INITIAL PEPTIDE MASS FINGERPRINTING ANALYSIS OF PROTEINS OBTAINED BY LYSIS OF COXIELLA BURNETII CELLS

Ľ. ŠKULTÉTY[1], L. HERNYCHOVÁ[2], R. TOMAN[1], M. KROČA[3], J. STULÍK[2], A. MACELA[2*]

[1]Institute of Virology, Slovak Academy of Sciences, 845 05 Bratislava, Slovak Republic; [2]Institute of Molecular Pathology and Proteome Centre for the Study of Intracellular Parasitism of Bacteria, Purkyně Military Medical Academy, 500 01 Hradec Králové, Czech Republic; [3]Central Military Medical Institute, 561 66 Techonin, Czech Republic

Received December 12, 3003; accepted March 3, 2004

Summary. – Whereas the complete genome of *Coxiella burnetii (C.b.)*, the etiological agent of Q-fever, has recently been published (Seshadri *et al.*, *Proc. Natl. Acad. Sci. USA*. **100**, 5455–5460, 2003), the *C.b.* proteome is still under study. Using the bioinformatic approach, we found in total 309 proteins on two dimensional electroctrophoretic images of *C.b.* whole cell lysates. Eighteen major protein species were subjected to peptide mass fingerprinting and identified as the products of 6 known open reading frames (ORFs): the chaperone DnaK (heat shock 70 K protein), chaperonin 60 K (GroEL protein, heat shock protein B), DnaJ-like protein djlA (mucoidy activation protein mucZ), elongation factor Ts (EF-Ts), ribosomal protein L7/L12, and chaperonin 10 K (GroES protein, heat shock protein A).

Key words: *Coxiella burnetii*; proteins; peptide mass fingerprinting; 2-D electrophoresis; MALDI-ToF

Introduction

Coxiella burnetii (C.b.), the etiological agent of Q fever, is one of the bacteria listed as biological warfare agents of category B. It is spread worldwide and is responsible for an acute and potentially severe diseases characterized by pneumonitis, hepatitis, and a significant incidence of neurological complications (Marrie, 1990). The bacterium is unique among *Rickettsiae* in that it undergoes a virulent (phase I) to low virulent (phase II) variation upon serial passages in embryonated hen eggs. This phase variation has been shown to cause noticeable modifications in both composition and structure of the lipopolysaccharide (Ftáček *et al*., 2000) and phospholipid (Domingues *et al*., 2002) components of the cell outer membrane. In this regard, however, there is a lack of information about the *C.b.* proteins, representing major immunoreactive antigens in serological diagnosis of Q fever. Therefore, we have started mapping of the *C.b*. proteins. We focused first on the proteins expressed in the *C.b.* strain RSA 493 in virulent phase I.

*Corresponding autor. E-mail: amacela@pmfhk.cz, fax: +420 49-5513018.

Abbreviations: *C.b* . = *Coxiella burnetii*; IEF = isoelectric focusing; pI = isoelectric point; IPG = immobilized pH gradient; 2-D = two dimensional; SDS-PAGE = polyacrylamide gel electrophoresis in the presence of sodium dodecyl sulfate; ORF = open reading frame; MALDI-ToF = matrix assisted laser desorption/ionization time of flight; M_r = relative molecular weight; MS = mass spectrometry; *m/z* = mass-to-charge ratio

Materials and Methods

Cultivation and purification of C.b. C.b. strain RSA 493 (Nine Mile), serologically in the virulent phase I (yolk sac passage 3 in our laboratory), was obtained from the WHO Collaborating Centre for Rickettsial Reference and Research at the Institute of Virology, Slovak Academy of Sciences, Bratislava, Slovak Republic. *C.b*. was propagated in embryonated hen eggs, killed with 0.45% phenol, and purified by differential centrifugation and ether treatment (Škultéty *et al*., 1998).

