Anal. Chem. **2001,** *73,* 4134−4144

Discrimination of Aerobic Endospore-forming Bacteria via Electrospray-Ionization Mass Spectrometry of Whole Cell Suspensions

Seetharaman Vaidyanathan,[†] Jem J. Rowland,[‡] Douglas B. Kell,[†] and Royston Goodacre[†,*]

Institute of Biological Sciences, and Department of Computer Sciences, University of Wales, Aberystwyth, Ceredigion SY23 3DD. Wales, UK

Direct injection electrospray ionization mass spectrometry (ESI-MS) without prior analyte separation was investigated for the analysis of whole cell suspensions of bacteria. Thirty-six strains of aerobic endospore-forming bacteria, consisting of six *Bacillus* species and one *Brevibacillus* species, were studied. ESI was performed in the positive ion mode on the bacterial suspensions. Several peaks in the range of 250−1500 *m*/*z* were observed to contribute to variations in the spectral information among the species. Application of cluster analysis to the spectral data showed that this ESI-MS technique was capable of discriminating strains of the species *B. subtilis*. This investigation demonstrates the feasibility of measuring liquid samples with minimal sample preparation that can be useful for discrimination at the subspecies level. A change in the cone potential in the electrospray ion source was found to influence the spectral information of representative strains of all of the seven species tested. This has implications with respect to optimizing the experimental conditions for discriminatory purposes, but could offer additional information with respect to microbial characterization. A comparison of the spectra of whole cell suspensions and cell-free supernatants showed a high degree of similarity between the two, which has consequences for ease of automation.

Rapid identification of microorganisms is critical in detecting biological hazards in food and the environment[1] and, of course, in medical microbiology[2]. Information on the physiological traits of microbial cells can also prove useful in assessing the concentration of biotechnologically relevant compounds that they may be producing or may be induced to produce.[3,4] Current techniques for microbial identification in routine operation require considerable time and effort. There is, therefore, a continuing need to develop methods for the identification and characterization of microorganisms that are rapid, reproducible, easy to operate, amenable to automation, and capable of high resolution.

Mass spectrometry has the capability to resolve rapidly individual components in a mixture based on the structural characterization specific to the component analyzed and is, therefore, well-suited to the analysis of complex biological systems. Mass spectrometric methods have, thus, been evaluated and shown to offer excellent potential for chemotaxonomy of microorganisms. Early studies dealt with the application of fast atom bombardment, laser desorption, plasma desorption,and pyrolysis mass spectrometry.[5,6] These techniques essentially provide information that relates to low-molecular-weight metabolites that could be used as potential biomarkers for chemotaxonomic purposes. However, the advent of "soft" ionization techniques has meant that larger macromolecules, such as proteins, can be measured intact, enabling a wider scope for the application of mass spectrometry within the rapid characterization of complex biological systems.[7,8]

Recently, there has been an explosion of interest in the use of soft ionization methods, such as matrix-assisted laser desorption ionization (MALDI) and electrospray ionization (ESI), for the mass spectrometric (MS) analysis of biomacromolecules, and the techniques are now important tools in functional genomics.[9,10] In particular, MALDI-MS is widely used for peptide and protein fingerprinting,[9] and ESI-MS, often coupled with liquid chromatography (LC), for cellular metabolite profiles.[11] Application of these techniques to the characterization of bacteria is also being widely investigated.[12] ESI-MS has typically been carried out by the analysis of isolated cellular components such as proteins, lipids, and nucleic acids.[13,14] By contrast, MALDI-MS has been shown to be useful in obtaining spectral patterns from whole

* Corresponding author. Phone and Fax: +44 (0)1970 621947. E-mail: rrg@aber.ac.uk.

[†] Institute of Biological Sciences.

[‡] Department of Computer Sciences.

(1) Altekruse, S. F.; Cohen, M. L.; Swerdlow, D. L. *Emerging Infect. Dis.* **1997**, *3*, 285−293.
(2) Magee, J. T. In *Topley and Wilson's Microbiology and Microbial Infections*; Collier, L., Ed., 2000, pp 45−64.
(3) Kell, D. B. *J. Theor. Biol.* **1991**, *152*, 49−51.
(4) Vaidyanathan, S.; Macaloney, G.; Vaughan, J.; McNeil, B.; Harvey, L. M. *Crit. Rev. Biotechnol.* **1999**, *19*, 277−316.
(5) Magee, J. T. In *Handbook of New Bacterial Systematics*; Goodfellow, M., O'Donnell, A. G., Eds.; Academic Press: London, 1993; pp 383−427.
(6) Fenselau, C. *Mass Spectrometry for the Characterization of Microorganisms*; American Chemical Society: Washington, DC, 1994.
(7) Siuzdak, G. *Mass Spectrometry for Biotechnology*; Academic Press: London, 1996.
(8) Gaskell, S. J. *J. Mass Spectrom.* **1997**, *32*, 677−688.
(9) Roepstorff, P. *Curr. Opin. Biotechnol.* **1997**, *8*, 6−13.
(10) Yates, J. R. *Trends Genet.* **2000**, *16*, 5−8.
(11) Snyder, A. P. *ACS Symposium Series* **1996**, *619*, 1−20.
(12) vanBaar, B. L. M. *FEMS Microbiol. Rev.* **2000**, *24*, 193-219.
(13) Fang, J.; Barcelona, M. J. *J. Microbiol. Methods* **1998**, *33*, 23−35.
(14) Krishnamurthy, T.; Davis, M. T.; Stahl, D. C.; Lee, T. D. *Rapid Commun. Mass Spectrom.* **1999**, *13*, 39−49.

10.1021/ac0103524 CCC: $20.00 © 2001 American Chemical Society
Published on Web 08/02/2001

microorganisms[15] and is being extensively investigated for chemotaxonomic applications (for a recent review, see ref 16). However, because MALDI-MS involves the mixing of a chemical matrix with the sample and the pulsed firing of a laser to effect desorption and ionization of the analyte, there are possible concerns that may limit the scope of its general applicability to microbial characterization. For example, (a) the dried sample-matrix preparation is inhomogeneous, requiring the selection of "sweet" spots on the target that would give the best ion formation; (b) the selection of matrix remains empirical; (c) matrix adduct ions are often formed that decrease effective mass resolution; (d) the power of the laser is known to exert a crucial influence, because significant ion formation occurs only within a limited window of the laser power range; and (e) the mass spectra are mostly obtained by positive ion detection, because negative ion formation is often relatively inefficient. In addition, liquid separation methods that would enable better resolution of spectral information are necessarily used in an off-line mode. Finally, and perhaps most significantly, there are still some concerns over the spectral reproducibility of MALDI-MS,[16−19] although Fenselau and colleagues[20] have shown that the variance of observed peaks in *Escherichia coli* from different investigators was within the domain of masses expected from proteomics.

By contrast, ESI-MS does not involve the use of a chemical matrix or laser and is, therefore, devoid of the limitations that these cause. Because the electrospray process is conducted at atmospheric pressure, the ionization of analytes enables on-line combination with separation techniques, such as liquid chromatography (LC), microdialysis, and capillary electrophoresis (CE), with a greater potential for automation.[8,11] In addition, the ease of observing multiply charged species enables the detection of large complex biomolecules (e.g., proteins or nucleic acids) at relatively lower mass ranges and, thus, with better accuracy than MALDI-MS. However, ESI-MS is sensitive to buffers, salts, or detergents that may interfere with the electrospray process, and these have to be carefully controlled.[21] ESI-MS has been shown to be a valuable tool for the reproducible analysis of complex biological samples, but to date the introduction of bacteria has largely been via specific cell fractions or lysates, namely, phospholipids,[13,22] lipopolysaccharides,[23] lipooligosaccharides,[24] muramic acids,[24,25] and proteins[11,14,26]. Moreover, these have usually been presented to the ESI-MS after liquid chromatographic (LC) separation.[13,14,23,25] The possibility of making analysis on liquid samples with little or no pretreatment would make this technique attractive for characterizing microorganisms. Intact ribosomes[27] and viruses[28] have been analyzed using ESI-MS. The ability of ESI-MS to provide information rich spectra in both the positive and negative ion modes *from whole bacterial suspensions* was first reported in 1999[29].

