



**HAL**  
open science

## Host cell-specific protein expression in

M. Postigo, A. Taoufik, L. Bell-Sakyi, C.P.J. Bekker, E. de Vries, W.I. Morrison, F. Jongejan

► **To cite this version:**

M. Postigo, A. Taoufik, L. Bell-Sakyi, C.P.J. Bekker, E. de Vries, et al.. Host cell-specific protein expression in. *Veterinary Microbiology*, 2008, 128 (1-2), pp.136. 10.1016/j.vetmic.2007.09.023 . hal-00532332

**HAL Id: hal-00532332**

**<https://hal.science/hal-00532332>**

Submitted on 4 Nov 2010

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

## Accepted Manuscript

Title: Host cell-specific protein expression *in vitro* in *Ehrlichia ruminantium*

Authors: M. Postigo, A. Taoufik, L. Bell-Sakyi, C.P.J. Bekker, E. de Vries, W.I. Morrison, F. Jongejan



PII: S0378-1135(07)00478-6  
DOI: doi:10.1016/j.vetmic.2007.09.023  
Reference: VETMIC 3836

To appear in: *VETMIC*

Received date: 13-7-2007  
Revised date: 21-9-2007  
Accepted date: 26-9-2007

Please cite this article as: Postigo, M., Taoufik, A., Bell-Sakyi, L., Bekker, C.P.J., de Vries, E., Morrison, W.I., Jongejan, F., Host cell-specific protein expression *in vitro* in *Ehrlichia ruminantium*, *Veterinary Microbiology* (2007), doi:10.1016/j.vetmic.2007.09.023

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

1 Host cell-specific protein expression *in vitro* in *Ehrlichia ruminantium*

2

3 M. Postigo <sup>a,b\*</sup>, A. Taoufik<sup>a</sup>, L. Bell-Sakyi<sup>b</sup>, C. P. J. Bekker<sup>c</sup>, E. de Vries<sup>a</sup>,

4

W. I. Morrison<sup>b</sup> and F. Jongejan<sup>a,d</sup>

5

6

7

8

9

<sup>a</sup> Utrecht Centre for Tick-borne Diseases, Faculty of Veterinary Medicine, Utrecht

10

University, P.O.Box 80165, 3508 TD Utrecht, The Netherlands

11

<sup>b</sup> Centre for Tropical Veterinary Medicine, Royal (Dick) School of Veterinary Studies,

12

University of Edinburgh, Easter Bush, Roslin, Midlothian EH25 9RG, Scotland, UK

13

<sup>c</sup> Department of Virology, Faculty of Veterinary Medicine, Utrecht University, P.O.Box

14

80165, 3508 TD Utrecht, The Netherlands

15

<sup>d</sup> Department of Veterinary Tropical Diseases, Faculty of Veterinary Science, University

16

of Pretoria, 0110 Onderstepoort, South Africa

17

18

**Veterinary Microbiology**

19

20

\*Corresponding author. Mailing address: Utrecht Centre for Tick-borne Diseases,

21

Faculty of Veterinary Medicine, Utrecht University, P.O.Box 80165, 3508 TD Utrecht,

22

The Netherlands Tel.: +31-30-2536932, fax: +31-30-2546723

23

E-mail address: [m.postigo@vet.uu.nl](mailto:m.postigo@vet.uu.nl)

24

25

26

27 **Abstract**

28

29 *Ehrlichia ruminantium*, a tick-transmitted pathogen, is the causative agent of heartwater  
30 in ruminants. In this study, a proteomic approach was used to identify host cell-specific  
31 *E. ruminantium* proteins encoded by the *map1* multigene family, expressed *in vitro* in  
32 bovine endothelial and tick cell cultures. Two-dimensional gel electrophoresis combined  
33 with mass spectrometry analysis was used to establish the identities of immunodominant  
34 proteins. Proteins extracted from *E. ruminantium*-infected endothelial cells were shown  
35 to be products of the *map1* gene, whereas tick cell-derived *E. ruminantium* proteins were  
36 products of a different gene, *map1-1*. The expressed proteins were found to be  
37 glycosylated. Differential expression of MAP1 family proteins *in vitro* in mammalian  
38 and tick cell cultures indicates that the *map1* multigene family might be involved in the  
39 adaptation of *E. ruminantium* to the mammalian host and vector tick.

40

41

42 **Keywords:** *Ehrlichia ruminantium*; *map1* multigene family; proteomics; protein  
43 expression; tick cell lines.

44

45

## 46 **Introduction**

47

48 The obligately intracellular rickettsial pathogen *Ehrlichia ruminantium*,  
49 transmitted by *Amblyomma* ticks, causes the economically important disease heartwater  
50 of domestic ruminants in sub-Saharan Africa and the Caribbean (Uilenberg, 1983).  
51 Within the mammalian host and tick vector, *E. ruminantium* organisms sequentially  
52 occupy different cellular environments, including mammalian neutrophils and reticulo-  
53 endothelial cells, and tick midgut and salivary glands (Prozesky and Du Plessis, 1987).  
54 To successfully invade and multiply within all these cell types, the pathogen requires life  
55 cycle stage-specific adaptations which are likely to be reflected in differential gene  
56 transcription and protein expression.

57 In *E. ruminantium*, the immunodominant major surface protein expressed in the  
58 mammalian host, MAP1, is encoded by a member of a multigene family comprising 16  
59 paralogs (van Heerden et al., 2004) which are differentially transcribed *in vitro* in  
60 endothelial and tick cell cultures (van Heerden et al., 2004; Bekker et al., 2005) and *in*  
61 *vivo* in tick midguts and salivary glands (Postigo et al., 2007). It is important to  
62 determine which proteins from the *E. ruminantium map1* cluster, other than MAP1 which  
63 is located on the surface of mammalian-stage elementary bodies (Jongejan et al., 1991),  
64 are actually expressed and if they are differentially expressed in tick and mammalian cell  
65 environments. In the present study a proteomic approach was used to identify *E.*  
66 *ruminantium* proteins, encoded by the *map1* cluster, which might be differentially  
67 expressed *in vitro* in bovine endothelial and tick cell cultures, and additional analysis was  
68 carried out to determine the glycosylation status of the expressed proteins.

69 **Materials and methods**

70

71 **Growth and harvest of *E. ruminantium* from *in vitro* cell cultures.** The CTVM  
72 subpopulation of the Gardel isolate of *E. ruminantium* (Uilenberg et al., 1985; Bekker et  
73 al., 2005) was cultured at 37 °C in bovine umbilical cord endothelial (BUE) cells; and at  
74 31 °C in non-vector *Ixodes scapularis* (IDE8), and vector *Amblyomma variegatum*  
75 (AVL/CTVM13), tick cell lines as described previously (Jongejan, 1991; Bell-Sakyi,  
76 2004). When *E. ruminantium*-infected BUE cultures showed about 90 % cytolysis, or at  
77 least 10 % of the tick cells were infected, the cells and supernate were harvested and  
78 centrifuged at 15000 g for 20 min at 4 °C. The resultant pellets containing both infected  
79 cells and free *E. ruminantium* organisms were frozen at -20 °C until used for protein  
80 extraction. Uninfected endothelial and tick cell cultures were harvested in the same way.

81

82 **Protein extraction.** Soluble and membrane-bound proteins were extracted by  
83 resuspending the thawed cell pellets in 10 mM Tris, 10 mM NaCl, 0.5 % Nonidet P40, 2  
84 % CHAPS and 1X of Complete® protease inhibitor (Roche). The suspensions were  
85 mixed on a shaker platform at 100 rpm for 45 min at 4 °C and then centrifuged at 15000 g  
86 for 20 min at 4 °C. The supernatants were recovered and passed through a protein  
87 desalting spin column (Pierce) according to the manufacturer's protocol. Desalted  
88 proteins were diluted 1:10 in PBS and the protein concentration measured in a Ceres UV  
89 900C ELISA reader (Biotek Instruments) according to the Bradford protein assay method  
90 (Bradford, 1976).

