Top-Down Proteomics for Rapid Identification of Intact Microorganisms

ARTICLE *in* ANALYTICAL CHEMISTRY · NOVEMBER 2005

Impact Factor: 5.64 · DOI: 10.1021/ac051419g · Source: PubMed

4 AUTHORS, INCLUDING:

Paul Kowalski
Bruker Corporation
14 PUBLICATIONS **199** CITATIONS

SEE PROFILE

Anal. Chem. **2005,** *77,* 7455−7461

Top-Down Proteomics for Rapid Identification of Intact Microorganisms

Plamen A. Demirev,*,† Andrew B. Feldman,† Paul Kowalski,‡ and Jeffrey S. Lin†

Johns Hopkins University Applied Physics Laboratory, Laurel, Maryland 20723, and Bruker Daltonics, Billerica, Massachusetts 01821

We apply MALDI−TOF/TOF mass spectrometry for the rapid and high-confidence identification of intact *Bacillus* spore species. In this method, fragment ion spectra of whole (undigested) protein biomarkers are obtained without the need for biomarker prefractionation, digestion, separation, and cleanup. Laser-induced dissociation (unimolecular decay) of higher mass (>5 kDa) precursor ions in the first TOF analyzer is followed by reacceleration and subsequent high-resolution mass analysis of the resulting sequence-specific fragments in a reflectron TOF analyzer. In-house-developed software compares an experimental MS/MS spectrum with in silico-generated tandem mass spectra from all protein sequences, contained in a proteome database, with masses within a preset range around the precursor ion mass. A *p*-value, the probability that the observed matches between experimental and in silico-generated fragments occur by chance, is computed and used to rank the database proteins to identify the most plausible precursor protein. By inference, the source microorganism is then identified on the basis of the identification of individual, unique protein biomarker(s). As an example, intact *Bacillus atrophaeus* and *Bacillus cereus* spores, either pure or in mixtures, were unambiguously identified by this method after fragmenting and identifying individual small, acid-soluble spore proteins that are specific for each species. Factors such as experimental mass accuracy and number of detected fragment ions, precursor ion charge state, and sequence-specific fragmentation have been evaluated with the objective of extending the approach to other microorganisms. MALDI−TOF/TOF-MS in a lab setting is an efficient tool for in situ confirmation/verification of initial microorganism identification.

To address the challenges in rapid and reliable detection and identification of bioterrorism agents, a number of MS-based sensor platforms have been developed recently.[1−5] Notably, MALDI-MS can successfully distinguish intact microorganisms, exploiting differences in the protein biomarker masses observed for each species.[6−19] The observed protein biomarkers are typically highly expressed proteins with housekeeping functions, such as ribosomal, chaperone, and translation/transcription factor proteins.[20,21] Several approaches based on fingerprint library matching of intact protein biomarkers, bioinformatics, or both have been applied for microorganism detection and identification by, for example, MALDI-MS.[22] However, for mixtures of microorganisms, these approaches have clear limitations.

Bottom-up proteomics methodologies[23,24] have been adapted for microorganism identification by MS.[25−31] These are based on the initial identification of individual proteins from their corre-

* Corresponding author. Phone: 443-778-7712. Fax: 443-778-6904. E-mail: plamen.demirev@jhuapl.edu.

† Johns Hopkins University.

‡ Bruker Daltonics.

(1) Bryden, W. A.; Benson, R. C.; Ecelberger, S. A.; Phillips, T.; Cotter, R. J.; Fenselau, C. *Johns Hopkins APL Technol. Dig.* **1995**, *16*, 296−310.

(2) English, R. D.; Warscheid, B.; Fenselau, C.; Cotter, R. J. *Anal. Chem.* **2003**, *75*, 6886−6893.

(3) Ecelberger, S.; Cornish, T.; Collins, B.; Lewis, D.; Bryden, W. *Johns Hopkins APL Technol. Dig.* **2004**, *25*, 14−19.

(4) Cornish, T. J.; Antoine, M. D.; Ecelberger, S. A.; Demirev, P. A. *Anal. Chem.* **2005**, *77*, 3954−3959.

(5) Czerwieniec, G. A.; Russell, S. C.; Tobias, H. J.; Pitesky, M. E.; Fergenson, D. P.; Steele, P.; Srivastava, A.; Horn, J. M.; Frank, M.; Gard, E. E.; Lebrilla, C. B. *Anal. Chem.* **2005**, *77*, 1081−1087.

(6) Holland, R. D.; Wilkes, J. G.; Rafii, F.; Sutherland, J. B.; Persons, C. C.; Voorhees, K. J.; Lay, J. O. *Rapid Commun. Mass Spectrom.* **1996**, *10*, 1227−1232.

(7) Claydon, M. A.; Davey, S. N.; Edwards Jones, V.; Gordon, D. B. *Nat. Biotechnol.* **1996**, *14*, 1584−1586.

(8) Krishnamurthy, T.; Ross, P. L.; Rajamani, U. *Rapid Commun. Mass Spectrom.* **1996**, *10*, 883−888.

(9) Welham, K. J.; Domin, M. A.; Scannell, D. E.; Cohen, E.; Ashton, D. S. *Rapid Commun. Mass Spectrom.* **1998**, *12*, 176−180.

(10) Wang, Z. P.; Russon, L.; Li, L.; Roser, D. C.; Long, S. R. *Rapid Commun. Mass Spectrom.* **1998**, *12*, 456−464.

(11) Arnold, R. J.; Reilly, J. P. *Rapid Commun. Mass Spectrom.* **1998**, *12*, 630−636.

(12) Demirev, P. A.; Ho, Y. P.; Ryzhov, V.; Fenselau, C. *Anal. Chem.* **1999**, *71*, 2732−2738.

(13) Hathout, Y.; Demirev, P. A.; Ho, Y. P.; Bundy, J. L.; Ryzhov, V.; Sapp, L.; Stutler, J.; Jackman, J.; Fenselau, C. *Appl. Environ. Microbiol.* **1999**, *65*, 4313−4319.

(14) Holland, R. D.; Duffy, C. R.; Rafii, F.; Sutherland, J. B.; Heinze, T. M.; Holder, C. L.; Voorhees, K. J.; Lay, J. O. *Anal. Chem.* **1999**, *71*, 3226−3230.

