

# Rapid Profiling of *E. coli* Proteins Up to 500 kDa from Whole Cell Lysates Using Matrix-assisted Laser Desorption/Ionization Time-of-flight Mass Spectrometry

Bathsheba E. Chong<sup>1</sup>, Daniel B. Wall<sup>1</sup>, David M. Lubman<sup>1\*</sup> and Shannon J. Flynn<sup>2</sup>

<sup>1</sup>Department of Chemistry, The University of Michigan, Ann Arbor, MI 48109-1055, USA

<sup>2</sup>Center for Microbial Ecology, Michigan State University, 540 Plant and Soil Science Building, East Lansing, MI 48824-1325, USA

SPONSOR REFEREE: J. R. Yates, Department of Molecular Biotechnology, University of Washington, Seattle, WA 98195-7730, USA

**Matrix-assisted laser desorption/ionization time-of-flight mass spectrometry was used to rapidly detect and profile large proteins from *Escherichia coli* whole cell lysates in the mass range 25–500 kDa. The bacterial samples were treated with guanidine hydrochloride and Triton X-100 to disrupt and solubilize the large inner membrane proteins. A sample preparation involving a nitrocellulose polymer film, and  $\alpha$ -cyano-4-hydroxycinnamic acid, sinapinic acid or caffeic acid as matrix was utilized to rapidly monitor the presence of induced and repressed protein synthesis in response to L-arabinose catabolism in *E. coli* cells. The results were compared to those of 1-D or 2-D gel electrophoresis. © 1997 John Wiley & Sons, Ltd.**

Received 30 September 1997; Accepted 30 September 1997

Rapid. Commun. Mass Spectrom. 11, 1900–1908 (1997)

No. of Figures: 5 No. of Tables: 5 No. of Refs: 33

Matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOFMS) has been shown to be a powerful tool for detection of proteins in complicated biological matrices.<sup>1–5</sup> In recent work, Cain *et al.*<sup>1</sup> reported the use of MALDI-TOFMS to rapidly differentiate bacteria based upon the analysis of proteins from whole bacterial cell lysates. The bacterial cells were disrupted by a simple sonication procedure. In related work, Lay<sup>2</sup> and co-workers used MALDI-MS to identify intact bacteria based on spectral patterns either by comparison to an archived reference or by co-analysis with cultures of known bacteria. In addition, Krishnamurthy *et al.*<sup>3</sup> have demonstrated a method employing MALDI-MS to directly analyze whole bacterial cells. By examining *Brucella* and *Yersinia* species along with a series of strains of several *Bacillus* species, the genus, species and strain-specific biomarkers from the measured molecular masses of the intact proteins were derived. Other recent applications of this technique include that of Tomer<sup>4</sup> in which MALDI-MS was used to determine an epitope on HIV-I (IIIB) p26 recognized by a monoclonal antibody as well as that of Jungblut<sup>5</sup> where MALDI-TOFMS was used to type new secondary metabolites of intact cyanobacteria.

Alternatively, high performance liquid chromatography (HPLC)<sup>6</sup> and capillary isoelectric focusing (cIEF)<sup>7</sup> have been coupled to MALDI-MS to separate and characterize proteins from bacterial cell lysates. As a result, specific protein profiles were generated. These methods were used to reduce suppression effects in protein mixtures and to provide two-dimensional separation and detection of bacterial proteins.

In the studies of MALDI-MS profiles of bacterial proteins thus far, only a limited number of proteins are observed and they are generally in the  $m/z$  range below 20 kDa. The expression of the bacterial genome results in the synthesis of around 2000 protein molecules,<sup>8</sup> most of which have masses above 20 kDa. Two-dimensional polyacrylamide gel electrophoresis (2-D PAGE) is widely used to examine bacterial protein content and has even been used in chemotaxonomic applications for detection of various strains of *legionella*<sup>9</sup> and *haemophilus-influenzae* proteins.<sup>10</sup> In the case of *Escherichia coli*, such gel analysis shows that relatively few proteins appear below 20 kDa and that the large majority of proteins studied are in the mass range 25–120 kDa. Indeed it is in this region that one might expect to find biologically significant proteins that define the function and identity of a particular bacterium. Thus, for applications involving identification and monitoring of bacteria in the environment and in clinical applications as well as other applications that involve monitoring changes in gene expression of bacterial species for biological and pharmaceutical studies, the detection of these large proteins is essential.

The purpose of this work is to describe a method for detection of large proteins (>25 kDa) from whole bacterial cell lysates. MALDI-MS has been shown capable of detecting proteins up to ~300 kDa<sup>11,12</sup> and polymers of >1000 kDa.<sup>13</sup> In this work, we describe a sample preparation method which allows large proteins to be made amenable to MALDI analysis. The proteins were released from the cell via chemical treatment with a combination of Triton X-100 and guanidine hydrochloride.<sup>14</sup> The crude extract was then analyzed by MALDI-TOFMS without further modification or clean up. This procedure has been applied herein for detection of large proteins from *E. coli* DH5 $\alpha$ , a derivative of

\*Correspondence to: D. M. Lubman, Department of Chemistry, The University of Michigan, Ann Arbor, MI 48109-1055, USA  
Contract/grant sponsor: National Science Foundation;  
Contract/grant number: DEB 9120006

the K-12 strain, grown under conditions which result in induction or repression of the arabinose genes. Induction of the genes required for the uptake and catabolism of pentose L-arabinose by *E. coli* is well documented.<sup>15,16</sup> Nine genes that are induced by arabinose have been identified and sequenced.<sup>17–22</sup> We have used the known molecular weights of the arabinose-specific proteins as a guide to distinguish MALDI-TOF protein profiles. The tabulated molecular weights (MW) of proteins observed in these experiments are compared to those observed in 1-D or 2-D gels in the mass range between 25 and 120 kDa.

