Application Note: 363

Analysis of a Complex Sample of Bacterial Whole Cell Lysate from *Haemophilus ducreyi* using the Finnigan LXQ

Kai Scheffler, Ph. D., Thermo Electron GmbH, Dreieich, Germany

Key Words

- Finnigan™ LXQ™
- Finnigan Surveyor Plus™ HPLC
- Linear ion trap
- Proteomics

Introduction

Haemophilus ducreyi is a gram-negative bacterium and a member of the *Pasteurellaceae* family. It is the causative agent of the genital ulcer disease chancroid. A detailed understanding of the pathogenesis of chancroid is an important research priority as chancroid has been shown to be an important cofactor in the heterosexual transmission of HIV. The genome of a human passage derivative *H. ducreyi* strain 35000 has recently been sequenced and is now publicly available.

The research focus in the past has often been on the glyconconjugate structures present on the surface of the bacterium, as the surface carbohydrates are believed to be involved in the attachment of the bacteria to the host cells. Bacterial membrane proteins are suspected to support the adhesion process and thus have been studied as well, also in a comparative analysis of different bacterial strains.

In the early 2000's, the elucidation of *H. ducreyi* proteome was initiated using the classical approach via 2D-gels (IEF and SDS-PAGE). To date, approximately 450 out of 1700 predicted distinct *H. ducreyi* proteins have been identified by applying modern proteomic techniques, such as high resolution mass spectrometry. In this study, we sought to leverage the power, speed and sensitivity of the Finnigan LXQ linear ion trap mass spectrometer to analyze a whole cell lysate of *H. ducreyi*. Two experimental parameters, gradient time and sample complexity, were varied to determine their effect on the number and quality of MS^n spectra generated and the number of proteins identified. Of particular interest was the identification of surface or membrane proteins, as these are likely to be critical to the pathogenesis of *H. ducreyi*.

Goal

Determine the feasibility of analyzing a complex mixture, represented by a whole cell lysate of *H. ducreyi*, on the Finnigan LXQ linear ion trap using a simple, 1D-SDS PAGE desalting and protein fractionation procedure. Study the effect of non-MS parameters, such as sample complexity and chromatographic resolution on the number and quality of the spectra generated and the resulting proteins identified, with particular emphasis on membrane and high molecular weight proteins.

Experimental

Sample Collection

The sample was prepared by lysing harvested *H. ducreyi* cells, which had been grown from a skim milk stock solution on chocolate agar plates in a candle jar, in 2D-gel sample buffer containing 7M urea, 2M thiourea, 65 mM DTT, and 4% CHAPS. Insoluble DNA and cell debris were removed by centrifugation at 12.000×*g*.

Sample Preparation (1D-SDS PAGE)

The lysate was mixed with 1D gel SDS sample buffer and run on a 4-20% SDS gel to reduce the high salt content. Approximately 100 µg of total protein was loaded per well. Electrophoresis was halted after the dye front had migrated about 2 cm into the gel. Coomassie staining was used to visualize protein bands.

Fractionation (Gel Slices)

The whole cell protein complement was prepared by excising the stained portion of one entire gel lane, dicing it into five smaller pieces (1.5×1.5 mm) and transferring each piece to a fresh test tube, giving 5 distinct samples. Fractionating in this way reduced the overall sample complexity of the whole cell complement. All gel pieces were completely destained followed by reduction/alkylation with 10 mM DTT at 56 °C for 1 hr and 55 mM iodoacetamide for 45 min in the dark at room temperature.

Enzymatic Digestion

Proteins in the gel sections were enzymatically digested overnight at 37 °C. The supernatant containing the peptide mixtures was transferred into five clean test tubes and extracted using 1% formic acid followed by an organic extraction using 50% ACN/5% formic acid.

HPLC Separation

Samples were separated on a capillary HPLC column with an integrated nanospray tip, packed with MagicMS C18, 5 µm particles, 300Å pores (Spectronex, Switzerland) using the Finnigan Surveyor Plus HPLC system. The flow rate was set at 130 µL/min but split to an actual flow rate of 600 nL/min using the splitting option of Thermo's PepFinder Kit.™ Gradients of acetonitrile in water, both

Thermo ELECTRON CORPORATION

containing 0.1% formic acid over two different linear time scales (45 min and 150 min) were tested (Table 1). To minimize baseline and cluster ion formation, all mobile phase solvents were LC-MS CHROMASOLV® grade (Sigma-Aldrich, St. Louis, MO, USA).