(15) Jarman, K. H.; Cebula, S. T.; Saenz, A. J.; Petersen, C. E.; Valentine, N. B.; Kingsley, M. T.; Wahl, K. L. *Anal. Chem.* **2000**, *73*, 1217−1223.

(16) Jones, J. J.; Stump, M. J.; Fleming, R. C.; Lay, J. O.; Wilkins, C. L. *Anal. Chem.* **2003**, *75*, 1340−1347.

(17) Williams, T. L.; Andrzejewski, D.; Lay, J. O.; Musser, S. M. *J. Am. Soc. Mass Spectrom.* **2003**, *14*, 342−351.

(18) Dickinson, D.; La Duc, M.; Haskins, W.; Gornushkin, I.; Winefordner, J.; Powell, D.; Venkateswaran, K. *Appl. Environ. Microbiol.* **2004**, *70*, 475−482.

(19) Whiteaker, J.; Warscheid, B.; Pribil, P.; Hathout, Y.; Fenselau, C. *J. Mass Spectrom.* **2004**, *39*, 1113−1121.

(20) Pineda, F.; Antoine, M.; Demirev, P.; Feldman, A.; Jackman, J.; Longenecker, M.; Lin, J. *Anal. Chem.* **2003**, *75*, 3817−3822.

(21) Demirev, P. A.; Feldman, A. B.; Lin, J. S. *Johns Hopkins APL Technol. Dig.* **2004**, *25*, 27−37.

(22) Fenselau, C.; Demirev, P. *Mass Spectrom. Rev.* **2001**, *20*, 157−171.

10.1021/ac051419g CCC: $30.25
Published on Web 10/21/2005

sponding tryptic peptides and, by inference, identification of the microorganism from which these proteins originate. As is well-established in bottom-up proteomics, proteolytic enzyme specificity, complementary to or concurrently with peptide sequence tag information (obtained by tandem mass spectrometry), improves the capability for unequivocal protein identification. In one such approach, rapid in situ (on a MALDI sample slide) proteolysis of proteins derived from intact viruses results in several tryptic peptide fragments for each biomarker protein.[32] Further developments of the bottom-up approaches include the identification of *Bacillus* spores (pure or in mixtures) by selective solubilization of the small acid-soluble spore protein (SASP) biomarkers and their subsequent in situ digestion with agarose-bead-immobilized trypsin. The resultant proteolytic peptides are analyzed by several different types of tandem mass spectrometers: a MALDI−TOF-MS with a curved-field reflectron,[33] a hybrid ion trap/time-of-flight mass spectrometer,[34] or an AP MALDI ion trap instrument.[35] In these experiments, SASP biomarkers are identified from the partial sequences of the proteolytic peptides combined with proteome-based database queries using the MASCOT search engine.

Here we present the first results in a top-down proteomics approach that employs MALDI−TOF/TOF-MS of whole (undigested) protein biomarkers for direct and rapid identification of individual *Bacillus* spore species, either pure or in a mixture (Figure 1). A major advantage of this method is that biomarker MS/MS spectra are obtained without the need for biomarker prefractionation, digestion, separation, and cleanup. The capability to identify an intact protein by deducing its partial amino acid sequence (a sequence tag) in an FTICR MS/MS experiment and subsequent homology search in a proteome database was first demonstrated by Mortz et al.[36] This top-down approach in proteomics has been further developed in a number of studies, all involving either FTICR[37−39] or quadrupole[40−43] ion traps. In analogy to bottom-up proteomics, unambiguous identification of one or more intact protein biomarkers by top-down proteomics allows successful microorganism identification (provided the proteome database contains both the respective protein sequences and the respective organism sources). For instance, biomarker proteins from *Bacillus cereus T* spores have been analyzed by high-resolution tandem FTICR MS.[44] Fragmentation-derived sequence tags and BLAST sequence similarity searches in a proteome database provide unequivocal identification of the major biomarker protein, observed in MALDI of intact spores,[13,45] as a SASP. Following individual protein identification, the spore species itself could be unambiguously identified.[44]

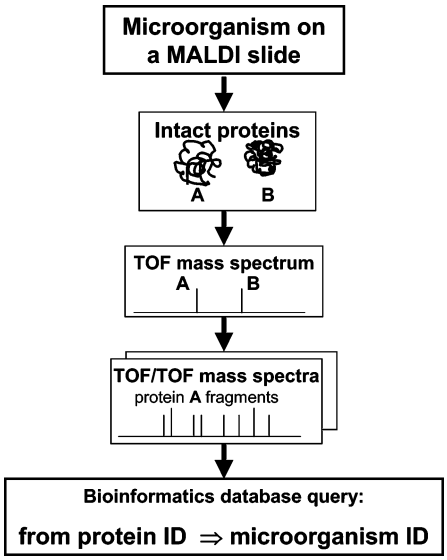

Figure 1. Top-down proteomics approach for identification of microorganisms from the MALDI−TOF/TOF mass spectra of whole unfractionated proteins.

Recent developments in TOF/TOF instruments[46] have considerably improved the efficiency of tandem MS of larger (>2 kDa) peptides and small proteins.[47−53] In the approach described here, mass/charge-selected precursor ions undergo laser-induced dissociation (LID)—metastable post-source decay[54−56]—in the field-