In this investigation, we report the application of direct ESI-MS, without prior analyte separation, for the reproducible discrimination to subspecies level of aerobic endospore-forming bacteria using whole cell suspensions. Furthermore, we examine the influence of the sample cone voltage on the spectral information and compare the spectra of whole cell suspensions with those of cell-free supernatants. Six species of the endospore-forming genus *Bacillus* and one of the genus *Brevibacillus* were studied. Members of the genus *Bacillus* are of considerable importance in the food processing industry and in the preparation of sterile products.[30] They are also of relevance as potential biological warfare agents.[31,32]

EXPERIMENTAL SECTION

Microorganisms and Cultivation. Thirty-six strains of aerobic endospore-forming bacteria,[33] confirmed by polyphasic taxonomic methods to belong to the *Bacillus* and *Brevibacillus* genera, were studied (Table 1). Vegetative cells were collected from axenic cultivations of the 36 strains on LabM blood agar base plates and incubated at 37 °C for 16 h. The biomass was carefully harvested using sterile plastic loops and suspended in sterile Milli-Q water. The cells were washed twice, before resuspending them to a final dry cell concentration of 100−200 ng μL^{-1}. Milli-Q water was used throughout so that the cell suspension was devoid of buffers, salts or detergents that would interfere with the electrospray process.[21]

For identification at the species level, all 36 strains were grown in replicates (between 2 and 5) and spectra were acquired for each replicate. The experiment was performed over a period of six months to randomize variations in the measurements due to possible confounding factors such as instrumental drift. For identification at the subspecies level, the seven strains of *B. subtilis* were grown seven times and the spectra were acquired for each of these replicates.

ESI-MS. ESI-MS was performed using an LCT mass spectrometer, supplied by Micromass Ltd. (Wythenshawe, Manchester, U.K.). Spectra were collected in the positive (ES+) ion mode. The washed microbial cells were suspended in 50% aqueous acetonitrile containing 0.1% formic acid. The cell suspension was then introduced into the LCT mass spectrometer, using a 100-μL-volume Hamilton gastight 7000 series syringe and a Harvard Apparatus pump at a flow rate of 5 μL min^{-1}. The mass spectrometer was operated at a capillary voltage of 3000 V; the

(15) Claydon, M. A.; Davey, S. N.; EdwardsJones, V.; Gordon, D. B. *Nature Biotechnol.* **1996**, *14*, 1584−1586.
(16) Lay, J. O. J. *Trends Anal. Chem.* **2000**, *19*, 507−516.
(17) Dalluge, J. J. *Fresenius' J. Anal. Chem.* **2000**, *366*, 701−711.
(18) Domin, M. A.; Welham, K. J.; Ashton, D. S. *Rapid Commun. Mass Spectrom.* **1999**, *13*, 222−226.
(19) Wang, Z. P.; Russon, L.; Li, L.; Roser, D. C.; Long, S. R. *Rapid Commun. Mass Spectrom.* **1998**, *12*, 456−464.
(20) Demirev, P. A.; Ho, Y. P.; Ryzhov, V.; Fenselau, C. *Anal. Chem.* **1999**, *71*, 2732−2738.
(21) Cole, R. B. *Electrospray Ionization Mass Spectrometry: Fundamentals, Instrumentation and Applications*; Wiley: New York, 1997.
(22) Black, G. E.; Snyder, A. P.; Heroux, K. S. *J. Microbiol. Methods* **1997**, *28*, 187−199.
(23) Li, J.; Thibault, P.; Martin, A.; Richards, J. C.; Wakarchuk, W. W.; van der Wilp, W. *J. Chromatogr. A* **1998**, *817*, 325−336.
(24) Gibson, B. W.; Phillips, N. J.; John, C. M.; Melaugh, W. *ACS Symposium Series* **1994**, *541*, 185−202.
(25) Wunschel, D. S.; Fox, K. F.; Fox, A.; Nagpal, M. L.; Kim, K.; Stewart, G. C.; Shahgholi, M. *J. Chromatogr. A* **1997**, *776*, 205−219.
(26) 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.
(27) Rostom, A. A.; Fucini, P.; Benjamin, D. R.; Juenemann, R.; Nierhaus, K. H.; Hartl, F. U.; Dobson, C. M.; Robinson, C. V. *Proc. Nat. Acad. Sci. U.S.A.* **2000**, *97*, 5185−5190.
(28) Fuerstenau, S. D.; Benner, W. H.; Thomas, J. J.; Brugidou, C.; Bothner, B.; Siuzdak, G. *Angew. Chem., Int. Ed.* **2001**, *40*, 542−544.
(29) Goodacre, R.; Heald, J. K.; Kell, D. B. *FEMS Microbiol. Lett.* **1999**, *176*, 17−24.
(30) Logan, N. A. *J. Med. Microbiol.* **1988**, *25*, 157−165.
(31) Barnaby, W. *The Plague Makers: The Secret World of Biolgoical Warfare*; Vision Paperbacks: London, 1997.
(32) Henderson, D. A. *Emerging Infect. Dis.* **1998**, *4*, 488−492.
(33) Goodacre, R.; Shann, B.; Gilbert, R. J.; Timmins, É. M.; McGovern, A. C.; Alsberg, B. K.; Kell, D. B.; Logan, N. A. *Anal. Chem.* **2000**, *72*, 119−127.

Table 1: The Thirty-six *Bacillus* Strains Studied

S no.	species	strain no.[a]	other strain nos.	identifier on plots
1	*B. sphaericus*	7134^T	DSM 28^T, G 1013	
2		B0219	S2	
3		B0408	F 2610/77	
4		B0769	NRS 592	
5		B1147	1691	
6	*B. subtilis*	B0014^T	DSM 10^T, ATCC 6051^T	e
7		B0044	Norris 1007	g
8		B0098	G 1007, Král, *B. niger*	c
9		B0099	G 1208, *B. niger*	d
10		B0410	F 2666/77	f
11		B0501	M (Malaysia) 79A	b
12		B1382	*B. globigii*	a
13	*B. licheniformis*	B0242	G 307	
14		B0252^T	G 46 (=B0089), NTCC 10341	
15		B0755	NRS 551	
16		B1081	Russell 28	
17		B1379	Schweppes	
18	*Br. laterosporus*	B0043	8/1/15/1 colonial variant	
19		B0115	G 308, NTCT 7579	
20		B0262	B0040 contaminant	
21		B0616	NRS 682, *B. orpheus*	
22	*B. cereus*	B0002^T	DSM 31^T, ATCC 14579, Ford 13	
23		B0550	NRS 721, *B. albolactis*	
24		B0702	V 26	
25		B0712	NRS 201, *B. siamensis*	
26		B0851	isolate 1	
27	*B. amyloliquefaciens*	B0168	Fukumoto F	
28		B0175	Campbell P, ATCC 23844	
29		B0177^T	Campbell F, ATCC 23350^T	
30		B0251	duplicate of B0172	
31		B0620	Harrman 1A35	
32	*B. megaterium*	B0010^T	DSM32^T, ATCC 14581^T	
33		B0056	UB 8/1/5/1	
34		B0057	UB 8/1/5/2	
35		B0076	G 659, NCTC 5637	
36		B0621	Hartman NRRL B-348	

[a] Superscript "T" indicates type strain.

sample and extraction cone voltages were maintained at 40 and 10 V, respectively, unless mentioned otherwise. The source and desolvation temperatures were maintained at 80 °C and 70 °C respectively, and the desolvation and nebulizer gas flow rates were kept at 350 l h^{-1} and 95 l h^{-1}, respectively.