Short – 45 min gradient			Long – 150 min gradient		
Time (min)	% A	% B	Time (min)	% A	% B
0	98	2	0	98	2
1	98	2	1	98	2
21	40	60	90	60	40
23	20	80	115	40	60
28	20	80	120	20	80
29	98	2	125	20	80
45	98	2	130	98	2
			150	98	2

Table 1: Gradient profiles

A: 0.1% formic acid in water, B: 0.1% formic acid in CH_3CN. Flow rate: 130 µL/min split to 600 nL/min

Mass Spectrometry

A Finnigan LXQ linear ion trap was operated in positive ion nanospray mode, using a Top 4 Data Dependent™ Double Play method as follows:

1. Full scan MS: 400–1400 *m/z* (survey scan) in enhanced scan mode for improved resolution.
2. MS/MS of the four most intense precursor ions from the survey scan.

Scan Settings

Charge state screening enabled
Charge state rejection enabled
Unassigned charge states: rejected
Charge state 1: not rejected
Charge state 2: not rejected
Charge state 3: not rejected
Charge states 4+: rejected

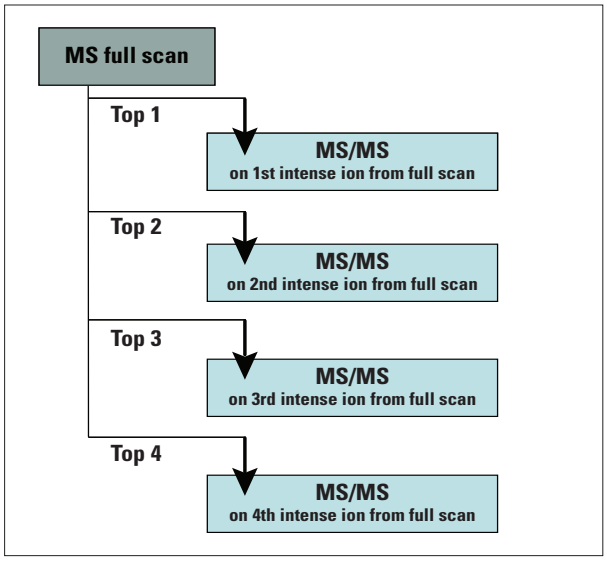

Figure 1: A schematic of the MS experiment

Protein Identification using BioWorks 3.2

For protein identification, the SEQUEST® algorithm within BioWorks™ 3.2 software was used to search an indexed form of the publicly available *Haemophilus ducreyi* FASTA database containing 1707 entries.
For indexing, the following parameters were set:

1) Database type: protein
2) Mass type: monoisotopic
3) MW range: 400–4000
4) Missed cleavage sites: 2
5) Differential Modifications:
   Met: Oxidation (+15.99492)
   Cys: Carboxyamidomethylation (+57.05130)
   Lys, Arg, Cys: Carbamylation (+43.00582)

(Note: Cysteine disulfide bridges were reduced by treatment followed by alkylation with iodoacetamide. Carbamylation of lysine, arginine, and cysteine residues is a side reaction of boiling in the SDS sample buffer in the presence of the urea used as a lysis buffer.)

BioWorks Result Filters:

Two different peptides were required for a protein to be considered a positive match:
Xcorr vs. Charge State: (+1)1.5, (+2) 2.0, (+3) 2.5
Peptide Probability: 5×10^{-2}

Results

1D-SDS PAGE Electrophoretic Fractionation

Since one aim of this application was to have a rapid desalting and fractionation of the whole cell lysate prior to LC-MS/MS analysis, the sample was allowed to migrate only 2 cm into the gel. This distance was sufficient to remove the salts from the protein bands and minimize dilution of the proteins in the gel matrix, yet still provide crude fractionation of the proteins based on electrophoretic mobility. Figure 2 shows the Coomassie stained 1D-SDS PAGE gel and the fractionation that was performed to yield the two different samples.

Fast (45 min) Gradient – Whole Gel Band

The complexity of the sample extract is evident in the base peak chromatogram shown in Figure 3a. Here, the whole cell lysate was separated using the fast (45 min) gradient. A SEQUEST search of 2,639 MS/MS spectra resulted in the identification of 79 proteins, with a partial list from BioWorks 3.2 shown in Figure 4, including seven membrane proteins and three proteins with a molecular weight above 100 kDa. Typically, higher molecular weight proteins (>200 kDa) as well as most membrane proteins are lost when using a 2D-gel fractionation separation method. Sequence coverage of ~80% was obtained for several of these proteins.

