



HAL
open science

Proteomic analysis, immunogenicity of secreted proteins from *Rhodococcus equi* ATCC 33701

Corinne Barbey, Aurélie Budin-Verneuil, Séverine Cauchard, Axel Hartke,
Claire Laugier, Vianney Pichereau, Sandrine Petry

► **To cite this version:**

Corinne Barbey, Aurélie Budin-Verneuil, Séverine Cauchard, Axel Hartke, Claire Laugier, et al.. Proteomic analysis, immunogenicity of secreted proteins from *Rhodococcus equi* ATCC 33701. *Veterinary Microbiology*, 2009, 135 (3-4), pp.334. 10.1016/j.vetmic.2008.09.086 . hal-00532518

HAL Id: hal-00532518

<https://hal.science/hal-00532518>

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: Proteomic analysis, immunogenicity of secreted proteins from *Rhodococcus equi* ATCC 33701

Authors: Corinne Barbey, Aurélie Budin-Verneuil, Séverine Cauchard, Axel Hartke, Claire Laugier, Vianney Pichereau, Sandrine Petry



PII: S0378-1135(08)00466-5
DOI: doi:10.1016/j.vetmic.2008.09.086
Reference: VETMIC 4219

To appear in: *VETMIC*

Received date: 7-8-2008
Revised date: 19-9-2008
Accepted date: 24-9-2008

Please cite this article as: Barbey, C., Budin-Verneuil, A., Cauchard, S., Hartke, A., Laugier, C., Pichereau, V., Petry, S., Proteomic analysis, immunogenicity of secreted proteins from *Rhodococcus equi* ATCC 33701, *Veterinary Microbiology* (2008), doi:10.1016/j.vetmic.2008.09.086

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 **Proteomic analysis and immunogenicity of secreted proteins from *Rhodococcus equi***

2 **ATCC 33701**

3
4
5
6
7
8 Corinne Barbey¹, Aurélie Budin-Verneuil², Séverine Cauchard¹, Axel Hartke², Claire
9 Laugier¹, Vianney Pichereau², and Sandrine Petry^{1*}

10 *AFSSA, Laboratoire d'Etudes et de Recherches en Pathologie Equine, IPC, Goustranville,*
11 *14430 Dozulé, France¹ and Laboratoire de Microbiologie de l'Environnement, USC INRA*
12 *2017, EA956, IFR ICORE 146, Université de Caen, 14032 Caen Cedex, France²*

13
14
15 _____
16 *Corresponding author. Mailing address: AFSSA, Laboratoire d'Etudes et de Recherches en
17 Pathologie Equine, IPC, Goustranville, 14430 Dozulé, France. Phone: 33 (0)2 31 79 22 76.
18 Fax: 33 (0)2 31 39 21 37. E-mail: s.petry@dozule.afssa.fr.

19 **Abstract**

20 *Rhodococcus equi* is one of the most important causes of mortality in foals between 1 to 6
21 months of age. Although rare, infection also occurs in a variety of other mammals including
22 humans, often following immunosuppression of various causes. Secreted proteins are known
23 to mediate important pathogen-host interactions and consequently are favored candidates for
24 vaccine development as they are the most easily accessible microbial antigens to the immune
25 system. Here, we describe the results of a proteomic analysis based on SDS PAGE,
26 immunoblot and mass spectrometry, which was carried out aiming the identification of
27 secreted proteins that are differently expressed at 30°C versus 37°C and at mid-exponential
28 versus early-stationary growth phase and antigenic proteins from *R. equi* ATCC 33701. A
29 total of 48 proteins was identified regardless of growth conditions. The cholesterol oxidase
30 ChoE appears to be the major secretory protein. Moreover, four proteins revealed high
31 homologies with the mycolyl transferases of the Ag85 complex from *Mycobacterium*
32 *tuberculosis*. The sequence analysis predicted that 24 proteins are transported by a signal
33 peptide-dependent pathway. Moreover, five antigenic proteins of *R. equi* were identified by
34 immunoblot, including a novel strongly immunoreactive protein of unknown function. In
35 conclusion, the elucidation of the secretome of *R. equi* identified several proteins with
36 different biological functions and a new candidate for developing vaccines against *R. equi*
37 infection in horse.

38

39 **Keywords:** *Rhodococcus equi*; Secretome; Mass spectrometry; Signal peptide; Antigenic
40 proteins

41 1. Introduction

42 *Rhodococcus equi* is the common cause of pyogranulomatous pneumonia with a high
43 mortality rate in foals up to 6 months of age (Muscatello et al., 2007). Even if the pathogenic
44 mechanisms of *R. equi* remain largely unknown, there is evidence that virulence is dependent
45 on a large plasmid with a pathogenicity island (Takai et al., 2000; Takai et al., 1991b),
46 containing nine virulence-associated protein (*vap*) genes (Muscatello et al., 2007). Among
47 them, *vapA* encodes a surface-expressed protein (Takai et al., 1991a; Tan et al., 1995), which
48 plays a crucial role in intracellular growth in macrophages and for disease development
49 (Giguère et al., 1999; Jain et al., 2003).

50 Secreted proteins are of high interest for functional investigations of bacterial pathogens
51 because they come into direct contact with host tissues. Particularly, secreted proteins of
52 intracellular pathogens such as *R. equi* represent the primary antigen targets of host immune
53 response (Kaufmann and Hess, 1999). Thus, they may be key molecules to induce immune
54 protection and, consequently, key candidates for vaccine development. Few data are available
55 about *R. equi* and only five secreted proteins were described such as the cholesterol oxidase
56 ChoE and the phospholipase C, two membrane-damaging exoenzymes, originally called equi
57 factors (Machang'u and Prescott, 1991a), and three virulence-associated proteins VapC, VapD
58 and VapE (Byrne et al., 2001) of still unknown function.

59 The goal of the present study was to identify proteins secreted by *R. equi* ATCC 33701 at
60 different temperatures (30°C, the optimal growth temperature, versus 37°C, the host
61 temperature) and at different growth phases (mid-exponential versus early-stationary phase of
62 growth). Bioinformatic study was carried out to predict the functions and to confirm the
63 extracellular localization of the identified proteins. Furthermore, immunogenic secreted
64 proteins expressed during equine rhodococcosis were characterized by using Western blot
65 probed with pooled sera from nine foals which died from *R. equi* infection.

66

67 **2. Materials and methods**68 *2. 1. Bacteria and culture conditions*

69 The bacterial strain used in experiments was *R. equi* ATCC 33701, a virulence plasmid-
70 bearing strain. A single colony was selected and grown in 10 ml of Luria Bertani (LB, Difco,
71 Le Pont de Claix, France) medium for 16 h at 30°C or 37°C, at pH 7.0 with shaking at 200
72 rpm. Cultures were inoculated at a 1:100 dilution with the preculture. Samples (500 ml) were
73 taken in mid-exponential (OD₆₀₀ of 0.4 to 0.5) and early-stationary (OD₆₀₀ of 0.7 to 0.8)
74 growth phases, respectively after 15 and 22 h of incubation with shaking at 200 rpm.

75 *E. coli* BL21(DE3) (Novagen, Merck, Darmstadt, Germany) and *E. coli* M15 (Qiagen,
76 Courtaboeuf, France) were used for protein expression.

77

78 *2. 2. Extracellular sample preparation for the proteome analysis*

79 Cultures were centrifuged at 20,000 g for 30 min at 4°C and the supernatants were sterile
80 filtered by using a 0.2-µm nitrocellulose filter (Millipore, Molsheim, France). Then a final
81 concentration of 10% (w/v) trichloroacetic acid (TCA, Sigma, St Quentin Fallavier, France)
82 was added to the filtrate supernatant, mixed well and placed on ice at 4°C for 16 h. The
83 precipitated proteins were harvested by centrifugation (20,000 g for 30 min at 4°C), washed
84 four times in cold methanol (Fluka, quality upgraded, St Quentin Fallavier, France) dried with
85 speed-vac and dissolved in sample solution consisting of 0.5 M Tris base pH 8.0 (Sigma).
86 Protein concentration was determined according to the Lowry method (Lowry et al., 1951),
87 using bovine serum albumine (BSA, Sigma) as a standard.