Lysis of C.b. cells. The cells (20 mg) were suspended in 10 ml lysis buffer consisting of 137 mmol/l NaCl, 10% glycerol, 1% p-octyl-β-D-glucopyranoside (Serva), 50 mmol/l sodium fluoride, 1 mmol/l sodium orthovanadate (Sigma-Aldrich), and one tablet

of Complete Mini-Protease Inhibitors (Boehringer). Ultrasonication (30s, Tesla, Czechoslovakia) and 10 cycles of freeze-thaw in liquid nitrogen have been used for disruption of *C.b.* cells. The undisrupted cells were removed by centrifugation at 12,578 x g for 15 mins at 4°C.

2-D electrophoresis. Bacterial proteins were precipitated in 20% (w/v) trichloroacetic acid (Sigma-Aldrich) in acetone containing 0.2% dithiothreitol (USB) at -18°C overnight (Görg *et al*., 1997). Then, they were solubilized in IEF buffer containing 9 mol/l urea (USB), 4% (w/v) CHAPS (USB), 70 mmol/l dithiothreitol, and 5% (v/v) carrier ampholytes pH 9–11 (Sigma). Protein concentration was determined by a modified bicinchoninic acid protein assay (Brown *et al.,* 1989). The solubilized proteins were separated on immobilized pH gradient (IPG) strips (Amersham, pH 3–10) in the first dimension. Either 75 μg of protein for analytical or 225 μg for preparative gels were loaded by in-gel rehydratation overnight. The rehydratation solution contained 2 mol/l thiourea (Sigma-Aldrich), 6 mol/l urea, 4% (w/v) CHAPS, 40 mmol/l Tris base (Sigma-Aldrich), 2 mmol/l tributyl phosphine (Fluka), 0.003% (w/v) bromphenol blue (LKB), 1% (v/v) pharmalytes pH 3–10 (Amersham) and 0.5% (v/v) pharmalytes pH 8–10.5.

Isoelectric focusing (IEF) was performed on a Multiphor II (Amersham Biosciences). In the second dimension, a gradient 9–16% polyacrylamide gel electrophoresis in the presence of sodium dodecyl sulfate (SDS-PAGE) was used on a Protean II xi Multi Cell system (BioRad). The 2-D electrophoresis reference map of *C.b.* strain RSA 493 was obtained by visualization of proteins with silver staining (Hochstrasser *et al.,* 1988). An improved silver staining, compatible with protein digestion and mass spectrometry (MS) analysis (Mortz *et al.,* 2001), was applied to preparative gels. The gels were scanned by a personal densitometer (Molecular Dynamics, Sunnyvale, CA, USA). The image analysis was carried out using the Melanie III software package (BioRad). The isoelectric points (pI) and the relative molecular weights (M_r) of proteins were approximated using polypeptide SDS-PAGE standards (BioRad).

In-gel digestion. Spots selected from the preparative gel were excised and destained for a few minutes in a freshly prepared mixture of 15 mmol/l potassium ferricyanide (Merck) in 50 mmol/l sodium thiosulfate (Sigma) followed by washing with deionized water (Mili-Q system) and equilibration in 50 mmol/l ammonium hydrogen carbonate pH 7.8 (Fluka) in 5% acetonitrile (Sigma-Aldrich). The gel pieces were dried in a Speed Vac (Eppendorf) and reswelled in 10–30 μl of the equilibration buffer to which 0.1 μg of the porcine trypsin (sequence grade modified, Promega) was added. After overnight incubation at 37°C, an aliquote (1 μl) of the supernatant was directly used for MALDI-ToF MS analysis. The remaining supernatant was removed and combined with the acetonitrile extracts of the gel pieces. The mixture was concentrated in the SpeedVac and stored at -20°C.