The rich quality of the MS/MS spectra generated by the LXQ is typified in Figure 5. Note the high number of fragment ions in the spectrum matching expected ions

from the candidate peptide sequence–in this case 15 amino acids in length with a modified lysine residue. Fragment ion masses in blue denote matched y-ions, pink denotes y-ions with a loss of NH_3 or water, red denotes b-ions, and orange denotes b-ions with a loss of NH_3 or water.

Long (150 min) Gradient – Whole Gel Band

The greater chromatographic resolution achieved by increasing the gradient from 45 to 150 min (Figure 3a and b) enabled the identification of nearly four times the number of proteins from the same sample. Using identical SEQUEST search criteria and results filters as used for the 45-min run, the 150-min gradient resulted in the identification of 222 proteins. Eleven of the identified proteins were membrane-derived, seven had a molecular weight larger than 100 kDa, and two very large supernatant proteins were found with a molecular weight greater than 450 kDa.

Long (150 min) Gradient – Five Individual Gel Slices

The complexity of the whole cell lysate sample was further reduced by both increasing the chromatographic gradient to 150 min and dividing the gel band into five sections. This combination reduced the sample complexity by roughly 80%, gave a 5-fold increase in the number of spectra searched and a 62% increase in the number of proteins identified compared to the whole cell lysate under the same conditions.

Each of the five gel sections was analyzed separately using identical SEQUEST search criteria and results filters as the previous data, but consolidated into a single MultiConsensus report, resulting in the identification of 360 proteins. Two proteins were identified with a molecular weight higher than 450 kDa (only found in the two top gel sections), 16 proteins were identified with a molecular weight higher than 100 kDa, and 12 membrane proteins were identified.

A summary of the results of the three different analyses appears in Table 2.

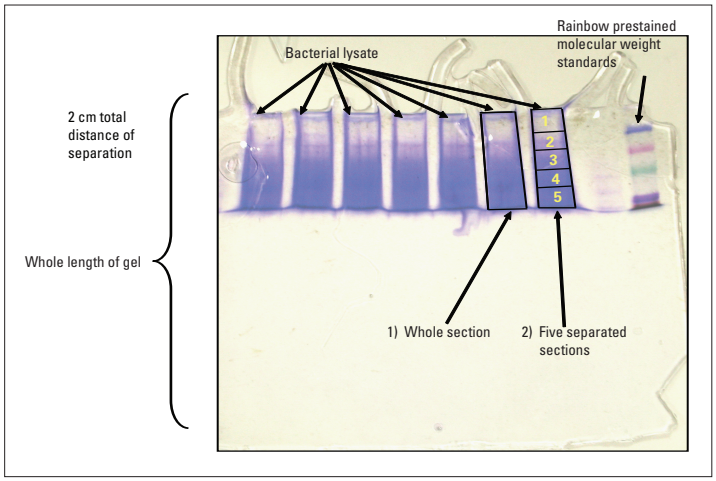

Figure 2: Coomassie stained 1D SDS-PAGE gel of *Haemophilus ducreyi* whole cell lysate