## EXPERIMENTAL

### Chemicals

Acetonitrile (99.9 + %, HPLC grade) and 2-propanol (99.9 + %, HPLC grade) were from Sigma-Aldrich (St. Louis, MO, USA; Milwaukee, WI, USA). Trifluoroacetic acid (TFA), sinapinic acid (SNP) or 3,5-dimethoxy-4-hydroxycinnamic acid, and caffeic acid (CAF) or 3,4-dihydroxycinnamic acid were purchased from Aldrich (Milwaukee, WI, USA). The acetone (HPLC grade) was obtained from Fisher (Fair Lawn, NJ, USA). Guanidine hydrochloride or guanidinium chloride (gu-HCl) and  $\alpha$ -cyano-4-hydroxycinnamic acid ( $\alpha$ -CHCA) came from Sigma (St. Louis, MO, USA). Triton X-100 was obtained from ICN (Aurora, OH, USA). The nitrocellulose (NC) used was of *Immobilon-NC Pure* from Millipore (Bedford, MA, USA) as well as the Amicon Ultrafiltration system. All chemicals were used without further purification for these experiments.

### Bacterial strain, medium, and growth conditions

*E. coli* K-12 strain DH5 $\alpha$  harboring the expression vector pBAD-GFP, a derivative of pBAD18,<sup>23</sup> containing the cycle three mutant of the green fluorescent protein (GFP)<sup>24</sup> was used throughout this study. The DH5 $\alpha$  strain was grown overnight at 37 °C in Luria broth (LB)<sup>25</sup> containing 100  $\mu$ g/ml ampicillin. The overnight culture was diluted 1:10 in three separate media: LB, LB plus 0.2% glucose, or LB plus 0.2% arabinose and incubated for four hours at 37 °C. Plate counts demonstrated that each culture grew to similar cell densities (data not shown). Expression of GFP from pBAD-GFP was confirmed by fluorescent microscopy (data not shown). Cells were harvested by centrifugation of 1 ml aliquots in an Eppendorf centrifuge at 14 k rpm for three minutes. The cell pellets were stored at –20 or –80 °C.

### Matrix, substrate and sample preparation

**Part 1.** A thin layer of nitrocellulose (NC) was applied to the stainless steel probe surface. The NC used was *Immobilon-NC Pure* from Millipore. In previous work, a NC film was shown to enhance molecular ion yield and reproducibility in MALDI-MS experiments.<sup>26</sup> The NC not only serves as an active purification method for low molecular weight impurities from the samples<sup>27</sup> but also provides an effective surface for protein adsorption. This NC film was prepared by applying 3  $\mu$ l of NC solution (10 mg/ml) to the probe surface and allowing it to air dry. The sample and matrix were subsequently applied to the NC surface as described below.

**Part 2.**  $\alpha$ -CHCA was the key matrix used in these experiments. It was prepared as a saturated solution in a 1:1 ratio of acetonitrile to distilled water containing 1% TFA. A few hours before the direct MALDI analysis, the bacterial cells were lysed by resuspension in 0.5 ml of 10 mM Tris buffer (pH 7.4), 2 M gu-HCl, and 2% Triton. This mixture was vortexed for at least 30 seconds, and stored at –20 °C for 20 min. Subsequently, the lysate was fractionated by centrifugation in an Eppendorf centrifuge at 15 k rpm for 20 min. The soluble fraction was stored at –20 °C while the insoluble fraction was resuspended in 0.5 ml of 10 mM Tris buffer (pH 7.4), 2 M gu-HCl, and 2% Triton. This chemical method of cell lysis was repeated to maximize the extraction of large soluble proteins. The two soluble fractions were then combined and stored at –20 °C along with the insoluble fractions for future analysis. Thereafter for MALDI analysis, 3  $\mu$ l of this soluble mixture was applied onto the probe tip, and allowed to air dry. This step was repeated by applying another 3  $\mu$ l of the mixture onto the first layer. The final step involved the application of 3  $\mu$ l of matrix solution onto the sample layers which in turn was allowed to air dry.  $\alpha$ -CHCA was also prepared in a 1:1 ratio of 2-propanol to distilled water containing 1% TFA.

**Part 3.** Sinapinic acid (SNP), or 3,5-dimethoxy-4-hydroxycinnamic acid, and CAF or 3,4-dihydroxycinnamic acid were also used for comparison purposes. These matrices were prepared in a saturated solution of a 1:1 ratio of acetonitrile to distilled water containing 1% TFA as well as a 1:1 ratio of 2-propanol to distilled water containing 1% TFA.

**Part 4.** In some experiments the bacterial samples were filtered via an Amicon Ultrafiltration system (Millipore) with a molecular weight cut-off of 30 kDa. The low molecular weight proteins (< 20 kDa) were filtered out of the lysate in order to minimize suppression effects in the MALDI process whereby the presence of the low molecular weight proteins may suppress the presence of the higher molecular weight proteins. MALDI-TOFMS analyses were then carried out with and without NC as well as with all three matrices.

### MALDI-TOFMS analysis

The TOFMS employed in these studies was a modified Wiley–McLaren design with a four-plate acceleration structure according to that described by Li and Whittal.<sup>28</sup> The TOF was designed for high voltage acceleration up to  $\pm$  20 kV (R. M. Jordan Co., Grass Valley, CA, USA). The laser source used to produce MALDI was a Continuum Minilite 10 Hz Nd:YAG laser system (Continuum, Santa Clara, CA, USA). This laser provided up to 4 mJ of 355 nm radiation for desorption. The laser beam was focused onto the probe tip at a 45° angle to the probe surface with a single 12.5 in f.l. quartz lens to a spot size of  $\sim$ 0.2 mm  $\times$  0.5 mm. A CCD camera and video monitor (Sony Model #DXC-107) was utilized to monitor the laser beam spot size and position on the probe surface. The resulting power density at this spot was estimated at  $\sim$ 5  $\times$  10<sup>6</sup> to 1  $\times$  10<sup>7</sup> W/cm<sup>2</sup> at 355 nm radiation in these experiments. The detector was a