88

89 *2. 3. SDS PAGE*

90 Proteins samples were denatured in Laemmli Buffer 6% (v/v) SDS (Eurobio, Les Ulis,
91 France), 16% (v/v) glycerol (Eurobio), 0.02% (v/v) β -mercaptoethanol (Sigma), and 0.08%
92 (w/v) bromophenol blue (Electran, BDH Laboratory Supplies, Poole, UK) in 100 mM Tris-
93 HCl Buffer pH6.8 by heating to 100°C for 5 min. Then samples (150 μ g in each well) were
94 separated by one-dimensional denaturing sodium dodecyl sulfate-polyacrylamide gel
95 electrophoresis (SDS-PAGE) with a 7% acrylamide (29:1 acrylamide/bisacrylamide, Eurobio)
96 stacking gel and a 12% acrylamide resolving gel. The Bio-Rad prestained broad standard
97 marker was used as a protein molecular weight marker. After electrophoresis using the
98 PROTEAN[®] II xi Cell (Bio-Rad Laboratories, Marnes-la-Coquette, France) at 150 V for 15 h,
99 the gel was stained with Coomassie Brilliant Blue R250 (Reactifs Ral, Martillac, France) to
100 visualize the proteins.

101

102 2. 4. *In gel-trypsin digestion*

103 Protein spots to be identified were manually excised from Coomassie stained gels and
104 excised gel fractions were washed in ultra-pure water. These fractions were then dehydrated
105 by several incubations in 50 mM ammonium bicarbonate (BICAM, Fluka) and acetonitrile
106 (CAN, Merck, Fontenay-sous-Bois, France), followed by rehydration in 50 mM BICAM. 0.5
107 μ g modified porcine trypsin (sequencing grade modified trypsin, Promega Corporation,
108 Charbonnières, France) was added to each sample, which was then incubated during 16 h at
109 37°C. The supernatant was recovered by centrifugation (5,000 g for 5 min). A second
110 extraction from these gel pieces was carried out by incubation in 50 mM BICAM/ACN (1/1 :
111 v/v) and then by incubation in 5% acetic acid (Carlo Erba, Val de Reuil, France). Peptide-
112 containing liquid fractions were pooled after each incubation and vacuum dried. Peptides
113 were resuspended in 20 μ l 0.1% acetic acid prior to storage at -70°C until mass spectrometry
114 (MS) analysis.

115

116 *2. 5. Trypsin digestion of crude extracts of secreted proteins*

117 Crude extracts of secreted proteins were obtained as described above from culture
118 supernatants at the early-stationary growth phase (OD_{600} of 0.7 to 0.8) and 30°C or 37°C.
119 Proteins were resuspended in 50 mM BICAM pH 8.0 containing 0.5 µg trypsin, incubated at
120 37°C during 16 h, before being vacuum dried and resuspended in 0.1% acetic acid. Peptides
121 were stored at -70°C until MS analysis.

122

123 *2.6. MS/MS analysis*

124 An electrospray ion trap mass spectrometer (LCQ DECA XP, Thermo Finnigan) coupled
125 on line with an HPLC (Surveyor LC, Thermo Finnigan) was used. Peptides were separated
126 and analysed mainly as previously described (Sauvageot et al., 2002). Recorded MS/MS
127 spectra were compared to theoretical fragmentations of a trypsinolysed database deduced
128 from the *R. equi* 103 DNA genomic sequence, using the Sequest software (version 2, Thermo
129 Finnigan).

130

131 *2. 7. Bioinformatic tools to analyze amino acid sequences*

132 Protein sequences were obtained thanks to the non-annotated genome sequence of *R. equi*
133 strain 103 produced by the *Rhodococcus equi* Sequencing Group at the Sanger Institute
134 (http://www.sanger.ac.uk/Projects/R_equi/). For unknown proteins, homologous proteins
135 were identified using the BLASTP program available at the website
136 <http://www.ncbi.nlm.nih.gov/BLAST/sequence>. Conserved protein domains were identified
137 by using CDD software based on NCBI
138 (<http://www.ncbi.nlm.nih.gov/Structure/cdd/cdd.shtml>).

139 The secretory (Sec-type) signal peptides were determined with the SignalP 3.0 algorithm
140 (<http://www.cbs.dtu.dk/services/SignalP/>) while the Twin-arginine (Tat-type) and the
141 lipoprotein signal peptides were determined with the TatP 1.0
142 (<http://www.cbs.dtu.dk/services/TatP/>) and the LipoP 1.0
143 (<http://www.cbs.dtu.dk/services/LipoP/>) algorithms respectively. TMHMM 2.0 algorithm was
144 used to predict transmembrane domains (<http://www.cbs.dtu.dk/services/TMHMM/>). The
145 different softwares cited below are all available from www.expasy.org.

146

147 2. 8. *Western blot analysis*

148 Two SDS PAGE gels were run in parallel. One was used for identification and the other
149 for immunoblot. 150 µg of proteins were loaded in each well to compare each culture
150 condition. Proteins resolved by SDS PAGE were electroblotted onto nitrocellulose
151 membranes (Amersham Biosciences, Orsay, France) with Towbin transfer buffer using a
152 Trans-Blot[®] SD semi-dry transfer cell (Bio-Rad Laboratories) at 15 V for 40 min. Blotted
153 membranes were blocked with 5% (w/v) skimmed milk powder in TBS buffer (10 mM Tris,
154 0.2 M NaCl, 0.05% (v/v) Tween 20) overnight at 4°C and incubated with pooled sera from
155 nine foals that had died of rhodococcosis at 1:200 dilution in blocking solution for 1 h at
156 37°C. Membranes were then washed three times for 15 min each with TBS buffer. They were
157 incubated with rabbit anti-horse IgG, conjugated to alkaline phosphatase, at 1:5000 dilution
158 (Bethyl Laboratories, Montgomery, TX, USA) for 1 h at 37°C and washed as described
159 above. Antigen-antibody reactions were developed with AP Conjugate Substrate kit (Bio-Rad
160 Laboratories) according to the manufacturer.

161

162 2. 9. *Cloning and expression of twelve genes*

163 The primers and vectors used for PCR amplification and cloning are listed (Table 1). The
164 primers were designed on the available *R. equi* genome sequence. PCR products obtained
165 with *R. equi* ATCC 33701 DNA were cloned into pQE30 (Qiagen) or pET-30 Ek/LIC vector
166 (Novagen) to obtain recombinant plasmids following the recommendations of the supplier
167 (Table 1). Recombinant pQE30 vectors were transformed into *E. coli* M15 while recombinant
168 pET-30 Ek/LIC vectors were transformed into *E. coli* BL21(DE3). The presence of
169 appropriate insert was confirmed by sequencing.

170 To assess expression of the cloned genes, recombinant *E. coli* M15 or *E. coli* BL21(DE3)
171 were grown overnight at 37°C in 10 ml of LB medium supplemented with 30 µg/ml of
172 kanamycin (Sigma) for the two *E. coli* strains and 100 µg/ml of ampicillin (Sigma) for *E. coli*
173 M15 only. One ml of the bacterial suspension was then subinoculated into 100 ml of fresh LB
174 medium and the bacteria were allowed to grow at 37°C until the optical density of the
175 medium at 600 nm reached 0.6. Thereafter, the expression of the cloned genes was induced by
176 addition of IPTG (ABgene, Courtaboeuf, France) to a final concentration of 1 mM. After a
177 further 3 h incubation at 37°C, the bacterial cells were collected by centrifugation for 10 min
178 at 5,000 g. The cell lysis was performed by incubation for 1 h with a lysis buffer containing
179 100 mM NaH₂PO₄ (Merck), 10 mM Tris base and 8 M urea (Eurobio). The cell debris were
180 removed by centrifugation at 10,000 g for 10 min. Supernatants were used for Western blot
181 analysis.

182

183 **3. Results**

184 *3. 1. Analysis of secreted proteins by SDS PAGE*

185 Secreted proteins were extracted from the supernatant of cultures of *R. equi* grown in the
186 poor protein containing LB medium in order to limit contamination by medium derived
187 polypeptides. Secreted proteins obtained at 30°C or 37°C and at mid-exponential or early-

188 stationary phase were analyzed by SDS PAGE (Fig. 1). According to standard marker
189 proteins, secreted polypeptides had molecular weight ranging from 7 to 100 kDa. At 30°C, the
190 migration patterns of secreted proteins did not appear to have significant differences
191 depending on the growth phase and at least 21 distinct bands were detected. At 37°C, three
192 additional bands around 29 kDa (bands No. 17 to 19) were present in supernatants harvested
193 in early-stationary phase compared to mid-exponential phase. Furthermore, regardless of the
194 growth phase, six additional bands between 7 and 20 kDa (bands No. 23 to 27 and 30) were
195 detected in supernatants from cultures grown at 37°C compared to 30°C. Finally, several
196 intensively Coomassie-stained bands were observed: four bands (No. 1, 3, 6 and 15) in all
197 tested supernatants and three additional bands (No. 25 to 27) expressed in supernatants
198 harvested at 37°C.