(23) Washburn, M. P.; Wolters, D.; Yates, J. R. *Nat. Biotechnol.* **2001**, *19*, 242−247.
(24) Aebersold, R.; Mann, M. *Nature* **2003**, *422*, 198−207.
(25) Liu, C. L.; Hofstadler, S. A.; Bresson, J. A.; Udseth, H. R.; Tsukuda, T.; Smith, R. D.; Snyder, A. P. *Anal. Chem.* **1998**, *70*, 1797−1801.
(26) Smith, R. D.; Anderson, G. A.; Lipton, M. S.; Pasa-Tolic, L.; Shen, Y. F.; Conrads, T. P.; Veenstra, T. D.; Udseth, H. R. *Proteomics* **2002**, *2*, 513−523.
(27) VerBerkmoes, N. C.; Bundy, J. L.; Hauser, L.; Asano, K. G.; Razumovskaya, J.; Larimer, F.; Hettich, R. L.; Stephenson, J. L. *J. Proteome Res.* **2002**, *1*, 239−252.
(28) Dworzanski, J. P.; Snyder, A. P.; Chen, R.; Zhang, H. Y.; Wishart, D.; Li, L. *Anal. Chem.* **2004**, *76*, 2355−2366.
(29) Warscheid, B.; Fenselau, C. *Proteomics* **2004**, *4*, 2877−2892.
(30) VerBerkmoes, N. C.; Hervey, W. J.; Shah, M.; Land, M.; Hauser, L.; Larimer, F. W.; Van Berkel, G. J.; Goeringer, D. E. *Anal. Chem.* **2005**, *77*, 923−932.
(31) Hu, A.; Tsai, P. J.; Ho, Y. P. *Anal. Chem.* **2005**, *77*, 1488−1495.
(32) Yao, Z.; Demirev, P.; Fenselau, C. *Anal. Chem.* **2002**, *74*, 2529−2534.
(33) Warscheid, B.; Fenselau, C. *Anal. Chem.* **2003**, *75*, 5608−5617.
(34) Warscheid, B.; Jackson, K.; Sutton, C.; Fenselau, C. *Anal. Chem.* **2003**, *75*, 5608−5617.
(35) Pribil, P. A.; Patton, E.; Black, G.; Doroshenko, V.; Fenselau, C. *J. Mass Spectrom.* **2005**, *40*, 464−474.
(36) Mortz, E.; O'Connor, P.; Roepstorff, P.; Kelleher, N.; Wood, T.; McLafferty, F. W.; Mann, M. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 8264−8267.
(37) Kelleher, N. L.; Lin, H. Y.; Valaskovic, G. A.; Aaserud, D. J.; Fridriksson, E. K.; McLafferty, F. W. *J. Am. Chem. Soc.* **1999**, *121*, 806−812.
(38) Kelleher, N. L. *Anal. Chem.* **2004**, *76*, 196A−203A.
(39) Meng, F. Y.; Forbes, A. J.; Miller, L. M.; Kelleher, N. L. *Mass Spectrom. Rev.* **2005**, *24*, 126−134.
(40) Cargile, B. J.; McLuckey, S. A.; Stephenson, J. L. *Anal. Chem.* **2001**, *73*, 1277−1285.
(41) Reid, G. E.; Shang, H.; Hogan, J. M.; Lee, G. U.; McLuckey, S. A. *J. Am. Chem. Soc.* **2002**, *124*, 7353−7362.
(42) Reid, G. E.; McLuckey, S. A. *J. Mass Spectrom.* **2002**, *37*, 663−675.
(43) Coon, J. J.; Ueberheide, B.; Syka, J. E. P.; Dryhurst, D. D.; Ausio, J.; Shabanowitz, J.; Hunt, D. F. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 9463−9468.
(44) Demirev, P.; Ramirez, J.; Fenselau, C. *Anal. Chem.* **2001**, *73*, 5725−5731.
(45) Hathout, Y.; Setlow, B.; Cabrera-Martinez, R. M.; Fenselau, C.; Setlow, P. *Appl. Environ. Microb.* **2003**, *69*, 1100−1107.
(46) Cotter, R. J. *Time-of-Flight Mass Spectrometry*; American Chemical Society: Washington, DC, 1997.
(47) Medzihradszky, K. F.; Campbell, J. M.; Baldwin, M. A.; Falick, A. M.; Juhasz, P.; Vestal, M. L.; Burlingame, A. L. *Anal. Chem.* **2000**, *72*, 552−558.
(48) Suckau, D.; Resemann, A.; Schuerenberg, M.; Hufnagel, P.; Franzen, J.; Holle, A. *Anal. Bioanal. Chem.* **2003**, *376*, 952−965.
(49) Suckau, D.; Resemann, A. *Anal. Chem.* **2003**, *75*, 5817−5824.
(50) Neubert, H.; Halket, J. M.; Ocana, M. F.; Patel, R. K. P. *J. Am. Soc. Mass Spectrom.* **2004**, *15*, 336−343.
(51) Lin, M.; Campbell, J. M.; Mueller, D. R.; Wirth, U. *Rapid Commun. Mass Spectrom.* **2003**, *17*, 1809−1814.
(52) Macht, M.; Asperger, A.; Deininger, S. O. *Rapid Commun. Mass Spectrom.* **2004**, *18*, 2093−2105.
(53) Liu, Z. Y.; Schey, K. L. *J. Am. Soc. Mass Spectrom.* **2005**, *16*, 482−490.
(54) Spengler, B. *J. Mass Spectrom.* **1997**, *32*, 1019−1036.
(55) Clipston, N. L.; Jai-nhuknan, J.; Cassady, C. J. *Int. J. Mass Spectrom.* **2003**, *222*, 363−381.

free region of the first TOF analyzer. The resulting sequence-specific fragments are reaccelerated and subsequently analyzed in the second high-resolution reflectron TOF mass analyzer. We utilize in-house-developed software to count the number of matches between an experimental MS/MS spectrum and in silico generated tandem mass spectra from all protein sequences, contained in a proteome database, with masses within a preset range around the precursor ion mass. The database proteins are ranked on the basis of the calculated probability (the *p*-value) that the observed matches between experimental and in silico-generated fragments occur by chance to identify the most plausible precursor protein. By inference, the source microorganism is then identified on the basis of the identification of individual, unique protein biomarker(s). As an example, intact *Bacillus atrophaeus* and *Bacillus cereus* spores, either pure or in a mixture, are unambiguously identified by MALDI−TOF/TOF after fragmenting and identifying individual small acid-soluble spore proteins specific for each species. Factors such as experimental mass accuracy and number of detected fragment ions, precursor ion charge state, and sequence-specific fragmentation have been evaluated with the objective of extending the approach to other potential biothreat agents.