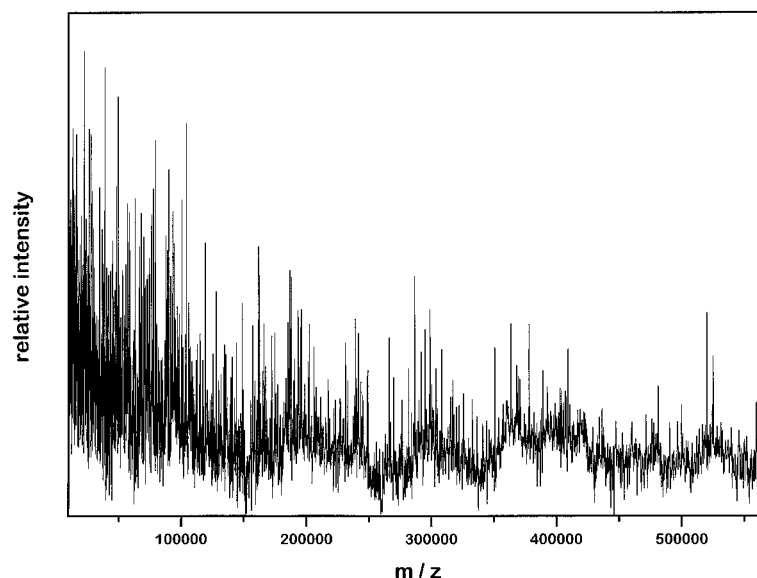


Figure 1. MALDI-TOF uninduced *E. coli* whole cell lysate spectrum (up to 500 kDa).

triple microchannel plate (MCP) detector (R. M. Jordan Co.) which adapted a CuBe conversion dynode with post-acceleration capability up to  $\pm 12$  kV<sup>29</sup> in front of the MCP. The total ion acceleration across the TOF device may thus be  $> 30$  kV. The post-acceleration stage enhances the efficiency for detection of heavy species, but at the expense of resolution. The voltages were similar for all experiments, as follows: repeller at 15 kV; extractor plate 1 at 10 kV; extractor plate 2 at 5 kV; ground; and post-acceleration at  $-10$  kV. The TOF used a 1 m flight tube and was pumped to a base pressure of  $2 \times 10^{-7}$  Torr. Pulsed delay extraction (PDE) was available as described in previous work<sup>30</sup> but was generally not used due to the higher mass range of samples in this work. Data was recorded using either a LeCroy 9350 AM or a LeCroy 9304 AM digital oscilloscope, and processed on a Gateway 586 166 MHz computer.

### 1-D gel method

For the 1-D gel electrophoretic analyses, whole cell extracts were prepared by resuspending the cell pellets in  $2\times$  sample buffer (20% glycerol, 8% SDS, 0.01% Bromophenol Blue, 2.5% DTT and 0.1 mg/ml EDTA) and boiling for five min. The lysate was then subjected to SDS-polyacrylamide gel electrophoresis (SDS-PAGE) using the mini-PROTEIN II and a 10% acrylamide tris-glycine precast ready gel (Bio-Rad Hercules, CA, USA).