199

200 3. 2. *Proteins identification*

201 30 bands were cut out of gels (bands No. 1 to 30, Fig. 1), subjected to in-gel trypsin
202 digestion and analyzed by LC-ESI tandem mass spectrometry. In parallel, the secretome of *R.*
203 *equi* was also analyzed without previous electrophoretic separation by injecting a
204 trypsinolysed crude preparation of secreted proteins into LC-ESI tandem mass spectrometer
205 (see Materials and Methods for details). These two approaches led to the identification of 44
206 and 16 proteins, respectively (Table 2). The method without previous electrophoretic
207 separation allowed us to identify fewer secreted proteins than the other method but led to the
208 rapid identification of dominant proteins present in the secretome. Since we found twelve
209 identical proteins in both methods, a total of 48 proteins was finally identified. All identified
210 proteins were encoded by the chromosome except for the Vap proteins. Because of the non-
211 annotated genome sequence of *R. equi*, homologous proteins were identified using BLASTP
212 program and the homology with the first hit was indicated for each protein (Table 2). Six

213 proteins arbitrarily named proteins X1 to X6 did not share any similarity with proteins from
214 other organisms. For the majority of the proteins (73%), a potential function could be
215 assigned due to the presence of conserved domains detected by NCBI CDD algorithm. All
216 proteins identified in the culture filtrate could be classified into several major functional
217 categories: pathogenicity (ChoE, VapA), stress adaptation (alkyl hydroperoxide reductase,
218 catalase, Cpn10, GroEL, CspA, probable Stage II sporulation protein), export (MspA), cell
219 wall biosynthesis (four probable mycolyl transferases arbitrarily named 1 to 4, N-acetyl
220 muramoyl-L-alanine amidase), reactivation of bacteria in stationary phase (probable
221 resuscitation promoting factor), energy metabolism (dihydrolipoyl dehydrogenase,
222 phosphomannomutase), fatty acid metabolism (lipid-transfer protein), general cell processes
223 and hypothetical proteins without known function.

224 The molecular weights of most proteins detected in SDS PAGE were consistent with the
225 theoretical values (Table 2). However, several proteins displayed a significantly higher mass
226 than expected. Six proteins were represented by several isoforms of different molecular
227 weights. The cholesterol oxidase ChoE was presented in six, VapD in four, VapA and
228 mycolyl transferase 3 (MT3) in three and elongation factor Tu and protein X4 in two different
229 fractions. By taking into account the growth conditions, the three additional bands around 29
230 kDa (No. 17 to 19, Fig. 1) in supernatants harvested in early-stationary phase compared to
231 mid-exponential phase at 37°C were composed of one isoform of mycolyl transferase 3, four
232 hypothetical proteins and a probable protease from *Rhodococcus* sp. strain RHA1. Moreover,
233 the six additional bands between 7 and 20 kDa (No. 23 to 27 and 30, Fig. 1) in supernatants
234 from cultures grown at 37°C compared to 30°C contained three isoforms of VapA, two
235 isoforms of VapD, protein X1 and probable mycolyl transferase 3 (no result was obtained for
236 the band No. 26). The unfractionated protein extract analysis revealed that VapA and VapD
237 were detected only at 37°C.

238 Concerning the most intensively Coomassie-stained bands, cholesterol oxidase and
239 probable mycolyl transferase 1 (No. 1, 3, 6 and 15, Fig. 1) were present in all tested
240 supernatants whereas protein X1, VapA and VapD (No. 25 to 27, Fig. 1) were only expressed
241 in supernatants harvested at 37°C. Cholesterol oxidase, VapA and VapD were also identified
242 among the unfractionated protein extract directly analyzed by mass spectrometry.

243

244 3.3. Protein structure and localization

245 SignalP 3.0 algorithm revealed the presence of cleavage site for signal peptidases (SPases)
246 for 24 out of the 48 identified proteins (Table 3), suggesting that these are secreted proteins.
247 Three proteins were recognized by the TatP 1.0 algorithm as harboring a twin-arginine (Tat)
248 signal peptide. The consensus sequence recognized by this algorithm is
249 RRX[FGAVML][LITMVF]. It contains the two invariant arginine residues and variable
250 amino-acids indicated between parentheses or by the X letter. Nine additional proteins
251 contained a complete or incomplete Tat motif with at least the two arginines in their signal
252 peptide but they were not recognized by the TatP 1.0 algorithm. The other 12 proteins
253 contained a signal peptide specific for the general protein secretion (Sec) pathway (Table 3).
254 In our study, the Sec and Tat signal peptides had a length varying from 26 to 63 amino acids.
255 Both signal peptides were composed of three domains, a positively charged N-terminal
256 domain (n-region), a hydrophobic domain (h-region) and a hydrophilic C-terminal domain (c-
257 region) usually carrying the consensus AXA at position -1 to -3 relative to the cleavage site
258 for SPases I (Tjalsma et al., 2000). This consensus was observed for twenty out of the 24
259 signal peptides (Table 3). If the 24 signal peptides were taken into consideration, the general
260 consensus sequence at position -3 to -1 relative to the SPases I cleavage site was (A/V)-X-
261 (A/K/T). The consensus sequence L-(A/S)-(A/G)-C at position -3 to +1 relative to the SPases

262 II cleavage site had never been found, suggesting the absence of lipoprotein signal peptides
263 (Table 3). At last, 24 proteins contained no obvious signal sequence.

264 According to the TMHMM 2.0 algorithm, we could predict a putative membrane
265 localization for two proteins. The polypeptide homologous to the hypothetical protein
266 RHA1_ro05477 contained, in addition to a signal peptide, two transmembrane domains at
267 central position. The polypeptide homologous to the hypothetical protein RHA1_ro03156,
268 which did not possess a signal peptide, contained a single transmembrane domain localized at
269 the N-terminus (Table 3). These two proteins may correspond to a membrane protein and a
270 surface protein, respectively. Furthermore, according to the NCBI CDD algorithm, the
271 probable mycolyl transferase 1 contained at its C-terminus three LGFP tandem repeats, which
272 may be involved in anchoring the protein to the cell wall.

273

274 3. 4. Antigenic profiles of *R. equi* secretome

275 Antigenic profiles were explored by immunoblot performed with pooled sera from nine *R.*
276 *equi* infected foals (Fig. 2). No band was detected in the control medium. Moreover, Western
277 blot control experiments using antisera from healthy foals did not reveal any immunogenic
278 protein (data not shown). Independent of growth conditions, immunogenic proteins were
279 present in four regions with molecular weights about 99, 54, 29 and 20 kDa, respectively. A
280 protein band of a monodimensional gel may harbor more than one polypeptide and probably
281 not all of them would correspond to immunoreactive proteins. Analysis of the four bands by
282 mass spectrometry revealed the presence of the following proteins: cholesterol oxidase,
283 catalase, hypothetical proteins RHA1_ro06188, RHA1_ro03156, RHA1_ro05216,
284 RHA1_ro05669, probable protease, probable secreted protein named PSP, VapA, VapD,
285 probable mycolyl transferases 3 and 1. In order to test which of them reacts actually with the
286 serum, the twelve proteins were first overexpressed in *E. coli* (Fig. 3A) and then used for a

287 Western blot analysis using the same pooled horse sera as in Fig. 2 (Fig. 3B). No *E. coli*
288 immunoreactive proteins were observed in the negative controls with empty pQE30 and pET-
289 30 vectors. In contrast, the serum reacted strongly with the PSP and the cholesterol oxidase.
290 The proteins VapA, VapD and the probable mycolyl transferase 1 were less immunoreactive
291 and the serum reacted only weakly with the proteins homologous to the hypothetical proteins
292 RHA1_ro06188 and RHA1_ro03156. The other recombinant proteins did not react with the
293 serum.

294

295 **4. Discussion**

296 In order to characterize new *R. equi* candidate antigens associated with rhodococcosis
297 among secreted proteins, we used a proteomic approach combined with serum profiling by
298 using sera from foals that had died of rhodococcosis. This approach enabled identification of
299 48 proteins released by the *R. equi* ATCC 33701 strain during growth. As expected, 24
300 polypeptides contained a signal peptide sequence required for their translocation across the
301 cytoplasmic membrane *via* the Sec-pathway (Keyzer et al., 2003) or the Tat-pathway (Lee et
302 al., 2006), the two most important secretion systems distributed in all bacterial genera.
303 Among them, half the proteins were predicted to possess a potential Tat signal peptide,
304 suggesting that the Tat pathway could be as significant as the Sec pathway in *R. equi*. In
305 contrast to the latter, the former pathway is involved in the secretion of exclusively folded
306 proteins and protein translocation has no requirement for ATP as an energy source but relies
307 solely on the proton motive force. It is interesting to note that all proteins identified in the
308 most intensively Coomassie-stained bands contained a Sec signal peptide (VapA and VapD)
309 or a Tat signal peptide (cholesterol oxidase, protein X1, mycolyl transferases 1 and 3).

