | 87.33 | 100 |
| | 1908.9966 | $^{70}$ITSTLGSNQDVVDRIYK$^{86}$ | | −3 | | 3.75 | 2803.3357 | 2619.2930 | $^{94}$QNEATLAVLEIDGFANGTPQFNNR$^{117}$ | | 5 | | 3.18 |
| | 2000.8737 | $^{236}$KSCSVSATPLGCDMLCSK$^{253}$ | | −9 | | 7.27 | | 2803.3241 | $^{2}$SSVFDEYEQLLAAQTRPNGAHGGGEK$^{27}$ | 3 | −1 | 5.94 | 12.31 |
| | 2392.0644 | $^{254}$ATETEEEDYNSAVPTRMVHGR$^{274}$ | | −8 | | 3.2 | | 2933.5120 | $^{362}$LGVEYAQAQGSSINEDMAAELKLTPAAR$^{389}$ | | 15 | | 12.37 |
| | 2409.1445 | $^{312}$VWFSVYGGLKPNTPSDTVQEGK$^{333}$ | | 3 | | 14.83 | | 3409.7783 | $^{86}$NHVALAGKQNEATLAVLEIDGFAN-GTPQFNNR$^{117}$ | | 16 | | 11.56 |
| | 2422.1445 | $^{461}$CPNSCVTGVYTDPYPLIFYR$^{480}$ | | 10 | | 13.22 | | 3489.7584 | $^{399}$VSEVTSSIDMPTQQVGVLTGLSEGGS-QALQGGSNR$^{433}$ | | 13 | | 46.61 |
| | 2726.3169 | $^{486}$GVFGTMLDGEQARLNPASAVFDSTSR$^{511}$ | | −1 | | 14.89 | 3797.7063 | 3797.7353 | $^{453}$AVANSMREAPNSAQGTPQSGPPPTP-GPSQDNDTDWGY$^{489}$ | 9 | 16 | 18.04 | 30.28 |
| | 2851.3971 | $^{435}$TATLHSPYTFNAFTRPGSIPCQASAR$^{460}$ | | 1 | | 18.11 | | | | | | | |
| | 3665.5961 | $^{237}$SCSVSATPLGCDMLCSKATETE-EEDYNSAVPTR$^{269}$ | | 7 | | 70.86 | | | | | | | |

[a] Abbreviation: HN, hemaggluiinin-nenraminidase; NC, nucleocapstd protein; M.c., Mascot score, which is given as $-10 \times \log_{10}(P)$, where $P$ is the probability that the observation represents a random event, typically greater than 70 ($p < 0.05$) indicates a statistically significant match; s.c, sequence coverage of the protein from matched tryptic digest fragments; a, values obtained without (w/o) prefractionation; b, values obtained with (w/) prefractionation.