EXPERIMENTAL SECTION

Mass Spectrometry. Intact *B. atrophaeus* (*Bacillus globigii*, *Bg*) and *B. cereus T* (*Bc*) spores, either pure or in a 1:1 mixture, were deposited from water suspensions (~10^9 spores/mL) on an AnchorChip sample slide (Bruker Daltonics, Billerica, MA). The estimated number of deposited microorganisms per sample spot was around 3×10^5. Trifluoroacetic acid (10%, 0.3 μL) was added to each spot to facilitate the SASP extraction from the spores.[33,45] The effect of two different MALDI matrixes, DHB (dihydroxybenzoic acid) and α-CHCA (α-cyano-4-hydroxycinnamic acid), on precursor ion yields and ultimately on the MS/MS spectral quality was tested. Around 0.3 μL of each matrix (20 mg/mL in acetonitrile/water, 1:1) was added to each sample spot. Positive ion MALDI−TOF and TOF/TOF spectra were acquired on an Ultraflex II MALDI−TOF/TOF instrument (Bruker Daltonics, Billerica, MA). The instrument performance, principle of operation and several applications were previously described.[48−50,52] The instrument was equipped with a high rep rate (200 Hz) diode-pumped all-solid-state SmartBeam laser (third harmonic at 355 nm, typical laser fluence 2.6×10^2 J/m^2). Spectra were acquired with delayed extraction at 8 kV accelerating voltage in the first leg of the TOF/TOF instrument. The respective precursor ion and its fragment ions were selected (isolated) in an ion gate, before reacceleration by the LIFT method[48] to a total of 25 kV for intact precursor ions in the second (reflectron) leg. Typically, 5000 single laser shot traces were summed for each MS/MS spectrum by manually scanning the laser beam across an individual sample spot (hunting for "sweet" spots). The MS and MS/MS spore spectra were calibrated by using pure peptides and proteins (somatostatin, bovine ubiquitin, horse heart cytochrome *c*). Matrix compounds and proteins were all obtained from Sigma Chemical Co. (St. Louis, MO) and were used without additional purification. All spore samples were grown in-house and purified according to well-established procedures.[13]

Database Searches. We used the combined SwissPROT/TrEMBL proteome databases[57] to extract precursor protein sequences with masses within a preset range around the precursor ion mass (±2.5 Da). The probability of a posttranslational protein modification, N-terminal Met truncation, was incorporated as already described.[20,58] In-house developed software (see Supporting Information for additional details) counted the number of matches between the experimental MS/MS spectrum and in silico-generated tandem mass spectra from all plausible precursor protein sequences. The in silico a-, b- and y-fragment ion masses were calculated from the respective protein sequences. A match was counted if the experimentally observed mass value coincided with the mass of an in silico fragment ion within the experimentally determined mass accuracy (tolerance) window Dm (= ±2.5 Da). A *p*-value, $0 < p < 1$, the statistical significance of the number of observed peak matches, was assigned to each database protein and used to rank the database precursor proteins to identify the most plausible one. A smaller *p*-value denotes higher probability of correct protein identification. The *p*-value depends on the number of matches, Δm, the number of predicted fragment ions and their mass distribution, and the number of fragment ion peaks in the spectrum.

RESULTS AND DISCUSSION

Tandem MS of intact protein biomarkers (top-down proteomics) typically requires complex instrumentation (most often electrospray ionization combined with FTICR[36−39]). Until now, protein mixtures had to be concentrated, purified, and separated prior to top-down MS analysis. In an MS/MS (tandem MS) experiment, a precursor peptide or protein cation is selected, isolated, and excited via interaction with neutral gas molecules (collisionally induced dissociation, CID[59−62]), surfaces (surface-induced dissociation, SID[63,64]), UV photons,[65] coherent[66] or non-coherent IR (blackbody infrared radiative dissociation, BIRD) radiation,[67,68] electrons (electron-capture dissociation, ECD[69]), or ions (electron-transfer dissociation, ETD[70]). The increased internal energy of the precursor ion above a threshold results in the ion dissociation into sequence-specific fragments, which in turn can

(56) Cotter, R. J.; Iltchenko, S.; Wang, D. X. *Int. J. Mass Spectrom.* **2005**, *240*, 169−182.

(57) Bairoch, A.; Apweiler, R.; Wu, C. H.; Barker, W. C.; Boeckmann, B.; Ferro, S.; Gasteiger, E.; Huang, H.; Lopez, R.; Magrane, M.; Martin, M. J.; Natale, D. A.; O'Donovan, C.; Redaschi, N.; Yeh, L. S. *Nucleic Acids Res.* **2005**, *33*, D154−159.

(58) Demirev, P.; Lin, J. S.; Pineda, F. J.; Fenselau, C. *Anal. Chem.* **2001**, *73*, 4566−4573.

(59) Jennings, K. R. *Int. J. Mass Spectrom.* **2000**, *200*, 479−493.

(60) Williams, E. R.; Furlong, J. J. P.; McLafferty, F. W. *J. Am. Soc. Mass Spectrom.* **1990**, *1*, 288−294.

(61) Cornish, T. J.; Cotter, R. J. *Anal. Chem.* **1993**, *65*, 1043−1047.

(62) Wysocki, V. H.; Tsaprailis, G.; Smith, L. L.; Breci, L. A. *J. Mass Spectrom.* **2000**, *35*, 1399−1406.

(63) Schey, K.; Cooks, R. G.; Grix, R.; Wollnik, H. *Int. J. Mass Spectrom. Ion Processes* **1987**, *77*, 49−61.

(64) Gamage, C. M.; Fernandez, F. M.; Kuppannan, K.; Wysocki, V. H. *Anal. Chem.* **2004**, *76*, 5080−5091.

(65) Thompson, M. S.; Cui, W. D.; Reilly, J. P. *Angew. Chem. Int. Ed.* **2004**, *43*, 4791−4794.

(66) Little, D. P.; Speir, J. P.; Senko, M. W.; O'Connor, P. B.; McLafferty, F. W. *Anal. Chem.* **1994**, *66*, 2809−2815.

(67) Price, W. D.; Schnier, P. D.; Williams, E. R. *Anal. Chem.* **1996**, *68*, 859−866.

(68) Dunbar, R. C. *Mass Spectrom. Rev.* **2004**, *23*, 127−158.

(69) Zubarev, R. A.; Kelleher, N. L.; McLafferty, F. W. *J. Am. Chem. Soc.* **1998**, *120*, 3265−3266.

(70) Syka, J. E. P.; Coon, J. J.; Schroeder, M. J.; Shabanowitz, J.; Hunt, D. F. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 9528−9533.