91

92 **SDS-PAGE and Western blot analysis.** Protein samples (5-10 µg), were solubilised in  
93 1X Laemmli sample buffer and heated for 5 min at 90°C. Once cooled, 15-20 µl samples  
94 were loaded onto 12.5% acrylamide gels, with a bisacrylamide/acrylamide ratio of 1:37.5  
95 (Bio-Rad Laboratories). Electrophoresis was performed in a Hoefer Scientific cell  
96 apparatus at 20mAmp per gel for 1 h at room temperature in a 50 mM Tris-glycine buffer.  
97 Separated proteins were immediately transferred to nitrocellulose (Whatman) membranes  
98 at 38mAmp per gel per hour in a LKB MultiphorII blotter unit (Pharmacia). Blots were  
99 stained with 8% Direct Blue 71 (0.1% in water) in 40% ethanol and 10% acetic acid  
100 solution, then scanned, destained in destaining solution (96% ethanol, 1M NaHCO<sub>3</sub>) and  
101 kept in deionised water. Non-specific binding was reduced by incubating the membranes  
102 for 1 hour at 37°C in blocking buffer consisting of 20 mM Tris-HCl (pH 9), 0.9% NaCl,  
103 0.05% Tween 20 (TBS), with 5% skimmed milk (Elk Campina, The Netherlands) added  
104 just before use (TBSM). The membranes were incubated with specific antibodies  
105 (diluted in TBSM) overnight at 4°C. Blots containing *E. ruminantium*-infected and  
106 uninfected endothelial and tick cell protein extracts were incubated with pre-infection and  
107 4 weeks post-infection sera, diluted 1:250, from a sheep immunised with supernatant  
108 from *E. ruminantium* (Gardel)-infected endothelial cell cultures, to identify *E.*  
109 *ruminantium* immunodominant proteins. Monoclonal antibodies 4F10B4 and 1E5H8,  
110 reactive with an *E. ruminantium* (Welgevonden) 32 kDa protein identified as MAP1  
111 (Jongejan and Thielemans, 1989; Jongejan et al., 1991), were used at dilutions of 1:2000  
112 and 1:200 respectively to identify MAP1 cluster proteins. Incubation with the appropriate  
113 horseradish peroxidase-conjugated anti-species immunoglobulin secondary antibody  
114 (DAKO) (diluted 1:2000 or according to the manufacturer's instructions) was carried out

115 for 1 h at room temperature. The membranes were washed three times with TBS for 5  
116 min after each incubation step. Finally the membranes were incubated with enhanced  
117 chemiluminescence detection reagents (Amersham Biosciences) and exposed to X-ray  
118 films (Hyperfilm, Amersham Biosciences).

119

120 **Two dimensional gel electrophoresis (2DE).** Protein samples (10-30  $\mu\text{g}$ ) were mixed  
121 with rehydration solution (7 M urea, 2 M thiourea, 4 % CHAPS, 0.3 % DTT, 0.5 % 3-10  
122 NL IPG buffer, 1 X of Complete® protease inhibitor and a trace of bromophenol blue)  
123 and resolved at 20 °C in the first dimension by isoelectric focusing (IEF) in an IPGphor  
124 (Amersham Pharmacia Biotech) using 7 cm long, pH 3 to 10, precast immobilised  
125 nonlinear pH gradient strips (Amersham Pharmacia Biotech). The IEF parameters were as  
126 follows: rehydration of the strips was carried out for 15 hours at 30 volts, followed by  
127 500 volts  $\times$  30 min, 1000 volts  $\times$  30 min and 5000 volts  $\times$  100 min. At the end of the  
128 IEF, the strips were equilibrated sequentially for 15 min in 5 ml each of equilibration  
129 buffers I (50 mM Tris-HCl [pH 8.8], 6 M urea, 2 % SDS, 30 % glycerol and 10 mg ml<sup>-1</sup>  
130 DTT) and II (50 mM Tris-HCl [pH 8.8], 6 M urea, 2 % SDS, 30 % glycerol and 25 mg  
131 ml<sup>-1</sup> of iodoacetamide). Subsequently, second-dimension SDS-polyacrylamide gel  
132 electrophoresis was performed as described above. The 2DE resolved gels were stained  
133 with either silver (Merck) or Coomassie blue (Bio-Rad) or used to perform Western blot  
134 analysis.

135

136 **Cloning and expression of recombinant proteins MAP1 and MAP1-1.** Expression of  
137 the recombinant MAP1-GST fusion protein (fragment F1R4, approximately 50 kDa) was



138 performed as described in Van Vliet et al. (1995). The full length of the *map1-1* gene  
139 was amplified by using a sense primer containing a *Bam*HI restriction site (MAP1-1F2,  
140 5'-GCGAGCCGGATCCGAAACCTGTAAGTTCAAAT), and an anti-sense primer  
141 containing a *Sal*I restriction site (MAP1-1R3, 5'-  
142 GCGAGCGTCGACGAAAGTAAACCTTACTCCA). The positions of these primers are  
143 respectively 15099-15116bp and 15851-15869bp on the Genbank accession nr.  
144 AF319940. The PCR product was cut with *Bam*HI and *Sal*I and ligated to pQE9 plasmid  
145 cut with the same enzymes. Ligation was done with T4 ligase and 5x ligase buffer  
146 (Promega). The recombinant pQE9 plasmid was transferred into *E. coli* strain M15.  
147 Positive clones were tested for incorporation of the correct insert by carrying out a PCR  
148 with specific primers. Inserts were checked by digestion with *Bam*HI & *Sal*I and  
149 sequencing. Expression of the recombinant MAP1-1 protein (approximately 30 kDa) was  
150 done using the QIA expressionist kit (QIAGEN) according to the manufacturer's  
151 instructions. Expression of MAP1-GST and MAP1-1 His-Tag recombinant proteins was  
152 confirmed by SDS-PAGE electrophoresis and immunoblotting.

153

154 **MALDItof MS analysis.** Protein spots excised from the gels were digested in gel with  
155 trypsin (Promega) in 50 mM ammonium bicarbonate (Sigma). Before Matrix-Assisted  
156 Laser Desorption/Ionization Time-Of-Flight (MALDI-TOF) analysis (carried out at the  
157 Moredun Research Institute, Scotland), peptides were concentrated using  $\mu$ C18-ZipTips  
158 (Millipore) and eluted directly on the MALDI-target in 1  $\mu$ l of a saturated solution of  $\alpha$ -  
159 cyanohydroxycinnamic acid in 50 % acetonitrile. Peptides were analysed using a  
160 Voyager DE-PRO MALDI-TOF mass spectrometer (Applied Biosystems) operated in

161 reflectron mode at 20 kV accelerating voltage. The resulting peptide mass fingerprints  
162 were subjected to an NCBI-nr database search using the Mascot search programmes,  
163 wherein MOWSE scores greater than 76 are considered significant  $p < 0.05$   
164 ([www.matrixscience.com](http://www.matrixscience.com)). Alignment of MAP1 and MAP1-1 was performed using the  
165 Clustal-X programme, and potential N- and O-glycosylation sites (showing G-scores  
166 higher than 0.4) were predicted using the NetNGlyc 1.0 and NetOGlyc 3.1 servers  
167 respectively (Julenius et al., 2005).

168

169 **Glycosylation analysis.** Glycoprotein staining was performed on proteins resolved by  
170 2DE using the Pro-Q Emerald 300 staining method according to the manufacturer's  
171 protocol (Molecular Probes). Images of the stained gels were captured using a UV  
172 transilluminator (UVP Bio imaging System). The gels were restained with silver nitrate  
173 according to the manufacturer's protocol to detect total protein present in the gels. A  
174 Candy Cane glycoprotein molecular weight standard (Molecular Probes) was used as a  
175 positive control for glycoprotein detection and protein size determinations. To further  
176 demonstrate glycosylation, protein extracts were treated using the GlycoProfile™ IV  
177 chemical deglycosylation kit (Sigma-Aldrich). Briefly, 10-20 µg of lyophilised total  
178 protein, extracted from *E. ruminantium*-infected endothelial and tick cell cultures, were  
179 mixed gently with trifluoromethanesulfonic (TFMS) acid and incubated at -20° C for 30  
180 min. After incubation, bromophenol blue and pyridine solution were added to the mix  
181 and neutralised deglycosylated glycoproteins were purified using the protein desalting  
182 spin column (Pierce) according to the manufacturer's protocol. Proteins were  
183 concentrated by TCA precipitation and resuspended in 10 mM Tris, 10 mM NaCl, 0.5 %

184 Nonidet P40, 2 % CHAPS and 1X of Complete® protease inhibitor (Roche).  
185 Subsequently, 2DE was carried out and deglycosylated proteins were stained with the  
186 Pro-Q Emerald 300 staining method and restained with silver nitrate according to the  
187 manufacturer's protocol.  
188

Accepted Manuscript

## 189 **Results**

190

191 **Immunodominant *E. ruminantium* proteins.** Proteins from uninfected or *E.*  
192 *ruminantium*-infected BUE, IDE8 and AVL/CTVM13 cells were separated by SDS-  
193 PAGE and transferred to nitrocellulose membranes. When probed with sera from sheep  
194 inoculated with endothelial cell-derived *E. ruminantium*, the post-inoculation serum  
195 reacted with three proteins, of approximately 29 kDa, 30 kDa and 32 kDa, extracted from  
196 *E. ruminantium*-infected endothelial cell cultures, but with only one protein in the same  
197 molecular weight region (~30 kDa) extracted from *E. ruminantium*-infected tick cells.  
198 The pre-inoculation serum did not reveal any protein bands (Fig. 1).