To obtain high mass accuracy, the LCT employs a reflectron time-of-flight (TOF) analyzer, and in this study, the mass range was set to 200−3000 m/z and spectra were collected for typically 2 min; these were then summed. The instrument was calibrated such that nominal mass measurement was achieved over the entire mass range.

Cluster Analyses. Data were exported from the MassLynx program version 3.4 (software provided by Micromass and running under Microsoft Windows NT on an IBM-compatible PC) and imported into Matlab version 5.3 (The MathWorks, Inc., Natick, MA), which also runs under Microsoft Windows NT on an IBM-compatible PC. Before analyses, the spectral data were normalized to percent total ion count, the m/z range from 250 to 1500 was used for the analysis, and the analysis was carried out at unit (amu) resolution.

Cluster analyses were carried out as detailed by us elsewhere.[33−35] The initial stage involved the reduction of the dimensionality of the MS data by principal components analysis (PCA). Discriminant function analysis (DFA) then discriminated between groups on the basis of the retained principal components (PCs) and the a priori knowledge of which spectra were replicates. The number of PCs used were those which accounted for at least 90% of the total spectral variance. Finally, the Euclidean distance between a priori group centers in DFA space was used to construct a similarity measure with the Gower similarity coefficient, S_G, and these distance measures were then processed by an agglomerative clustering algorithm to construct a dendrogram.

Influence of Sample Cone Voltage on the Spectral Information. The ESI mass spectra of representative strains from each of the seven bacterial species (type strains, typically) were acquired at different sample cone voltages from 10 to 150 V in incremental steps of 10 V, with all other conditions being maintained as detailed above. In all cases, the extraction cone voltage was maintained at 10 V. The microorganisms were grown in triplicates, and the spectra were acquired for each replicate at each sample cone voltage. The replicate spectra were averaged and corrected for variance between replicate measurements (within-group variance), using the following formula,

$$\mathbf{S} = A[1 - (\mathbf{V}/A)]$$

where, **S** is the matrix of variance-corrected average spectra, A is

(34) Timmins, É. M.; Howell, S. A.; Alsberg, B. K.; Noble, W. C.; Goodacre, R. *J. Clin. Microbiol.* **1998**, *36*, 367−374.

(35) Goodacre, R.; Timmins, É. M.; Burton, R.; Kaderbhai, N.; Woodward, A. M.; Kell, D. B.; Rooney, P. J. *Microbiology* **1998**, *144*, 1157−1170.

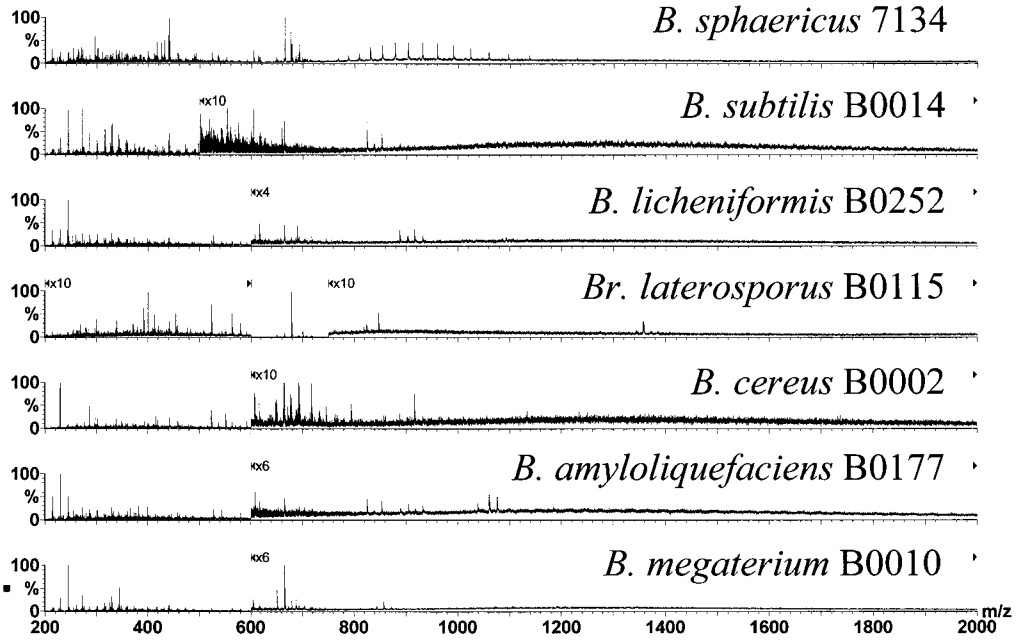

Figure 1. Whole cell suspension positive ion ESI mass spectra of representative strains of the six *Bacillus* species and the single *Brevibacillus* species studied in the 200−2000 *m*/*z* range. Regions of the spectra that are relatively low in intensity are magnified for clarity (the maginfication factor is shown for the regions magnified in each spectrum).

the average spectrum, and **V** is the matrix of variance of the replicate measurements.

The variance-corrected average spectra were then used for assessing changes in the spectral information. Variance was also calculated among the species for each sample cone voltage (between-group variance) using the average spectra that were corrected for the within-group variance. Data analysis of these spectra was carried out in the m/z range of 200−2000.

Comparison of Spectra of Whole Cells and Centrifuged Supernatants. The ESI mass spectra of the same representative strains used above, suspended in the analysis solvent (50% acetonitrile in water containing 0.1% formic acid) were acquired in triplicate at three representative sample cone voltages, namely, 40, 80, and 120 V, all other conditions being maintained as detailed above. The suspensions were centrifuged to remove cells/cell debris, and the spectra of the supernatants were then acquired at the three representative cone voltages for each representative strain, also in triplicate. Cluster analysis, as detailed above, was used to compare the spectral information between the two treatments.

RESULTS AND DISCUSSION

Raw Data. Simple visual inspection of the mass spectra of type strains from the seven species shows obvious differences in the spectral pattern between the different species of bacilli in the 200−2000 m/z range (Figure 1). Some regions of the spectra have been magnified to enable visualization of peaks that would otherwise be unnoticeable in the presence of other dominant ones. It is clear that these spectra are information-rich, and although analytes are detected across the whole mass range, the majority of them appear below m/z 700. Some peaks can be seen to be conserved in all of the spectra. For instance, analyte(s) at m/z values of 440, 523, and 665 are seen in all seven species; conversely, there are regions where the spectral information is apparently unique to a particular species. A prominent example is the charge distribution observed in the mass spectrum of *Bacillus sphaericus* between m/z 700 and 1200. Another clear example is that of the peaks in the m/z range 1000−1100 in the spectrum of *Bacillus amyloliquefaciens*. It is noteworthy that although the concentration of the biomass was kept identical for all the microorganisms that were analyzed, the signal intensity (total ion counts, TIC) for *Brevibacillus laterosporus* was observed to be relatively higher, by an order of magnitude, because a strong signal at an m/z of 679 can be seen to dominate the *Br. laterosporus* spectrum. It is possible that this analyte is more efficiently ionized by the electrospray process and may be quenching the ionization of other analytes.

The charge distribution observed in the m/z range 700−1200 in the *B. sphaericus* spectrum is likely to be due to a large macromolecule, such as a protein. To estimate the molecular weight of this macromolecule from the mass spectrum, the maximum entropy routine[36,37] supplied by Micromass was used to calculate the charges on this distribution, and from this, the molecular weight of the macromolecule was estimated to be ∼32 kDa (Figure 2A). To confirm this, MALDI-MS was employed on the same bacterial sample, and a peak around 32 kDa was observed (data not shown). *B. sphaericus* are known to carry an S-layer protein coat on their surface.[38] S-layers represent the outermost cell envelope component of many bacteria and archaea and are composed of regularly repeated units of a single protein or glycoprotein, usually in the molecular weight range 40 to 200 kDa.[39] S-layer protein components of 30−40 kDa have also been

(36) Skilling, J. In *Maximum Entropy and Bayesian Methods*; Skilling, J., Ed.; Kluwer Academic Publishers: Dordtrecht, 1989; pp 45−52.
(37) Skilling, J. In *Maximum Entropy and Bayesian Methods*; Erickson, G., Rychert, J. T., Smith, C. R., Eds.; Kluwer Academic Publishers: Dordtrecht, 1998; p 14.
(38) Ilk, N.; Kosma, P.; Puchberger, M.; Egelseer, E. M.; Mayer, H. F.; Sleytr, U. B.; Sara, M. *J. Bacteriol.* **1999**, *181*, 7643−7646.
(39) Sleytr, U. B.; Beveridge, T. J. *Trends Microbiol.* **1999**, *7*, 253−260.