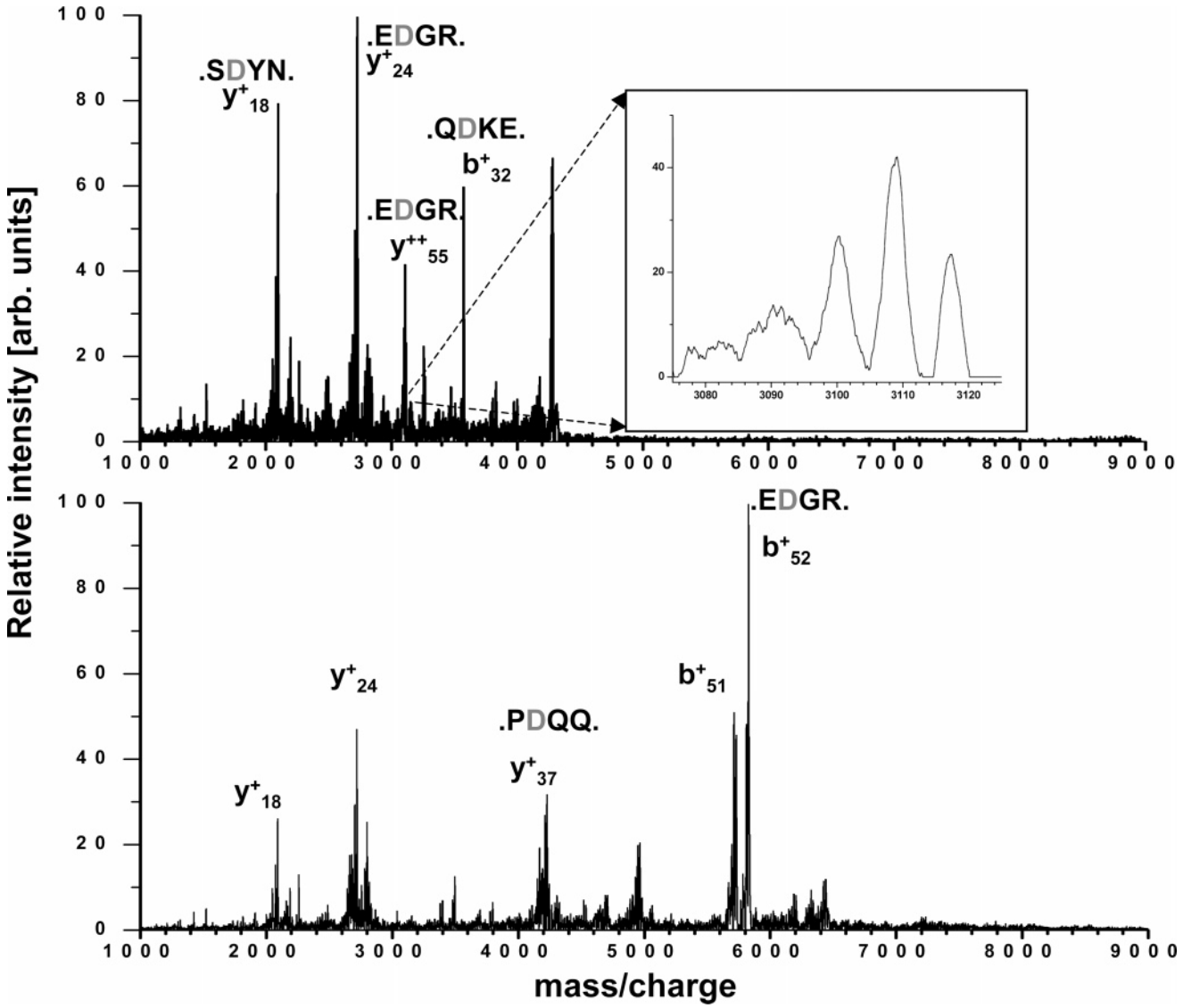

Figure 2. LID−TOF/TOF mass spectra of doubly charged (top) and singly charged (bottom) ubiquitin molecular ions. Partial sequences where cleavages occur are indicated, as well. The singly charged (doubly charged) precursor results in 65 fragment ions with distinct mass/charge in the range from 500 to 8565 (56 fragment ions in the range from *m*/*z* 500 to 4282). The inset represents the expanded region around the y_{55}^{2+} fragment ion.

be correlated to the precursor ion sequence. Here we utilize for top-down proteomics MALDI−TOF/TOF-MS, which is more tolerant to impurities; does not require biomarker extraction, purification, and separation; and can be directly applied to intact microorganisms.[22,71]

It is well-established that the amount of transferred internal energy required for efficient dissociation of a large polyatomic system scales with the system size (the kinetic shift effect[72,73]). The energy transfer efficiency of the "slow heating" activation methods[74] (such as low energy CID or BIRD) thus determines the type of instruments, Penning or Paul ion traps with time scales on the order of hundreds of milliseconds to seconds, for observation of the dissociation of larger (>3 kDa) polypeptide ions.[42,75] The nonergodic ion activation methods (such as ECD and ETD) may not require long observational time scales, but so far, they have been coupled solely to ion trap instruments. Here we demonstrate that unimolecular (metastable) decay of large (>5 kDa) peptide and protein precursor ions, excited during their formation in the MALDI source, yields abundant sequence-specific fragmentation (Figure 2). This process of protein fragment generation has been termed laser-induced dissociation (LID[48]). LID bears many similarities to the post-source decay of MALDI-generated peptide ions,[54] used in conjunction with reflectron TOF instruments. In LID, the protein fragment ions together with intact precursor ions are isolated on the basis of their time-of-flight in the first TOF. The LID rate is a function of the incident desorption laser fluence. The fluence in a LID experiment is typically a factor of 2−5 higher as compared to the near-threshold fluence at which a high-resolution spectrum of the precursor protein ion is obtained. Although elucidating the exact nature of the energy excitation processes in LID was beyond the scope of this study, we argue

(71) Lay, J. *Mass Spectrom. Rev.* **2001**, *20*, 172−194.
(72) Lifshitz, C. *Eur. J. Mass Spectrom.* **2002**, *8*, 85−98.
(73) Laskin, J.; Futrell, J. H. *Mass Spectrom. Rev.* **2005**, *24*, 135−167.
(74) McLuckey, S. A.; Goeringer, D. E. *J. Mass Spectrom.* **1997**, *32*, 461−474.
(75) Jockusch, R. A.; Schnier, P. D.; Price, W. D.; Strittmatter, E. F.; Demirev, P. A.; Williams, E. R. *Anal. Chem.* **1997**, *69*, 1119−1126.