MS. The samples were mixed in an Eppendorf tube with the same volume of a matrix solution of 2,5-dihydroxybenzoic acid (10 mg/ml) in 20% acetonitrile containing 0.5% trifluoroacetic acid. Alternatively, a matrix solution of α-cyano-4-hydroxycinnamic acid (10 mg/ml) in 50% acetonitrile containing 0.5% trifluoroacetic acid was applied. The mixtures, 1 μl of each, were applied to a MALDI sample plate, allowed to air-dry, and inserted into the mass spectrometer. Peptide mass fingerprint spectra were recorded with a Voyager DE STR (Perseptive Biosystems, Framingham, MA, USA) equipped with a delayed extraction. The spectra were obtained in a positive reflectron mode with the 20 kV acceleration voltage, 75% grid voltage, 0.02 wire voltage, 100 ns delay time and low mass gate at 500 *m/z*. External calibrations were carried out with standards – angiotensin I, angiotensin II, adrenocorticotropic hormone (clip 1–17) and (clip 18–39) (Laser Bio Labs, France) covering the 1000–2500 Da mass range. Each mass spectrum was obtained by averaging 100 laser shots. ProteinProspector program (University of California, San Francisco Mass Spectrometry Facility, CA, USA) was applied to match the experimentally determined peptide mass values against the *C.b.* database. A MALDI-ToF peak list was searched against this database using variation of parameters (a basic set up: monoisotopic mass resolution, 100 ppm error, carbamidomethylated Cys, 1 missed cleavage, minimum 4 peaks required to match, wide open window of pI and M_r).

Results and Discussion

In this work, we attempted to develop a suitable system for separation, isolation and identification of the *C.b.* proteins and to initiate creation of a *C.b.* protein database using the bioinformatic approach. The proteome analysis was aimed at generation of the reference protein map of the bacterium and selection of useful molecular determinants for its identification and typing. Moreover, the protein database can be a useful tool for screening immunoreactive bacterial proteins, searching for genes controlling expression of virulence traits and virulence factors themselves. For a more practical reason, mapping of bacterial surface proteins can be utilized for both rapid detection and identification of the microbial taxons.

C.b. cells strain RSA 493 were disrupted by lysis and the proteins were isolated. IEF followed by SDS-PAGE were used for their separation. Silver stained reference map was then compared to the preparative gel by the Mellanie III software and all protein spots have been detected. Thus, 309 proteins (Fig 1.) differing one from another in their pIs and M_r were found.

The spots containing protein species were cut out from the preparative gel, peptides were generated from them by trypsin digestion and the mass values (*m/z*) of peptides were determined by MALDI-ToF. An example of a MALDI-ToF spectrum is shown on Fig. 2. Matching of the real peptide mass values against the theoretical peptides listed in the SWISS-PROT/TrEMBL protein sequence databases created the most optimal hypothetical protein list. The peptide sequences were found in the *C.b.* RSA 493 ORF libraries generated from the known complete *C.b.* genome (Seshadri *et al.*, 2003).

Due to a low amount of proteins in minor spots it was not possible to obtain their sequences. At present, 18 protein