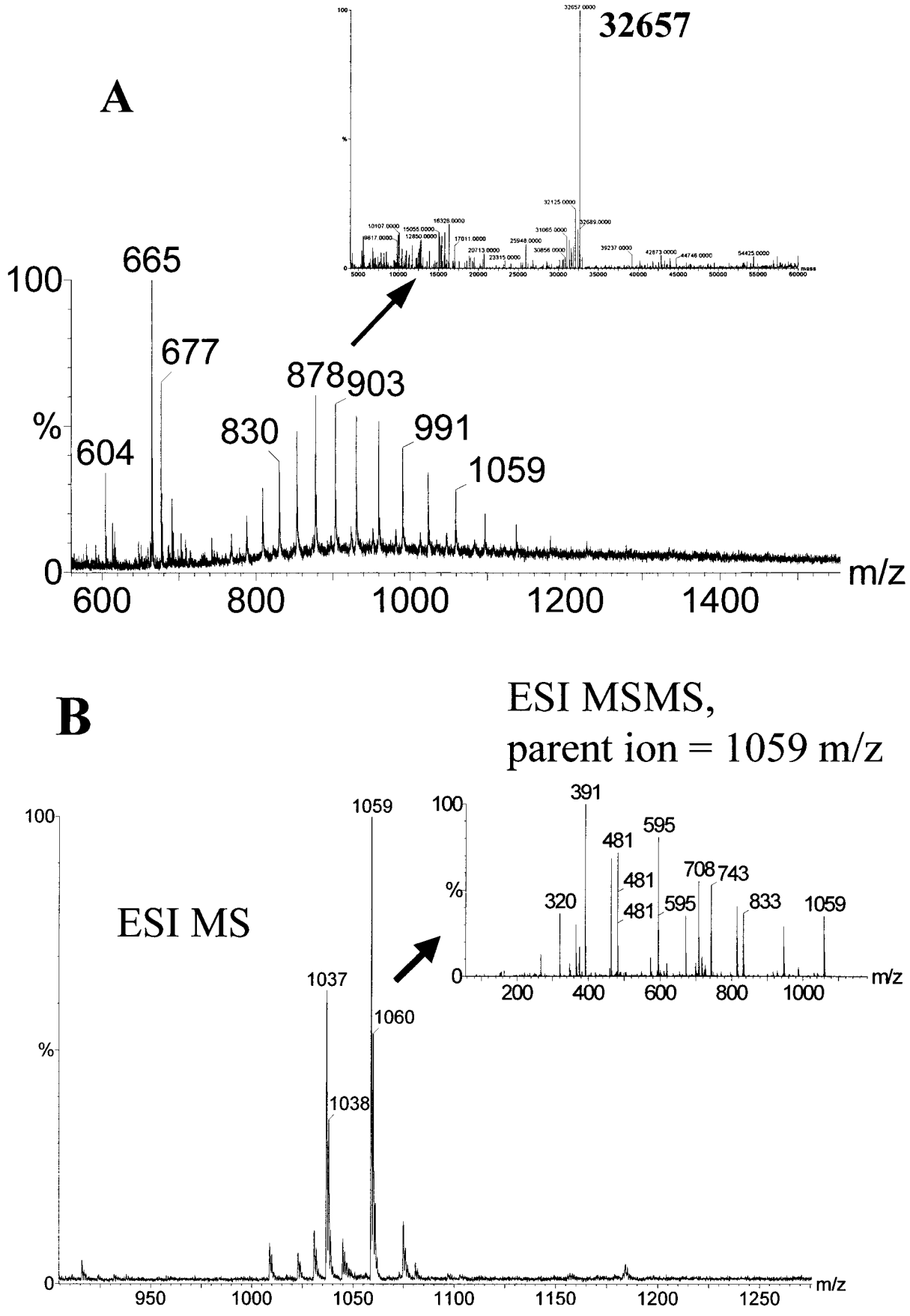

Figure 2. (A) Spectrum of *B. sphaericus* showing the charge distribution and the results of calculating the molecular weight of the molecule, using maximum entropy, in the inset. (B) Spectrum of *B. amyloliquefaciens* showing the prominent peaks in the 1000−1100 *m*/*z* region. The daughter ions on tandem mass spectrometric fragmentation of the parent ion at 1059 *m*/*z* are shown in the inset.

reported.[38,40] It is, therefore, possible that the characteristic charge distribution observed in the spectrum of this species is that of an S-layer protein. Bacterial S-layer proteins have been characterized using ESI mass spectra,[40] but only after isolating the relevant protein(s), and using extensive sample preparation steps. If the observed charge distribution is that of an S-layer protein, it is noteworthy that such proteins can be observed in ESI-MS, even with minimal sample preparation, as in the present study.

We attempted to characterize the prominent peak in the *B. amyloliquefaciens* spectrum in the type strain, at m/z 1059, using tandem mass spectrometry. An ESI-quadrupole-time-of-flight MS instrument (QToF Micromass Ltd., Manchester, U.K.) was used for the purpose operating in a manner very similar to that used above. Argon was employed as the collision gas at an energy of 60 eV. The resultant daughter ion spectrum is shown in Figure 2B. Although no definitive assignment could be made on the origins of the peak at m/z 1059, the spectral pattern of the MS/MS spectrum is suggestive of a peptide, or a lipopeptide, because lipopeptides of similar mass units have been observed with MALDI-MS analysis of *B. subtilis*, albeit on cell lysates.[41] Observation of such product ion (MS/MS) patterns can improve specificity and sensitivity of the analysis and provide additional structural information, increasing the scope of bacterial identification, perhaps even in crude samples. Indeed, ESI-MS/MS has been shown to be informative and potentially useful for microbial characterization,[42] albeit on cell lysates.

Discrimination at the Species Level Using Cluster Analyses. One approach to the assessment of the spectral information for discriminatory purposes is the use of PCA using the entire spectral information to extract spectral variance and perform DFA on the PC scores prior to hierarchical cluster analysis (HCA). Such an approach has been successfully employed for the analysis of PyMS, FT-IR, and Raman spectra,[33−35] and the same methodology was adopted in the present study. The data analysis was performed on MS spectra of single *amu* resolution and the most information-rich m/z range of 250−1500 was used. The number of PCs to include for DFA was based on the percent explained variance such that at least 90% of the variance in the data was used (typically, the first 15 PCs were employed). The first 3 dimensions in DFA space were then used for HCA.