199

### 200 **Identification of 2DE-separated *E. ruminantium* proteins by Western blot analysis.**

201 In order to identify which proteins of the MAP1 multigene family were expressed *in*  
202 *vitro*, proteins derived from uninfected and *E. ruminantium*-infected endothelial and tick  
203 cell cultures were resolved by 2DE and silver stained. Differences in the protein profiles  
204 of samples from uninfected and *E. ruminantium*-infected sources were apparent by  
205 comparing the stained gels, especially in the area of 24 to 37 kDa size range, where  
206 MAP1 family proteins are expected to migrate since their predicted molecular weights  
207 range from 24 to 35 kDa (van Heerden et al, 2004) (Fig. 2a-f).

208 The main differences included a group of proteins around 30 kDa, with *pI* values between  
209 4.5 and 6.0, present only in infected samples of the three culture systems (BUE, IDE8 and  
210 AVL/CTVM13). Comparison of the pattern of proteins from *E. ruminantium*-infected  
211 endothelial cell cultures with those derived from tick cells revealed a higher density of

212 proteins, migrating between *pI* 4.5 and 5.5, with molecular weights of around 30kDa, in  
213 extracts of infected endothelial cells (Fig. 2a, 2b) compared with samples derived from *E.*  
214 *ruminantium*-infected tick cells, where a single row of proteins of around 30 kDa,  
215 migrating widely between *pI* 4.5 and 6.0, was observed (Fig. 2c-f).  
216 Western blotting analysis showed that the row of approximately 30 kDa proteins in each  
217 cell type (Fig. 3a) was recognised by post-infection sheep serum and not by pre-infection  
218 serum, indicating that they represented immunodominant *E. ruminantium* proteins (Fig.  
219 3b, 3c). The same sera did not recognise any spots of this size in uninfected endothelial  
220 and tick cell protein extracts (data not shown). In addition, the MAP1-reactive  
221 monoclonal antibody 4F10B4 reacted with the rows of proteins in both endothelial and  
222 tick cells infected with *E. ruminantium* (Fig. 3d) whereas only spots present in infected  
223 endothelial cells were recognised by the MAP1-reactive monoclonal antibody 1E5H8  
224 (Fig. 3e).

225

226 **Expression of recombinant proteins and monoclonal antibody specificity.** Since the  
227 MAP1 protein has been found to be predominant in *E. ruminantium*-infected endothelial  
228 cell cultures (Jongejan and Thielemans, 1989; Jongejan et al., 1991) and transcripts of the  
229 *map1-1* gene predominated in infected tick cells (Bekker et al., 2005), expression of  
230 recombinant MAP1 and MAP1-1 proteins was carried out to define the specificity of  
231 monoclonal antibodies 4F10B4 and 1E5H8. Expression of recombinant proteins was  
232 confirmed by SDS-PAGE (Fig. 4a). Immunoblotting of recombinant proteins with the  
233 monoclonal antibodies demonstrated specificity of 1E5H8 antibodies for the MAP1  
234 protein, whereas 4F10B4 reacted with both MAP1 and MAP1-1 recombinant proteins

235 (Fig. 4b). These results indicated that the immunodominant *E. ruminantium* proteins  
236 expressed in endothelial and tick cells *in vitro* were not identical.

237

238 **Identification by MALDI tof MS analysis of host cell-specific protein expression by**

239 **the *E. ruminantium map1* cluster.** In order to definitively identify the most abundant

240 proteins expressed in the three culture systems, the three most prominent spots within the

241 30 kDa region of gels prepared from infected BUE cells (spots B1, B2, B3) and tick cells

242 (spots Is1, Is2 and Is3 in IDE8 and Av1, Av2, and Av3 in AVL/CTVM13) were excised

243 and submitted to MALDItof MS analysis (Fig. 3a). All fingerprint analyses showed

244 MOWSE scores greater than 76 (between 86 and 113 for B1, B2 and B3; between 83 and

245 127 for Is1, Is2, and Is3; and between 85 and 116 for Av1, Av2, and Av3) and therefore

246 were considered a significant match for the proteins. The proteins expressed in BUE

247 cells were encoded by the *map1* gene, while the proteins extracted from *E. ruminantium*-

248 infected tick cell cultures, both IDE8 and AVL/CTVM13, were all found to be products

249 of the *map1-1* gene. Similar results were obtained using *E. ruminantium*-infected

250 cultures from different passage level and protein batch extractions (results not shown).

251 The identified peptides are indicated on an alignment of MAP1 and MAP1-1 sequences

252 for all nine recognised spots (Fig. 5). A slightly different set of peptides was identified

253 for each spot which could have resulted from differential post-translational modification.

254 Since predicted *N*- and *O*-linked glycosylation sites (shown within boxes in Fig. 5) were

255 found to be present in these regions of the proteins, they were further analysed for

256 glycosylation.

257

258 **Identification of glycoproteins.** Glycoprotein staining revealed that all three forms of  
259 MAP1 expressed in *E. ruminantium*-infected endothelial cells, and all three forms of  
260 MAP1-1 expressed in both *E. ruminantium*-infected tick cell lines were glycoproteins  
261 (Fig. 6a; although all three spots from AVL/CTVM13 are not well-reproduced in the  
262 photograph, they were clearly visible in the original gel). Glycosylation was confirmed  
263 by staining of two positive control proteins of 42 kDa and 82 kDa included in the Candy  
264 Cane molecular marker (Fig. 6a, 6b). Chemical deglycosylation of proteins extracted  
265 from *E. ruminantium*-infected cultures further demonstrated the glycosylated nature of  
266 these proteins. *E. ruminantium*-infected BUE, IDE8 and AVL/CTVM13 protein samples  
267 treated with TFMS acid no longer stained for glycoproteins while the control 42 kDa and  
268 82 kDa glycoproteins of the Candy Cane molecular marker remained positive (results not  
269 shown). Furthermore, when these gels were restained with silver nitrate, the rows of  
270 spots that were previously located in the area between *pI* 4.5 and 6.0 (Fig. 7a), were  
271 found to have shifted to the right (between *pI* 6.0 and 9.0) (Fig. 7b). To confirm that the  
272 spots in the post-deglycosylation gels were in fact deglycosylated *E. ruminantium* MAP1  
273 family proteins, blots were prepared and reacted with monoclonal antibody 4F10B4. The  
274 most prominent spot in treated *E. ruminantium*-infected BUE and all the spots in treated  
275 *E. ruminantium*-infected IDE8 and AVL/CTVM13 were recognised by 4F10B4,  
276 indicating that the spots were deglycosylated forms of MAP1 family proteins (Fig. 7c).  
277 The reactivity of the monoclonal antibody with the deglycosylated forms of the proteins  
278 was in general weaker than that observed previously with the corresponding  
279 glycoproteins.  
280

281

282 **Discussion**

283           The present study identified the antigenic *E. ruminantium* MAP1 family proteins  
284 predominantly expressed in infected bovine endothelial and tick cell cultures. Reaction  
285 with immune serum in one-dimensional Western blots revealed several proteins between  
286 29 and 32 kDa in *E. ruminantium*-infected endothelial cells and a single band, of  
287 approximately 30 kDa, in *E. ruminantium*-infected tick cells. In 2DE blots, the ~30 kDa  
288 proteins from infected endothelial and tick cells were strongly recognised by immune  
289 serum and MAP1-reactive monoclonal antibody 4F10B4, while monoclonal antibody  
290 1E5H8 only reacted with the proteins expressed in endothelial cells. Expression of  
291 recombinant proteins and Western blotting with the same monoclonal antibodies  
292 indicated that the protein spots identified in endothelial and tick cell systems represented  
293 different proteins, since recombinant MAP1 reacted with both 4F10B4 and 1E5H8, while  
294 recombinant MAP1-1 only reacted with 4F10B4. MALDItoF MS analysis revealed that  
295 the row of proteins around 30 kDa in extracts from endothelial cells represented  
296 expressed products of *map1*, whereas those detected in extracts from both tick cell lines  
297 were products of the *map1-1* gene.