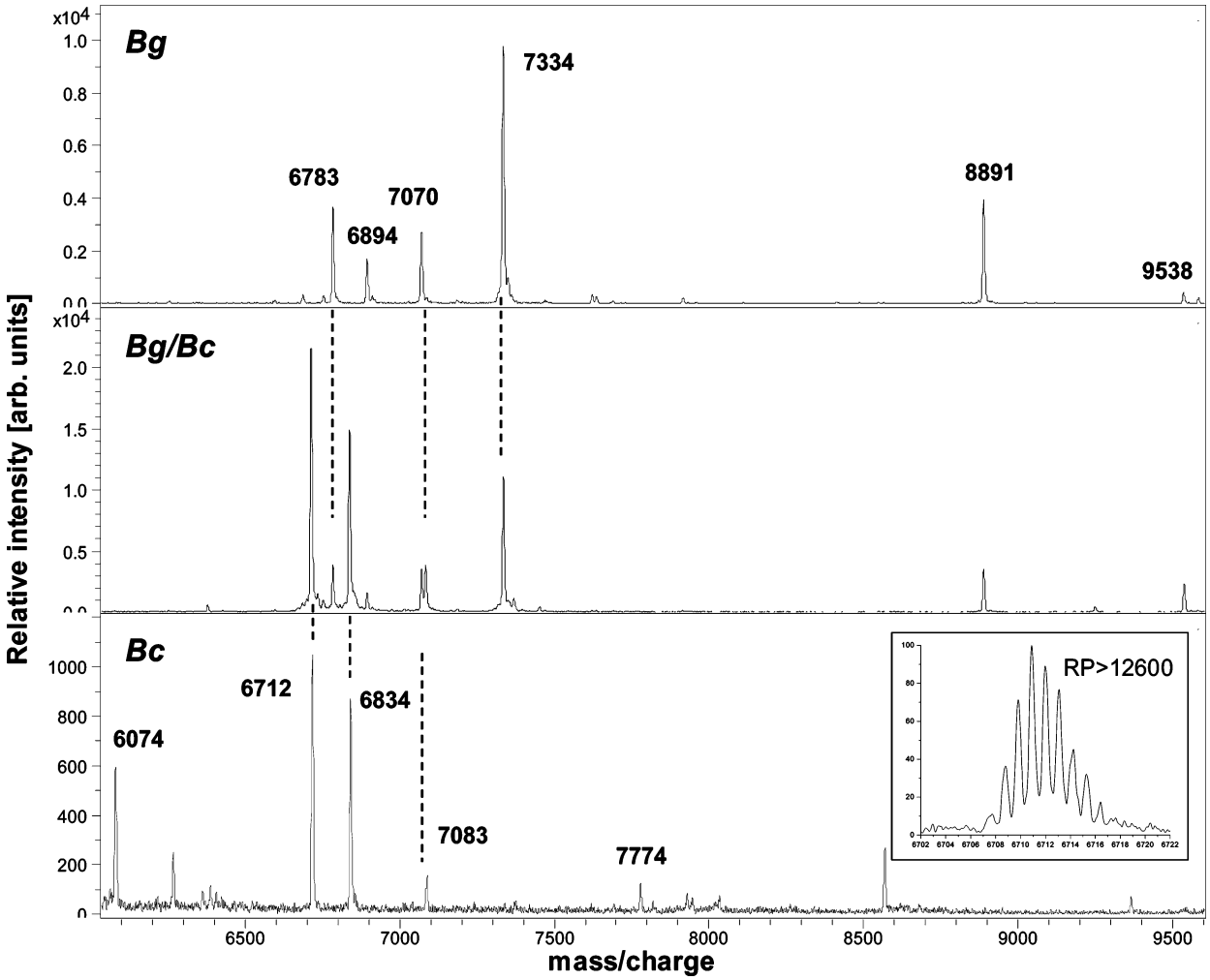

Figure 3. MALDI−TOF mass spectra of intact *Bacillus* spores. Top, pure *B. globigii* (*Bg*); middle, binary *B. globigii*/*B. cereus* mixture; bottom, pure *B. cereus* (*Bc*) spores. The inset at the bottom represents the expanded molecular ion region around the ion at mass/charge 6712.

that upon increased laser fluence, a larger amount of neutral matrix material is being ablated. In turn, the increased particle density in the expanding plume would favor multiple low-energy collisions between intact protein ions and matrix molecules, reminiscent of CID conditions (e.g., in an ion trap instrument). This argument is supported by the similarity of the protein MS/MS spectra observed in both cases, LID and CID (vide infra). In the TOF/TOF instrument, the MALDI yield of doubly charged protein molecular ions (including spore-specific biomarkers) is enhanced by α-CHCA, as compared to DHB; however, better MS/MS spectra (higher fragment ion yields) for the singly charged precursor ions are obtained using DHB as a matrix. This effect may be due both to the more intense overall precursor ion signal as well as to more efficient precursor ion excitation (increased amount of ablated matrix). We have also compared the LID and the CID−TOF/TOF mass spectra of the same precursor ion (singly charged ubiquitin) on the same instrument (CID was performed with argon as a collision gas, other experimental conditions similar to already described[52]). In contrast to small peptides, LID−TOF/TOF is much more efficient for larger (>5 kDa) peptides, as compared to CID, and only LID−TOF/TOF spectra will be subsequently discussed.