Figure 3 summarizes the results of these analyses performed with a priori knowledge based on strain numbers (i.e., 36 groups) rather than on the species (this would obviously bias the results and artificially force species separation). When the analysis was done on all the samples, three clusters resulted: one comprised all of the *B. sphaericus* strains, another all of the *Br. laterosporus* samples, and the third large cluster consisted of the rest of the bacilli. The spectra of *B. sphaericus* and *Br. laterosporus* show prominent signals that differentiate them from those of the others (Figure 1), and these large differences reflect themselves in the clustering pattern; it is likely that these differences skew the PC scores in the direction of these two groups, making it difficult to observe the differences due to the other species that were

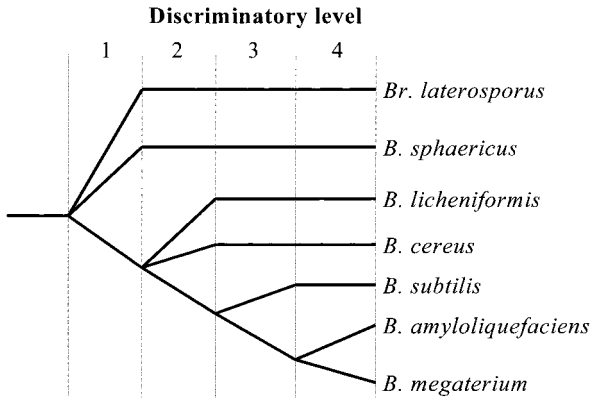

Figure 3. Sequential clustering pattern observed on data analysis of the different *Bacillus* strains studied using a priori knowledge of strain numbers. The clustering pattern at the four discriminatory levels is shown. Details of the analysis are given in the text.

analyzed. When the analysis was carried out again excluding the *B. sphaericus* and *Br. laterosporus* spectra, the *Bacillus licheniformis* strains and the *Bacillus cereus* strains were separated out into two groups, and the third group comprised the remaining bacilli (*B. amyloliquefaciens*, *Bacillus megaterium*, and *B. subtilis*). On analysis of just this group *B. subtilis* was recovered. *B. amyloliquefaciens* and *B. megaterium* formed a tighter cluster that was separated only when analyzed in isolation from the spectra from the other five species.

These *Bacillus* species have previously been subjected to the API biochemical tests.[43] *B. cereus* belongs to group I; *Br. laterosporus*, to group II; *B. sphaericus*, to group III; and *B. subtilis*, *B. licheniformis*, *B. amyloliquefaciens*, and *B. megaterium* were grouped together in group IV, although phylogenetic analysis on the 16S rDNA sequences[44] reveals five clusters that comprise (1) *B. sphaericus*, (2) *Br. laterosporus*, (3) *B. subtilis*, *B. licheniformis*, and *B. amyloliquefaciens*, (4) *B. megaterium*, and (5) *B. cereus*. As is to be expected, the species level discrimination that we observe using ESI-MS (Figure 3) is more similar to that obtained from the biochemical characterizations than those from genetic analysis.

Discrimination at Subspecies Level. To study the fine discrimination capability of the ESI-MS technique, the seven strains of *B. subtilis* were studied further. Each species was cultured seven times and a single replicate from each culture was analyzed. To challenge the analysis, the replicate measurements from the seven strains of *B. subtilis* were divided such that there were two groups from each strain; thus, DFA was performed with a priori knowledge that there were 14 groups to discriminate. Figure 4 shows the DFA and HCA dendrogram from this analysis, and it can be clearly seen that all of the strains were successfully recovered together and that seven different groups were formed. The fact that the duplicate groups cluster together indicates that the spectral information differs reproducibly, even at the subspecies level.

(40) Mauri, P. L.; Pietta, P. G.; Maggioni, A.; Cerquetti, M.; Sebastianelli, A.; Mastrantonio, P. *Rapid Commun. Mass Spectrom.* **1999**, *13*, 695−703.
(41) Leenders, F.; Stein, T. H.; Kablitz, B.; Franke, P.; Vater, J. *Rapid Commun. in Mass Spectrom.* **1999**, *13*, 943−949.
(42) Xiang, F.; Anderson, G. A.; Veenstra, T. D.; Lipton, M. S.; Smith, R. D. *Anal. Chem.* **2000**, *72*, 2475−2481.
(43) Logan, N. A.; Berkeley, R. C. W. *J. Gen. Microbiol.* **1984**, *130*.
(44) Goodacre, R.; Shann, B.; Gilbert, R. J.; Timmins, É. M.; McGovern, A. C.; Alsberg, B. K.; Logan, N. A.; Kell, D. B. Proceedings of the 1997 ERDEC Scientific Conference on Chemical and Biological Defense Research; Aberdeen Proving Ground: 1998; pp 257−265.

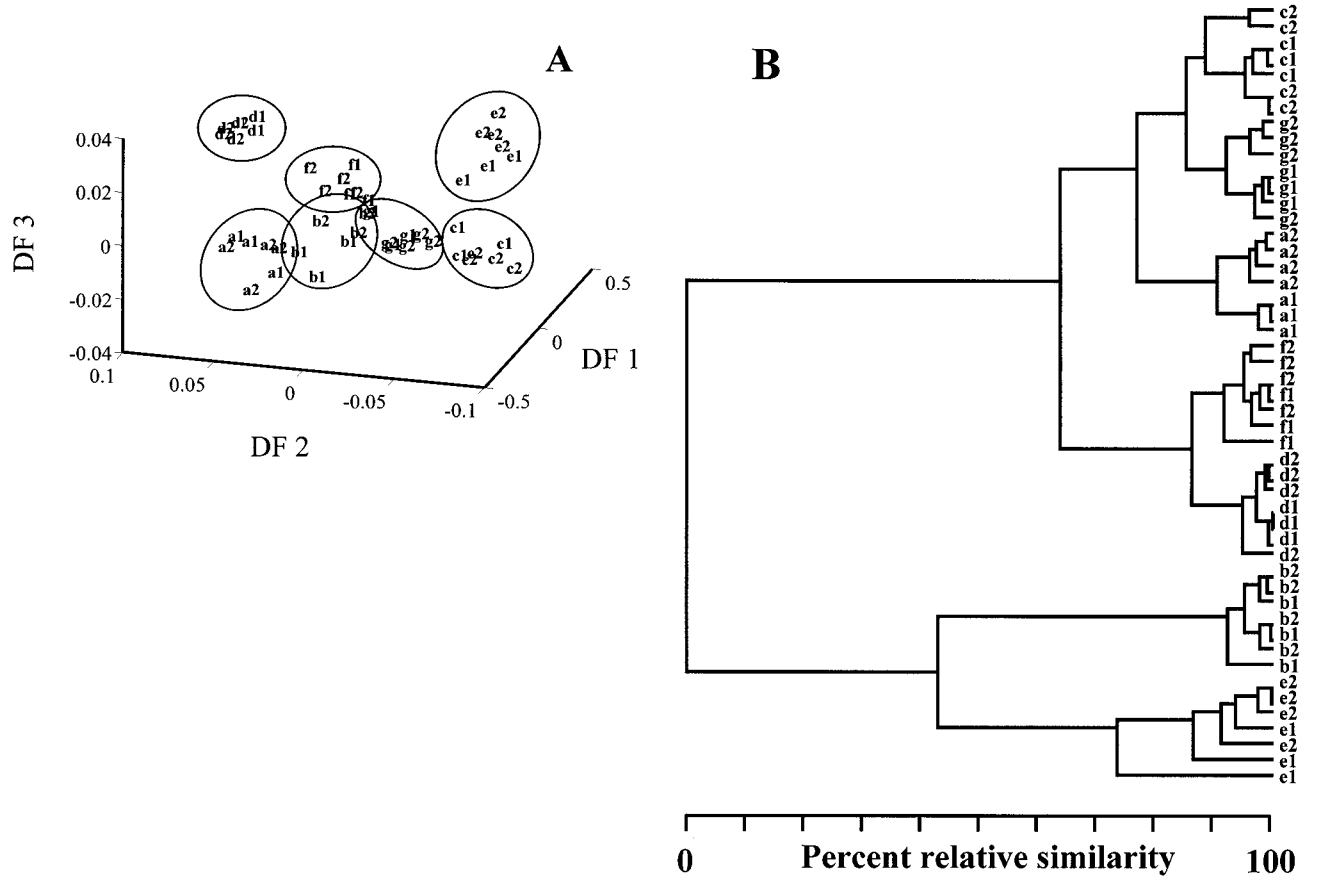

Figure 4. Discrimination at subspecies level: (A) a pseudo-3D plot of the first three discriminant functions and (B) dendrogram based on the positive ion ESI-MS spectra (seven replicates) of the seven *B. subtilis* strains. The seven replicates of each strain were divided into two groups (suffixed 1 and 2) and analyzed as detailed in the text. The strains can be identified from Table 1.