310 The majority of the 24 remaining proteins without identifiable signal peptides contained
311 proteins usually considered as strictly cytoplasmic and their presence in the secretome

312 requires further investigation. Abundant cytoplasmic proteins are often found in the secretome
313 of other pathogens such as *M. tuberculosis* (Malen et al., 2007), *Listeria monocytogenes*
314 (Trost et al., 2005) or *Staphylococcus aureus* (Sibbald et al., 2006) and extracellular roles
315 have been suggested. For example, the cytoplasmic elongation factor Tu (EF-Tu) is described
316 as a surface protein, which may exhibit an unexpected and novel function, the adhesion of
317 bacteria to host tissue (Dallo et al., 2002). A similar activity has been reported for GroEL in
318 several pathogens (Bergonzelli et al., 2006). Two *R. equi groEL2*- based DNA vaccines were
319 used in immunization of mice and were found to elicit a predominantly Th1 biased immune
320 response (Phumoonna et al., 2008; Vanniasinkam et al., 2004). But their protective efficacy
321 against *R. equi* infection was not demonstrated. The mode of secretion for these proteins is
322 unknown but they can be released by cell lysis or by hitherto unknown non-classical secretion
323 pathways or signal peptide-independent pathways (Bendtsen et al., 2005). Such non-classical
324 systems have been reported in *M. tuberculosis* and *L. monocytogenes* for the export of
325 virulence factors without signal peptide (Bendtsen et al., 2005). For many years, it has been
326 known that *M. tuberculosis* secretes antigenic proteins without apparent signal peptide such as
327 ESAT-6 (early secretory antigenic target) (Bendtsen et al., 2005).

328 The cholesterol oxidase was, under all conditions tested, the major secreted protein of *R.*
329 *equi* ATCC 33701. Despite its membranolytic activity, this enzyme was demonstrated not to
330 be an important virulence factor in the natural host (Pei et al., 2006). It may be important in
331 providing carbon sources for *R. equi* inside macrophages, as they are for other intracellular
332 bacteria such as *Mycobacterium avium* and *M. tuberculosis*. The variants or isoforms of
333 cholesterol oxidase may be attributed to biologically important post-translational
334 modifications, ranging from chemical modifications, such as glycosylations for isoforms of
335 higher molecular weight than the calculated one, to proteolytic cleavage for polypeptides with
336 molecular weight significantly lesser than the expected one. VapA, an atypical lipoprotein, is

337 another virulence factor involved in the intracellular growth in macrophages and identified
338 among secreted proteins in our experimental conditions, contrary to previous studies which
339 detected VapA in the cell pellet only (Byrne et al., 2001; Kohler et al., 2003). Its presence in
340 the supernatant can be explained by the fact that lipoproteins are vulnerable to proteolytic
341 cleavage relatively close to the N-terminal cysteine and are consequently released as soluble
342 protein in the culture filtrate (Malen et al., 2007). Moreover, it has been suggested that some
343 lipoproteins may be alternatively processed by signal peptidase I or II, and this mechanism
344 may explain their dual localization in the extracellular environment or as lipoproteins in the
345 cell wall (Malen et al., 2007). In our work, VapA was observed in three isoforms, in
346 accordance with the study of Tan et al. (1995). Indeed, in that work, the polypeptides isolated
347 from the 18-22 kDa and 17.5 kDa fractions were hydrophobic and lipid modified whereas the
348 15-kDa protein did not show this modification. VapD was detected in four isoforms which
349 correspond most probably to degradation products. VapA and VapD were observed only at
350 37°C in our study which is in accordance with previous findings (Byrne et al., 2001).

351 Some proteins identified in the present study were predicted to be involved in the cell wall
352 biosynthesis. Indeed, four out of at least 13 *R. equi* probable mycolyl transferases predicted by
353 genome analysis (Sydor et al., 2007) were detected in our study. These four enzymes were
354 homologous to the mycobacterial antigen 85 complex (Ag85 complex), major secretion
355 products of *M. tuberculosis* (Content et al., 1991; Wiker and Harboe, 1992). This Ag85
356 complex is made of three mycolyl transferases (FbpA, FbpB and FbpC), pivotal for the
357 synthesis of the mycolic acid layer in the cell wall (Belisle et al., 1997). The probable mycolyl
358 transferases 1 and 2 are encoded by the gene *fbpA* (Sydor et al., 2007) and *fbpB* (Pei et al.,
359 2007) of *R. equi*, respectively. In *R. equi*, FbpA could be involved in capsule anchorage since
360 a *fbpA* mutant is characterized by a non-mucoid phenotype (Sydor et al., 2007). Since
361 synthesis of the capsule is constitutive in *R. equi*, it is not surprising that FbpA was one of the

362 most abundant proteins of the secretome regardless of growth conditions (Band No 3, Fig 1).

363 It should be noted that recent virulence studies in a murine model showed that both FbpA and

364 FbpB might not be involved in *R. equi* pathogenesis (Pei et al., 2007; Sydor et al., 2007).

365 Immunological characterization of *R. equi* secreted proteins by sera from foals that had

366 died of rhodococcosis revealed antigenic reactivity with four 99, 54, 29 and 20 kDa

367 immunogenic regions, similar to the immunoblots obtained by Kohler et al. (2003). By using

368 recombinant proteins, we demonstrated that five proteins (VapA, VapD, cholesterol oxidase,

369 probable mycolyl transferase 1 and the PSP) were indeed involved in the immunogenicity of

370 these four regions. Contrary to previous studies which identified Vap proteins as strongly

371 immunoreactive proteins (Byrne et al., 2001; Kohler et al., 2003), the two most strongly

372 immunoreactive proteins detected in the present study were the cholesterol oxidase and the

373 PSP. Machang'u and Prescott (1991b) showed that antibodies against the cholesterol oxidase

374 associated to the phospholipase C of *R. equi* did not prevent foals from developing lung

375 abscesses following experimental infection. However, this result has to be examined by taking

376 into account the infective dose, the dose and the nature of antigens which correspond to a

377 combination of two proteins, the immunization and infection routes and the type of adjuvant.

378 Indeed, recent studies demonstrated that these various parameters may influence the

379 polarization of immune response which must be a cellular type response to confer protection

380 against *R. equi* infection (Lunn and Townsend, 2000). The strong immune response to the

381 PSP suggests that this protein was readily accessible to the immune system *in vivo* and

382 consequently supports the fact that it was actively secreted despite the lack of a signal peptide.

383 This protein needs to be further characterized. Finally, the latter protein, cholesterol oxidase

384 and VapA that has been shown highly immunogenic (Kohler et al., 2003) in foals and

385 protective in a murine model (Haghighi and Prescott, 2005) may represent promising antigen

386 candidates for the vaccine development against equine rhodococcosis.

387

388 **5. Conclusion**

389 To our knowledge, this is the first report of a comprehensive proteomic analysis of
390 secreted proteins of *R. equi*. The associated immunological characterization allowed us to
391 identify several proteins recognized by the host immune system, including a novel strongly
392 immunoreactive protein without known function. Studies to examine their immunoprotective
393 potential against *R. equi* infections are in progress.

394

395 **Acknowledgements**

396 This research was supported by grants from the Région Basse-Normandie, Haras
397 Nationaux and ADREF (Association pour le Développement de la Recherche Equine en
398 France).

399 We thank Doctor J. Cauchard (AFSSA Dozulé, France) for providing sera from nine foals
400 that had died of rhodococcosis.

401 **References**

402

403 Belisle, J.T., Vissa, V.D., Sievert, T., Takayama, K., Brennan, P.J., Besra, G.S., 1997. Role of
404 the major antigen of *Mycobacterium tuberculosis* in cell wall biogenesis. *Science* 276,
405 1420-1422.

406 Bendtsen, J.D., Kiemer, L., Fausboll, A., Brunak, S., 2005. Non-classical protein secretion in
407 bacteria. *BMC Microbiol.* 58, 1-13.

408 Bergonzelli, G.E., Granato, D., Pridmore, R.D., Marvin-Guy, L.F., Donnicola, D., Cortesy-
409 Theulaz, I.E., 2006. GroEL of *Lactobacillus johnsonii* La1 (NCC 533) is cell surface
410 associated: potential role in interactions with the host and the gastric pathogen
411 *Helicobacter pylori*. *Infect. Immun.* 74, 425-434.

412 Byrne, B.A., Prescott, J.F., Palmer, G.H., Takai, S., Nicholson, V.M., Alperin, D.C., Hines,
413 S.A., 2001. Virulence plasmid of *Rhodococcus equi* contains inducible gene family
414 encoding secreted proteins. *Infect. Immun.* 69, 650-656.

