ELSEVIER

Journal of Chromatography A, 925 (2001) 123–132

JOURNAL OF CHROMATOGRAPHY A

www.elsevier.com/locate/chroma

Automated mass analysis of low-molecular-mass bacterial proteome by liquid chromatography–electrospray ionization mass spectrometry

Kevin Y. Dunlop, Liang Li*

Department of Chemistry, University of Alberta, Edmonton, Alberta, Canada T6G 2G2

Received 7 March 2001; received in revised form 28 May 2001; accepted 6 June 2001

Abstract

Due to the presence of a large number of proteins in cell extracts, ion chromatograms of cell extracts obtained by electrospray ionization mass spectrometry (ESI-MS) can be quite complicated. It is found that the elevated baseline in an ion chromatogram contains many protein signals. One deficiency of current commercially available LC–ESI-MS data interpretation software is found to be the lack of functional operation that allows automated mass spectral integration and interpretation over signals hidden in the baseline. This current limitation can be overcome by a technique that involves the introduction of artificial pulses to an ion chromatogram by removing the solvent mixer in the HPLC pump. These artificial pulses are treated as chromatographic peaks by the software, thereby allowing automated spectral integration over the duration of a pulse. The reliability of mass analysis from the integrated spectra is shown to be dependent on spectral interpretation parameters such as mass spectral baseline threshold. The application of this method is demonstrated for rapid detection and mass analysis of low-molecular-mass proteins from cell extracts of *Escherichia coli* or *Bacillus globigii*.

Keywords: Bacteria; *Escherichia coli*; *Bacillus globigii*; Proteins

1. Introduction

With the advent of mass spectrometric ionization techniques that enable ionization of intact biological molecules, bacterial proteins are becoming a major focus of study for bacteria identification. An average bacterial cell contains approximately 10% lipids and more than 50% proteins in dried mass [1]. As a result, proteins exhibit great potential for use as biomarkers for sensitive detection. Considering the large differences found in genomes of a number of microorganisms, the bacterial proteome (i.e., proteins expressed by the genome of a bacteria) is expected to be greatly different from each other. Comparative analysis of a sufficient number of proteins or a subset of proteome, instead of the entire proteome, is expected to be adequate for unambiguous identification of bacteria. One approach to identifying bacteria by mass spectrometry (MS) is to measure the masses of a set of proteins from a bacterium, followed by correlation of these masses with those in a proteome database. The correlation results can be statistically analyzed to arrive at a unique match with a known bacterium in the database. For protein mass measure-

*Corresponding author. Tel.: +1-780-492-3254; fax: +1-780-492-8231.

E-mail address: liang.li@ualberta.ca (L. Li).

0021-9673/01/$ – see front matter
PII: S0021-9673(01)01030-5

ment, matrix-assisted laser desorption/ionization time-of-flight (MALDI-TOF) MS has been used for the analysis of either intact bacterial cells or cell lysates [2–16]. In addition, LC–electrospray ionization (ESI)-MS has also been recently applied for the analysis of low-mass bacterial proteins [17–19].

Although current instrumentation in LC–ESI-MS is able to rapidly produce an ion chromatogram, there is little information gained from the chromatogram alone for bacterial proteome analysis. We need to extract mass information from the ion chromatogram. Specifically, it is necessary to average the individual mass spectra produced over a selected period in the chromatogram, followed by deconvolution of the multiply charged ion peaks found in the averaged spectrum and then determination of the mass of one or several proteins in the spectrum. At present, it is this portion of data analysis that requires the greatest amount of user intervention. For relatively simple ion chromatograms containing distinct peaks arising from one or few components, current commercial data analysis software from major mass spectrometer manufacturers can automatically handle the spectral integration and mass interpretation. However, for a more complicated separation, such as in the case of bacteria whole cell lysate analysis, these data analysis software lack the capability of automatically integrating spectra over a defined time (e.g., integrating spectra every minute, not just over the chromatographic peaks).

There has been very little work done on exploring ways of completely automating mass interpretation. One recent study examined both manual and automated production of deconvoluted molecular masses produced from LC–ESI-MS analysis of human plasma filtrates [20]. The automated analysis, utilizing a third party program (SHERPA 3.1.1), was considered a helpful aspect of the overall study due to the complex nature of the analyte, and the resulting multifaceted nature of the chromatogram. However, the automated approach produced 48% false-positive molecular mass determinations. It was pointed out that different search settings in the program did not lead to significant changes in the reliability of the results. Another study developed a computer program to automatically deconvolute mass spectra obtained from LC–ESI-MS runs [21].