298           *E. chaffeensis* and *E. canis*, the causative agents of human and canine monocytic  
299 ehrlichiosis respectively, and the pathogens most closely related to *E. ruminantium*,  
300 possess gene clusters encoding outer membrane proteins, commonly referred to as the  
301 *p28*- and *p30-Omps* respectively, which are orthologous to the MAP1 multigene family  
302 (van Heerden et al., 2004). Singu et al. (2006) demonstrated that the P28/P30 proteins  
303 expressed *in vitro* in infected canine macrophage cell cultures included the orthologs of  
304 the *map1* (*p28-Omp19* and *p30-1*) and *map1+1* (*p28-Omp20* and *p30-20*) genes whereas



305 in vector (*Amblyomma americanum*) and non-vector (*I. scapularis*) tick cell lines, the  
306 ortholog of *map1-1* (*p28-Omp14* and *p30-10*) was detected. Since many *map1* gene  
307 paralogs appear to be transcribed in different tick cell lines (Bekker et al., 2005), it is  
308 interesting that only one paralog appears to be dominantly translated in tick cell culture  
309 systems. In contrast, in endothelial cell cultures, in which all 16 *map1* paralogs are  
310 transcribed (van Heerden et al., 2004; Bekker et al., 2005), immune serum recognised  
311 several proteins in the 29–32 kDa range in Western blots and many *E. ruminantium*-  
312 specific protein spots were present in the 24–37 kDa range in 2DE gels, although we  
313 were only able to subject the three major ~30 kDa spots to MALDItoF MS analysis. It is  
314 possible that some of the additional spots observed in the silver-stained gels and blots  
315 represented different MAP1 family proteins, since 14 out of 16 proteins of the MAP1  
316 family are predicted to have signal peptides and to be situated at the surface of the  
317 bacterial cell (Collins et al., 2005). Recently Ge and Rikihisa (2007) reported that 19 out  
318 of 22 *p28* family proteins were expressed at the protein level in *E. chaffeensis* cultured in  
319 human monocytic leukemia THP-1 cells; the *p28-Omp14* (ortholog of *map1-1*) was not  
320 detected in these cultures. Altogether these results indicate that *Ehrlichia* pathogens  
321 consistently exhibit a differential pattern of expression of multigene family surface  
322 proteins between host and vector cells *in vitro*.

323 An interesting observation was the identification in 2DE gels of MAP1 and  
324 MAP1-1 proteins in rows of spots with the same molecular size but different *pI*, which  
325 could be explained by the presence of post-translational modifications. Different degrees  
326 of glycosylation at a single site in a single protein can result in “trains” of protein spots  
327 that separate on the basis of different isoelectric point and/or molecular mass in 2DE gels

328 (Sickmann et al., 2002). Moreover, location of the detected peptides in the alignment of  
329 MAP1 and MAP1-1 proteins showed predicted glycosylation sites in the different  
330 sequences which might be responsible for the differences in migration detected in 2DE  
331 gels. By applying chemical deglycosylation in combination with immunoblotting for  
332 definitive detection of *pI*-shifted spots, we confirmed the glycosylated nature of MAP1  
333 family proteins.

334 Glycosylation of immunodominant proteins has been reported in *Ehrlichia* spp. (McBride  
335 et al., 2000, 2003, 2007; Doyle et al., 2006), and evidence of glycosylation of P28-OMP1  
336 proteins in *E. chaffeensis* has been demonstrated (Singu et al., 2005). Doyle et al. (2006)  
337 reported substantially lower immunoreactivity of immune sera against nonglycosylated  
338 synthetic peptides from *E. canis* gp36 and *E. chaffeensis* gp47 than against the  
339 corresponding recombinant glycosylated proteins, suggesting that glycans are important  
340 epitope determinants. The major surface proteins MSP1a and MSP1b of the closely-  
341 related pathogen *Anaplasma marginale* have been shown to be glycosylated and it was  
342 suggested that the glycosylation of MSP1a plays a role in the adhesion of the organism to  
343 tick cells (Garcia-Garcia et al., 2004).

344 Expression of the MAP1 protein was not detected in the present study in  
345 organisms grown in the tick cell lines, although transcription of *map1* has been detected  
346 in salivary glands of infected ticks during feeding and in organisms growing in tick cell  
347 cultures (Bekker et al., 2002, 2005; Postigo et al., 2007). MAP1 expression by tick cells  
348 *in vitro* cannot be ruled out, as the detection techniques used were of relatively low  
349 sensitivity. The IDE8 cells resemble haemocytes (Munderloh et al., 1996) while  
350 AVL/CTVM13 comprises a variety of cells including haemocytes and could contain

351 differentiated midgut cells since it was established from moulting larvae (Bell-Sakyi,  
352 2004), but neither line contains recognisable differentiated salivary gland cells, which  
353 might be necessary for development *in vitro* of *E. ruminantium* salivary gland stages  
354 expressing MAP1. On the other hand, organisms grown in both vector and non-vector  
355 tick cell lines expressed MAP1-1 protein, in line with the detection of abundant  
356 transcripts of the *map1-1* gene in *E. ruminantium*-infected midguts of *A. variegatum* ticks  
357 (Postigo et al. 2007) and in vector and non-vector tick cell lines (Bekker et al., 2002,  
358 2005). Since both midgut and tick cell line stages have low infectivity for sheep  
359 (Waghela et al., 1991; Bell-Sakyi et al., 2002), these results suggest that the MAP1-1  
360 protein may be associated with colonisation and replication of *E. ruminantium* in the tick  
361 midgut, rather than development of mammal-infective stages in the salivary glands.  
362 Ganta et al. (2007) reported that *E. chaffeensis* grown in tick cells induced a cellular and  
363 humoral immune response in experimentally-infected mice that was distinct from the  
364 response following infection with the pathogen grown in mammalian cells. This response  
365 was apparently related to the shift in gene expression from the tick cell-specific *omp14* to  
366 the macrophage-specific *omp19*. The different humoral immune responses observed in  
367 sheep immunised with *E. ruminantium* derived from tick and mammalian cell cultures  
368 (Bell-Sakyi et al., 2002) might similarly be related to the differences in MAP1 family  
369 protein expression between the host cell types that we have described in the present  
370 study.

371 The glycoproteins MAP1 and MAP1-1 and their orthologs in related pathogens have  
372 consistently been detected differentially in mammalian and tick cell culture systems  
373 respectively, regardless of transcription of other *map1* and orthologous genes in these

374 cultures. These results indicate the need for validation at the protein level of *E.*  
375 *ruminantium* gene expression data, using proteomics technologies; and most  
376 significantly, support the hypothesis that MAP1 family proteins and their orthologs are  
377 important for *Ehrlichia* species in host adaptation and intracellular survival.

378

379

### 380 **Acknowledgements**

381

382 The research described in this manuscript was supported by two consecutive grants from  
383 the European Union (ICA4-CT-2000-30026 and INCO-CT-2005-003713). The authors  
384 would like to thank Ulrike Munderloh for kindly providing the IDE8 cell line, Stuart  
385 Smith for access to unpublished results and Ana Yatsuda for excellent technical  
386 assistance.

387

388

### 389 **References**

390

391 Bekker, C. P. J., Bell-Sakyi, L., Paxton, E. A., Martinez, D., Bensaid, A. and Jongejan,  
392 F., 2002. Transcriptional analysis of the major antigenic protein 1 multigene  
393 family of *Cowdria ruminantium*. *Gene* 285, 193-201.

394

395 Bekker, C. P. J., Postigo, M., Taoufik, A., Bell-Sakyi, L., Ferraz, C., Martinez, D. and  
396 Jongejan, F., 2005. Transcription analysis of the major antigenic protein 1  
397 multigene family of three *in vitro*-cultured *Ehrlichia ruminantium* isolates. *J.*  
398 *Bacteriol.* 187, 4782-4791.

399

400 Bell-Sakyi, L., 2004. *Ehrlichia ruminantium* grows in cell lines from four ixodid tick  
401 genera. *J. Comp. Path.* 130, 285-293.