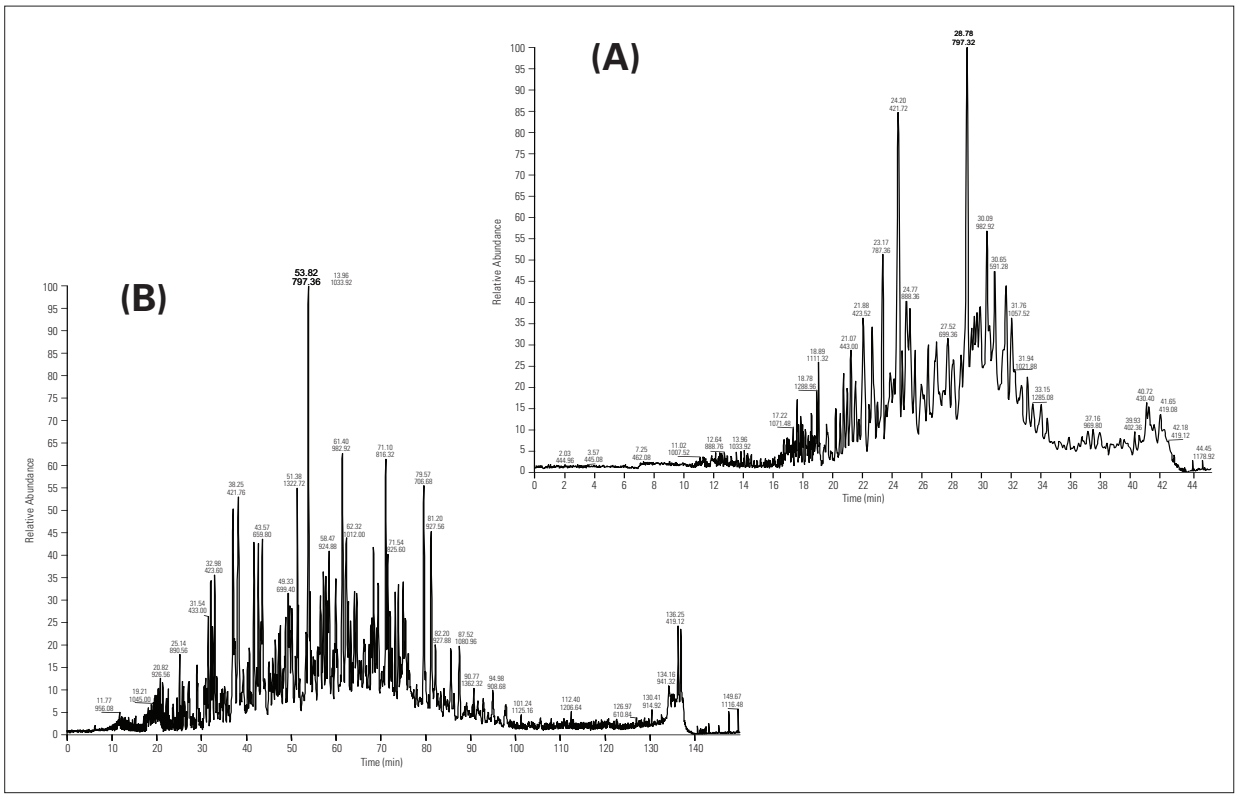

Figure 3: Base peak chromatogram, *Haemophilus ducreyi* whole cell lysate (aqueous extract), A: 45-min run, B: 150-min run

LXQ Spectral Quality, Scan Speed (Cycle Time) and Sensitivity

The upper panel of Figure 6 shows the scan sequence for the specific time window 42.23 to 42.52 min from the 150-min gradient run of the whole cell lysate. The black vertical lines represent full MS scans, each of which is followed by four MS/MS scans (red vertical lines). The height of the lines represents the ion intensities in the corresponding spectra. All MS/MS scans are labeled with numbers 1 to 32. In this brief time window of only 0.29 min (17.4 sec), the LXQ performed nine full scans plus 32 MS/MS scans.

To demonstrate sensitivity (i.e. how many peptides of varying abundance were detected in the full scans during this time frame), nine full scan MS spectra were summed and averaged as shown in the lower panel of Figure 6. The numbers in the top and bottom panels correspond, showing that a total of 32 MS/MS spectra from different precursor ions were selected for fragmentation. This data clearly demonstrates the scan speed (top panel) and the sensitivity (lower panel) with both abundant ions (such as peaks 5, 2, 13, and 14), and trace ions from the full MS spectrum. The averaged full scan MS spectrum in the lower panel demonstrates the power of Dynamic Exclusion,™ as each of the peaks labeled with red boxes were selected for fragmentation. The peak at *m/z* 587 (labeled with an asterisk) was not selected in this time window because it had been selected previously and was on the exclusion list.

A further example of the quality of the spectra generated by the LXQ appears in Figure 7, with the full scan spectrum (upper panel) at RT 42.27 expanded from Figure 6, plus the MS/MS spectra of precursor ions 5, 6, 7, and 8, which were selected for fragmentation (lower panel).