We present a method of generating mass data from LC–ESI-MS of bacteria cell extracts in a fully automated fashion. The method relies on the production of an oscillating signal in the total ion chromatogram (TIC), and subsequent analysis by the auto-integration feature found in Agilent's ChemStation. The artificial chromatographic peaks are produced when the static mixer on the Agilent 1100 series pump is removed. The reduced mixing is believed to produce sections richer in organic content, which subsequently produce an enhanced signal when proteins are sprayed. Using this method, the TIC is divided into a series of narrow peaks with a baseline width of approximately 1 min. In this work, the utility of this technique is demonstrated for the analysis of two bacterial extracts, namely *Escherichia coli* (*E. coli*) and *Bacillus globigii* (*B. globigii*). Results from manual and automated data analysis are compared.

2. Experimental

2.1. Materials

Escherichia coli (ATCC 9637) and *B. globigii* were a gift from the Edgewood RDE Center at Aberdeen Proving Ground, MD, USA. Spectrophotometric-grade trifluoroacetic acid (TFA) was purchased from Sigma–Aldrich Canada (Oakville, Canada). HPLC-grade acetonitrile and glacial acetic acid were from Fisher Scientific Canada (Edmonton, Canada). Water was obtained from a Milli-Q Plus purification system (Millipore, Bedford, MA, USA).

2.2. Sample preparation

Bacterial extracts were prepared by a solvent suspension method. About 7 mg of lyophilized bacteria were suspended in 1 ml 0.1% (v/v) aqueous TFA, vortexed for 3–5 min and centrifuged. A total of five extractions were performed, producing a total volume of 5 ml. The supernatants from each of the bacterial samples were combined and filtered by Microcon-3 filters with a molecular mass cut-off of 3000 (Amicon, Oakville, Canada).

2.3. LC–ESI-MS

Solvent delivery and separations were performed on an Agilent (Palo Alto, CA, USA) 1100 series

HPLC equipped with an autosampler. All connections were made with 0.127 mm I.D. polyether ether ketone (PEEK) tubing and finger tight fittings (Upchurch Scientific). Chromatographic analysis was performed with a reversed-phase Vydac C_8 column (150 mm×2.1 mm I.D.; particle size, 5 μm; pore size, 300 Å). A flow-rate of 200 μl/min was used for all samples in the LC–ESI-MS analysis. Gradient elution was performed with solvent A (0.05%, v/v, aqueous TFA) and B (0.05%, v/v, TFA in acetonitrile). The following gradient profile was used for both bacteria samples (min:%B), 0:2, 10:20, 40:40, 45:55, 60:75, 75:2 at 200 μl/min. Each HPLC run consisted of a 30-μl injection of bacterial extract. This volume of extract contains an approximately 30 μg of proteins from *E. coli* or 1 μg of proteins from *B. globigii* based on Coomassie Plus Protein Assay Reagent (Pierce, Rockford, IL, USA) absorbance at 590 nm.

The HPLC effluent was analyzed with an Agilent 1100 series MSD quadrupole mass spectrometer. Agilent ChemStation software provided control of both the HPLC and mass spectrometer. The Igor Pro software package (WaveMetrics, Lake Oswego, OR, USA) was used to reprocess the acquired data.

Post-column addition of glacial acetic acid was delivered at 100 μl/min in order to overwhelm the signal suppressing properties of TFA in the mobile phase. The acid was added to the column effluent using a PEEK ''Y'' connector (Upchurch) and a syringe pump (Cole Parmer). The ''Y'' was connected to the electrospray interface with 30 cm of PEEK tubing. The length of the tubing exiting the ''Y'' and entering the ESI source was crucial for adequate mixing of the acid with the column effluent.

A variable fragmentation voltage was used during the ESI detection. The optimized voltage ramp was found to be: m/z 500=60 V, 1000=120 V, and 3000=220 V. In addition, the following parameters were used for MS detection of the chromatographic eluent: gain, 5; step size, 0.15; drying gas, 10 l/min; drying gas temperature, 350°C; nebulizer pressure, 25 p.s.i. (1 p.s.i. 6894.76 Pa); capillary voltage, 4500 V.

2.4. Protein identification

The SWISS-PROT and TrEMBL databases were searched using the Sequence Retrieval System. The molecular mass measured by LC–ESI-MS was entered into the information field, along with the organism *Escherichia coli*. Protein names and descriptions are listed in Table 1 only if the database molecular masses matched the experimental masses to within ±0.05%.