402

- 403 Bell-Sakyi, L., Paxton, E. A., Wright, P. and Sumption, K. J., 2002. Immunogenicity of  
404 *Ehrlichia ruminantium* grown in tick cell lines. *Exp. Appl. Acarol.* 28, 177-185.  
405
- 406 Bradford, M.M. (1976). A rapid and sensitive method for the quantitation of microgram  
407 quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.*  
408 72, 248-254.  
409
- 410 Collins, N. E., Liebenberg, J., De Villiers, E. P., Brayton, K. A., Louw, E., Pretorius, A.,  
411 Faber, F. E., van Heerden, H., Josemans, A. I., van Kleef, M., Steyn, H. C., van  
412 Strijp, M. F., Zwegarth, E., Jongejan, F., Maillard, J. C., Berthier, D., Botha, M.,  
413 Joubert, F., Corton, C. H., Thomson, N. R., Allsopp, M. T. and Allsopp, B. A.,  
414 2005. The genome of the heartwater agent *Ehrlichia ruminantium* contains  
415 multiple tandem repeats of actively variable copy number. *PNAS* 102, 838-843.  
416
- 417 Doyle, K. C., Nethery, K. A., Popov, V. L. and McBride, J. W., 2006. Differentially  
418 expressed and secreted major immunoreactive protein orthologs of *Ehrlichia*  
419 *canis* and *E. chaffeensis* elicit early antibody responses to epitopes on  
420 glycosylated tandem repeats. *Infect. Immun.* 74, 711–720.  
421
- 422 Ganta, R. R., Cheng, C., Miller, E. C., McGuire, B. L., Peddireddi, L., Sirigireddy, K. R.  
423 and Chapes, S. K., 2007. Differential clearance and immune responses to tick cell-  
424 derived versus macrophage culture-derived *Ehrlichia chaffeensis* in mice. *Infect.*  
425 *Immun.* 75, 135-145.  
426
- 427 Garcia-Garcia, J. C., de la Fuente, J., Bell-Eunice, G., Blouin, E. F. and Kocan, K. M.,  
428 2004. Glycosylation of *Anaplasma marginale* major surface protein 1a and its  
429 putative role in adhesion to tick cells. *Infect. Immun.* 72, 3022-3030.  
430
- 431 Ge Y. and Rikihisa Y., 2007. Surface-exposed Proteins of *Ehrlichia chaffeensis*. *Infect*  
432 *Immun.* doi:10.1128/IAI.00188-07.  
433

- 434 Jongejan, F. and Thielemans, M. J. C., 1989. Identification of an immunodominant  
435 antigenically conserved 32-kilodalton protein from *Cowdria ruminantium*. Infect.  
436 Immun. 57, 3243-3246.  
437
- 438 Jongejan, F., 1991. Protective immunity to heartwater (*Cowdria ruminantium* infection)  
439 is acquired after vaccination with *in vitro*-attenuated rickettsiae. Infect. Immun.  
440 59,729-731.  
441
- 442 Jongejan, F., Thielemans, M. J. C., de Groot, M., van Kooten, P. J. S. and van der Zeijst,  
443 B. A. M., 1991. Competitive enzyme-linked immunosorbent assay for heartwater  
444 using monoclonal antibodies to a *Cowdria ruminantium*-specific 32- kilodalton  
445 protein. Vet. Microbiol. 28, 199-211.  
446
- 447 Julenius, K., Mølgaard, A., Gupta, R. and Brunak, S., 2005. Prediction, conservation  
448 analysis and structural characterization of mammalian mucin-type O-  
449 glycosylation sites. Glycobiology 15, 153-164.  
450
- 451 McBride, J. W., Doyle, C. K., Zhang, X., Cardenas, A. M., Popov V. L., Nethery, K. A.  
452 and Woods E. M., 2007. Identification of a glycosylated *Ehrlichia canis* 19-  
453 kilodalton major immunoreactive protein with a species-specific serine-rich  
454 glycopeptide epitope. Infect. Immun. 75, 74–82.  
455
- 456 McBride, J. W., Comer, J. E. and Walker, D. H., 2003. Novel immunoreactive  
457 glycoprotein orthologs of *Ehrlichia* spp. Ann. N Y Acad. Sci. 990, 678-684.  
458
- 459 McBride, J. W., Yu, X. J. and Walker, D. H., 2000. Glycosylation of homologous  
460 immunodominant proteins of *Ehrlichia chaffeensis* and *Ehrlichia canis*. Infect.  
461 Immun. 68, 13-18.  
462  
463

- 464 Munderloh, U. G., Blouin E. F., Kocan, K. M., Ge, N. L., Edwards, W. L. and Kurtti T.  
465 J., 1996. Establishment of the tick (Acari:Ixodidae)-borne cattle pathogen  
466 *Anaplasma marginale* (Rickettsiales:Anaplasmataceae) in tick cell culture. J.  
467 Med. Entomol. 33, 656-664.  
468
- 469 Postigo, M., Taoufik, A., Bell-Sakyi, L., de Vries, E., Morrison, W. I. and Jongejan. F.,  
470 2007. Differential transcription of the major antigenic protein 1 multigene family  
471 of *Ehrlichia ruminantium* in *Amblyomma variegatum* ticks. Vet. Microbiol. 122,  
472 298-305.  
473
- 474 Prozesky, L. and J. L. Du Plessis 1987. Heartwater. The development and life cycle of  
475 *Cowdria ruminantium* in the vertebrate host, ticks and cultured endothelial cells.  
476 Onderstepoort J. Vet. Res. 54: 193-6.  
477
- 478 Sickmann, A., Mreyen, M. and Meyer, H. E., 2002. Identification of modified proteins by  
479 mass spectrometry. IUBMB Life 54, 51-57.  
480
- 481 Singu, V., Liu, H., Cheng, C. and Ganta, R. R., 2005. *Ehrlichia chaffeensis* expresses  
482 macrophage- and tick cell-specific 28-kilodalton outer membrane proteins. Infect.  
483 Immun. 73, 79-87.  
484
- 485 Singu, V., Peddireddi, L., Sirigireddy, K. R., Cheng, C., Munderloh, U. and  
486 Ganta, R. R., 2006. Unique macrophage and tick cell-specific protein expression  
487 from the p28/p30-outer membrane protein multigene locus in *Ehrlichia*  
488 *chaffeensis* and *Ehrlichia canis*. Cell. Microbiol. 8, 1475–1487.  
489
- 490 Uilenberg, G. (1983). Heartwater (*Cowdria ruminantium* infection): current status. Adv.  
491 Vet. Sci. Comp. Med. 27, 427-480.  
492  
493

- 494 Uilenberg, G., Camus, E. and Barré, N., 1985. Quelques observations sur une souche de  
495 *Cowdria ruminantium* isolée en Guadeloupe (Antilles francaises). Rev. Elev.  
496 Med. vet. Pays trop. 38, 34-42.  
497
- 498 van Heerden, H., Collins, N. E., Brayton, K. A., Rademeyer, C. and Allsopp, B. A., 2004.  
499 Characterization of a major outer membrane protein multigene family in *Ehrlichia*  
500 *ruminantium*. Gene 330, 159-168.  
501
- 502 van Vliet, A. H. M., van der Zeijst, B. A. M., Camus, E., Mahan, S. M., Martinez, D. and  
503 Jongejan., F., 1995. Use of a specific immunogenic region on the *Cowdria*  
504 *ruminantium* MAP1 protein in a serological assay. J. Clin. Microbiol. 33, 2405-  
505 2410.  
506
- 507 Waghela, S. D., Rurangirwa, F. R., Mahan, S. M., Yunker, C. E., Crawford, T. B.,  
508 Barbet, A. F., Burridge, M. J. and McGuire T. C., 1991. A cloned DNA probe  
509 identifies *Cowdria ruminantium* in *Amblyomma variegatum* ticks. J. Clin.  
510 Microbiol. 29, 2571-2577.  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523



524

525 **Figure legends**

526

527 **Figure 1:** Western blots of 12.5% SDS-PAGE gels containing proteins extracted from *E.*  
528 *ruminantium*-infected bovine endothelial (BUE) and tick (IDE8 and AVL/CTVM13)  
529 cells. Blots were probed with pre- (1) and post- (2) infection sera of sheep inoculated  
530 with endothelial cell-derived *E. ruminantium*. Molecular size marker (M) in kilodaltons.

531

532 **Figure 2:** Two-dimensional SDS-PAGE gels of *E. ruminantium* grown *in vitro*. Total  
533 protein extracts from uninfected and *E. ruminantium*-infected BUE (a, b), IDE8 (c, d) and  
534 AVL/CTVM13 (e, f) cell cultures respectively, were subjected to 2DE analysis and  
535 silver-stained. The region of interest (approx. 24-37 kDa) is surrounded by a box in all  
536 panels. M: Molecular masses in kDa.