LID of intact protein ions results in sequence-specific backbone cleavages, generating a-, b-, and y-ion fragments. Up to two small neutral group (NH_3/H_2O) losses are typically observed. Cleavages on the C-terminal side of aspartic or glutamic acid residues dominate the spectra (Figure 2). This type of charge-remote fragmentation is similar to fragmentation patterns observed in CID of peptides and small intact proteins in low charge states.[41,76,77] When the number of charges is much smaller than the number of potential protonation sites, there are no mobile protons, and the side chain of the acidic amino acid provides the proton next to the cleavage site.[62,76] For doubly charged ubiquitin precursors ions, nonsymmetric unimolecular fragmentation into a doubly charged and a neutral fragment is also observed, a series of doubly charged sequence-specific y-fragments (y_{48}^{2+}, y_{52}^{2+}, y_{53}^{2+}, y_{54}^{2+}, y_{55}^{2+}, y_{58}^{2+}, y_{62}^{2+}, y_{64}^{2+}). Although metastable decay of singly charged proteins in a TOF/TOF experiment has been reported before,[51] this is the first observation of nonsymmetric fission of doubly charged large (>5 kDa) protein ions. The fragment-ion charge state is evident from the difference of $m/z \sim 8.5$ (two consecutive small neutral group losses) discerned between individual fragment ion peaks as well as the fwhm of these fragment peaks (Figure 2). Around 60 fragment ion peaks are each observed in the MS/MS spectra of either singly or doubly charged bovine ubiquitin. The y-ions

(76) Paizs, B.; Suhai, S. *Mass Spectrom. Rev.* **2005**, *24*, 508−548.
(77) Reid, G. E.; Wu, J.; Chrisman, P. A.; Wells, J. M.; McLuckey, S. A. *Anal. Chem.* **2001**, *73*, 3274−3281.

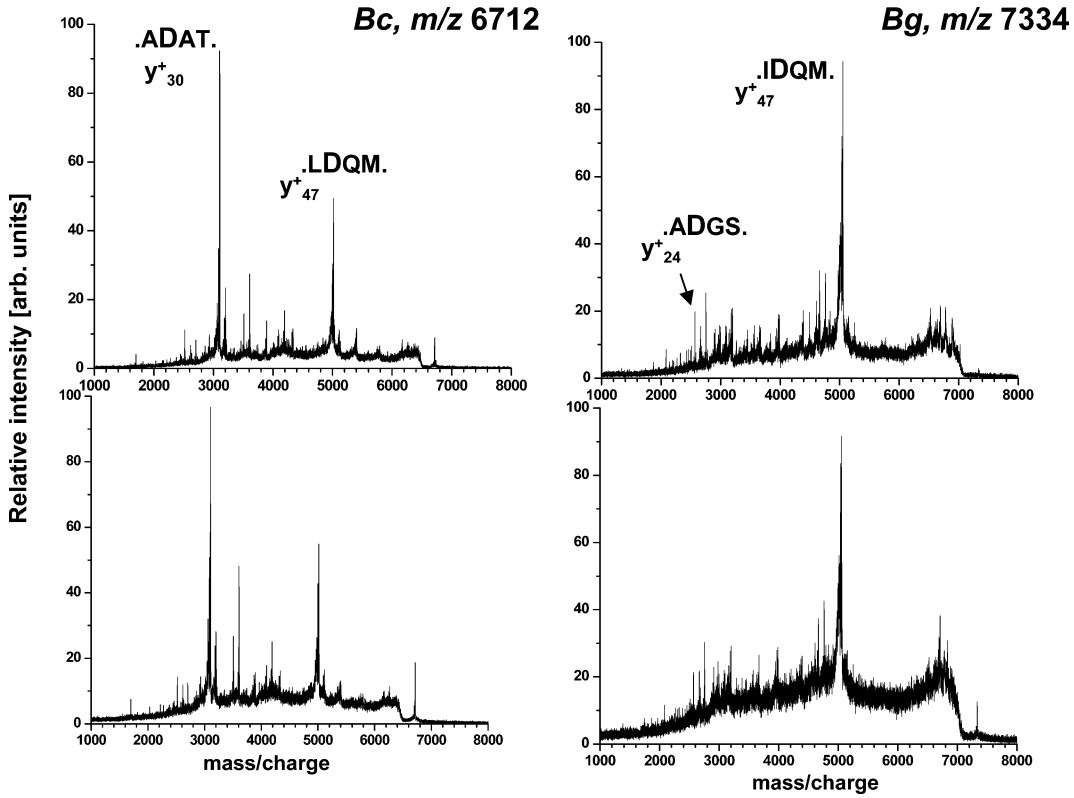

Figure 4. MALDI−TOF/TOF mass spectra of the major biomarkers for *B. globigii* and *B. cereus* spores obtained either from the respective pure species (top) or from the binary spore mixture (bottom).

from doubly charged ubiqitin are more dominant, as compared to singly charged ubiquitin (in which more abundant a- and b-ions are also observed). The overall LID fragmentation efficiency for doubly charged protein ions, as compared to the respective singly charged precursor, is certainly higher than reported here (singly charged fragment ions with m/z higher than the m/z of the respective doubly charged precursor ion are not detected in the current TOF/TOF experimental setup). For the proteins with MW from 3 to 9 kDa tested here, the number of individual fragment ions varies between 30 and 100 (depending on precursor ion intensity, laser fluence, matrix, etc). On the basis of the number of matched fragment ions, the calculated p-values for all proteins tested here are better than 10^{-4} (e.g., 2.5×10^{-9} and 9×10^{-9} for singly and doubly charged ubiquitin, respectively).

The major biomarker ions between m/z 6000 and 9000 observed in MALDI spectra of intact *Bacillus* spores (either pure or in a mixture, Figure 3) are all unambiguously identified as SASPs using this tandem MS method. Figure 4 illustrates the tandem MS spectrum of the major *Bc* SASP at m/z 6712. The instrument resolving power for these precursor ions is better than 12 600 (Figure 3, inset), which is remarkable given the inherent spatial nonhomogeneity of the sample surface (an intact spore has characteristic dimensions on the order of 1 μm). Representative MS/MS spectra of the major SASP biomarkers for *Bc* and *Bg* obtained either from a pure spore sample or from a mixture, are presented in Figure 4. The similarity between the different spectra for the same precursor ion demonstrates the reproducibility of the MALDI−TOF/TOF method. This result suggests that library fingerprint matching[11,14] can be applied to TOF/TOF spectra of biomarkers from intact microorganisms, as well. Such an approach will be advantageous for cases when the microorganism proteome is not available, when nonprotein or only a few protein biomarker peaks are observed in an ordinary TOF spectrum, for example, virus coat proteins or protein toxins.