3. Results and discussion

Considerable effort was spent on optimizing the separation conditions used for LC–ESI-MS of bacteria extracts. When the conditions were not optimized, very few proteins could be detected due to the ion suppression effect associated with ESI that reduced the protein ion signals. We used the *E. coli* extract that had been previously examined by MALDI for LC optimization. This extract should contain over 500 proteins with molecular masses less than 20 000. However, in LC–ESI-MS we usually detect close to 40 proteins. One major reason for this difference is that ion suppression in ESI is much more severe compared to MALDI. In LC–off-line MALDI–MS, each LC fraction often contains more than five proteins, whereas in LC–ESI-MS an averaged spectrum from each peak usually contains only one major protein and several smaller protein signals that are difficult to discern from the background signals.

A thorough analysis of the ion chromatograms obtained from bacteria extracts reveals a number of proteins eluting in regions where no chromatographic peaks are present (i.e., the baseline). These protein signals can be recovered by averaging spectra collected over a certain period of elution. However, current analysis software, such as ChemStation (Agilent Technologies), DataAnalysis (Bruker), and Xcalibur (Finnigan), does not possess the capability to average mass spectra over a preset time window for the entire chromatogram. In the case of ChemStation, the automatic peak finder in the analysis software would either skip over regions where no apparent chromatographic peaks are present, or would break a region down into too many sections. If a chromatographic peak is segmented too much, the peak cannot be properly averaged for the mass spectrum. This current deficiency makes it difficult to automatically find peaks, average the mass spectra

Table 1
Protein masses of protonated species from *E. coli* determined by manual and automated interpretation of TIC with static mixer removed* and tentative protein assignments based on the molecular mass

	Manual	Auto 1500	Auto 2000	Auto 2500	Auto 3000	Mass of possible Match	Name and description
	2431.0	2431.0	2431.0	2431.0	2431.0	–	–
	3027.1	**3027.1**				–	–
	3140.4					–	–
	4465.1					–	–
	5096.5	5095.8	5096.4	5095.8	5096.4	5096	30A ribosomal protein S22
	5173.7	5173.9	5174.0			5173	NIKE protein (fragment)
		5469.7				5472	Hypothetical M_r 5400 protein in replication origin region
	6255.3	6255.2	6254.9			6254	MOB9 ORF
	6315.6					6315	50S ribosomal protein L32
	6411.2	**6411.1**	**6411.3**			6411	50S ribosomal protein L30
	6856.7	**6856.9**				6856	Carbon storage regulator
	7273.4	**7274.1**	**7274.9**	**7274.7**		7273	50S ribosomal protein L29
	7333.0	**7332.5**	**7332.7**			7332	Cold shock-like protein CSPE
	7707.5	7707.5	7707.5	7707.5	7707.5	7704	Class I intergrase (fragment)
	7867.5					7871	50S ribosomal protein L31
	8322.6					8323	Major outer membrane lipoprotein
	8325.9					8325	Protein YJBJ or BFPA (fragment)
	8752.4			**8752.2**		8753	Hypothetical M_r 8800 in PRC-PPHA intergenic region
	8834.8					8832	Putative inner membrane protein WZC
	8876.7	**8875.2**				8875	50S ribosomal protein L28
	9064.1	9064.1	9064.0	9064.1	9064.1	9065	Protein HDEB
	9190.5	9191.8	**9191.6**	**9191.2**		9190/9191	30S ribosomal protein S16/SHFC protein
	9226.6	9226.6	9226.6	9226.6	9226.6	9225/9227	DNS-binding protein HU-BETA/KLEC protein
		9326.5	**9325.4**			9330	Putative inner membrane protein WZC (fragment)
	9387.4	9387.3				9386	Hypothetical M_r 9400 protein SOHB-TOPA intergenic region
	9536.1	9536.1	9536.1	9536.1	9536.1	9535	DNA-binding protein HU-ALPHA
	9739.8	9739.8	9739.8	9739.8	9739.7	9737	DR hemagglutinin AFA-III operon regulatory protein AFAF
	10 300	10 299				10 300	Putative acylphosphatase
			10 867			10 865/10 866	Hypothetical M_r 10 900 protein in PFKB-CEDA intergenic region/TraQ protein
	11 124					11 126	CII protein
	11 186	11 185	**11 186**		**11 177**	11 185	50S ribosomal protein L24
				11 253		11 255	Hypothetical M_r 11 300 protein in HIPB-UXAB intergenic region
	11 689	11 688	11 688	11 688		11 686	Phosphate starvation — inducible protein PSIF
	11 782	11 782	11 782	11 782	11 782	11 781	E-HLYC protein
	11 976	11 977	11 977	11 977	11 959	11 967	Putative CURLI production protein CSGC
	12 483	**12 483**	**12 484**	**12 484**	**12 485**	12 484	Plasmid R751, complete genome
	13 128	13 128	13 129	13 128	13 128	13 127	ORF30 protein
		13 143	**13 142**			13 144	β-Galactosidase
	15 410	15 409				15 408	DNA-binding protein H-NS
	15 692	**15 693**	15 693	15 693		15 689	N utilization substance protein B
				16 096		16 093	PTS system, fructose-like-1 IIA component
	16 686	**16 686**	16 685	16 685	16 685	16 683	Hypothetical M_r 16 700 protein in TNAB–BGLB intergenic region
	17 515	17 515	**17 516**			17 513	Putative electron transport protein YGFS
	18 162	18 162	18 162	18 162	18 162	18 156	Lipoprotein signal peptidase
Total	38	32	26	20	14		
False		3	3	2	0		
Manual matched (%)		76	61	47	37		