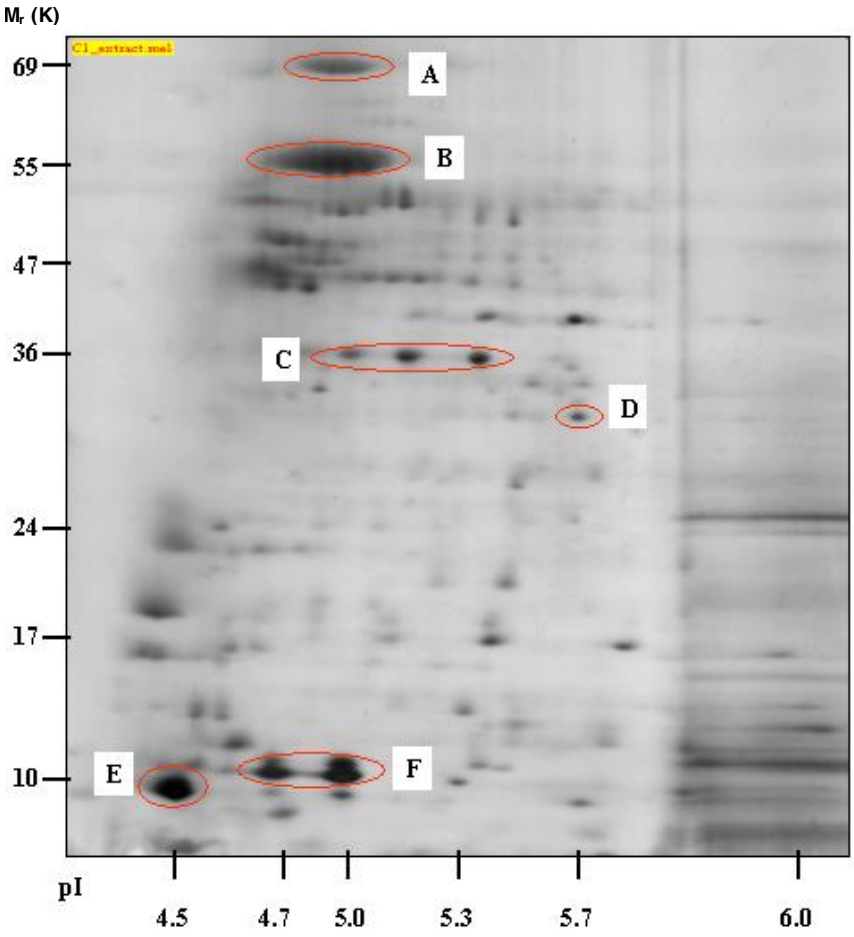

Fig. 1

Silver stained 2-D electrophoresis reference map of *C.b.* strain RSA 493.

Only a part of the 2-D electrophoresis reference map corresponding to the pI range of 4.4–6.2 and to the M_r range of 8–69 K is shown. The identified protein spots are encircled.

species in the regions A (pI = 4.9–5.1, M = 69 K), B (pI = 4.8–5.2, M = 55.2K), C (pI = 5.1–5.4, M = 36.4 K), D (pI = 5.7, M = 32 K), E (pI = 4.5, M = 9.4 K), and F (pI = 4.7–5.0, M = 10 K) were detected (Table 1). They represent the protein products of 6 different ORFs. Multiplicity of the protein identity among protein species can predict potential post-translational modifications resulting in proteins with roughly the same M_r and different pI detectable by the 2-D electrophoresis employed in this study. The chaperone DnaK (heat shock 70 K protein), chaperonin 60 K (GroEL protein, heat shock protein B), DnaJ-like protein djlA (mucoidy activation protein mucZ), elongation factor Ts (EF-Ts), ribosomal protein L7/L12, and chaperonin 10 K (GroES protein, heat shock protein A) were identified thus far. In Table 1, the experimental pI and M_r values are compared with the theoretical ones derived from the hypothetical proteins. Major differences between the observed and expected M_r values of proteins were found with the proteins having experimental M_r values below 18 K. These discrepancies might be associated with post-translational modifications of proteins mentioned above. However, a possibility of incorrectly identified start and end sites of the putative ORFs should also be taken into consideration.

A structural constituent of ribosome – ribosomal protein L7/L12 and cytoplasmic elongation factor Ts, are essential components of translational machinery (Seshadri *et al.*, 2003; Seshadri *et al.,* 1999). The identified chaperones and chaperonins work as protein-folding machines (Bhutani and Udgaonkar, 2002). Thus, the GroEL protein folds nascent polypeptide and misfold proteins, while the GroES protein functions as a co-chaperonin. The chaperone protein DnaK serves for protein assembly and translocation across the membrane (Bhutani and Udgaonkar, 2002) and interacts with