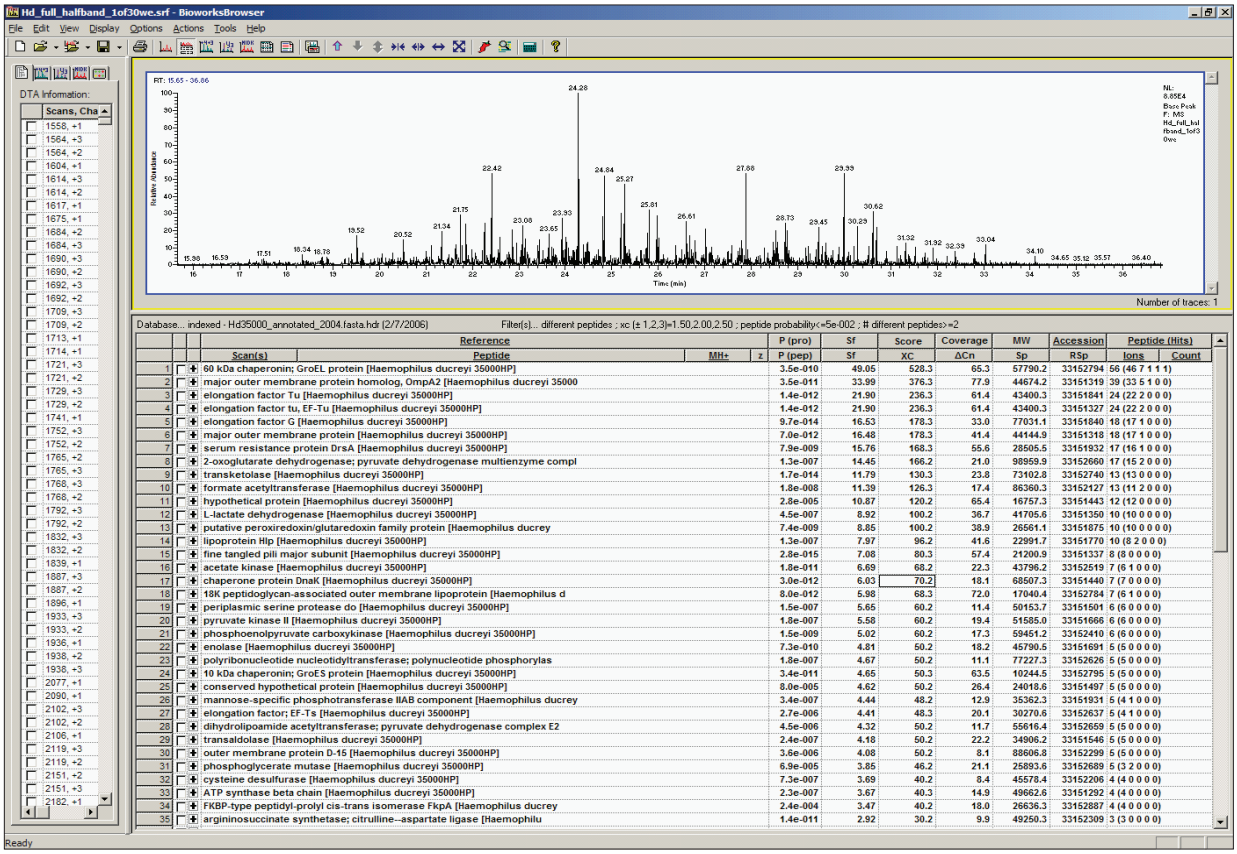

Figure 4: BioWorks search results, proteins 1–35, 45-min run

Sample Preparation	Gradient	# of MS/MS spectra searched	# of proteins identified	# of membrane proteins	# of proteins with MW > 100 kDa
Whole gel band	45 min	2,639	79	7	3
Whole gel band	150 min	22,372	222	11	9 (Two proteins > 450 kDa)
5 gel sections (Consolidated)	150 min	107,486	360	12	18 (Two proteins > 450 kDa)