Bold masses are listed second in the automated interpretation peak list.

and then produce a list of deconvoluted masses. However, if an artificial peak were introduced in a region where it previously did not exist, the mass spectra could then be averaged and the resulting spectrum deconvoluted. The introduction of artificial peaks to the chromatogram can be accomplished by removing the static mixer in the HPLC pump.

Fig. 1 shows how the removal of the static mixer influences the pattern of the total ion chromatogram (TIC). Fig. 1A is the TIC of *E. coli* 9637 cell extract obtained with the static mixer removed and Fig. 1B is that obtained with the static mixer installed. As Fig. 1A shows, removal of the static mixer from the flow path of the HPLC pumps produces an oscillating TIC signal. The source of the relatively evenly spaced peaks is thought to arise from the production of segmented flow in which the ionization process is augmented and diminished as the segments are sprayed. The effect was only observed as proteins eluted from the column. This was verified by conducting blank gradient runs with the static mixer removed and with post-column addition of glacial acetic acid. The electrospray current did not display the distinctive oscillations. The artificial peaks shown in Fig. 1A have peak widths at their base of about 60 s. The mass spectra contained in the peak were averaged and background subtracted. The resulting mass spectra were then deconvoluted in order to determine the $(M+H)^+$ mass for inclusion in the mass tables for bacterial analysis. In the following discussion, this mode of operation is termed *automated spectral integration*.

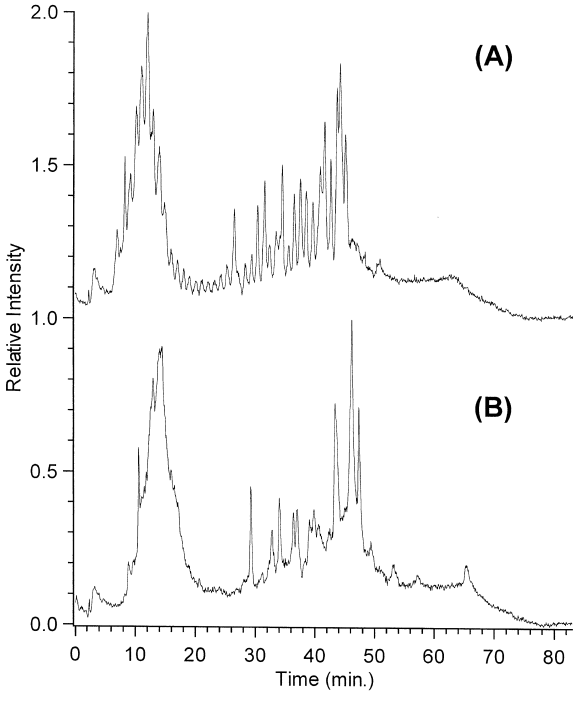

Fig. 1. TIC of *E. coli* 9637 cell extract obtained with (A) static mixer removed and (B) static mixer installed in the HPLC pumps.