For *Bg* and *Bc* spores, the p-values and the identification of the respective major SASPs from their LID−MS/MS spectra are given in Table 1. The enhanced discriminatory power of the tandem MS approach is illustrated by the large difference—at least 10^{12}—in the p-values between the first and the next ranked protein candidate from the database for the two major SASP biomarkers in, for example, *Bc* (m/z 6712) and *Bg* (m/z 7034) spore spectra (whether pure or in a mixture). Combining TOF and TOF/TOF-MS provides additional practical constraints (e.g., the accurate molecular weight of the precursor ion) for improved bioinformatics-based protein identification. The same SASP (same protein sequence) can be present in several "near-neighbor" *Bacillus* species. For instance, the "6834" SASP biomarker is found not only in *Bc*, but also in *Bacillus thuringiensis* and *Bacillus anthracis* spores (likewise, "7084" SASP is characteristic for both *Bc* and *B. anthracis*). Therefore, it is important to obtain independent and unambiguous identification of the SASP, which are unique for each species, especially when a mixture is being analyzed. We stress that the "6712" SASP uniquely differentiates *Bc* spores from *Bg* spores and, most importantly, from *B. anthracis*.[21,45] The latter species also contains a unique SASP (sspB, SwissPROT accession number Q81KU1) with a mass of 6680 Da. The "6680" SASP has a sequence that differs in two amino acids from the sequence of the "6712" SASP, thus resulting in a 32-Da difference between the two SASPs, unique for *B. anthracis* and *Bc*, respectively. This mass difference is sufficient to differentiate between the two *Bacillus* spore species by MALDI−TOF-MS.[21,45] Furthermore, the fragment masses of these two intact proteins would also differ

Table 1. *Bacillus* Spore Biomarker ID by MALDI−TOF/TOF-MS of Intact Proteins

sample	major biomarker MH^+ (Th)	no. of database entries (±2.5 Th)	no. of matches (no. of observed peaks)	protein ID	*p*-value
Bc	6712	58	51 (68)	P0A4F4 *B. cereus* SASP-2	1.32×10^{-21}
	6834	88	41 (55)	Q63FF2 *B. cereus* sspA	4.52×10^{-17}
	7083	83	24 (54)	Q739S3 *B. cereus* SASP	9.09×10^{-5}
Bg	7070	98	20 (38)	*B. globigii* SASP[a]	2.25×10^{-5}
	7334	105	46 (68)	*B. globigii* SASP[a]	2.81×10^{-16}
Bc/*Bg* mixture	6712	58	77 (104)	P0A4F4 *B. cereus* SASP-2	1.81×10^{-30}
	7334	105	58 (100)	*B. globigii*	8.13×10^{-16}

[a] Sequence data for *Bg* spore SASP are taken from ref 19.

substantially, presenting an additional powerful tool for differentiation between these two closely related species. Rapid confirmation/rejection of an initial "positive" alert signal is a very important step in every technology for biothreat agent detection and identification. The MALDI−TOF/TOF-MS of intact protein biomarkers provides such a high confidence orthogonal approach for the in situ confirmation/verification of initial MALDI−TOF-MS detection of biothreats (e.g., *Bacillus* spores). MALDI−TOF/TOF-MS of intact protein biomarkers obviates the need for additional sample preparation and reduces considerably the time for confirmatory analysis: in a MALDI−TOF/TOF instrument, switching from MS to MS/MS mode can take just a few seconds.

We have also performed TOF/TOF−MALDI-MS of several lower mass (from 1 to 5 kDa) intact nonprotein *Bacillus* spore biomarkers (e.g., lipopeptides[78,79]). These structure-specific spectra are very informative and can be also used in differentiating between various *Bacillus* spore species (unpublished results). Thus, combining the identification by MS/MS of several individual biomarkers, the organism itself can be unambiguously identified. In top-down proteomics the probability for positive matching of a protein but false matching of a microorganism is lower as compared to bottom-up shotgun proteomics.[28,30] In both cases, correct microorganism identification relies on the identification of unique versus degenerate biomarker ions. In this context, degenerate means either intact proteins (i.e., orthologues) or tryptic peptides[28] found in two or more microorganisms. However, it is clear that the ratio of unique versus degenerate biomarkers would be much higher for intact proteins than for tryptic peptides.

CONCLUSION AND PROSPECTS

High-resolution MALDI−TOF/TOF mass spectrometry in a lab setting combined with bioinformatics is an efficient tool for in situ high confidence verification of the initial identification of an intact microorganism by MALDI−TOF-MS, particularly when the microorganism is present in a mixture. The method is based on identification (and thus, confirmation of the presence) of individual protein biomarker peaks, based on their tandem mass spectra. It provides a much higher level of identification specificity for individual microorganisms, particularly in complex outdoor environments with high biological background. This approach can also be used for identification of unknown biomarker peaks in intact microorganisms. Presumably, top-down proteomics combined with bioinformatics can be successful in the detection of genetically engineered microorganisms, as well. We plan to extend this top-down proteomics approach to intact protein biomarkers larger than 13 kDa for rapid virus or protein toxin identification. In this line, we try to improve the generation of intense high (>3+) charge-state protein precursor ions by selecting appropriate MALDI matrixes and substrates. Expanding the bioinformatics algorithms by, for example, incorporation of experimentally derived fragmentation rules, Bayesian probability calculations, and de novo derivation of sequence tags from the spectra can further improve the success of that approach. Finally, instrument improvements that include better mass accuracy, automated data-dependent acquisition, and design of scaled-down and field-portable tandem TOF mass spectrometers would certainly expand the uses of MALDI-MS for microorganism identification and characterization.

(78) Hathout, Y.; Ho, Y. P.; Ryzhov, V.; Demirev, P.; Fenselau, C. *J. Nat. Prod.* **2000**, *63*, 1492−1496.
(79) Madonna, A. J.; Voorhees, K. J.; Taranenko, N. I.; Laiko, V. V.; Doroshenko, V. M. *Anal. Chem.* **2003**, *75*, 1628−1637.

ACKNOWLEDGMENT

We thank James Crookston and Jason Quizon (Applied Physics Lab) for providing the microorganisms used in this study. Financial support for this research was provided by APL independent R&D grants. This work was presented in part at the 53rd Annual Conference of the American Society for Mass Spectrometry, San Antonio, TX, June 2005.

SUPPORTING INFORMATION AVAILABLE

Material as stated in the text. This material is available free of charge via the Internet at http://pubs.acs.org.

Received for review August 8, 2005. Accepted September 20, 2005.

AC051419G