537

538 **Figure 3:** Total protein extracts from *E. ruminantium* grown in endothelial or tick cell  
539 cultures resolved by 2DE. (a) Gels were Coomassie blue-stained; spots analysed by  
540 MALDItof MS are indicated as B1, B2, B3 for BUE cells and Is1, Is2, Is3 and Av1, Av2  
541 and Av3 for IDE8 and AVL/CTVM13 cells respectively. Blots from 2DE gels were  
542 probed with pre-infection (b) and post-infection (c) sheep serum; and monoclonal  
543 antibodies 4F10B4 (d) and 1E5H8 (e). M: Molecular masses in kDa.

544

545 **Figure 4:** SDS-PAGE electrophoresis of uninduced (0) and after 4 hours induced (4) *E.*  
546 *coli* cultures. MAP1-GST (~50 kDa) and MAP1-1 His-Tag (~30 kDa) recombinant  
547 proteins are indicated with asterisks (Panel A). Extracts containing recombinant proteins  
548 were transferred to membranes and reacted with monoclonal antibodies 4F10B4 and  
549 1E5H8. U: Uninduced *E. coli* cultures (Panel B).

550

551

552

553

554

555

556 **Figure 5:** Amino acid sequence alignment of *E. ruminantium* MAP1 (Acc.no.  
557 CAI28368) and MAP1-1 (Acc.no. CAI28367) proteins identified by MS analysis.  
558 Identical residues between the two proteins are shown in the top row. The identified  
559 peptide sequences in B1, B2, B3 (bovine endothelial cells) (MAP1 proteins), Is1, Is2, Is3  
560 (IDE8, the non-vector line) and Av1, Av2 and Av3 (AVL/CTVM13, the vector line)  
561 (MAP1-1 proteins) are shaded black, grey and light blue respectively. Predicted *N*- and  
562 *O*-linked glycosylation sites are enclosed in boxes.

563

564 **Figure 6:** Glycoprotein and total protein staining of *E. ruminantium* 2DE gels. Total  
565 protein extracts from *E. ruminantium* grown in BUE, IDE8 and AVL/CTVM13 cell  
566 cultures were resolved by 2DE. Gels were stained for glycoproteins (a) and re-stained  
567 with silver nitrate for total protein comparisons (b). Spots of MAP1 (BUE) and MAP1-1  
568 (IDE8 and AVL/CTVM13) proteins are surrounded by boxes. Glycoprotein molecular  
569 weight standards are indicated on the right.

570

571 **Figure 7:** 2DE gels of total protein extracts from *E. ruminantium*-infected endothelial  
572 (BUE) and tick (IDE8 and AVL/CTVM13) cell cultures before (a) or after chemical  
573 deglycosylation (b). Deglycosylated protein samples were transferred to membranes and  
574 reacted with monoclonal antibody 4F10B4 (c). The spots of interest are surrounded by  
575 boxes in all panels. M: Molecular masses in kDa. *pI*: Isoelectric point.