After the automated spectral integration, the shapes of the proposed charge profiles in each averaged spectrum were then examined manually to see if they matched the expected profiles produced by the electrospray process. A mass spectral profile was deemed reasonable if it displayed an approximate Gaussian distribution and if the majority of the peaks were greater than three times the background. Alternatively, instead of manually interpreting the deconvoluted mass spectra, an attempt was made to completely automate the process of mass interpretation. In this case, the first and second masses proposed by ChemStation's deconvolution program were taken. Automated data analysis required 5 min using a 600-MHz Pentium III with 256 MB RAM. In order to compare the results with the manual mass interpretation, identical parameters except the level of noise cutoff were used in the deconvolution program for the automated mass interpretation.

Table 1 summarizes the masses found for both manual and automated interpretation of the deconvoluted mass spectra along with the proteins tentatively identified based on their molecular masses. For the automated interpretation with the noise cutoff parameter set at 1500 counts, 29 masses matched those from manual interpretation (i.e., 76% match). There are three masses not matching with the manual interpretation results (i.e., 8% false-positive). The effect of a larger noise cutoff was examined to gauge whether it had any influence in reducing the number of false masses. If the noise cutoff is increased from 1500 to 2000 counts, 23 matches and three false-positives are found. If the noise cutoff is increased further to 2500 counts, 18 masses match those in the manual column with two false-positives. At a threshold cutoff of 3000 counts, the number of matches has decreased to 14; however, the false-positives

have been eliminated. It is clear that raising the noise cutoff can reduce false-positive results.

A closer examination of the false masses detected by the automated mass interpretation procedure reveals the possibility of improved performance in the procedure. Fig. 2 shows three integrated spectra from the pulsed peaks at different retention times in the TIC shown in Fig. 1A. Fig. 2A gives an example of a single component found in the integrated spectra. In this case, the automated mass interpretation procedure readily detected the mass of the protein. However, multiple components (usually two) were also detected in some chromatographic peaks and if the intensities were high enough, the automated analysis was able to pick up the second component. As an example, Fig. 2B shows the spectrum from two components with very intense protein ion peaks. They are from $(M+H)^+$ of $A=$ 9739.8, and $B=6856.7$ (see Table 1). The second component found by automated interpretation is listed in Table 1 as a bold mass. Fig. 2C shows an example where the signal-to-noise ratio is low. In this case, the automated procedure would pick up components A and B. However, closer examination of the mass spectrum shown in Fig. 2C reveals that these two components should not be considered for inclusion in Table 1. The signal-to-noise ratio is too low for the protein masses to be reliable.

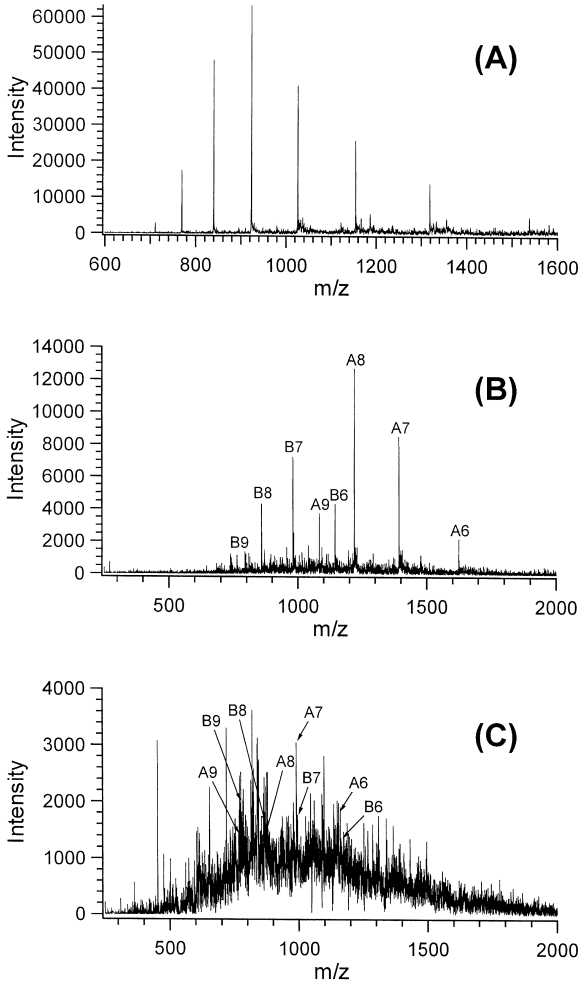

Fig. 2. Representative ESI spectra of (A) single component easily detected with automated interpretation, (B) two components with m/z of $(M+H)^+$ of $A=9739.8$ and $B=6856.7$ (see Table 1), and (C) false matches using the automated spectral interpretation. Automated interpretation lists component A first and B second. Numbers next to letters refer to positive charge state.