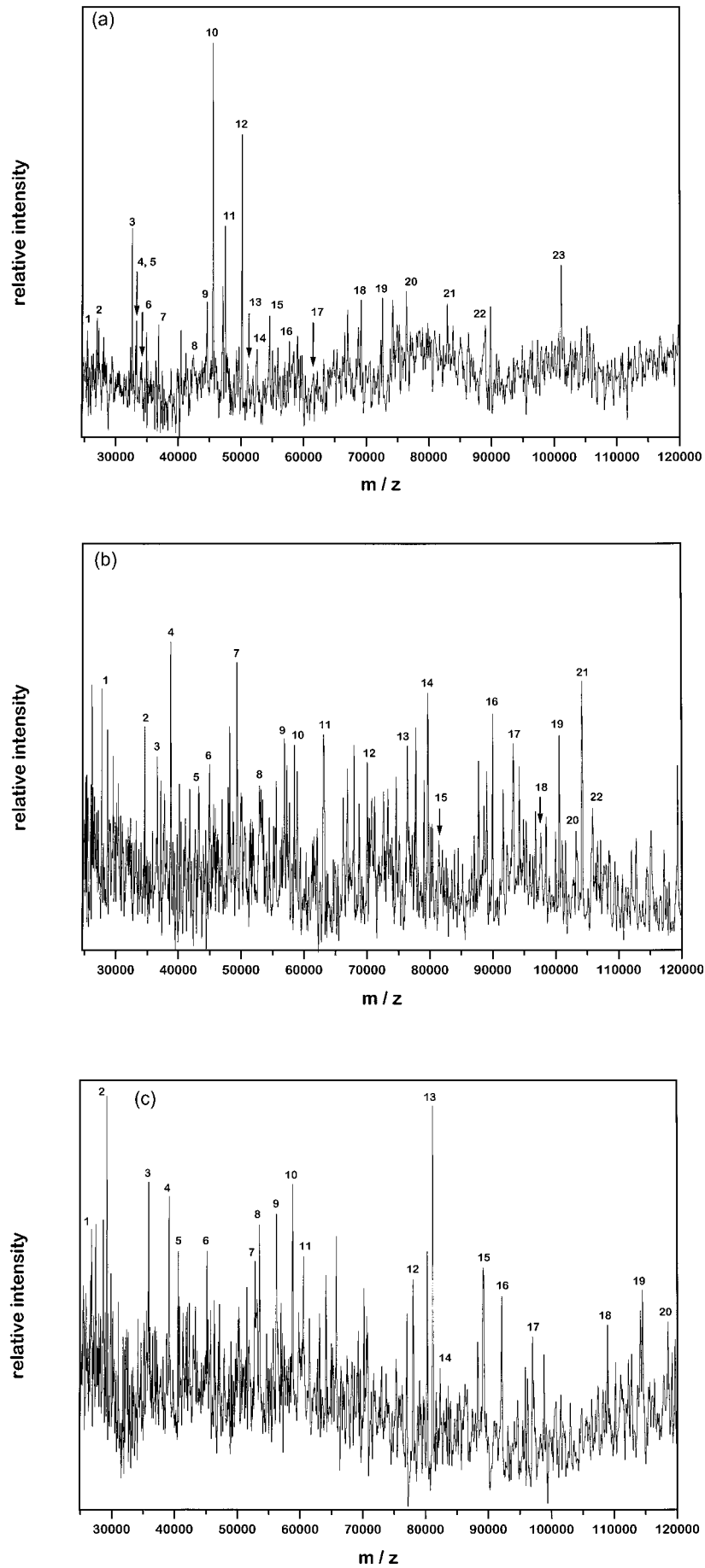
### 2-D gel method

For the 2-D gel electrophoresis, the dried whole cell lysate was re-swelled overnight in the following mixture: 50  $\mu$ l of the whole cell lysate was mixed with 350  $\mu$ l of 7 M urea, 2 M Thiourea, 0.8% CHAPS, 1% Triton, 0.6% Pharmalyte 3-10, and 0.01% Bromophenol Blue. Apart from this alteration, the standard procedure as described in the 2-D gel electrophoresis manual from Pharmacia Biotech (Piscataway, NJ, USA) was followed.

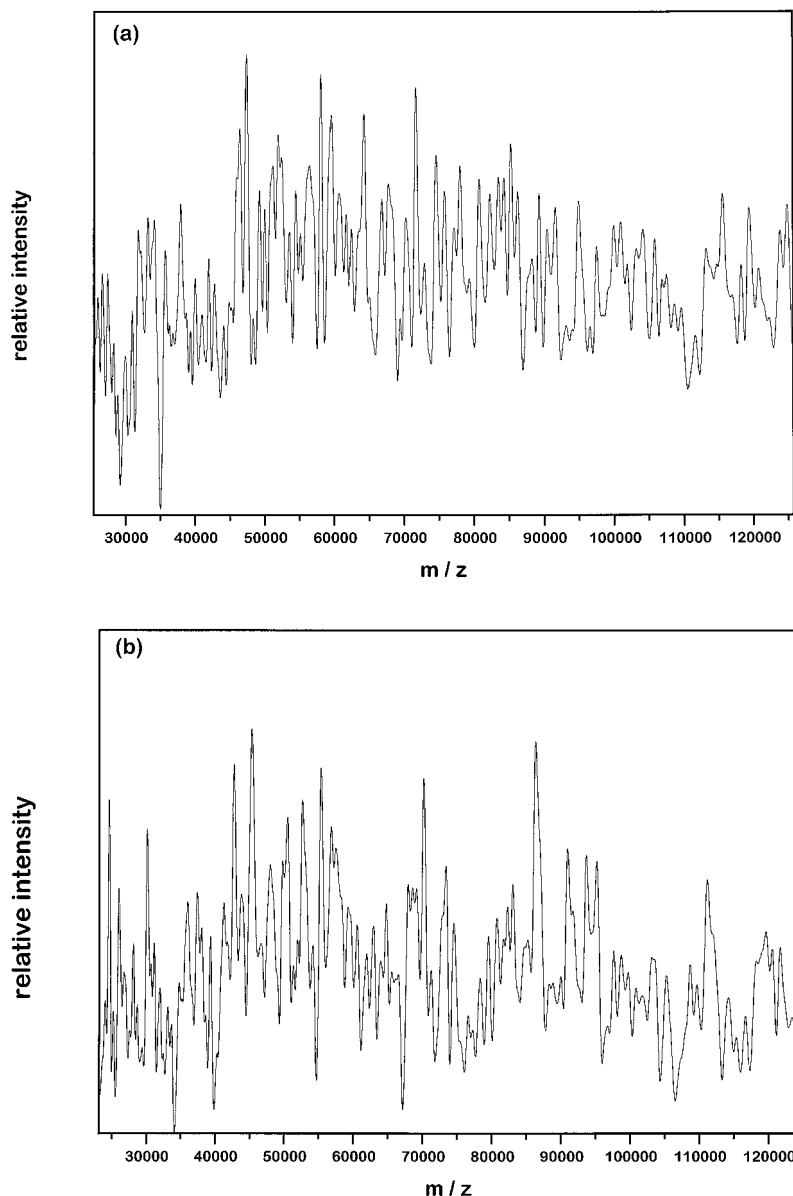
## RESULTS AND DISCUSSION

In Fig. 1 is shown the MALDI-TOF spectrum of a whole cell lysate of the uninduced *E. coli* DH5 $\alpha$  which was prepared using the Triton/gu-HCl combination to extract large proteins. The sample mixture was vortexed for at least 30 seconds before placing it on the NC-coated probe tip. MALDI-TOFMS analyses were performed at a wavelength of 355 nm using  $\alpha$ -CHCA matrix and the ions were detected in the positive mode by applying a post-acceleration voltage of  $-10$  kV on the CuBe dynode in order to post-accelerate large ions. The detector voltage was set at  $-3$  kV. As shown in Fig. 1 a large number of ion peaks are observed across the entire TOF range with ions up to 500 kDa being detected. As expected from 2-D gel results there are a large number of components between 25 and 120 kDa. This is further illustrated in Fig. 2 which shows an expanded view of the MALDI-TOF spectra of the whole cell lysates of *E. coli* in this mass range. Likewise, there are only a limited number of components observed in the  $m/z$  range below 25 kDa as expected from 2-D gel analysis. The three *E. coli* whole cell lysate spectra as shown in Fig. 2 for (a) induced, (b) uninduced and (c) repressed were readily distinguishable from one another. It should be noted that without a NC thin film on the probe fewer ion peaks were observed. In addition, two other matrices, SNP and CAF were also investigated, of which both gave distinguishable spectra for the induced sample as well as the uninduced sample. Figure 3 (a) and (b) shows MALDI-TOF spectra of the induced and uninduced samples using CAF as matrix; spectra for the SNP are not shown. Of the three matrices,  $\alpha$ -CHCA produced the best spectral quality for these bacterial samples. Throughout these experiments, an external standard was adopted to minimize the suppression effects on large protein detection. Due to these suppression effects, fewer peaks were observed when an internal standard was used. Since the background noise was not homogeneous for these whole cell lysate experiments, the signal-to-noise (S/N) ratios of the protein peaks were not computed.