415 Content, J., de la Cuvellerie, A., De Wit, L., Vincent-Levy-Frebault, V., Ooms, J., De Bruyn,
416 J., 1991. The genes coding for the antigen 85 complexes of *Mycobacterium*
417 *tuberculosis* and *Mycobacterium bovis* BCG are members of a gene family: cloning,
418 sequence determination, and genomic organization of the gene coding for antigen 85-
419 C of *M. tuberculosis*. *Infect. Immun.* 59, 3205-3212.

420 Dallo, S.F., Kannan, T.R., Blaylock, M.W., Baseman, J.B., 2002. Elongation factor Tu and E1
421 beta subunit of pyruvate dehydrogenase complex act as fibronectin binding proteins in
422 *Mycoplasma pneumoniae*. *Mol. Microbiol.* 46, 1041-1051.

423 Giguère, S., Hondalus, M.K., Yager, J.A., Darrach, P., Mosser, D.M., Prescott, J.F., 1999. Role
424 of the 85-kilobase plasmid and plasmid-encoded virulence-associated protein A in
425 intracellular survival and virulence of *Rhodococcus equi*. *Infect. Immun.* 67, 3548-
426 3557.

427 Haghghi, H.R., Prescott, J.F., 2005. Assessment in mice of vapA-DNA vaccination against
428 *Rhodococcus equi* infection. *Vet. Immunol. Immunopathol.* 104, 215-225.

429 Jain, S., Bloom, B.R., Hondalus, M.K., 2003. Deletion of vapA encoding Virulence
430 Associated Protein A attenuates the intracellular actinomycete *Rhodococcus equi*.
431 *Mol. Microbiol.* 50, 115-128.

432 Kaufmann, S.H., Hess, J., 1999. Impact of intracellular location of and antigen display by
433 intracellular bacteria: implications for vaccine development. *Immunol. Lett.* 65, 81-84.

434 de Keyzer, J., van der Does, C., Driessen, A.J., 2003. The bacterial translocase: a dynamic
435 protein channel complex. *Cell. Mol. Life Sci.* 60, 2034-2052.

436 Kohler, A.K., Stone, D.M., Hines, M.T., Byrne, B.A., Alperin, D.C., Norton, L.K., Hines,
437 S.A., 2003. *Rhodococcus equi* secreted antigens are immunogenic and stimulate a type
438 1 recall response in the lungs of horses immune to *R. equi* infection. *Infect. Immun.*
439 71, 6329-6337.

440 Lee, P.A., Tullman-Ercek, D., Georgiou, G., 2006. The bacterial twin-arginine translocation
441 pathway. *Annu. Rev. Microbiol.* 60, 373-395.

442 Lowry, O.H., Rosebrough, N.J., Farr, A.L., Randall, R.J., 1951. Protein measurement with the
443 Folin phenol reagent. *J. Biol. Chem.* 193, 265-275.

444 Lunn, D.P., Townsend, H.G., 2000. Equine vaccination. *Vet. Clin. North Am. Equine Pract.*
445 16, 199-226.

446 Machang'u, R.S., Prescott, J.F., 1991a. Purification and properties of cholesterol oxidase and
447 choline phosphohydrolase from *Rhodococcus equi*. *Can. J. Vet. Res.* 55, 332-340.

- 448 Machang'u, R.S., Prescott, J.F., 1991b. Role of antibody to extracellular proteins of
449 *Rhodococcus equi* in protection against *R. equi* pneumonia in foals. *Vet. Microbiol.*
450 26, 323-333.
- 451 Malen, H., Berven, F.S., Fladmark, K.E., Wiker, H.G., 2007. Comprehensive analysis of
452 exported proteins from *Mycobacterium tuberculosis* H37Rv. *Proteomics* 7, 1702-
453 1718.
- 454 Muscatello, G., Leadon, D.P., Klayt, M., Ocampo-Sosa, A., Lewis, D.A., Fogarty, U.,
455 Buckley, T., Gilkerson, J.R., Meijer, W.G., Vazquez-Boland, J.A., 2007. *Rhodococcus*
456 *equi* infection in foals: the science of 'rattles'. *Equine Vet. J.* 39, 470-478.
- 457 Pei, Y., Dupont, C., Sydor, T., Haas, A., Prescott, J.F., 2006. Cholesterol oxidase (ChoE) is
458 not important in the virulence of *Rhodococcus equi*. *Vet. Microbiol.* 118, 240-246.
- 459 Pei, Y., Parreira, V., Nicholson, V.M., Prescott, J.F., 2007. Mutation and virulence
460 assessment of chromosomal genes of *Rhodococcus equi* 103. *Can. J. Vet. Res.* 71, 1-7.
- 461 Phumoonna, T., Barton, M.D., Vanniasinkam, T., Heuzenroeder, M.W., 2008. Chimeric
462 vapA/groEL2 DNA vaccines enhance clearance of *Rhodococcus equi* in aerosol
463 challenged C3H/He mice. *Vaccine* 26, 2457-2465.
- 464 Sauvageot, N., Pichereau, V., Louarme, L., Hartke, A., Auffray, Y., Laplace, J.M., 2002.
465 Purification, characterization and subunits identification of the diol dehydratase of
466 *Lactobacillus collinoides*. *Eur. J. Biochem.* 269, 5731-5737.
- 467 Sibbald, M.J., Ziebandt, A.K., Engelmann, S., Hecker, M., de Jong, A., Harmsen, H.J.,
468 Raangs, G.C., Stokroos, I., Arends, J.P., Dubois, J.Y., van Dijl, J.M., 2006. Mapping
469 the pathways to staphylococcal pathogenesis by comparative secretomics. *Microbiol.*
470 *Mol. Biol. Rev.* 70, 755-788.
- 471 Sydor, T., von Barga, K., Becken, U., Spuerck, S., Nicholson, V.M., Prescott, J.F., Haas, A.,
472 2007. A mycolyl transferase mutant of *Rhodococcus equi* lacking capsule integrity is
473 fully virulent. *Vet. Microbiol.* 128, 327-341.
- 474 Takai, S., Hines, S.A., Sekizaki, T., Nicholson, V.M., Alperin, D.A., Osaki, M., Takamatsu,
475 D., Nakamura, M., Suzuki, K., Ogino, N., Kakuda, T., Dan, H., Prescott, J.F., 2000.
476 DNA sequence and comparison of virulence plasmids from *Rhodococcus equi* ATCC
477 33701 and 103. *Infect. Immun.* 68, 6840-6847.
- 478 Takai, S., Koike, K., Ohbushi, S., Izumi, C., Tsubaki, S., 1991a. Identification of 15- to 17-
479 kilodalton antigens associated with virulent *Rhodococcus equi*. *J. Clin. Microbiol.* 29,
480 439-443.
- 481 Takai, S., Sekizaki, T., Ozawa, T., Sugawara, T., Watanabe, Y., Tsubaki, S., 1991b.
482 Association between a large plasmid and 15- to 17-kilodalton antigens in virulent
483 *Rhodococcus equi*. *Infect. Immun.* 59, 4056-4060.
- 484 Tan, C., Prescott, J.F., Patterson, M.C., Nicholson, V.M., 1995. Molecular characterization of
485 a lipid-modified virulence-associated protein of *Rhodococcus equi* and its potential in
486 protective immunity. *Can. J. Vet. Res.* 59, 51-59.
- 487 Tjalsma, H., Bolhuis, A., Jongbloed, J.D., Bron, S., van Dijl, J.M., 2000. Signal peptide-
488 dependent protein transport in *Bacillus subtilis*: a genome-based survey of the
489 secretome. *Microbiol. Mol. Biol. Rev.* 64, 515-547.
- 490 Trost, M., Wehmhoner, D., Karst, U., Dieterich, G., Wehland, J., Jansch, L., 2005.
491 Comparative proteome analysis of secretory proteins from pathogenic and
492 nonpathogenic *Listeria* species. *Proteomics* 5, 1544-1557.
- 493 Vanniasinkam, T., Barton, M.D., Heuzenroeder, M.W., 2004. The immunogenicity of
494 *Rhodococcus equi* GroEL2-based vaccines in a murine model. *Vet. Immunol.*
495 *Immunopathol.* 98, 91-100.
- 496 Wiker, H.G., Harboe, M., 1992. The antigen 85 complex: a major secretion product of
497 *Mycobacterium tuberculosis*. *Microbiol. Rev.* 56, 648-661.