The example shown in Fig. 2C illustrates the deficiency of the current software for automated mass interpretation. The software relies on the use of an absolute count of noise cutoff to select the ion peaks for deconvolution. In this case, the use of cutoff 1500 or 2000 will still allow the selection of noise peaks for protein mass calculation, resulting in false-positives. However, if we set the noise cutoff level too high, protein peaks will go undetected in other spectra where protein ion peaks have high S/N, but low absolute ion counts. This is illustrated in Fig. 3 where even the lowest threshold examined, 1500, is too high and two protein components are not detected. In this example, only two charge states for each component are present above the 1500 thres-

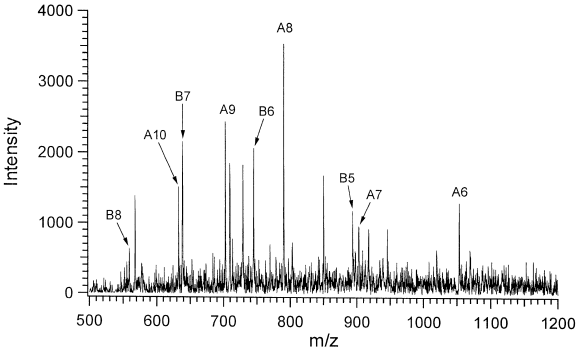

Fig. 3. ESI spectrum showing protein ion peaks having high S/N but low absolute ion counts.

hold setting (A8, A9 and B6, B7). A program requirement for deconvolution is that there be a minimum of three peaks in each set. As a result, the software does not detect this group of components. Table 1 shows protonated species at m/z 4465.1 and 6315.6 present in the manual column and absent from all three automated interpretation columns.

The above examples suggest that if a threshold value based on the S/N, instead of the absolute ion count, were used for spectral interpretation in the automated procedure, it would be possible to eliminate false-positives and increase the number of detected masses. Unfortunately this feature is not present in the current software and hopefully any future release will include this function.

While we have shown that the incorporation of artificial pulses in the TIC allows the current software to carry out automated spectral integration over the entire TIC where proteins are detected, it is conceivable that such an operation can also be done in a conventional TIC. The introduction of improved software that has the capability of integrating spectra over any defined time window, regardless whether there is a chromatographic peak or not, would be a powerful feature. For performance comparison, we carried out *manual spectral integration* on the TIC. In this case, the TIC obtained with the mixer installed (i.e., Fig. 1B) was studied by manually averaging spectra over a selected time window. In effect, this operation mimics the artificial peaks created with the static mixer removed. Mass spectra were averaged over one min interval as well as offsetting the averaging by 30 s and starting at 30 s and averaging over 60 and 90 s.

As was the case with automated spectral integration, averaged mass spectral profiles were examined manually and automatically. Table 2 summarizes the masses found for both manual and automated interpretation of peaks found by examining the conventional TIC over fixed time intervals. In the case of manual mass interpretation, the largest number of masses was observed when integration started at 0 min and averaged over 60 s. A total of 44 masses were found when examining the data in this fashion, of which 39 match those found using the other two methods of spectral averaging (Table 2, columns 2 and 3). If the time period averaged was made too wide (e.g., 90 s instead of 60 s), slightly fewer masses were found. The majority of the masses found using the 90-s window were also found when the time period was offset by 30 s; however, eight did not match and did not correspond to any of the 44 found without offsetting the start time.

When automated interpretation of the charge profiles was undertaken on spectra averaged for 60 s starting at 0 min, 73% matched those masses found by manual interpretation of the same data. Note that deconvolution becomes difficult when the S/N is low and there are essentially peaks at every mass unit. Without the option for a threshold value based on S/N, there will always be a relatively large contribution of false-positive masses included in the results.

To further illustrate the application of this data analysis method, a Gram-positive bacterium, *B. globigii,* was chosen for LC–ESI-MS. The solvent suspension method used to extract the cell contents was not as effective at rupturing the thicker cell walls of *B. globigii*. As a result, the chromatograms exhibit a much weaker signal compared to that of *E. coli*. The effect of removing the static mixer, however, is quite apparent and can be seen in Fig. 4. Even though the signal from the Gram-positive *B. globigii* is much flatter and weaker than *E. coli*, automated spectral integration with manual examination of the data found 28 masses. At a threshold of 1000 counts, automated spectral integration using the first two masses found by ChemStation produced 16 matches (i.e., 57% of those found manually) with no false-positives. As in the case of *E. coli*, the main reason of detecting less number of *B. globigii* proteins using automated analysis is due to the use of a threshold based on absolute counts. If the threshold were based on S/N, most proteins detected by manual operation would be counted for by the automated analysis procedure.