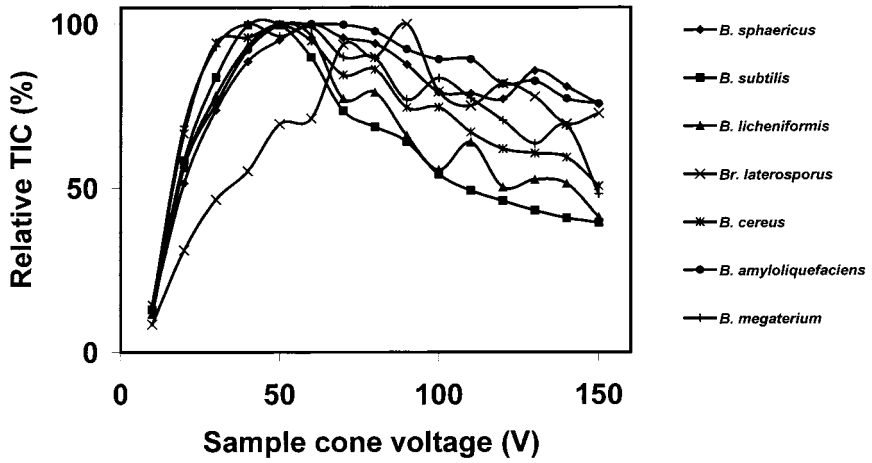

Figure 5. Influence of sample cone voltage on the percent relative total ion counts (TIC) for the representative strains of the seven bacterial species. Each data point is the mean of triplicate measurements.

Although good discrimination is observed, on the basis of the overall spectral pattern, further studies are clearly required to identify and ascertain the biomarkers that are responsible for the discrimination and to assess the conditions under which such biomarkers can be detected. However, these results suggest that the ESI-MS spectra of bacterial suspensions can provide discriminatory information, even with minimal sample preparation, and can thus be potentially very useful for identification purposes.

Influence of Sample Cone Voltage on the Spectral Information. In ESI-MS, the pressure differential necessary for the combination of ionization at atmospheric pressure with the low pressures required for mass analysis and ion detection is usually achieved with a sample cone/extraction (skimmer) cone arrangement.[21] Potentials applied between the two cones serve both to focus the ion beam through the intermediate pressure region between the sample and skimmer cones and to accelerate ions through this intermediate pressure region. Investigations on chemical analysis have shown that raising the cone potential improves the signal-to-noise ratios, which has been attributed to improved ion focusing,[45] and that changing the capillary/skimmer

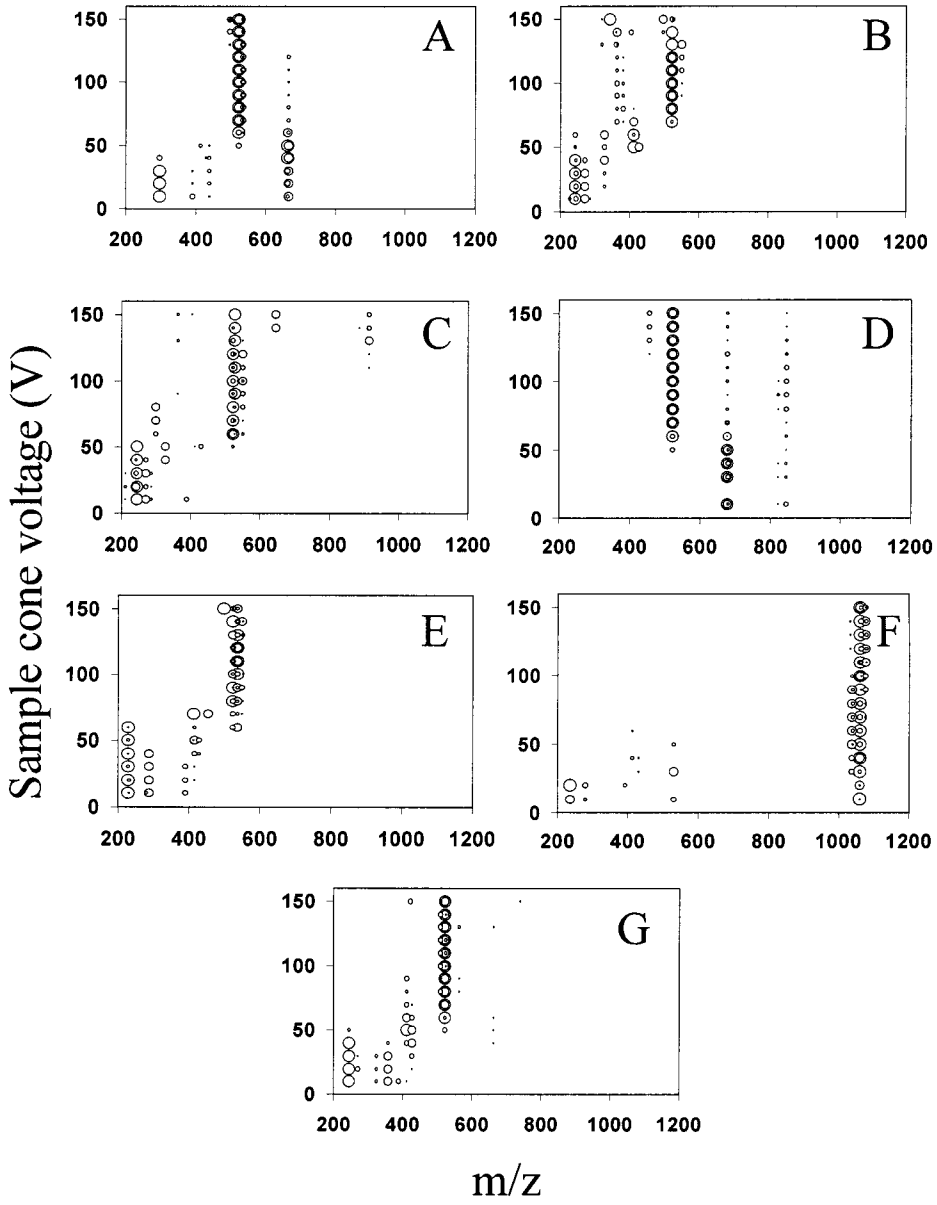

Figure 6. Influence of sample cone voltage on the spectral information for the representative strains of each of the seven bacterial species investigated: A, *B. sphaericus*; B, *B. subtilis*; C, *B. licheniformis*; D, *Br. laterosporus*; E, *B. cereus*; F, *B. amyloliquefaciens*; and G, *B. megaterium*. The first five dominant peaks for each spectrum is plotted at each cone voltage and for each bacterial species. Circles in the plots correspond to the dominant peaks, with their sizes corresponding to their relative intensities, within each spectrum.

potential difference causes collision-induced dissociation by increasing the energy of collisions with background gas.[46] Moreover, shifts in the charge-state distributions of proteins at raised cone potentials have also been reported.[47,48] The influence of cone potential on the microbial spectral information was, therefore, studied by acquiring mass spectra at different sample cone voltages (SCVs) while keeping the extraction cone voltage the same (10 V). Figure 5 shows the relative total ion counts (rTIC) for the seven representative species studied, normalized as a percent of the maximum for each species over all of the SCVs. It is clear that with the exception of *Br. laterosporus*, the rest of the microorganisms give maximum ion counts at a SCV of 30−70 V.

The influence of sample cone voltage (from 10 to 150 in 10 V steps) on the five most prominent analytes (m/z peaks), for each of the species is depicted in Figure 6. The spectra are derived from the average spectra of triplicate acquisitions, corrected for variance among replicate acquisitions as detailed in the Experimental Section. The sizes of the circles indicate the relative dominance of each peak within the spectrum; the larger the circle, the more dominant the signal. For example, the highest peak intensity in the *B. sphaericus* spectrum (Figure 6A), at a sample cone voltage of 40 V is the signal at 665 m/z, with peaks at 440 and 296 m/z being less dominant.