Table 2: Summary of results

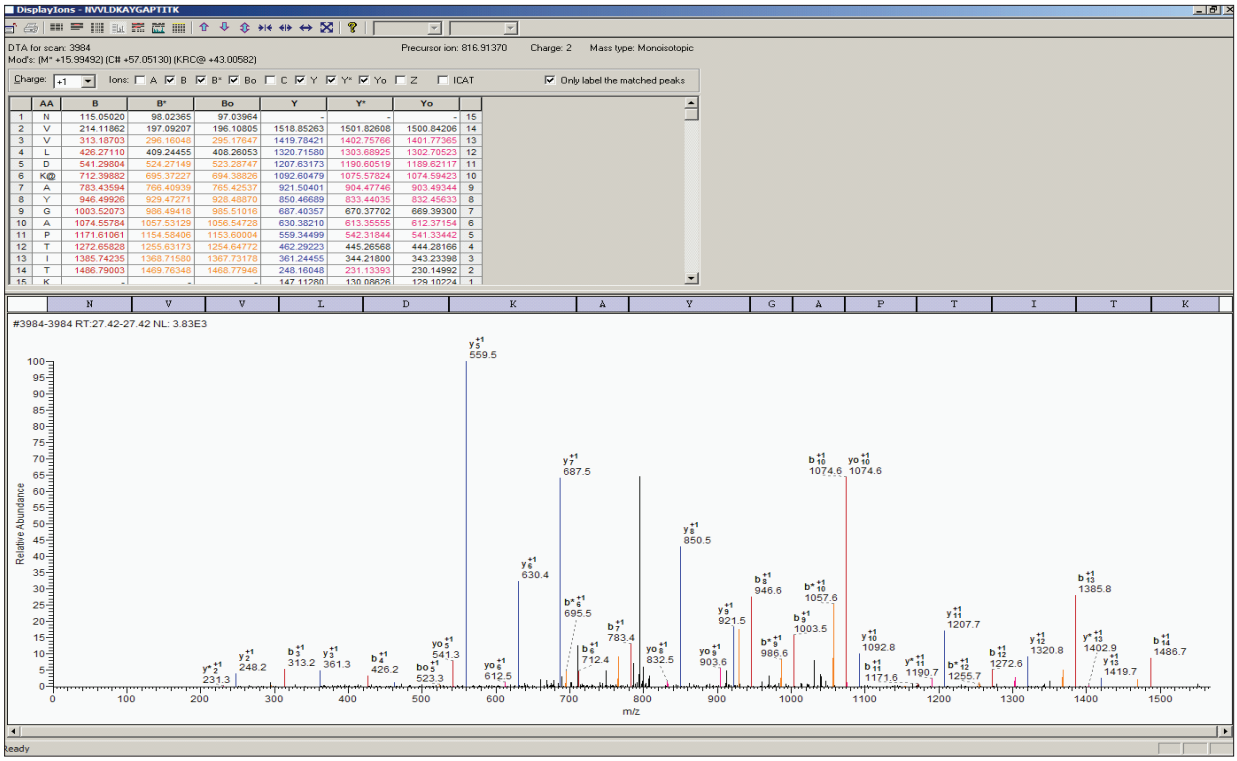

Figure 5. Example of assigned MS/MS spectrum of peptide NVVLDK*AYGAPTITK

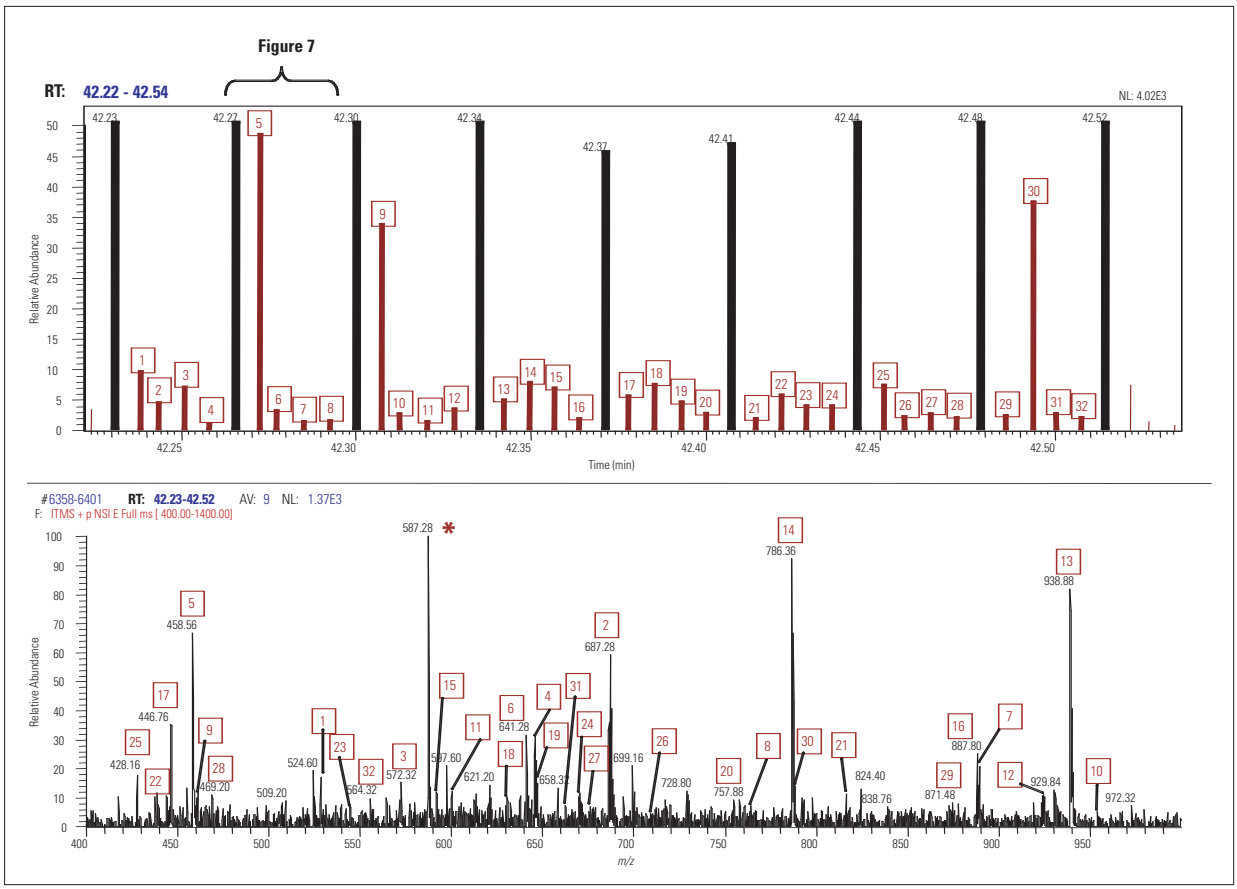

Figure 6: Scan sequence for a specific time window: 42.23 — 42.52 min., whole gel band, 150 minute gradient.
Upper panel: Full scan (black lines) followed by four MS/MS scans (red lines). Lower panel: Average of nine full scan MS spectra.