4. Conclusions

Automated spectral integration and interpretation has been successfully applied to the analysis of total ion chromatograms of bacterial lysates. The proposed method layers an oscillating signal over the normal signal produced by the ionization process. The signal enhancement is believed to arise from the production of solvent plugs that augment the current

Table 2
Protein masses of protonated species from *E. coli* determined by manual and automated interpretation of the conventional TIC (i.e., static mixer installed)[a]

Manual spectral integration			Automated spectral integration
(0–1 min)	(0–1.5 min)	(0.5–1.5 min)	(0–1 min)
2430.1	2431.1	2430.9	2431.1
2982.0	2982.0	2982.1	2473.2
3027.2	3027.3	4363.4	2982.0
4363.8	4363.4	4464.6	3027.2
4464.0	4465.0	5020.8	3275.3
5005.0	5020.7	5096.6	4464.9
5020.6	5096.8	5381.7	5097.0
5096.9	5379.9	6255.6	6255.7
5381.8	6255.5	6316.4	6315.5
6255.4	6316.4	6414.6	6856.9
6316.4	6411.3	6856.9	7430.7
6414.9	6414.4	7274.6	7707.7
6856.8	6856.8	7333.2	7722.5
7272.6	7274.6	7707.7	7867.9
7274.3	7333.2	7722.2	7903.7
7333.1	7707.7	7847.2	8325.7
7637.6	7867.7	7868.6	8752.9
7708.0	8326.0	8326.0	8834.7
7722.8	8834.6	8751.9	8877.5
7867.9	8876.8	8834.6	8992.7
7870.4	8992.8	8876.3	9064.1
8326.0	9064.2	9064.2	9191.5
8877.2	9191.4	9191.5	9226.7
9064.2	9226.7	9226.7	9536.2
9191.1	9536.2	9428.1	9540.7
9226.8	9739.7	9536.2	9634.9
9536.2	9755.6	9540.3	9739.7
9540.0	10 087	9739.8	11 123
9634.9	11 132	10 300	11 186
9739.8	11 186	11 186	11 688
9755.7	11 688	11 688	11 781
11 124	11 781	11 781	11 850
11 133	12 771	11 977	11 977
11 186	13 127	12 484	12 006
11 688	15 409	12 772	12 770
11 781	15 692	13 128	13 049
11 977	16 686	15 409	13 362
12 770	18 162	18 162	15 410
13 128			15 693
15 409			18 162
15 693			18 773
16 686			
18 163			
18 772			

[a] The masses listed were found by deconvoluting mass spectra that were averages of the listed time frame.

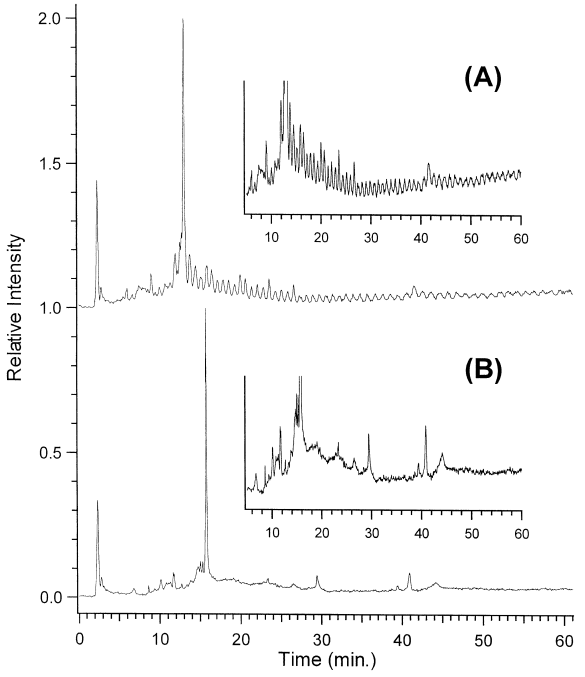

Fig. 4. TIC of *B. globigii* cell extract obtained with (A) static mixer removed and (B) static mixer installed in the HPLC pumps.

as proteins elute from the column. Using *E. coli* ATCC 9637 as an example, a noise threshold value of 1500 counts with the automated procedure found 32 masses, whereas manually interpreting the same data revealed 38 masses. The automated method with a threshold of 1500 counts found 76% of those masses observed in the manual case. The completely automated approach is rapid and capable of easily producing a set of masses in which the majority corresponds to those found by manual means. Increasing the noise threshold was found to decrease the number of false-positives.