498 **Figure legends**

499 **Fig. 1.** Separation of secreted proteins from *R. equi* by SDS-PAGE. Proteins extracted from
500 *R. equi* ATCC 33701 grown in LB medium at 30°C (lanes 2, 3) or 37°C (lanes 4, 5) and at the
501 mid-exponential (lanes 2, 4) or the early-stationary (lanes 3, 5) growth phase. Lane 1 contains
502 proteins extracted from non inoculated LB medium (negative control). All lanes are loaded
503 with 150 µg of protein. The spots 1 to 30 indicate bands analyzed by LC ESI MS/MS. The
504 numbers on the left indicate molecular masses (in kilodaltons).

505

506 **Fig. 2.** Detection of immunogenic proteins in the culture supernatant of *R. equi*. Immunoblot
507 of secreted proteins from *R. equi* ATCC 33701 was probed with pooled sera from nine foals
508 that had died of rhodococcosis. Lanes were loaded with 150 µg of secreted proteins *R. equi*
509 ATCC 33701 as follows: lane 1, non inoculated LB medium (negative control medium); lane
510 2, 30°C mid-exponential phase; lane 3, 30°C early-stationary phase; lane 4, 37°C mid-
511 exponential phase; lane 5, 37°C stationary phase. The numbers on the left indicate molecular
512 masses (in kilodaltons).

513

514 **Fig. 3.** Overexpression and Western blot analysis of recombinant *R. equi* proteins. A) SDS-
515 PAGE of lysates of *E. coli* M15 or *E. coli* BL21(DE3) expressing the recombinant proteins.
516 Lane 1, molecular mass marker; lane 2, empty pQE30 vector; lane 3, *E. coli* lysate containing
517 recombinant cholesterol oxidase; lane 4, *E. coli* lysate containing recombinant catalase; Lane
518 5, empty pET-30 vector; Lanes 6 to 15, the following recombinant proteins: hypothetical
519 proteins RHA1_ro06188, RHA1_ro03156, RHA1_ro05216, RHA1_ro05669, probable
520 protease, PSP (probable secreted protein), VapA, VapD, probable mycolyl transferase 3,
521 probable mycolyl transferase 1. B) Western blot showing reactivity of sera with the twelve
522 recombinant proteins. Sera were obtained from foals that had died of rhodococcosis.

Fig. 1

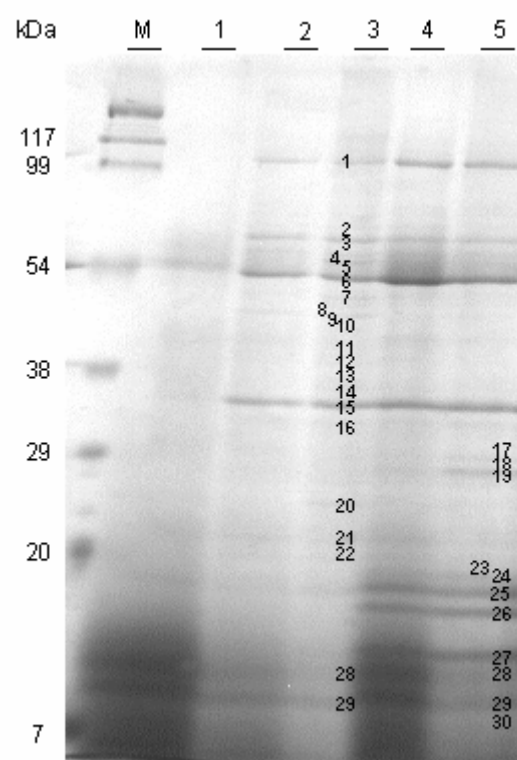
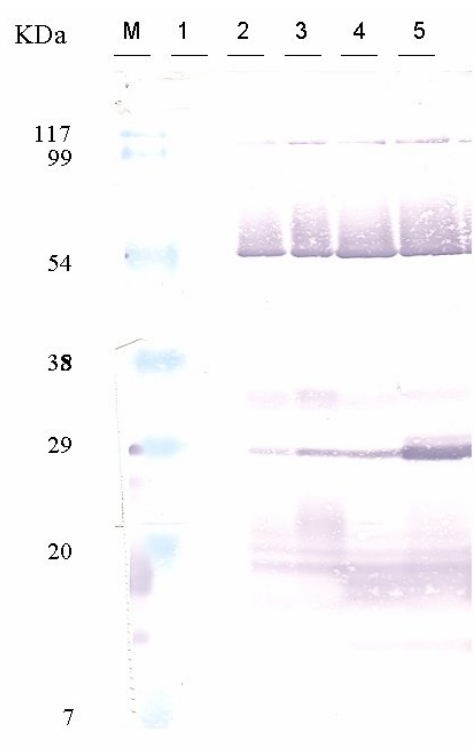


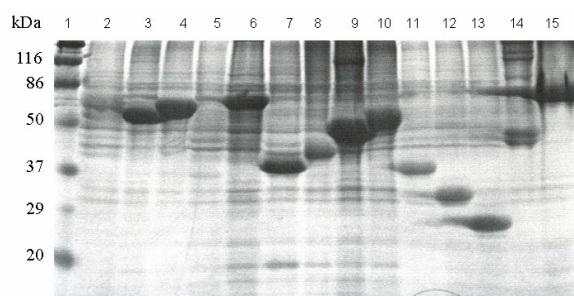
Fig. 2



Accepted Manuscript

Fig. 3

A)



B)

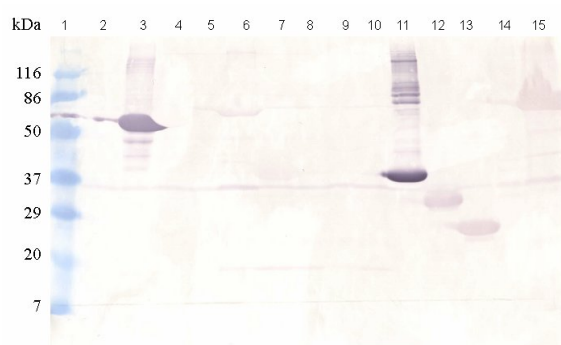


Table 1. Primers and vectors used in gene cloning

Gene of	Primers ^a	Vector
catalase	5' ATGACGTCCATTCAGCCACGT 3' 5' GTTCTGCCGACCGAACTACC 3'	pQE30
ChoE	5' TCCGCCTCCGGAGCTCCAGCC 3' 5' TCTCATAAGCTTCGACGCGT 3'	pQE30
hypothetical protein RHA1_ro03156	5' <u>GACGACGACAAGATCCAGCAGGAGAACACCTC</u> 3' 5' <u>GAGGAGAAGCCCCGGTCTAGCCGTA</u> ACTGACGACGAC 3'	pET-30 Ek/LIC
hypothetical protein RHA1_ro05216	5' <u>GACGACGACAAGATCGTGGCTGCAATGACCTCC</u> 3' 5' <u>GAGGAGAAGCCCCGGTCTACTCGCAGGCGATCCCCGTC</u> 3'	pET-30 Ek/LIC
hypothetical protein RHA1_ro05669	5' <u>GACGACGACAAGATCCAGCACAAGACGATCACACTC</u> 3' 5' <u>GAGGAGAAGCCCCGGTTTACCGGAGACCCA</u> ACTTGCT 3'	pET-30 Ek/LIC
hypothetical protein RHA1_ro06188	5' <u>GACGACGACAAGATACCGCTGCCGGCGTCGAT</u> 3' 5' <u>GAGGAGAAGCCCCGGTCTAGCCGAGTGTGGCGAG</u> 3'	pET-30 Ek/LIC
probable antigen 85 protein precursor (mycolyl transferase 3)	5' <u>GACGACGACAAGATCTCGACGCCCGGCGCGACCGAC</u> 3' 5' <u>GAGGAGAAGCCCCGGTCTACCGGGGAGTGCCCATG</u> 3'	pET-30 Ek/LIC
probable antigen 85 complex protein (mycolyl transferase 1)	5' <u>GACGACGACAAGATGGCACAGTCCCTCGGCACCCCC</u> 3' 5' <u>GAGGAGAAGCCCCGGTTCAGGCGATGGTCGCCTTTCC</u> 3'	pET-30 Ek/LIC
probable protease	5' <u>GACGACGACAAGATCGAACCCGCCGCTCGGGT</u> 3' 5' <u>GAGGAGAAGCCCCGGTATTACCGATGTCGGTCAGTG</u> 3'	pET-30 Ek/LIC
probable secreted protein (PSP)	5' <u>GACGACGACAAGATTGGTGTGGCGGACGCGG</u> 3' 5' <u>GAGGAGAAGCCCCGGTCTACGACAGCTCCCATCC</u> 3'	pET-30 Ek/LIC
VapA	5' <u>GACGACGACAAGATGACCGTTCTTGATTCCGGTAGC</u> 3' 5' <u>GAGGAGAAGCCCCGGTCTAGGCGTTGTGCCAGCTACC</u> 3'	pET-30 Ek/LIC
VapD	5' <u>GACGACGACAAGATGCAGGAGCTAGCTGGCACCAAG</u> 3' 5' <u>GAGGAGAAGCCCCGGTCTACTCCCACCCGCCAGTGCC</u> 3'	pET-30 Ek/LIC

^a The underlined letters in primers are complementary extensions compatible to sequences in the pET-30 Ek/LIC cloning vector.