Fig.1

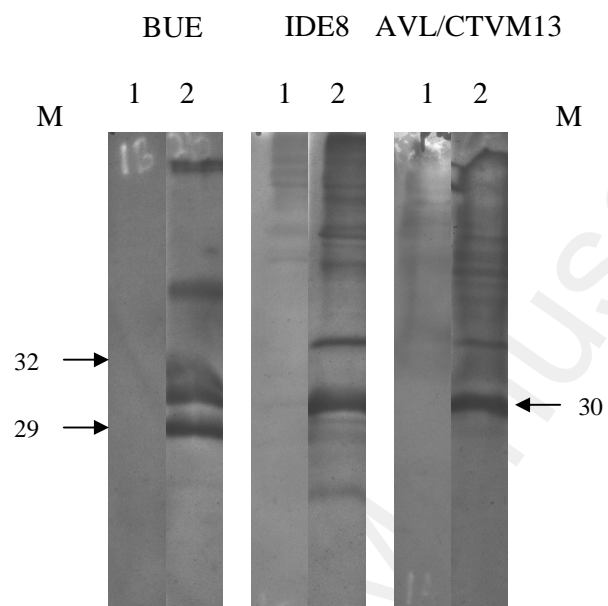


Fig. 2

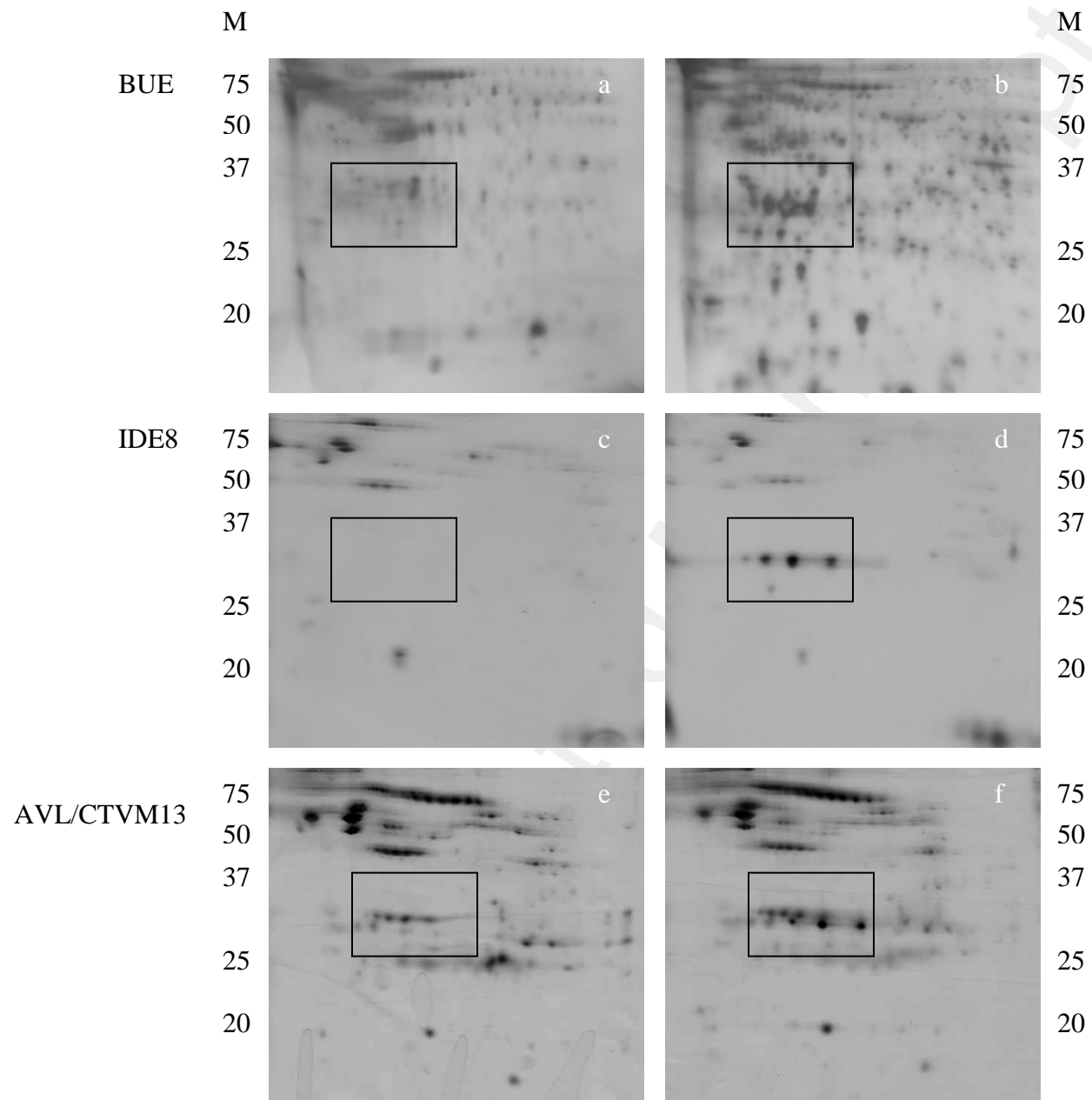


Fig. 3

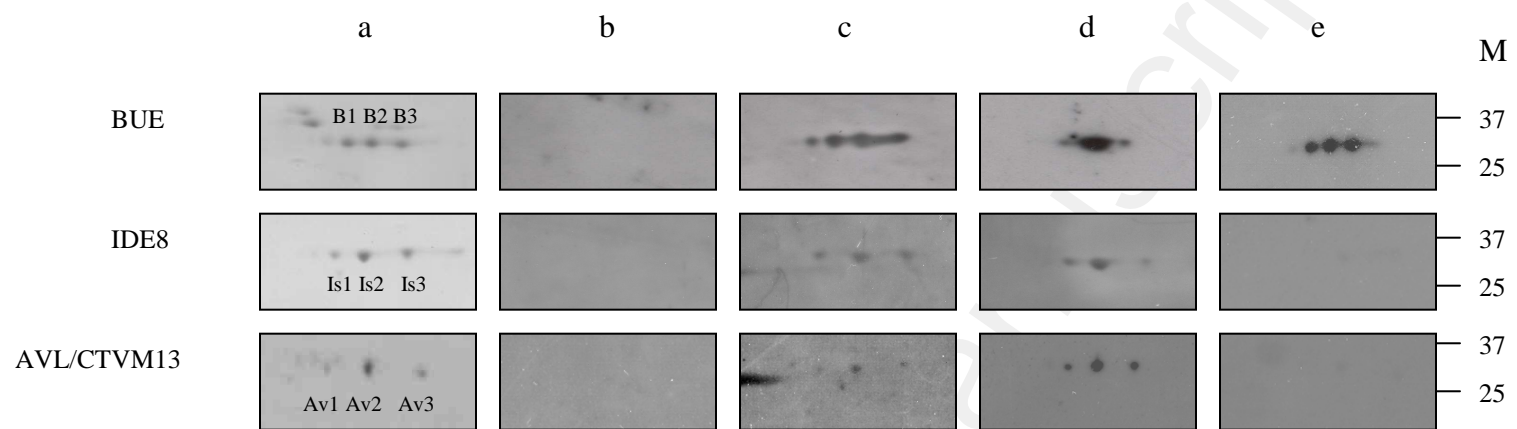


Fig. 4

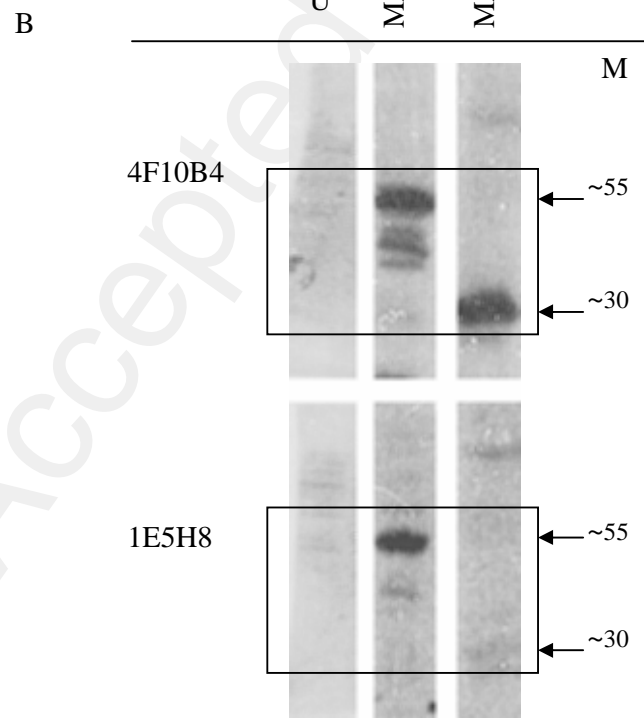
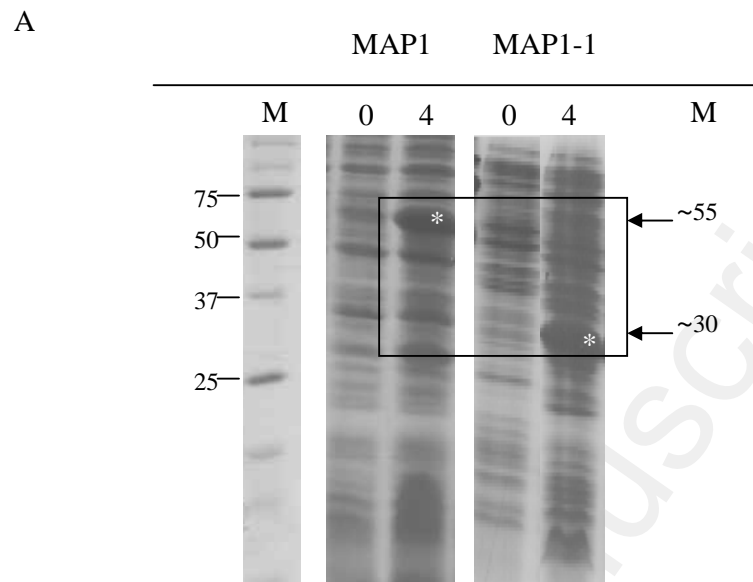