dx.doi.org/10.1021/ac102846q |Anal. Chem. 2011, 83, 1717–1725

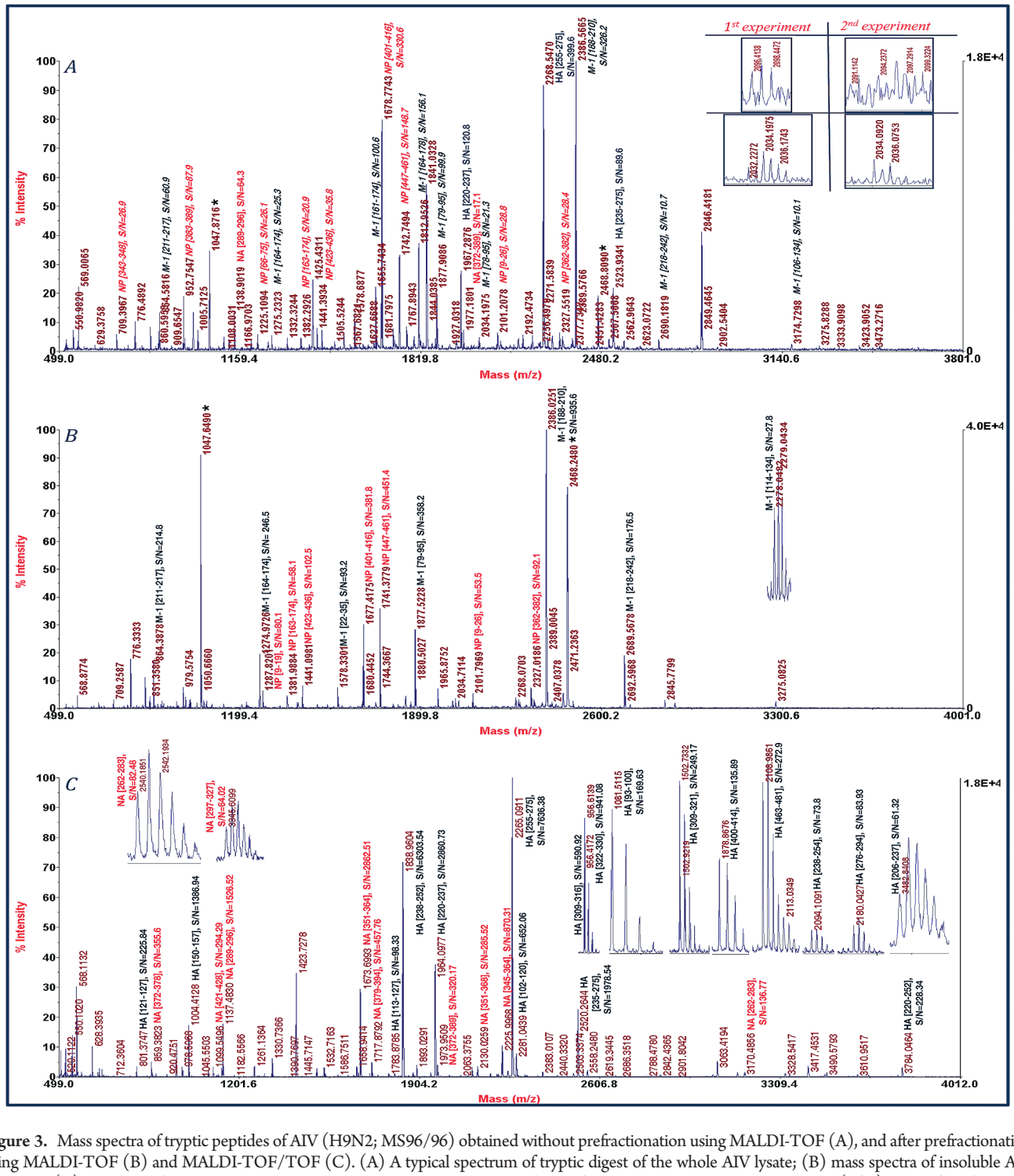



**Figure 3.** Mass spectra of tryptic peptides of AIV (H9N2; MS96/96) obtained without prefractionation using MALDI-TOF (A), and after prefractionation using MALDI-TOF (B) and MALDI-TOF/TOF (C). (A) A typical spectrum of tryptic digest of the whole AIV lysate; (B) mass spectra of insoluble AIV materials; (C) soluble surface proteins. Peaks are labeled with residue number of identified proteins and signal-to-noise ratio (S/N). The S/N of relevant peaks is calculated by root-mean-square (rms) calculation method. The observed internal calibrants are represented with asterisks (*).

For higher-throughput analysis, the identities of proteins in supernatant fraction were also validated by employing MALDI-TOF/TOF-MS analysis in which observed peptides matched to each HA and NA within 20 ppm, as precursor peptides, were manually selected to yield their fragment ions (Supporting Information Figure S-3 and S-4). Such supplementary MALDI-TOF/TOF-MS analysis, switching from MS to MS/MS, can provide more information to confirm AIV subtype.

Additionally, to investigate whether analysis of only one of two surface proteins could achieve complete subtyping of AIV, the maximum number of search results for MALDI-TOF/TOF analysis was set to 50. Of these, Mascot yielded 39 identical subtypes of H9N2 in HA analysis and 28 in NA. Interestingly, however, all the others identified different AIV subtypes (Supporting Information Figure S-5 and S-6), indicating that complementary analysis of the two surface proteins is imperative


dx.doi.org/10.1021/ac102846q |*Anal. Chem.* 2011, 83, 1717–1725

**Table 2. Mascot Search Results of the Whole AIV Lysate (A), the Pellet Fraction (B), and the Al V Supernatant Fraction (C)**