*E. coli* cells were grown on LB or LB with arabinose



**Figure 2.** MALDI-TOF (a) induced, (b) uninduced and (c) repressed *E. coli* whole cell lysate spectra ( $m/z$  25–120 kDa).



**Figure 3.** MALDI-TOF (a) induced and (b) uninduced *E. coli* whole cell lysate spectra with CAF as matrix ( $m/z$  25–120 kDa).

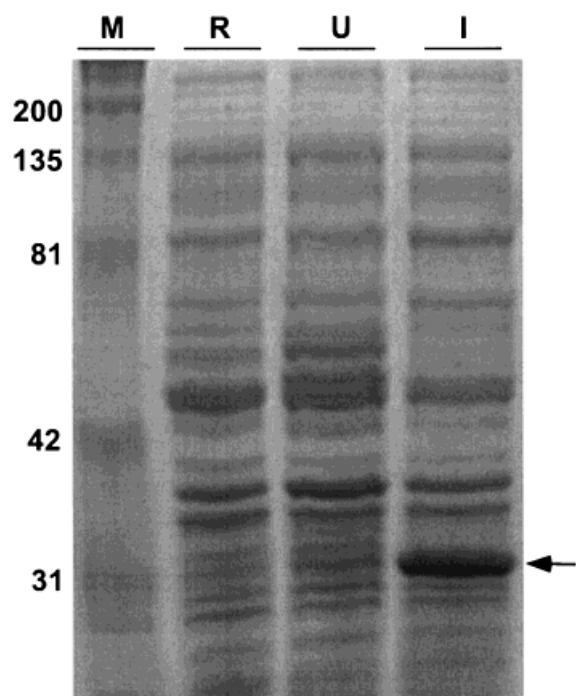
to induce the arabinose genes as described in the experimental section, and a reporter gene, GFP, under the control of the arabinose pBAD promoter was used to confirm induction (see experimental section). The molecular weights calculated from the sequence of the nine arabinose genes and GFP are shown in Table 1. Differences are less distinguishable between the protein profiles from the uninduced and repressed samples by SDS-PAGE (Fig. 4) than by MALDI-TOFMS. In this 1-D gel of the *E. coli* whole cell lysates GFP can be clearly distinguished as a large band migrating at about 30 kDa in the induced sample, however, the arabinose gene products are not readily distinguished. In Fig. 2 (a) the ion peaks that correspond to the  $m/z$  of these *ara* gene products appear more frequently in the induced sample than either the (b) uninduced or (c) repressed sample. The calculated MW of GFP is near 27 kDa,<sup>24</sup> and a peak (number 2) at 26 979.2  $m/z$  is detected only in the induced lysate. Several proteins with a  $m/z$  corresponding to the *ara* gene products were also

observed only in the induced lysates; AraF (33 kDa),<sup>17</sup> AraC (33 kDa)<sup>21</sup> and AraH (34 kDa)<sup>17</sup> labeled as peak numbers 4, 5 and 6 respectively in Fig. 2 (a). The appearance of another band around 51 to 54 kDa could indicate the presence of AraE (51 kDa)<sup>22</sup> and AraG (54 kDa).<sup>17</sup> Again, these may correlate with peak numbers 12 and 13 in Fig. 2 (a). Except for AraA<sup>19</sup> and AraH (see Table 1), peaks consistent with the  $m/z$  of the Ara proteins are not observed in the uninduced and repressed mass spectra. The details of these large peaks are further summarized in Table 2 (induced), Table 3 (uninduced), and Table 4 (repressed).

As anticipated, both 1-D SDS-PAGE and MALDI-TOFMS have detected many common protein MWs in all three bacterial whole cell lysates since they were from the same strain. For example, three of the common bands of the induced, uninduced and repressed samples migrate at approximately 36 kDa, 52 kDa and 81 kDa on 1-D SDS-PAGE (Fig. 4). These may coincide with the MALDI peaks in these spectra

**Table 1. Comparison of protein peaks in Fig. 2 MALDI-TOF (a) induced, (b) uninduced and (c) repressed *E. coli* mass spectra**

Protein	Calculated MW	MW from MALDI		
		(a) Induced (Peak #)	(b) Uninduced (Peak #)	(c) Repressed (Peak #)
AraA	56 109 <sup>19</sup>	not observed	56 786.9 (9)	56 138.4 (9)
AraB	61 163 <sup>19</sup>	61 616.9 (17)	not observed	not observed
AraC	33 315 <sup>21</sup>	33 282.8 (5)	not observed	not observed
AraD	25 504 <sup>19</sup>	25 350.3 (1)	not observed	not observed
AraE	51 683 <sup>22</sup>	51 111.8 (13)	not observed	not observed
AraF	33 200 <sup>17</sup>	33 106.5 (4)	not observed	not observed
AraG	54 947 <sup>17</sup>	54 567.5 (15)	not observed	not observed
AraH	34 309 <sup>17</sup>	34 213.9 (6)	34 496.6 (2)	not observed
AraJ	42 898 <sup>18</sup>	42 420.9 (8)	not observed	not observed
GFP	~27 000 <sup>24</sup>	26 979.2 (2)	not observed	not observed
<i>Some common ion peaks in all three E. coli extracts:</i> <sup>31</sup>				
D040.7	36 731	36 771.1 (7)	36 794.8 (3)	—
H021.7	35 911	—	—	35 904.4 (3)
I049.0	52 518	52 494.1 (14)	—	—
G046.5	52 578	—	52 924.4 (8)	52 782.4 (7)
H096.8	81 682	82 896.7 (21)	—	—
F080.0	81 559	—	81 371.4 (15)	81 121.8 (13)



**Figure 4.** SDS-PAGE separation of whole cell extracts from *E. coli* K-12 DH5 $\alpha$  strain grown under repressed (R), uninduced (U) and induced (I) conditions for the arabinose genes. The proteins were resolved on 10% acrylamide. A molecular marker was run in the far left lane labeled M (kaleidoscope prestained standard, Bio-Rad Hercules, CA, USA). The molecular mass of each protein marker (kDa) is shown on the left. The arrow on the right indicates the position of GFP.