(45) Crescenzi, C.; Dicorcia, A.; Marchese, S.; Samperi, R. *Anal. Chem.* **1995**, *67*, 1968−1975.
(46) Voyskner, R. D.; Pack, T. *Rapid Commun. Mass Spectrom.* **1991**, *5*, 263−268.
(47) Ashton, D. S.; Beddel, C. R.; Cooper, D. J.; Green, B. N.; Oliver, R. W. A. *Org. Mass Spectrom.* **1993**, *28*, 721−728.
(48) Hunt, S. M.; Sheil, M. M.; Belov, M.; Derrick, P. J. *Anal. Chem.* **1998**, *70*, 1812−1822.

Overall, it can be seen that the spectral information alters with changes in the SCV for all of the seven species tested. Note that in some cases, the dominant peaks in a given bacterial spectrum change with the cone voltage. For instance, the most dominant peak in the *B. sphaericus* spectrum, at 665 *m*/*z*, at lower cone voltages decreases in intensity and loses out in dominance to the peak at 523 *m*/*z* at higher cone voltages. This indicates that the cone potential does have a significant influence on the spectral information. Signals at the lower end of the spectrum are more dominant at lower cone voltages than at higher SCVs; therefore, increasing the cone potential increases the chances of observing ions with a higher *m*/*z*, as was noted by Hunt et al.,[48] for polymer analysis. However, the opposite is seen for *Br. laterosporus* (Figure 6D), for which the signal at 523 *m*/*z* is absent at lower SCVs but starts to dominate the spectrum once the SCV is increased and the signal at 679 *m*/*z* decreases in intensity.

It was also observed that the charge-state distribution in the *B. sphaericus* spectrum (Figure 2A), characteristic of a protein, shifted to higher *m*/*z* values as the cone potential was increased but disappeared beyond 80 V (data not shown). Other workers have also reported on shifts in charge-state distributions[47,48] and increased ion dissociation[46] with increase in cone potential in the analysis of polymers and biopolymers.

Given that different spectra were observed at different SCVs, it is important to assess the effect this may have on the discriminatory ability of ESI-MS. The variance spectra were calculated for each of the seven representative species as detailed above, and the between-species variance spectra for representative cone voltages were computed and plotted in Figure 7. In agreement with our above observations, it can be seen that the spectral variance is higher in the lower end of the spectrum at lower cone voltages and that the variance increases at the higher end of the spectrum as the cone voltage increases, except for the signal at 523 *m*/*z*. In comparison with Figure 6, it can also be noted that the variance at 523 *m*/*z* is likely to be due to a quantitative difference in the relative signal intensities between the species. It should be remembered that the spectral peaks were normalized with respect to the total intensity for each spectrum. This suggests that once identified, the relative levels of specific analytes (relative peak intensities in the ESI-MS spectrum) may be an additional factor that could be useful for microbial characterization.

Changes in the spectral information, influenced by changes in the cone potential, indicate that greater effort may be required in optimizing "generally applicable ESI-MS conditions" so that the spectral patterns can be employed for routine discrimination purposes. Once established, then the application of supervised, intelligent machine-learning approaches[49] can be useful for such purposes, and this will be a subject of further investigation. By contrast, it is also true that there is greater scope for using spectral information from a combination of different cone voltages on the same microorganism, because these differences may enhance the discriminatory ability of this technique, and this will also be an area for future study.

Comparison of the Spectra of Whole Cells and Centrifuged Supernatants. Suspension of microbial cells in a solvent such as acidic aqueous acetonitrile (as employed here) is a procedure used routinely in "soft" ionization mass spectrometric investiga-

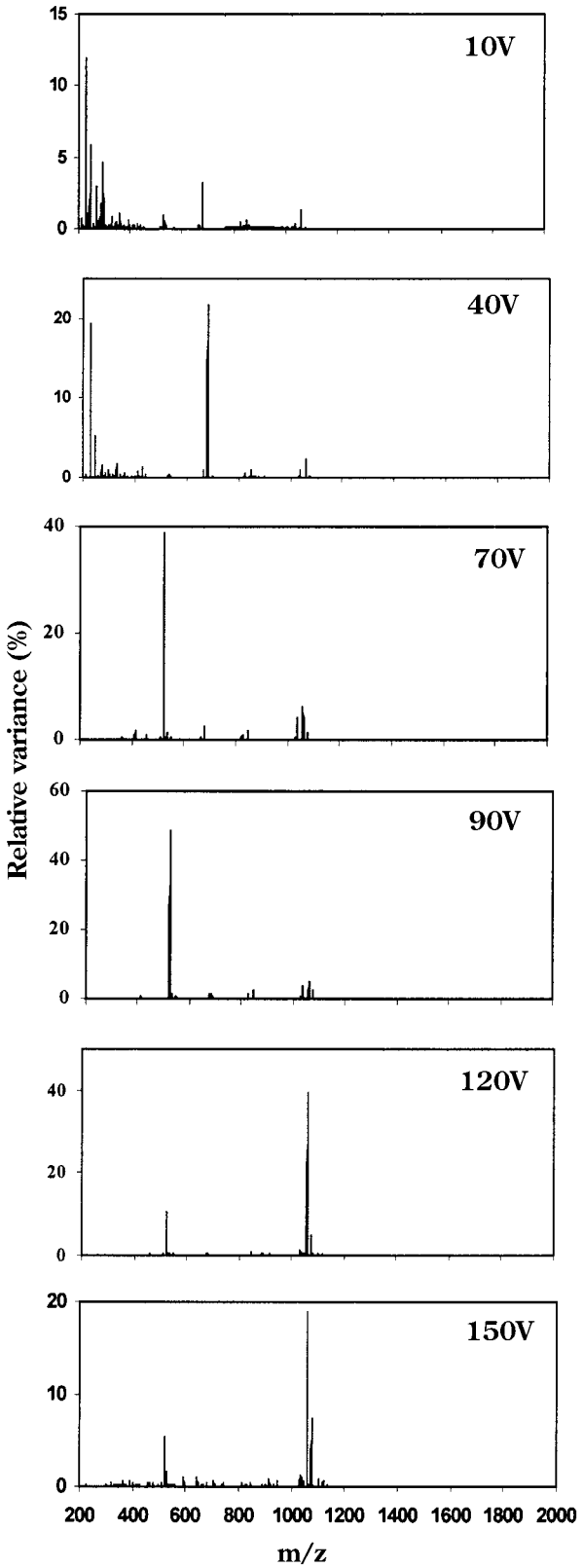

Figure 7. The relative spectral variance between the representative strains of the seven bacterial species, at six representative cone voltages. The variance calculated is relative to the total variance for the entire spectrum, at the given cone voltage.

(49) Beavis, R. C.; Colby, S. M.; Goodacre, R.; Harrington, P. B.; Reilly, J. P.; Sokolow, S.; Wilkerson, C. W. In *Encyclopedia of Analytical Chemistry*; Meyers, R. A., Ed.; John Wiley & Sons, Ltd: Chichester, 2000; pp 11558−115597.