Fig. 5

```

      *           20           *           40           *           60           *           80           *           100
ML F   I   L F I   N M N K K I   S L I S L S F L P G V S F S D V I Q E D   Y I S A K Y P   H F K S E   T V F G L K K   G
B1 M L L F F T S T I V N L F V I I R --- C N M N C K K I F I T S T L I S L V S F L P G V S F S D V I Q E D --- S S P A G S V Y I S A K Y M P T A S H F G R M S I K E D --- S R N T Q T V F G L R K D W D G V   95
B2 M L L F F T S T I V N L F V I I R --- C N M N C K K I F I T S T L I S L V S F L P G V S F S D V I Q E D --- S S P A G S V Y I S A K Y M P T A S H F G R M S I K E D --- S R N T Q T V F G L R K D W D G V   95
B3 M L L F F T S T I V N L F V I I R --- C N M N C K K I F I T S T L I S L V S F L P G V S F S D V I Q E D --- S S P A G S V Y I S A K Y M P T A S H F G R M S I K E D --- S R N T Q T V F G L R K D W D G V   95
I s1 M L F F T Y F I I L K L F I I F Y K R F I N M N Y K K I L V R S A L I S L M S F L P Y Q S F A E P V S S N N I G N E N A R E G F Y I S A K Y N P S I P H F R K F S A E E T P V Y G R D S P T K K V F G L K K --- E G S   105
I s2 M L F F T Y F I I L K L F I I F Y K R F I N M N Y K K I L V R S A L I S L M S F L P Y Q S F A E P V S S N N I G N E N A R E G F Y I S A K Y N P S I P H F R K F S A E E T P V Y G R D S P T K K V F G L K K --- E G S   105
I s3 M L F F T Y F I I L K L F I I F Y K R F I N M N Y K K I L V R S A L I S L M S F L P Y Q S F A E P V S S N N I G N E N A R E G F Y I S A K Y N P S I P H F R K F S A E E T P V Y G R D S P T K K V F G L K K --- E G S   105
A v1 M L F F T Y F I I L K L F I I F Y K R F I N M N Y K K I L V R S A L I S L M S F L P Y Q S F A E P V S S N N I G N E N A R E G F Y I S A K Y N P S I P H F R K F S A E E T P V Y G R D S P T K K V F G L K K --- E G S   105
A v2 M L F F T Y F I I L K L F I I F Y K R F I N M N Y K K I L V R S A L I S L M S F L P Y Q S F A E P V S S N N I G N E N A R E G F Y I S A K Y N P S I P H F R K F S A E E T P V Y G R D S P T K K V F G L K K --- E G S   105
A v3 M L F F T Y F I I L K L F I I F Y K R F I N M N Y K K I L V R S A L I S L M S F L P Y Q S F A E P V S S N N I G N E N A R E G F Y I S A K Y N P S I P H F R K F S A E E T P V Y G R D S P T K K V F G L K K --- E G S   105

      *           120           *           140           *           160           *           180           *           200           *
T S D   F T D S F   N N   G F G I G Y M G P R E E   Y F K N P N   Y   V N T L M N A C Y
B1 K T P S S D S G N N S I I F T E K D Y S F Y E N N P F L G F A G A I G Y S M N G P R I E F E V S Y E T F D V K N P G G N Y K N D A H M Y C A L D T G T P G S T C G A T I N S S V M V K N E N L T D I A L M L N A C Y   202
B2 K T P S S D S G N N S I I F T E K D Y S F Y E N N P F L G F A G A I G Y S M N G P R I E F E V S Y E T F D V K N P G G N Y K N D A H M Y C A L D T G T P G S T C G A T I N S S V M V K N E N L T D I A L M L N A C Y   202
B3 K T P S S D S G N N S I I F T E K D Y S F Y E N N P F L G F A G A I G Y S M N G P R I E F E V S Y E T F D V K N P G G N Y K N D A H M Y C A L D T G T P G S T C G A T I N S S V M V K N E N L T D I A L M L N A C Y   202
I s1 I T R K Y S D ----- F T R T D I S F E G Q N N F I S G F S G S I G Y I M D G P R V E I E A A Y Q K F N P K N E - A N E T D T S D Y Y K H Y G L S R --- A E A M A D K K Y V V L T N N G V T F S S L M F N A C Y   201
I s2 I T R K Y S D ----- F T R T D I S F E G Q N N F I S G F S G S I G Y I M D G P R V E I E A A Y Q K F N P K N E - A N E T D T S D Y Y K H Y G L S R --- A E A M A D K K Y V V L T N N G V T F S S L M F N A C Y   201
I s3 I T R K Y S D ----- F T R T D I S F E G Q N N F I S G F S G S I G Y I M D G P R V E I E A A Y Q K F N P K N E - A N E T D T S D Y Y K H Y G L S R --- A E A M A D K K Y V V L T N N G V T F S S L M F N A C Y   201
A v1 I T R K Y S D ----- F T R T D I S F E G Q N N F I S G F S G S I G Y I M D G P R V E I E A A Y Q K F N P K N E - A N E T D T S D Y Y K H Y G L S R --- A E A M A D K K Y V V L T N N G V T F S S L M F N A C Y   201
A v2 I T R K Y S D ----- F T R T D I S F E G Q N N F I S G F S G S I G Y I M D G P R V E I E A A Y Q K F N P K N E - A N E T D T S D Y Y K H Y G L S R --- A E A M A D K K Y V V L T N N G V T F S S L M F N A C Y   201
A v3 I T R K Y S D ----- F T R T D I S F E G Q N N F I S G F S G S I G Y I M D G P R V E I E A A Y Q K F N P K N E - A N E T D T S D Y Y K H Y G L S R --- A E A M A D K K Y V V L T N N G V T F S S L M F N A C Y   201

      220           *           240           *           260           *           280           *           300           *           320
D I T E G P   P Y C A G I G D L S   N K Y Q G K G I S Y I P E S F I G G H V I G N   T   L A   A S T L D   F G E G R F F
B1 D I T L E G M P V S P Y V C A G I G T D L V S I N A T N P K L S Y Q G K L G I S Y S I N P E A S I F I G G H F H R V I G N E F K D I T T S K I F T S T G K L A T A A S P G F A S A T L D V C H F G I E I G G R F V F   309
B2 D I T L E G M P V S P Y V C A G I G T D L V S I N A T N P K L S Y Q G K L G I S Y S I N P E A S I F I G G H F H R V I G N E F K D I T T S K I F T S T G K L A T A A S P G F A S A T L D V C H F G I E I G G R F V F   309
B3 D I T L E G M P V S P Y V C A G I G T D L V S I N A T N P K L S Y Q G K L G I S Y S I N P E A S I F I G G H F H R V I G N E F K D I T T S K I F T S T G K L A T A A S P G F A S A T L D V C H F G I E I G G R F V F   309
I s1 D I T A E G V P F I P Y A C A G I G A D L I S I F D D I N L K F A Y Q G K I G I S Y P I T P E I S A F I G G Y H V I G N K Y N K V P V K L E V T --- L I D A P Q S T S A S V T L D A G Y F G G E L G V R F T F   304
I s2 D I T A E G V P F I P Y A C A G I G A D L I S I F D D I N L K F A Y Q G K I G I S Y P I T P E I S A F I G G Y H V I G N K Y N K V P V K L E V T --- L I D A P Q S T S A S V T L D A G Y F G G E L G V R F T F   304
I s3 D I T A E G V P F I P Y A C A G I G A D L I S I F D D I N L K F A Y Q G K I G I S Y P I T P E I S A F I G G Y H V I G N K Y N K V P V K L E V T --- L I D A P Q S T S A S V T L D A G Y F G G E L G V R F T F   304
A v1 D I T A E G V P F I P Y A C A G I G A D L I S I F D D I N L K F A Y Q G K I G I S Y P I T P E I S A F I G G Y H V I G N K Y N K V P V K L E V T --- L I D A P Q S T S A S V T L D A G Y F G G E L G V R F T F   304
A v2 D I T A E G V P F I P Y A C A G I G A D L I S I F D D I N L K F A Y Q G K I G I S Y P I T P E I S A F I G G Y H V I G N K Y N K V P V K L E V T --- L I D A P Q S T S A S V T L D A G Y F G G E L G V R F T F   304
A v3 D I T A E G V P F I P Y A C A G I G A D L I S I F D D I N L K F A Y Q G K I G I S Y P I T P E I S A F I G G Y H V I G N K Y N K V P V K L E V T --- L I D A P Q S T S A S V T L D A G Y F G G E L G V R F T F   304

```

Fig. 6

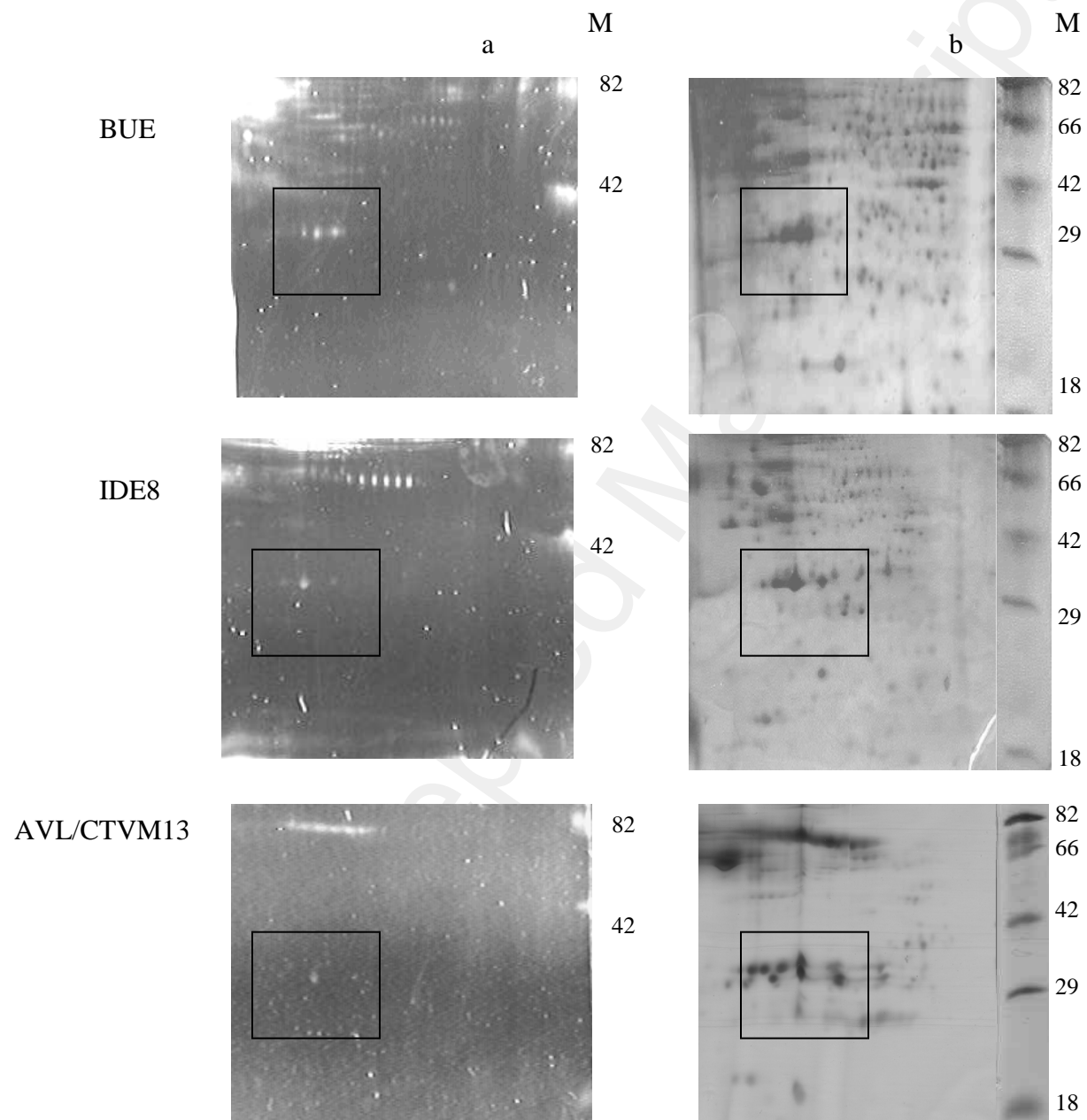




Fig. 7