(Fig. 2) as shown in Table 1. For instance, both the induced and uninduced spectra show an ion peak near  $m/z$  36.7 kDa. From the *Gene-protein database of E. coli: Edition 4*,<sup>31</sup> this peak could be D040.7, a protein alpha-numeric designation. On the other hand, the repressed sample exhibits a peak near  $m/z$  35.9 kDa which could correspond to the H021.7 protein. Alpha-numeric (A-N) designations are assigned to individual protein spots based on their migrations on 2-D gels.<sup>32</sup> In Table 3 several protein MWs as determined from the 2-D gel (not shown) done in our laboratory of the

**Table 2. Key to protein peaks in Fig. 2 (a) MALDI-TOF induced *E. coli* spectrum**

Peak	MW from MALDI	Alpha-Numeric <sup>a</sup>	MW from 2-D Gel <sup>b</sup>	Calculated MW <sup>a</sup>
1	25 350.30	G028.1 ( <i>araD</i> )	25 484	25 504 <sup>19</sup>
2	26 979.2	GFP <sup>c</sup>	—	27 000 <sup>24</sup>
3	32 594.1	G028.0	32 402	—
4	33 106.5	F032.0 ( <i>araF</i> )	51 517	33 200 <sup>17</sup>
5	33 282.8	H030.2 ( <i>araC</i> )	33 384	33 315 <sup>21</sup>
6	34 213.9	( <i>araH</i> )	34 623	34 309 <sup>17</sup>
7	36 771.1	D040.7	36 731	—
8	42 420.9	( <i>araJ</i> )	?	42 898 <sup>18</sup>
9	44 605.1	?	?	—
10	45 407.3	G043.8	45 174	—
11	47 412.6	F046.6	47 199	—
12	50 220.1	B046.7	50 171	—
13	51 111.8	( <i>araE</i> )	?	51 683 <sup>22</sup>
14	52 494.1	I049.0	52 518	—
15	54 567.5	( <i>araG</i> )	?	54 947 <sup>17</sup>
16	57 677.5	E050.0	57 370	—
17	61 616.9	D055.0 ( <i>araB</i> )	61 139	61 163 <sup>19</sup>
18	69 070.4	B066.0	69 121	—
19	72 468.9	?	?	—
20	76 479.6	C080.0	76 225	—
21	82 896.7	?	?	—
22	89 002.6	?	?	—
23	101 134.8	F113.0	102 960	—

<sup>a</sup> This match-up is an approximation. The protein's alpha-numeric and 2-D gel MW are from Ref. 31.

<sup>b</sup> This match-up is an approximation. The arabinose induced protein peaks are also compared to the MWs from known sequences. See references.

<sup>c</sup> GFP is a green fluorescent protein under the control of the pBAD promoter. It was used to demonstrate the induction of arabinose genes in this sample.

uninduced *E. coli* whole cell lysate are also compared to those of MALDI and the literature.<sup>31</sup> As anticipated, there are a substantial number of proteins located in the mass range of 25 kDa and higher. These large proteins are especially concentrated in the alpha-numeric G and H regions of the electrophorogram. The molecular mass of the protein markers used extends from 14.4 to 94 kDa. This 2D electrophorogram is quite comparable to that seen in the literature,<sup>31</sup> thus

**Table 3. Key to protein peaks in Fig. 2 (b) MALDI-TOF uninduced *E. coli* spectrum**

Peak	MW from MALDI	Alpha-Numeric <sup>a</sup>	MW from 2-D Gel <sup>a</sup>	MW from 2-D Gel <sup>b</sup>
1	27 701.9	C032.7	27 614	27 500
2	34 496.6	H031.3	34 297	35 200
3	36 794.8	G036.0	36 614	36 500
4	38 699.4	B037.0	38 922	38 400
5	43 225.6	G043.0	43 078	43 000
6	44 995.2	?	?	44 600
7	49 368.2	F045.4 or F045.5	49 284	49 000
8	52 924.4	F050.4	52 578	52 600
9	56 786.9	H047.3 (or <i>araA</i> ?)	56 750	56 500
10	58 403.5	?	?	?
11	63 099.8	G061.0	63 148	62 500
12	70 042.1	B082.0	70 089	?
13	76 337.9	C080.0	76 225	?
14	79 741.0	G073.5	79 678	79 700
15	81 371.4	F080.0	81 559	?
16	89 916.3	?	?	?
17	93 302.4	?	?	?
18	97 675.4	G117.0	97 217	—
19	100 585.1	?	?	—
20	103 154.5	?	?	—
21	104 124.4	G097.0	104 903	—
22	105 910.9	G095.0	105 778	—

<sup>a</sup> This match-up is an approximation. The protein's alpha-numeric and 2-D gel MW are from Ref. 31.

<sup>b</sup> This match-up is an approximation. In this column, the protein peaks are compared to the results from 2-D gel done in-house on the uninduced *E. coli* whole cell lysate. The molecular mass of protein marker used ranges from 14.4 to 94 kDa.