Table 2. Identification of secreted proteins by mass spectrometry on SDS PAGE (bands noted by a number) or in unfractionated protein extract

Band No. ^a	Sp score ^b	Description	Identified protein ^c			Protein domain ^c	Observed MW (kDa) ^f	Theoretical MW (kDa) ^g
			Bacteria ^d (% Identity)		NCBI Accession No.			
1	1442	Cholesterol oxidase ChoE	<i>R. equi</i> 103	(100)	CAC44897	choline dehydrogenase / related flavoproteins	100	54.5
2	1627	Cholesterol oxidase ChoE	<i>R. equi</i> 103	(100)	CAC44897	choline dehydrogenase / related flavoproteins	70	54.5
	2009	30S ribosomal protein	<i>R. sp</i> RHA1	(95)	YP_118107	ribosomal protein S1 like RNA binding domain	70	54.0
3	798	Probable antigen 85 complex protein (mycolyl transferase 1)	<i>R. sp</i> RHA1	(71)	YP_704140	putative esterase	65	62.8
4		No result					60	
5	1388	Catalase	<i>R. equi</i> 103	(100)	AY779519.1	catalase	55	55.0
	1058	Cholesterol oxidase ChoE	<i>R. equi</i> 103	(100)	CAC44897	choline dehydrogenase / related flavoproteins	55	54.5
	2104	Hypothetical protein RHA1_ro06188	<i>R. sp</i> RHA1	(75)	YP_706123		55	58.5
6	1440	Cholesterol oxidase ChoE	<i>R. equi</i> 103	(100)	CAC44897	choline dehydrogenase / related flavoproteins	54	54.5
7	676	Elongation factor Tu	<i>N. farcinica</i>	(95)	YP_121291	EF Tu family (II and III domains)	50	43.5
8	1007	Elongation factor Tu	<i>N. farcinica</i>	(95)	YP_121291	EF Tu family (II and III domains)	48	43.5
9	922	DNA-directed RNA polymerase α -subunit	<i>R. sp</i> RHA1	(98)	YP_706097	RNA polymerase Rbp3 / Rpb11 dimerisation domain / RpoA insert domain / α -chain C terminal domain	46	38.1
	647	Hypothetical protein RHA1_ro05154	<i>R. sp</i> RHA1	(77)	YP_705093		46	21.0
	1035	Probable stage II sporulation protein	<i>R. sp</i> RHA1	(51)	ABG95848	stage II sporulation protein	46	54.8
	448	Short-chain dehydrogenase/reductase	<i>S. tropica</i>	(67)	YP_001159019	short chain dehydrogenase / 3 ketoacyl (acyl carrier protein) reductase	46	33.0
10	1004	Cholesterol oxidase ChoE	<i>R. equi</i> 103	(100)	CAC44897	choline dehydrogenase / related flavoproteins	45	54.5
	1056	Hypothetical protein RHA1_ro01045	<i>R. sp</i> RHA1	(54)	YP_701030	secretory lipase	45	36.5
	1795	Hypothetical protein RHA1_ro04854	<i>R. sp</i> RHA1	(60)	YP_704797	cutinase	45	39.0
11	715	Alkyl hydroperoxide reductase	<i>R. sp</i> RHA1	(87)	YP_703236	peroxiredoxin family	42	21.0
	703	Cell division initiation protein	<i>R. sp</i> RHA1	(78)	YP_701065	DivIVA protein (cell division initiation protein)	42	31.0
	714	Probable thioesterase	<i>R. sp</i> RHA1	(66)	YP_707177	thioesterase protein	42	17.0
	957	Not matched (protein X3)					42	24.0
	645	Short-chain dehydrogenase/reductase	<i>S. tropica</i>	(67)	YP_001159019	short chain dehydrogenase / 3 ketoacyl (acyl carrier protein) reductase	42	33.0
12	1895	Not matched (protein X4)					40	37.0
	1589	Hypothetical protein RHA1_ro05704	<i>R. sp</i> RHA1	(64)	YP_705641	glucose dehydrogenase	40	37.2
13	1435	Probable antigen 85-A precursor (mycolyl transferase 2)	<i>R. sp</i> RHA1	(43)	YP_704946	putative esterase	38	36.8
	1906	Not matched (protein X4)					38	37.0
	739	Not matched (protein X5)					38	37.0
14	701	Probable antigen 85 protein precursor (mycolyl transferase 3)	<i>R. sp</i> RHA1	(57)	YP_702108	putative esterase	37	34.5
	816	Hypothetical protein RHA1_ro04050	<i>R. sp</i> RHA1	(56)	YP_704005	N-acetylmuramoyl-L-alanine amidase domain	37	68.0
15	1678	Cholesterol oxidase ChoE	<i>R. equi</i> 103	(100)	CAC44897	choline dehydrogenase / related flavoproteins	36	54.5
16	450	Amidase AmiC_1	<i>M. ulcerans</i>	(49)	YP_906566	amidase	34	49.2
17	701	Probable antigen 85 protein precursor (mycolyl transferase 3)	<i>R. sp</i> RHA1	(57)	YP_702108	putative esterase	32	34.5
18	1672	Hypothetical protein RHA1_ro03156	<i>R. sp</i> RHA1	(35)	YP_703117	MspA (porin)	29	30.2
	1379	Hypothetical protein RHA1_ro05216	<i>R. sp</i> RHA1	(66)	YP_705155		29	30.2
	1151	Hypothetical protein RHA1_ro05669	<i>R. sp</i> RHA1	(67)	YP_705606	transglycosylase-like domain	29	36.0
19	1803	Probable protease	<i>R. sp</i> RHA1	(63)	YP_705835		28	42.8
	1124	Probable secreted protein (PSP)	<i>R. erythropolis</i>	(46)	CAC37046		28	29.8
20	817	Probable mycolyl transferase (mycolyl transferase 4)	<i>R. sp</i> RHA1	(65)	YP_704013	putative esterase	25	33.0
	1625	Hypothetical protein nfa45220	<i>N. farcinica</i>	(73)	YP_120737	transglycosylase like domain	25	8.0
	2135	Hypothetical protein RHA1_ro06358	<i>R. sp</i> RHA1	(69)	YP_706293		25	8.0
	840	Cold shock protein A	<i>R. sp</i> RHA1	(98)	YP_704299	cold shock protein	25	7.0
	713	Not matched (protein X6)					25	53.3
21	2579	Not matched (protein X2)					22	18.5