| search score rank | protein from species (accession number) | match no.[a] (s.c.(%))[b] | score[c] (e.v.)[d] | search score rank | protein from species (accession number) | match no.[a] (s.c.(%))[b] | score[c] (e.v.)[d] |
|---|---|---|---|---|---|---|---|
| A-lst | mixture 1 total store, 123; matches, 25; expect, $5.8 \times 10^{-6}$ | | | B | mixture 2 total score, 141; matches, 15; expect, $7.1 \times 10^{-9}$ | | |
| 1 | matrix protein L (gi\|117163694) | 11 (52%) | 110 (0.00012) | 1 | matrix protein L (gi\|87247013) | 7 (46%) | 98 (0.0001) |
| | (A/duck/YonEzhou/02/20(H8N4)) | 14 (33%) | 87 (0.022) | | (A/duck/Taiwan/WB29/99(H6N1)) | 8 (25%) | 96 (0.00018) |
| | nucleocapsid protein (gi\|149704995) | | | | nucleocapsid (gi\|5732293) | | |
| | (A/chicken/Taiwan/2838 V/00(H6NI)) | | | | (A/chicken/Korea/25232-006/96(H9N2) | | |
| | mixture 2 total score, 123; matches, 25; expect, $5.8 \times 10^{-6}$ | | | | mixture 2 total score, 140; matches, 15; expect, $7.1 \times 10^{-9}$ | | |
| 2 | matrix protein 1 (gi\|117163694) | 11 (52%) | 110 (0.00012) | 2 | matrix protein L (gi\|58429267) | 7 (44%) | 97 (0.00014) |
| | (A/duck/YonEzhou/02/20(H8N4)) | 14 (31%) | 87 (0.024) | | (A/chicken/Korea/S15/03(H6N1)) | 8 (25%) | 96 (0.00018) |
| | nucleocapsid protein (gi\|291491041) | | | | nucleocapsid (gi\|5732293) | | |
| | (A/aqiialic bird/Korea/w216/2007)(H5N2) | | | | (A/chicken/Korea/25232-006/96(H9N2) | | |
| | A-2nd | | | C | mixture 1 total score, 224; matches, 29; expect, $4.6 \times 10^{-16}$ | | |
| 1 | matrix protein 1 (gi\|70905123) | 10 (50%) | 110 (0 00012) | 1 | hemagglutinin (gi\|269826347) | 18 (33%) | 184 ($4.8 \times 10^{-12}$) |
| | (A/swine/Bakum/5/95(HlNl)) | | | | (A/chkken/Korea/MS96-CE6/l996(H9N2) | | |
| | matrix protein 1 (gi\|8I936940) | 10 (40%) | 103 (0.0006) | | neuraminidase subtype 2 (gi\|7861787) | 11 (25%) | 100 (0.0012) |
| | {A/swine/NetherlancV 12/1985(H1N1)) | | | | (A/chicken/Korea/MS96/96(H9N2)) | | |
| | | | | 2 | hcmagglutinin precursor (gi\|600700S) | 16 (38%) | 178 ($1.9 \times 10^{-11}$) |
| | | | | | (A/chicken/Korea/38349-p96323/96(H9N2)) | | |
| | matrix protein 1 (gi\|77632744) | 10 (40%) | 103 (0.0006) | | hemaggiulinin (gi\|158429239) | 15 (29%) | 139 ($1.5 \times 10^{-7}$) |
| | (A/swine/Bakum/5/95(H IN 1)) | | | | (A/chicken/Korea/S18/03(H9N2)) | | |
| | matrix protein 1 (gi\|225733806) (A/swine/Belgium/WVL5/i989(l[lNl)) | 10 (40%) | 103 (0.0006) | | | | |

[a] Numbers denote matching tryptic digest fragments in the peptidc mass fingerprinting. [b] Amino acid sequence coverage (s.c.) of the protein from matched tryptic digest fragments. [c] Mascot score is given as $-10 \times \log_{10}(P)$, where P is the probability that the observed match is a random event, typically greater than 70 ($p < 0.05$) indicates a statistically significant match. [d] Expected value (e.v.) is the number of matches with equal or better scores that are expected to occur by chance; the peptide mass fingerprinting was carried out using Mascot search program, interrogating the NCBInr database with mass tolerance of 20 ppm and no taxonornical restriction on species.



dx.doi.org/10.1021/ac102846q |Anal. Chem. 2011, 83, 1717–1725

to precisely subtype AIV. Although MALDI-TOF/TOF analysis can thus unambiguously subtype AIV, inconsistencies in AIV lineage determination may arise from the antigenic variation that develops during viral passage in allantoic fluid.[37]

Turning to the pellet fraction, Table 2B shows a mixture of M-1 and NP of the same AIV type were reliably identified. As expected, smaller numbers of tryptic fragments were identified because of absorption of hydrophobic proteins in the Amicon-nUltra tube. However, given their high cross-strain sequence similarity,[39] such data could be also useful in typing AIV.