**Table 4. Key to protein peaks in Fig. 2 (c) MALDI-TOF repressed *E. coli* spectrum**

Peak	MW from MALDI	Alpha-Numeric <sup>a</sup>	MW from 2-D Gel <sup>a</sup>
1	26 680.4	H027.0	26 618
2	29 055.1	I026.0	28 667
3	35 904.4	H021.7	35 911
4	39 103.4	B037.0	38 922
5	40 496.8	A036.1 or B036.1	40 446
6	45 108.8	G043.8	45 174
7	52 782.4	G046.5	52 578
8	53 488.9	?	?
9	56 138.4	G054.1	56 180
10	58 787.8	?	?
11	60 475.6	D057.0	60 422
12	77 922.8	C078.0	77 137
13	81 121.8	F080.0	81 559
14	82 240.4	?	?
15	89 227.1	?	?
16	92 151.3	?	?
17	97 038.1	D101.5	96 963
18	108 911.6	E106.0	108 192
19	114 485.2	?	?
20	118 547.7	E133.0	117 664

<sup>a</sup> This match-up is an approximation. The protein's alpha-numeric and 2-D gel MW are from Ref. 31.

implying that the maximum number of expressed proteins that are expected are present in the sample.

An Amicon Ultrafiltration system (Millipore) with a molecular weight cut-off of 30 kDa was employed to remove smaller proteins (<20 kDa) from the *E. coli* whole cell lysate to eliminate possible suppression effects from these components in the MALDI process. In comparison, Fig. 5 shows a MALDI-TOF spectrum of the filtered *E. coli* induced lysate. This spectrum is similar to the one found in Fig. 2 (a) except that fewer ion peaks are observed after filtration. It is likely that

some proteins are trapped in the filter membrane using this filtration system. Nevertheless, the induced mass spectrum in Fig. 5 can still be differentiated readily from the filtered uninduced (not shown) and repressed (not shown) mass spectra. Some of the *ara* gene products detected in this filtered sample include AraC (33 kDa), AraJ (42 kDa), AraF (51 kDa) and AraG (54 kDa) as indicated in Table 5. Naturally, these peaks are also observed in the unfiltered induced sample, Fig. 2 (a).

Clearly large proteins can be solubilised by the Triton/gu-HCl procedure<sup>14</sup> and observed in the MALDI mass spectrum. Chaotropics, such as gu/HCl and urea, do not cleave covalent bonds. Hence, they disrupt only secondary, tertiary and quaternary structures while leaving the primary structure intact. At high concentration gu-HCl or urea would allow water molecules to penetrate into the interior of proteins, thereby disrupting the hydrophobic interactions that normally stabilize the native conformation. This is the first step in solubilization of these proteins. Nonionic detergents, like Triton, will form micelles when the critical micelle concentration is achieved. Micelles are the aggregation of groups of detergent molecules. Suspension of nonpolar compounds, such as proteins, in water in this fashion is termed solubilization. In a study by Bornsen and co-workers,<sup>33</sup> when working with complex biological materials in MALDI-MS, detergent is necessary to prevent proteins, especially at low concentrations (<5 × 10<sup>-7</sup> M), from being rapidly adsorbed onto accessible surfaces. This study further reports that in protein analysis, detergent helps to prevent samples from agglomeration as well as to enhance their solubility.

In conclusion, MALDI-TOFMS along with gu-HCl/Triton treatment of bacterial samples provide a simple and rapid method for assessing the large proteins in a



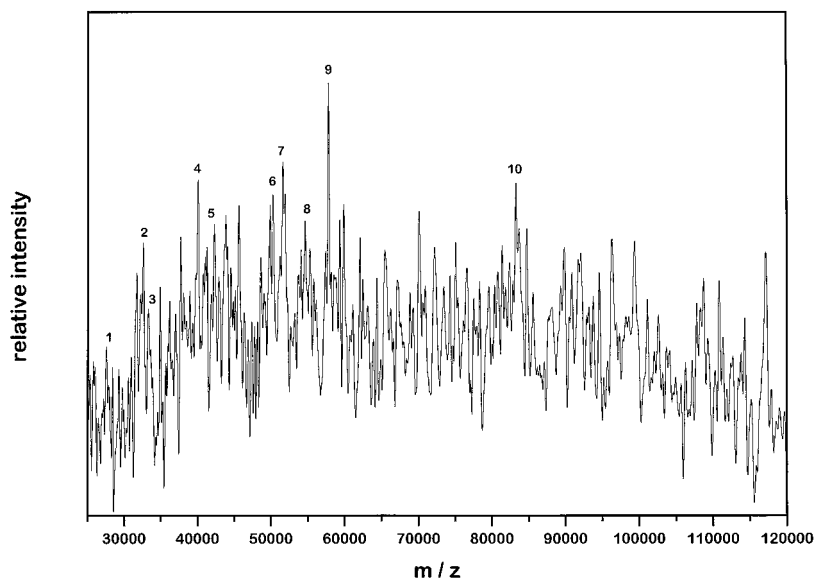


Figure 5. MALDI-TOF induced *E. coli* filtered spectrum ( $m/z$  25–120 kDa).

Table 5. Key to protein peaks in Fig. 5 MALDI-TOF induced *E. coli* filtered spectrum

Peak	MW from MALDI	Alpha-Numeric <sup>a</sup>	MW from 2-D Gel <sup>b</sup>	Calculated MW <sup>b</sup>
1	27 370.5	C032.7	27 614	—
2	32 581.9	G028.0	32 402	—
3	33 326.4	H030.2 ( <i>araC</i> )	33 384	33 315 <sup>21</sup>
4	40 026.8	A036.1 or B036.1	40 446	—
5	42 260.2	( <i>araJ</i> )	?	42 898 <sup>18</sup>
6	50 253.6	B046.7	50 171	—
7	51 566.3	( <i>araE</i> )	?	51 683 <sup>22</sup>
8	54 544.2	( <i>araG</i> )	?	54 947 <sup>17</sup>
9	57 698.5	E050.0	57 370	—
10	83 187.2	?	?	—

<sup>a</sup> This match-up is an approximation. The protein's alpha-numeric and 2-D gel MW are from Ref. 31.