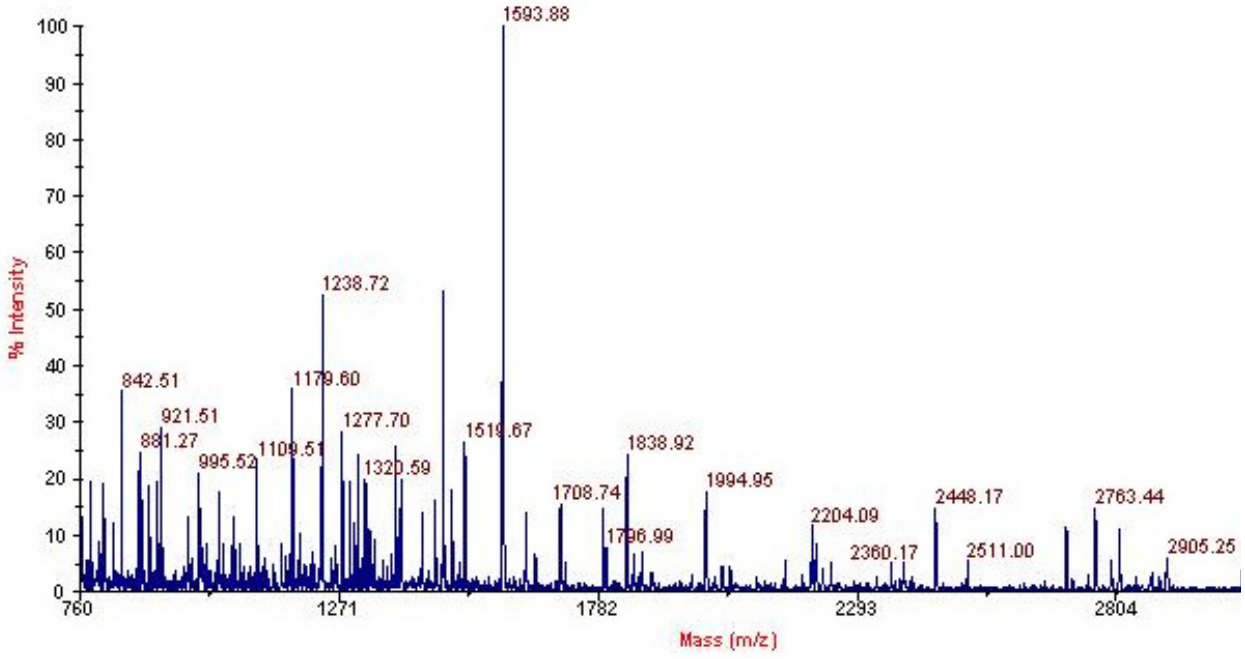

Fig. 2

Positive-ion MALDI-ToF mass spectrum of chaperonin 60 K from *C.b.* strain RSA 493

DnaJ-like protein djlA (Zuber *et al.*, 1995). In addition, association of DnaK with outer membrane as well as cytoplasm has been reported (Macellaro *et al.,* 1998). A 60 K heat shock protein has been localized, for example, in the cytoplasm of *Borrelia burgdorferi,* but its significant fraction was also associated with the cell membrane (Scopio *et al.,* 1994). Most likely, the protein has a dual association also in *C.b.* as it is immunogenic (Vodkin and Williams, 1988). For this reason, it might serve as an efficacious vaccine against *C.b.* and other pathogenic microorganisms that express the conserved antigen. The proteins detected in this work have been found also in several bacteria analyzed in our laboratories thus far, e.g. in the intracellular bacterium *Francisella tularensis* (Hubálek *et al*., 2003). Therefore, for an unambiguous bacterial identification by peptide mass fingerprinting, more studies focused on precise amino acid sequence analysis should be done.

In conclusion, this work represents our initial attempt to create a proteome database of *C.b.* and to find unique protein/peptide molecules – biomarkers that could be used for a rapid and unambiguous detection of the bacterium. Further work following this goal is in progress.

Acknowledgements. The authors acknowledge excellent technical assistance of Mmes M. Benkovičová, A. Firychová, and J. Michaličková. This work was supported by the following grants: the grant LN00A033 from the Ministry of Education, Youth and Sports, Czech Republic, the grants MO 9079201301 and 9079201303 from the Ministry of Defence, Czech Republic, and the grant No. 2/3021/23 of the Scientific Grant Agency of the Ministry of Education of the Slovak Republic and the Slovak Academy of Sciences.

References

Bhutani N, Udgaonkar JB (2002): Chaperonins as protein-folding machines. *Curr. Sci.* **83**, 1337–1351.

Brown RE, Jarvis KL, Hyland KL (1989): Protein measurement using bicinchoninic acid: elimination of interfering substances. *Anal. Biochem.* **180**, 136–139.

Domingues P, Palkovič P, Toman R (2002): Analysis of phospholipids from *Coxiella burnetii* by fast atom bombardment mass spectrometry. A rapid method for differentiation of virulent phase I and low virulent phase II cells. *Acta Virol.* **46**, 121–124.

Ftáček P, Škultéty Ľ, Toman R (2000): Phase variation of *Coxiella burnetii* strain Priscilla: influence of this phenomenon on biochemical features of its lipopolysaccharide. *J. Endotoxin Res*. **6,** 369–376.