We have demonstrated that isolation of the surface soluble proteins enables us to obtain more precise information on subtype of AIV. Using this approach, Web-based protein search engines can be directly applied, thereby allowing for the timely identification of AIV based on four major proteins. This can be a more accessible approach, compared to those of recent studies in which whole virus digest and statistical algorithm combined with MALDI FT-ICR was used for typing and subtyping influenza virus.[10−13]

## CONCLUSIONS

Identification of AIV and NDV using MALDI-MS can be greatly improved by simple application of a prefractionation step utilizing detergent-based isolation combined with differential centrifugation. This allows unambiguous information on virulence or subtype to be obtained, by facilitating the detection of more peptide fragments. With AIV in particular, typing/subtyping can be completely achieved, based on analysis of four major proteins, without the need for intricate statistical procedures or expensive technologies such as Fourier transform MS. Hence, our approach offers a new opportunity to improve MS-based surveillance strategies for AIV detection.

## ASSOCIATED CONTENT

**Supporting Information.** Additional material as described in the text. This material is available free of charge via the Internet at http://pubs.acs.org.

## AUTHOR INFORMATION

### Corresponding Author

*E-mail: jungts@gnu.ac.kr. Tel.: 82-55-751-5822. Fax: 82-55-762-6733.

## ACKNOWLEDGMENT

We thank Dr. Dong-won Bae and Dr. Myeong-sang Kwon for expert assistance in MALDI-TOF/TOF analysis. This work was supported by National Research Foundation of Korea (NRF) Program Grant 2008-0061829 and by World Class University (WCU) Project Grant R32-10253.