<sup>b</sup> This match-up is an approximation. The arabinose induced protein peaks are also compared to the MWs from known sequences. See references.

bacterial whole cell lysate. This combination will enable MALDI-TOFMS to rapidly monitor changes in gene expression in bacteria as demonstrated in this work. A continuation of this endeavour includes the separation and isolation of these arabinose induced gene products as well as characterization of other large proteins. Ultimately, this technique will be applied to analyze complex biological systems such as wastewater microbial communities, bioaerosols in the environment and immunochemistry.

### Acknowledgements

We gratefully acknowledge support of this work under National Science Foundation Grant DEB 9120006 to the Center for Microbial Ecology at Michigan State University.

### REFERENCES

1. T. C. Cain, D. M. Lubman and W. J. Weber, Jr., *Rapid Commun. Mass Spectrom.* **8**, 1026 (1994).
2. R. D. Holland, J. G. Wilkes, F. Rafii, J. B. Sutherland, C. C. Persons, K. J. Voorhees and J. O. Lay, Jr., *Rapid Commun. Mass Spectrom.* **10**, 1227 (1996).
3. T. Krishnamurthy, P. L. Ross and U. Rajamani, *Rapid Commun. Mass Spectrom.* **10**, 883 (1996).

4. C. E. Parker, D. I. Papac, S. K. Trojak and K. B. Tomer, *J. Immunol.* **157** N1, 198 (1996).
5. M. Erhard, von Dohren and P. Jungblut, *Nature Biotechnol.* **15**, 901 (1997).
6. X. Liang, K. Zheng, M. G. Qian and D. M. Lubman, *Rapid Commun. Mass Spectrom.* **10**, 1219 (1996).
7. F. Foret, O. Muller, J. Thorne, W. Gotzinger and B. L. Karger *J. Chrom. A* **716** N1-2, 157 (1995).
8. F. C. Neidhardt, in *Escherichia coli and Salmonella typhimurium: Cellular and Molecular Biology*, American Society for Microbiology, Washington, DC (1987), p.3.
9. A. Verissimo, P. V. Morais, A. Diogo, C. Gomes and M. S. Dacosta, *Int. J. System. Bacteriol.* **46** N1, 41 (1996).
10. P. Cash, E. Argo and K. D. Bruce, *Electrophoresis* **16** N1, 135 (1995).
11. R. J. Cotter, *Anal. Chem.* **64**, 1027A (1992).
12. R. C. Caprioli, A. Malorni and G. Sindona, *Mass Spectrometry in Biomolecular Sciences*, Kluwer Academic, Boston, (1996), p.41.
13. D. C. Schriemer and L. Li, *Anal. Chem.* **68**, 2721 (1996).
14. D. J. Hettwer, *Ph.D. Dissertation*, Dept. of Chem. Eng., The Univ. of Michigan, Ann Arbor, MI, DISS 20795 (1987).
15. E. Englesberg, *J. Bacteriol.* **81**, 996 (1961).
16. E. Englesberg, R. L. Anderson, R. Weinberg, N. Lee, P. Hoffee, G. Huttenhauer and H. Boyer, *J. Bacteriol.* **84**, 137 (1962).
17. J. B. Scripture, C. Voelker, S. Miller, R. T. O'Donnell, L. Polgar, J. Rade, B. F. Horazdovsky and R. W. Hogg, *J. Mol. Bacteriol.* **197**, 37 (1987).
18. T. Reeder and R. Schleif, *J. Bacteriol.* **173**, 7765 (1991).
19. N. Lee, W. Gielow, R. Martin, E. Hamilton and A. Fowler, *Gene* **47**, 231 (1986).
20. J. Patrick and N. Lee, *J. Biol. Chem.* **16**, 4312 (1968).
21. C. G. Miyada, A. H. Horwitz, L. G. Cass, J. Timko and G. Wilcox, *Nuc. Acids Res.* **8**, 5267 (1980).
22. M. C. Maiden, M. C. Jones-Mortimer and P. J. Henderson, *J. Biol. Chem.* **263**, 8003 (1988).
23. L. Guzman, D. Medlin, M. J. Carson and J. Beckwith, *J. Bacteriol.* **177**, 4121 (1995).
24. A. Cramer, E. A. Whitehorn, E. Tate and P. C. Stemmer, *Nature Biotech.* **14**, 315 (1996).
25. J. Sambrook, E. F. Fritsch and T. Maniatis, *Molecular Cloning A Laboratory Manual*. Second ed. Harbor Press, Cold Spring (1989).
26. L. M. Preston, K. K. Murray and D. H. Russell, *Biol. Mass Spectrom.* **22**, 544 (1993).
27. Y.-H. Liu, J. Bai, X. Liang and D. M. Lubman, *Anal. Chem.* **67**, 3482 (1995).
28. R. M. Whittall and L. Li, *Anal. Chem.* **67**, 1950 (1995).
29. J. Bai, Y.-H. Liu, T. C. Cain and D. M. Lubman, *Anal. Chem.* **66**, 3423 (1994).
30. Y. D. Zhu, L. He, J. R. Srinivasan and D. M. Lubman, *Rapid Commun. Mass Spectrom.* **11**, 987 (1997).



31. R. A. VanBogelen and F. C. Neidhardt, *Electrophoresis* **12**, 995 (1991).
32. S. Pederson, P. L. Bloch, S. Reeh and F. C. Neidhardt, *Cell* **14**, 179 (1978).
33. K. O. Bornsen, M. A. S. Gass, G. J. M. Bruin, J. H. M. von Adrichem, M. C. Biro, G. M. Kresbach and M. Ehrat, *Rapid Commun. Mass Spectrom.* **11**, 603 (1997).