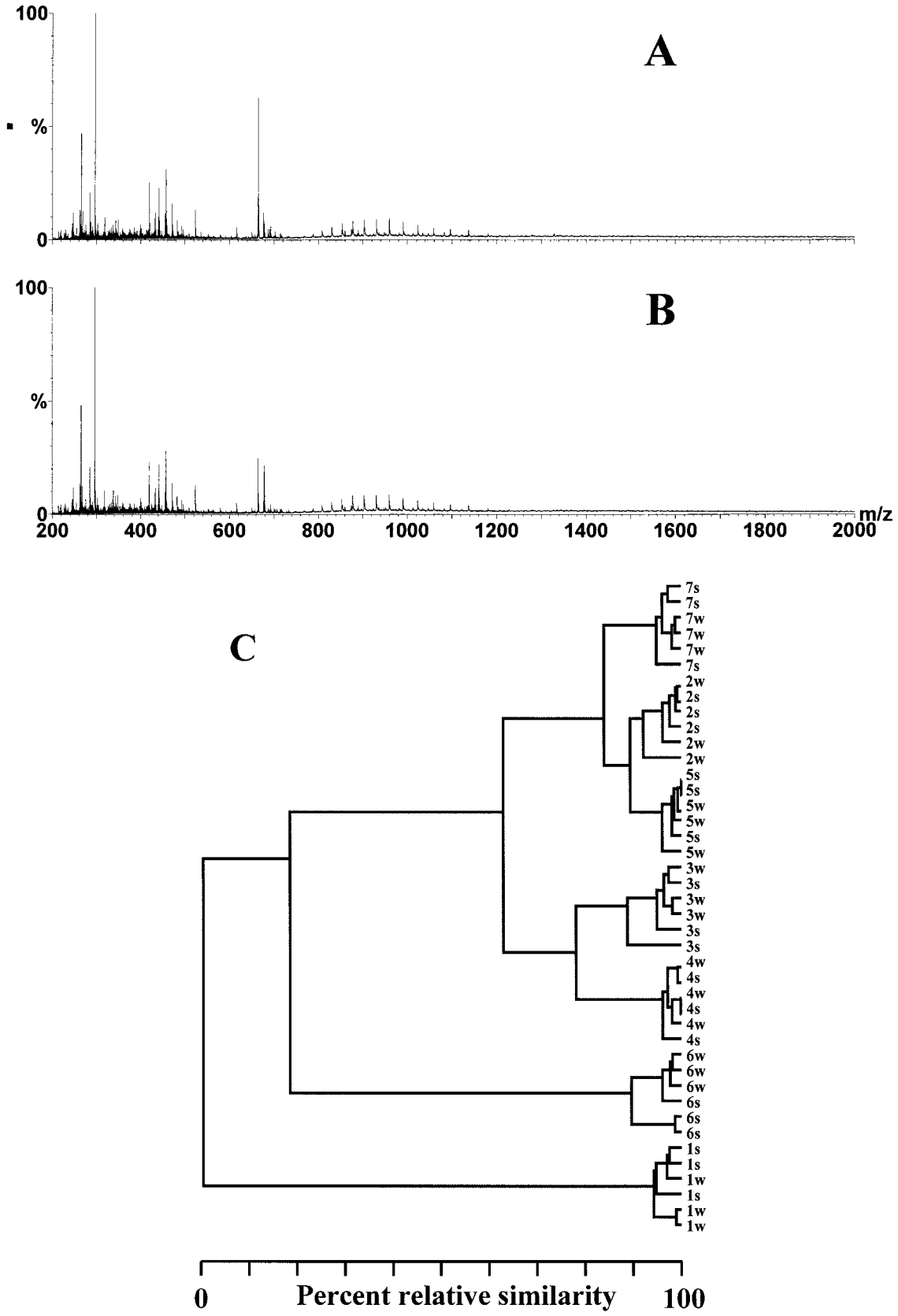

Figure 8. Representative (*B. sphaericus* 7134) positive ion ESI-MS spectra of (A) whole cell suspension and (B) cell-free supernatant. (C) The HCA dendrogram derived form the spectra of the representative strains of the seven bacterial species studied (1, *B. sphaericus*; 2, *B. subtilis*; 3, *B. licheniformis*; 4, *Br. laterosporus*; 5, *B. cereus*; 6, *B. amyloliquefaciens*; and 7, *B. megaterium*) at an SCV of 40 V; w, whole cell suspension; s, cell-free supernatant.

tions of microorganisms. This is particularly the case in MALDI-MS studies, in which whole cell suspensions have been analyzed directly without separating the cellular residue. By contrast, ESI-MS is usually carried out using cell-free supernatants after analyte

separation by LC. Some workers[50] report that partial lysis of the cells occurs as a result of the acidic conditons employed in such techniques and that this results in the release of proteins and peptides from the cells that are being detected on MS analysis. Others,[18,51] however, have shown that there is little evidence of cellular damage by lysis, even after such a treatment, and that the signals observed are that of cell wall or membrane associated proteins.[52,53] Although the origins of the spectral contributors in such techniques largely remain an issue for further investigation, the approach of suspending the cells in weak acidic solvents offers a more rapid method, as compared to the laborious, elaborate approach of cell lysis. Note, of course, that a lysis-based method often results in intracellular salts leaking into the solvent medium, thus necessitating a salt removal step prior to mass spectrometry. The analysis of bacterial suspensions in weakly acidic solvents is, therefore, an attractive option for rapid microbial characterization using mass spectrometry.

To compare whole cell suspensions with cell free extracts, we acquired the spectra of the seven representative strains by direct injection ESI-MS at three different sample cone voltages (40, 80, and 120 V) by both approaches. Figure 8 shows representative spectra from whole cell suspensions (Figure 8A) and supernatants (Figure 8B), and it is evident from visible inspection that these are, indeed, very similar. A more comprehensive indication of these similarites can be obtained from the HCA analysis of the spectra. The dendrogram obtained (Figure 8C) from the spectra collected at an SCV of 40 V shows that the spectra of whole cell suspensions (w) and cell-free supernatants (s) are very similar for all of the seven species tested; the similarity index for each of the seven groups is greater than 90%. Very similar results were found for the other two cone voltages tested (data not shown). These results clearly suggest that the spectra of whole cell suspensions are constituted chiefly of material leached out into the suspended solvents and that the cells themselves do not contribute much to the spectral information. This finding has implications with respect to the automation of analysis, because cell-free extracts are more likely to be stable over longer periods of analysis, as compared to whole-cell suspensions.[14]

CONCLUSIONS

This investigation demonstrates that the spectral information available from direct injection ESI-MS of whole cell suspensions without prior analyte separation may be highly informative and can be useful in microbial identification with discrimination capability at the subspecies level. As with all ESI-MS methods, the spectral information is influenced by the cone potential that is employed to extract the ions into the analyzer. Cone potential is only one of the several instrumental parameters that could influence the spectral information. There is, therefore, a requirement for identifying and optimizing influential parameters, such as the cone potential, before the technique can be applied routinely for discriminatory purposes. However, the combined spectral pattern of bacterial suspensions at different cone potentials can offer additional information that can be useful for devising a sensitive method for microbial identification and characterization. In addition, it is important to investigate the influence of microbial growth conditions on the spectral information and optimize sample preparation conditions in order to arrive at an optimal strategy for obtaining reproducible strain-specific spectral information with minimal sample preparation. These aspects are presently being investigated. The biochemical contributors to the spectra and the sensitivity of the signals to physiological variations are also subjects of further investigation. Finally, this study indicates that the spectra of whole cell suspensions are very similar to those of cell-free supernatants, which has implications for automation and high-throughput screening of microbial cultures by metabolite profiling.

ACKNOWLEDGMENT

We thank the Engineering and Biological Systems Committee of the UK BBSRC for financial support. We are indebted to Mr. Jim Heald for his advice with ESI-MS and are very grateful to Dr. Niall Logan for providing us with the bacterial strains.

Received for review March 27, 2001. Accepted June 15, 2001.

AC0103524

(50) Krishnamurthy, T.; Rajamani, U.; Ross, P. L.; Jabhour, R.; Nair, H.; Eng, J.; Yates, J.; Davis, M. T.; Stahl, D. C.; Lee, T. D. *J. Toxicol., Toxin Rev.* **2000**, *19*, 95−117.

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

(52) Dai, Y.; Li, L.; Roser, D. C.; Long, S. R. *Rapid Commun. Mass Spectrom.* **1999**, *13*, 73−78.

(53) 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.