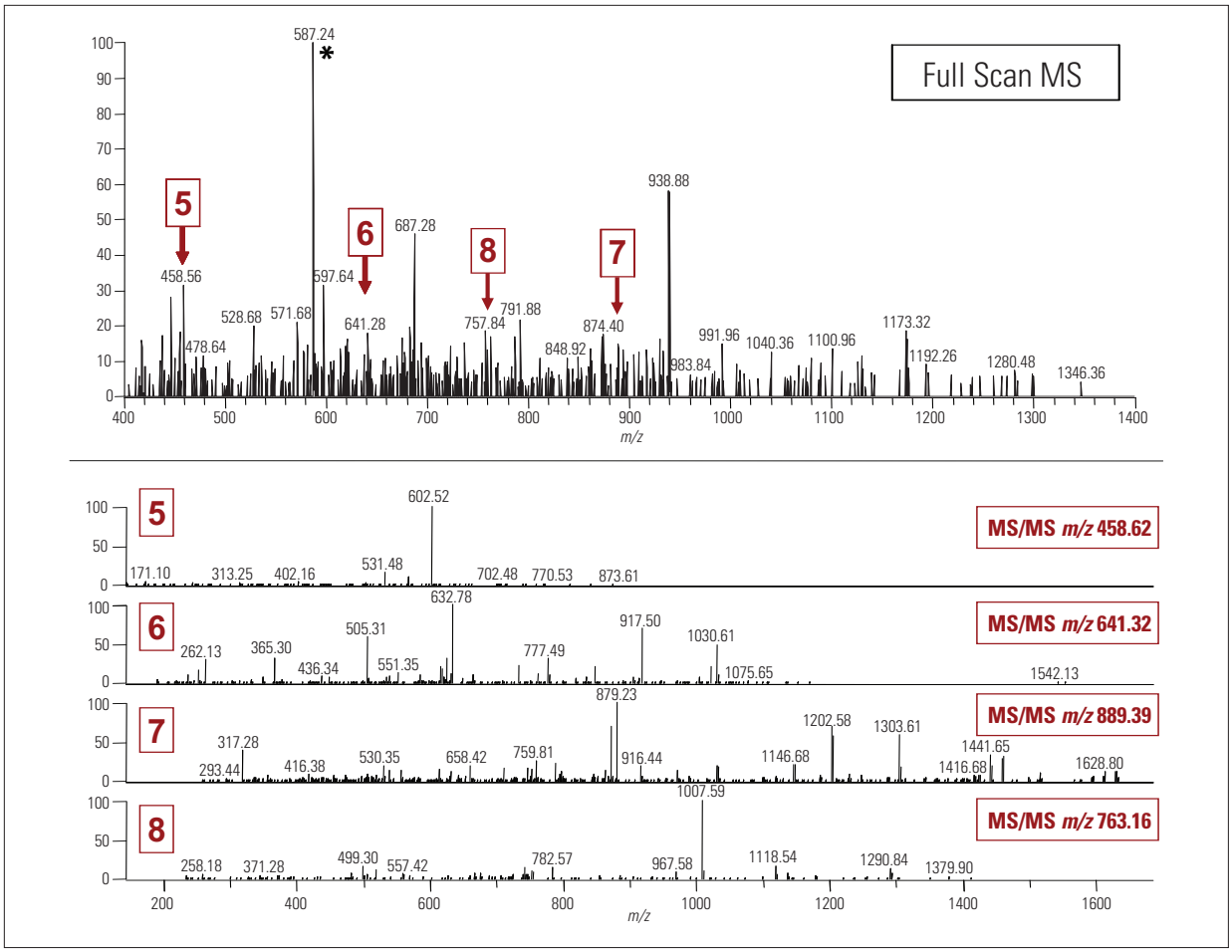

Figure 7: Full scan at RT 42.27 from Figure 6 plus the MS/MS spectra of precursor ions 5-8.
Upper panel: Full scan at RT 42.27 from Figure 6. Lower panel: MS/MS spectra 5-8.

Conclusions

This study was performed to demonstrate a rapid and sensitive protocol for the analysis of a whole cell lysate using a quick and simple 1D-SDS PAGE fractionation step followed by LC/MS^n analysis using the Finnigan LXQ linear ion trap mass spectrometer. The data clearly shows both the speed and sensitivity of the Finnigan LXQ, ideal for analyzing complex proteomics samples. Lengthening the chromatographic gradient as well as reducing the complexity of the sample via gel fractionation increased the number of SEQUEST-searchable MS^n spectra and subsequent proteins identified.

In this experiment, seventy-nine proteins were identified in a 45-min run, 222 proteins were identified by extending the gradient to 150 min, and 360 proteins were found in total by dividing the whole cell lysate gel band into five distinct slices and using the 150-min gradient. A variety of small and large membrane proteins, which are linked to pathogenesis, and large molecular weight proteins, which are often lost using 2D-gels, were identified. By employing the longer gradients, two very large supernatant proteins (>450 kDa) were also identified.

References

Spinola, S.M.; Bauer, M.E.; Munson, R.S., Jr.; Immunopathogenesis of *Haemophilus ducreyi* Infection (Chancroid). *Infection and Immunity* **2002**, *70*, 1667-1676.

Janowicz, D.; Luke, N.R.; Fortney, K.R.; Katz, B.P.; Campagnari, A.A.; Spinola, S.M.; Expression of OmpP2A and OmpP2B is not required for pustule formation by *Haemophilus ducreyi* in human volunteers. *Microb. Pathog.* **2006**, *40(3)*, 110-115.