Implementation of the proposed method allows for altering the mass spectral current without significantly varying the actual separation or the number of protein masses that are detected. The method is particularly advantageous in cases with lengthy elevated baselines where numerous proteins are known to elute. This work calls for future improvement in the LC–ESI-MS software that should include the capability of averaging spectra over a defined chromatographic time frame and the capability of interpreting mass spectra by selecting ion peaks based on S/N, rather than the absolute ion count. However, until such versions are made available, the method presented in this work offers an alternative to the lengthy and operator-dependent task of manually identifying peaks and interpreting the results.

Acknowledgements

This work was partially funded by the US Army ERDEC. We thank Drs. Peter Synder and Dennis Roser of ERDEC for the bacteria samples and helpful discussion during the course of this work.

References

[1] T.C. Cain, D.M. Lubman, W.J. Weber, Rapid Commum. Mass Spectrom. 8 (1994) 1026.
[2] D.J. Evason, M.A. Claydon, D.B. Gordon, Rapid Commun. Mass Spectrom. 14 (2000) 669.
[3] Y. Dai, L. Li, D.C. Roser, S.R. Long, Rapid Commun. Mass Spectrom. 13 (1999) 73.
[4] R.D. Holland, C.R. Duffy, F. Rafii, J.B. Sutherland, T.M. Heinze, C.L. Holder, K.J. Voorhees, J.O. Lay Jr., Anal. Chem. 71 (1999) 3226.
[5] R.J. Arnold, J.A. Karty, A.D. Ellington, J.P. Reilly, Anal. Chem. 71 (1999) 1990.
[6] E.C. Lynn, M.C. Chung, W.C. Tsai, C.C. Han, Rapid Commun. Mass Spectrom 13 (1999) 2022.
[7] F. Leenders, T.H. Stein, B. Kablitz, P. Franke, J. Vater, Rapid Commun. Mass Spectrom. 13 (1999) 943.
[8] D.B. Wall, D.M. Lubman, S.J. Flynn, Anal. Chem. 71 (1999) 3894.
[9] Z. Wang, L. Russon, L. Li, D.C. Roser, S.R. Long, Rapid. Commun. Mass Spectrom. 12 (1998) 456.
[10] J.H.M. van Adrichem, K.O. Bornsen, H. Conzelmann, M.A.S. Gass, H. Eppenberger, G.M. Kresbach, M. Ehrat, C.H. Leist, Anal. Chem. 70 (1998) 923.
[11] B.E. Chong, D.B. Wall, D.M. Lubman, S.J. Flynn, Rapid Commun. Mass Spectrom. 11 (1997) 1900.
[12] T. Krishnamurthy, P.L. Ross, U. Rajamani, Rapid Commun. Mass Spectrom. 10 (1996) 883.
[13] T. Krishnamurthy, P.L. Ross, Rapid Commun. Mass Spectrom. 10 (1996) 1992.
[14] M.A. Claydon, S.N. Davey, V. Edwards-Jones, D.B. Gordon, Nat. Biotech. 14 (1996) 1584.
[15] X. Liang, K. Zheng, M.G. Qian, D.M. Lubman, Rapid Commun. Mass Spectrom. 10 (1996) 883.
[16] T.C. Cain, D.M. Lubman, W.J. Weber Jr., Rapid Commun. Mass Spectrom. 8 (1994) 1026.
[17] J.L. Dalluge, R. Prasad, BioTechniques 28 (2000) 156.

[18] T. Krishnamurthy, M.T. Davis, D.C. Stahl, T.D. Lee, Rapid Commun. Mass Spectrom 13 (1999) 39.
[19] J.E. MacNair, G.J. Opiteck, J.W. Jorgenzon, M.A. Moseley, Rapid Commun. Mass Spectrom 11 (1997) 1279.
[20] M. Raida, P. Schultz-Knappe, G. Heine, W. Forssmann, J. Am. Soc. Mass Spectrom. 20 (1999) 45.
[21] J.O. Pearcy, T.D. Lee, J. Am. Soc. Mass Spectrom. 12 (2001) 599.