Gorg A, Obermaier C, Boguth G, Csordas A, Diaz JJ, Madjar JJ (1997): Very alkaline immobilized pH gradients for two-dimensional electrophoresis of ribosomal and nuclear proteins. *Electrophoresis* **18**, 328–337.

Table 1. Sequence coverage of identified proteins of *Coxiella burnetii* strain RSA 493 with their characteristics

Excised spots	Protein	M_r (K)		pI		Sequence coverage
		Theoretical	Measured	Theoretical	Measured	
A	Chaperone protein DnaK (Heat shock 70 K protein)	70.8	69.0	5.1	4.9–5.1	29%
B	Chaperonin, 60 K (GroEL protein, Heat shock protein B)	58.3	55.2	5.1	4.8–5.2	52%
C	DnaJ-like protein djlA (Mucoidy activation protein mucZ)	31.3	36.4	10	5.1–5.4	48%
D	Elongation factor Ts (EF-Ts)	31.8	32	5.9	5.7	37%
E	Ribosomal protein L7/L12	13.2	9.4	4.7	4.5	68%
F	Chaperonin, 10 K (GroES protein, Heat shock protein A)	10.5	10.0	5.2	4.7–5.0	83%

Hochstrasser DF, Merril CR (1988): 'Catalysts' for polyacrylamide gel polymerization and detection of proteins by silver staining. *Appl. Theoret. Electrophor.* **1**, 35–40.

Hubálek M, Hernychová L, Havlasová J, Kasalová I, Neubauerová V, Stulík J, Macela A, Lundqvist M, Larsson P (2003): Towards proteome database of *Francisella tularensis*. *J. Chromatogr. B.* **787**, 149–177.

Macellaro A, Tujulin E, Hjalmarsson K, Norlander L (1998): Identification of a 71-kilodalton surface-associated Hsp70 homologue in *Coxiella burnetii*. *Infect. Immun.* **66**, 5882–5888.

Marrie TJ (1990): Acute Q fever. In Marrie TJ (Ed.): *Q Fever.* Vol. I. *The Disease*. Boca Raton, FL, CRC Press, pp. 125–160.

Mortz E, Krogh TN, Vorum H, Gorg A (2001): Improved silver staining protocols for high sensitivity protein identification using matrix-assisted laser desorption/ ionization-time of flight analysis. *Proteomics* **1**, 1359–1363.

Scopio A, Johnson P, Laquerre A, Nelson DR (1994): Subcellular localization and chaperone activities of *Borrelia burgdorferi* Hsp60 and Hsp70. *J. Bacteriol.* **176**, 6449–6456.

Seshadri R, Hendrix LR, Samuel JE (1999): Differential expression of translational elements by life cycle variants of *Coxiella burnetii. Infect. Immun.* **67**, 6026–6033.

Seshadri R, Paulsen IT, Eisen JA, Read TD, Nelson KE, Nelson WC, Ward NL, Tettelin H, Davidsen TM, Beanan MJ, Deboy RT, Daugherty SC, Brinkac LM, Madupu R, Dodson RJ, Khouri HM, Lee KH, Carty HA, Scanlan D, Heinzen RA (2003): Complete genome sequence of the Q-fever pathogen *Coxiella burnetii*. *Proc. Natl. Acad. Sci. USA* **100**, 5455–5460.

Škultéty Ľ, Toman R, Pätoprstý V (1998): A comparative study of lipopolysaccharides from two *Coxiella burnetii* strains considered to be associated with acute and chronic Q fever. *Carbohyd. Polymers* **35**, 189–194.

Vodkin MH, Williams JC (1988): A heat shock operon in *Coxiella burnetti* produces a major antigen homologous to a protein in both mycobacteria and *Escherichia coli*. *J. Bacteriol.* **170**, 1227–1234.

Zuber M, Hoover TA, Court DL (1995): Analysis of a *Coxiella burnetti* gene product that activates capsule synthesis in *Escherichia coli*: requirement for the heat shock chaperone DnaK and the two-component regulator RcsC. *J. Bacteriol.* **177**, 4238–4244.