	1033	Probable lipoprotein	<i>R. sp</i> RHA1	(60)	YP_704138	peptidase M23	22	18.0
	881	Probable resuscitation-promoting factor	<i>R. sp</i> RHA1	(67)	YP_704901	transglycosylase like domain	22	17.5
22	1543	Peptidyl-prolyl cis-trans isomerase	<i>R. sp</i> RHA1	(88)	YP_704938	cyclophilin	20	22.0
23	678	VapA (intracellular multiplication)	<i>R. equi</i> 103	(100)	BAB16621	<i>R. equi</i> virulence-associated protein	19	16.2
24	762	VapA (intracellular multiplication)	<i>R. equi</i> 103	(100)	BAB16621	<i>R. equi</i> virulence-associated protein	18	16.2
25	793	VapA (intracellular multiplication)	<i>R. equi</i> 103	(100)	BAB16621	<i>R. equi</i> virulence-associated protein	17	16.2
	1604	Probable antigen 85 protein precursor (mycolyl transferase 3)	<i>R. sp</i> RHA1	(57)	YP_702108	putative esterase	17	34.5
	666	Not matched (protein X1)					17	11.0
26		No result					16	
27	1045	VapD	<i>R. equi</i> 103	(100)	BAB16624	<i>R. equi</i> virulence-associated protein	14	14.0
28	1705	VapD	<i>R. equi</i> 103	(100)	BAB16624	<i>R. equi</i> virulence-associated protein	12	14.0
	1109	Hypothetical protein RHA1_ro05477	<i>R. sp</i> RHA1	(37)	YP_705415		12	18.5
	1055	Chaperonin Cpn10	<i>M. sp</i> JLS	(94)	ZP_01280279	chaperonin Cpn10	12	11.0
29	2117	VapD	<i>R. equi</i> 103	(100)	BAB16624	<i>R. equi</i> virulence-associated protein	10	14.0
	1607	Chaperonin Cpn10	<i>M. sp</i> JLS	(94)	ZP_01280279	chaperonin Cpn10	10	11.0
	581	Phosphomannomutase	<i>R. sp</i> RHA1	(82)	YP_706251	Phosphoglucomutase / phosphomannomutase	10	48.6
	930	Probable extracellular protein	<i>R. sp</i> RHA1	(37)	YP_700670		10	14.7
30	1475	VapD	<i>R. equi</i> 103	(100)	BAB16624	<i>R. equi</i> virulence-associated protein	9	14.0
unfractionated protein extract	450	Catalase	<i>R. equi</i> 103	(100)	AY779519.1	catalase		55.0
	669	Chaperonin Cpn10	<i>M. sp</i> JLS	(94)	ZP_01280279	chaperonin Cpn10		11.0
	1898	Cholesterol oxidase ChoE	<i>R. equi</i> 103	(100)	CAC44897	choline dehydrogenase / related flavoproteins		54.5
	793	Cold shock protein A	<i>R. sp</i> RHA1	(98)	YP_704299	cold shock protein		7.2
	895	Dihydrolipoyl dehydrogenase	<i>R. sp</i> RHA1	(89)	YP_706197	pyridine nucleotide-disulphide oxidoreductase		21.4
	545	Elongation factor G	<i>R. sp</i> RHA1	(95)	YP_701891	EF G subfamily (II and IV domains)		76.9
	630	Elongation factor Tu	<i>N. farcinica</i>	(94)	YP_121291	EF Tu family (II and III domains)		43.5
	1205	Chaperonin GroEL2	<i>R. equi</i> 103	(100)	AAK95493	chaperonin GroEL		56.5
	687	Hypothetical protein RHA1_ro04050	<i>R. sp</i> RHA1	(56)	YP_704005	N-acetylmuramoyl-L-alanine amidase domain		68.0
	1268	Hypothetical protein RHA1_ro05477	<i>R. sp</i> RHA1	(37)	YP_705415			21.1
	1843	Hypothetical protein RHA1_ro06188	<i>R. sp</i> RHA1	(75)	YP_706123			58.5
	625	Probable nonspecific lipid-transfer protein	<i>R. sp</i> RHA1	(89)	YP_704432	thiolase domain		40.1
	2337	Probable antigen 85 protein precursor (mycolyl transferase 3)	<i>R. sp</i> RHA1	(57)	YP_702108	putative esterase		34.5
	1339	Not matched (protein X2)						18.4
	976	VapA	<i>R. equi</i> 103	(100)	BAB16621	<i>R. equi</i> virulence-associated protein		16.2
	549	VapD	<i>R. equi</i> 103	(100)	BAB16624	<i>R. equi</i> virulence-associated protein		14.0

^a The band numbers were shown in Fig. 1.

^b Sp score given by Sequest for the most significant peptide (Sp \geq 450) found using a DNA database prepared from the non annotated genome sequence of *R. equi* 103.

^c Best matching proteins with identities \geq 35% were taken into consideration.

^d *M. sp* JLS, *Mycobacterium sp* JLS; *M. ulcerans*, *Mycobacterium ulcerans*; *N. farcinica*, *Nocardia farcinica*; *R. sp* RHA1, *Rhodococcus sp* RHA1; *R. erythropolis*, *Rhodococcus erythropolis*; *S. tropica*, *Salinispora tropica*.

^e Protein domain as deduced from sequence analysis with the NCBI conserved domain algorithm.

^f Deduced from the gels.

^g The molecular weight of proteins without signal peptide.

Table 3. Subcellular localization of proteins based on sequence analysis

Identified proteins	Export pathway ^a	Signal Peptide sequence (SP) ^b	TMH ^c (except for SP)
Probable extracellular protein	Sec	MRKVTLAAGLVAAAGLVVPATA ANA	
Protein X2	Sec	MQKLRTIVVAAAVALPVVAAP ANA	
Hypothetical protein RHA1_ro05477	Sec	MKLRKFAATSALVIAAMGVGAGT AYA	140-162 169-191
VapD	Sec	MVRARAFGRFLTFFLLAVAVIATVSMGG ANA	
VapA	Sec	MKTLHKTVSKAIAATAVAAAAAMI PAGVANA	
Hypothetical protein RHA1_ro04854	Sec	MGRRVERMVGTAALGGLVAAATVFASGGV AAA	
Hypothetical protein nfa45220	Sec	MAKFDIKRTIGTVAATAALVAAPFALSTGT TANA	
Protein X4	Sec	VLKTTPLRLWLLVAVLSVVTMASCAGLAP QAQA	
Hypothetical protein RHA1_ro05669	Sec	VSPFARINTARSPLLYVVVAALFATLLVGGGL VVT	
Probable resuscitation-promoting factor	Sec	MSGRHRKPTTTGRVAVKAVTGAIMGVGGVAFAGT TANA	
Probable mycolyl transferase 4	Sec	MRTHRPGSTHRLKRLIGLVAAALLPALTA VTTTANA	
Hypothetical protein RHA1_ro05216	Sec	VPAPLHLKSCRPLSFHRVRSSTLLGAAVAATVAASLAACGSDSAGNTLTVSPSSSTTT PAAK	
Probable mycolyl transferase 2	Tat	MWGS DPRRAARRRL RRGVV GASLAALVALPLAGE AAA	
Cholesterol oxidase ChoE	Tat	MTDSRANRADATRGVASVSR RREFLAGAGL TAGAIALSSMSTS ASA	
Probable mycolyl transferase 1	Tat	MRSGLERSYRAGTEGR WRRRALS VAAALVMP IAAGITSVT AGA	
Hypothetical protein RHA1_ro04050	Tat	LPH RR PKNSIVLAAVAAVAVATPF AVQA	} ^d
Hypothetical protein RHA1_ro01045	Tat	MRRSIGESL RR ATAVGAALFALTALVPA VASA	
Probable mycolyl transferase 3	Tat	MRVTQKSR RLKAL IGATTAAVAALPVFVGAG VASA	
Hypothetical protein RHA1_ro06188	Tat	MNRR TFLGAAG VGALGISALPLVGNRSWNPLAA QARA	
Probable protease	Tat	MRIPLAPRAAVHSF ARRVAV VGSSALLVLP LSASAQA	
Hypothetical protein RHA1_ro05704	Tat	VSS TPADREGSFVRR STL GAGA AVAAAIIGIGV APTACA	
Probable lipoprotein	Tat	MKGC MAGSKWRLGVRR MVAGVVAAAGL TAGVAVMSAPAASA	
Protein X1	Tat	MIR RWWRCSSKSSMMPRR KNGPIRK PQAGSFANAF CRFI ATA	
Probable stage II sporulation protein	Tat	MRSE PMATPRLRIS ADRVL RRR PG RRP SR SVGRLAALGLV PAL AVGA	
Hypothetical protein RHA1_ro03156	None		61-83 (N-terminal)
Alkyl hydroperoxide reductase	None		
Amidase AmiC_1	None		
Catalase	None		
Cell division initiation protein	None		
Chaperonin Cpn10	None		
Cold shock protein A	None		
Dihydrolipoyl dehydrogenase	None		
DNA-directed RNA polymerase α -subunit	None		
Elongation factor G	None		
Elongation factor TU	None		
Chaperonin GroEL2	None		
Hypothetical protein RHA1_ro05154	None		
Hypothetical protein RHA1_ro06358	None		
Peptidyl-prolyl cis-trans isomerase	None		
Phosphomannomutase	None		
Probable nonspecific lipid-transfer protein	None		
Probable thioesterase	None		
Protein X3	None		
Protein X5	None		
Protein X6	None		
Probable secreted protein (PSP)	None		
30S ribosomal protein S1	None		
Short-chain dehydrogenase/reductase	None		

^a Sec, protein secretion system; Tat, twin-arginine translocation system.

^b Signal peptide as deduced from sequence analysis with the SignalP 3.0 and TatP 1.0 algorithms (www.expasy.org). Only results with score $\geq 0,8$ were taken into consideration. The consensus motif recognized by TatP1.0 algorithm RRX[FGAVML][LITMVF] is underlined. The two invariant arginines are enlarged. The recognition sequence of the SPases I at position -1 to -3 relative to the cleavage site is indicated in bold.

^c TMH, transmembrane helix as predicted from sequence analysis with the TMHMM 2.0 algorithms (www.expasy.org). Their position in the sequence is indicated.

^d These signal peptides contain a complete or incomplete Tat motif, not recognized by the TatP1.0 algorithm.