## REFERENCES

(1) Siuzdak, G. *J. Mass Spectrom.* **1998**, *33*, 203–211.
(2) Trauger, S.; Junker, T.; Siuzdak, G. *Top. Curr. Chem.* **2003**, *225*, 265–282.
(3) Maxwell, K. L.; Frappier, L. *Microbiol. Mol. Biol. Rev.* **2007**, *71*, 398–411.
(4) Fuerstenau, S. D.; Benner, W. H.; Thomas, J. J.; Brugidou, C.; Bothner, B.; Siuzdak, G. *Angew. Chem., Int. Ed. Engl.* **2001**, *40*, 541–544.
(5) Thomas, J. J.; Falk, B.; Fenselau, C.; Jackman, J.; Ezzell, J. *Anal. Chem.* **1998**, *70*, 3863–3867.
(6) Yao, Z. P.; Demirev, P. A.; Fenselau, C. *Anal. Chem.* **2002**, *74*, 2529–2534.
(7) Kim, Y. J.; Freas, A.; Fenselau, C. *Anal. Chem.* **2001**, *73*, 1544–1548.
(8) Swatkoski, S.; Russell, S.; Edwards, N.; Fenselau, C. *Anal. Chem.* **2007**, *79*, 654–658.
(9) Downard, K. M.; Morrissey, B.; Schwahn, A. B. *Mass Spectrom. Rev.* **2009**, *28*, 35–49.
(10) Schwahn, A. B.; Wong, J. W.; Downard, K. M. *Anal. Chem.* **2009**, *81*, 3500–3506.
(11) Schwahn, A. B.; Wong, J. W.; Downard, K. M. *Analyst* **2009**, *134*, 2253–2261.
(12) Schwahn, A. B.; Wong, J. W.; Downard, K. M. *J. Virol. Methods* **2010**, *165*, 178–185.
(13) Schwahn, A. B.; Wong, J. W.; Downard, K. M. *Anal. Chem.* **2010**, *82*, 4584–4590.
(14) Protein Prospector. http://prospector.ucsf.edu/prospector/cgi-bin/msform.cgi?form=msfitstandard. (Accessed October 11, 2010).
(15) Mascot. http://matrixscience.com. (Accessed October 11, 2010).
(16) PeptIdent. http://ca.expasy.org/tools/peptident.html.
(17) Subbarao, K.; Joseph, T. *Nat. Rev. Immunol.* **2007**, *7*, 267–278.
(18) Annesley, T. M. *Clin. Chem.* **2003**, *49*, 1041–1044.
(19) Koomen, J. M.; Zhao, H.; Li, D.; Abbruzzese, J.; Baggerly, K.; Kobayashi, R. *Rapid Commun. Mass Spectrom.* **2004**, *18*, 2537–2548.
(20) Kiselar, J. G.; Downard, K. M. *Biochemistry* **1999**, *38*, 14185–14191.
(21) Wisniewski, J. R.; Zougman, A.; Nagaraj, N.; Mann, M. *Nat. Methods* **2009**, *6*, 359–362.
(22) Sun, W.; Gao, S.; Wang, L.; Chen, Y.; Wu, S.; Wang, X.; Zheng, D.; Gao, Y. *Mol. Cell. Proteomics* **2006**, *5*, 769–776.
(23) Kussmann, M.; Roepstorff, P. *Methods Mol. Biol.* **2000**, *146*, 405–424.
(24) Corbel, M. J.; Rondle, C. J. *J. Hyg.* **1970**, *68*, 81–96.
(25) Seto, J. T.; Becht, H.; Rott, R. *Med. Microbiol. Immunol.* **1973**, *159*, 1–12.
(26) Stegmann, T.; Morselt, H. W.; Booy, F. P.; van Breemen, J. F.; Scherphof, G.; Wilschut, J. *EMBO J.* **1987**, *6*, 2651–2659.
(27) Radyukhin, V.; Fedorova, N.; Ksenofontov, A.; Serebryakova, M.; Baratova, L. *Arch. Virol.* **2008**, *153*, 1977–1980.
(28) Zhang, J.; Pekosz, A.; Lamb, R. A. *J. Virol.* **2000**, *74*, 4634–4644.
(29) Glee, P. M.; Hazen, K. C. *Colloids Surf., B* **1995**, *5*, 181–188.
(30) Bondos, S. E.; Bicknell, A. *Anal. Biochem.* **2003**, *316*, 223–231.
(31) Smith, C. *Nat. Methods* **2005**, *2*, 71–77.
(32) Haslam, E. A.; Cheyne, I. M.; White, D. O. *Virology* **1969**, *39*, 118–129.
(33) Tan, L. T.; Xu, H. Y.; Wang, Y. L.; Qin, Z. M.; Sun, L.; Liu, W. J.; Cui, Z. Z. *J. Clin. Microbiol.* **2008**, *46*, 750–753.
(34) Huang, Z.; Panda, A.; Elankumaran, S.; Govindarajan, D.; Rockemann, D. D.; Samal, S. K. *J. Virol.* **2004**, *78*, 4176–4184.
(35) Smirnov, I. P.; Zhu, X.; Taylor, T.; Huang, Y.; Ross, P.; Papayanopoulos, I. A.; Martin, S. A.; Pappin, D. J. *Anal. Chem.* **2004**, *76*, 2958–2965.
(36) Baldwin, M. A. *Mol. Cell. Proteomics* **2004**, *3*, 1–9.
(37) Lee, C. W.; Swayne, D. E.; Linares, J. A.; Senne, D. A.; Suarez, D. L. *J. Virol.* **2005**, *79*, 11412–11421.
(38) Li, S.; Guo, Z.; Wu, H. F.; Liu, Y.; Yang, Z.; Woo, C. H. *Anal. Bioanal. Chem.* **2010**, *397*, 2465–2476.
(39) Sullivan, J. S.; Selleck, P. W.; Downton, T.; Boehm, I.; Axell, A. M.; Ayob, Y.; Kapitza, N. M.; Dyer, W.; Fitzgerald, A.; Walsh, B.; Lynch, G. W. *J. Mol. Genet. Med.* **2009**, *3*, 217–224.
(40) Saito, T.; Lim, W.; Suzuki, T.; Suzuki, Y.; Kida, H.; Nishimura, S. I.; Tashiro, M. *Vaccine* **2001**, *20*, 125–133.
(41) Panda, A.; Huang, Z.; Elankumaran, S.; Rockemann, D. D.; Samal, S. K. *Microb. Pathog.* **2004**, *36*, 1–10.