Elias, J.E.; Haas, W.; Faherty, B.K.; Gygi, S.P.; Comparative evaluation of mass spectrometry platforms used in large-scale proteomics investigations. *Nature Methods*. **2005**, *2(9)*, 667-675.

Kolker, E.; Purvine, S.; Galperin, M.Y.; Stolyar, S.; Goodlett, D.R.; Nesvizhskii, A.I.; Keller, A.; Xie, T.; Eng, J.K.; Yi, E.; Hood, L.; Picone, A.F.; Cherny, T.; Tjaden, B.C.; Siegel, A.F.; Reilly, T.J.; Makarova, K.S.; Palsson, B.O.; Smith, A.L.; Initial proteome analysis of model microorganism *Haemophilus influenzae* strain Rd KW20. *J. Bacteriol.* **2003**, *185(15)*, 4593-4602.

Sun,W; Wu, S; Wang, X; Zheng, D; Gao,Y.; A systematical analysis of tryptic peptide identification with reverse phase liquid chromatography and electrospray ion trap mass spectrometry. *Genomics Proteomics Bioinformatics* **2004**, *2(3)*, 174-183.

Acknowledgement

The *Haemophilus ducreyi* sample was kindly provided by Dr. Deborah Post at the Buck Institute's Institute for Chemistry in Novato, CA.

In addition to these offices, Thermo Electron Corporation maintains a network of representative organizations throughout the world.

Australia
+61 2 8844 9500
Austria
+43 1 333 50340
Belgium
+32 2 482 30 30
Canada
+1 800 532 4752
China
+86 10 5850 3588
France
+33 1 60 92 48 00
Germany
+49 6103 4080
India
+91 22 2778 1101
Italy
+39 02 950 591
Japan
+81 45 453 9100
Latin America
+1 512 251 1503
Netherlands
+31 76 587 98 88
Scandinavia
+46 8 556 468 00
South Africa
+27 11 570 1840
Spain
+34 91 657 4930
Switzerland
+41 61 48784 00
UK
+44 1442 233555
USA
+1 800 532 4752

www.thermo.com

AN62095_E 04/06S

Thermo ELECTRON CORPORATION