Molecular cloning of a gene encoding the immunogenic 21 kDa protein of *Cowdria ruminantium*

S. M. Mahan,¹† T. C. McGuire,² S. M. Semu,[†] M. V. Bowie,¹ F. Jongejan,³ F. R. Rurangirwa⁴ and A. F. Barbet¹

Author for correspondence: S. M. Mahan. Tel. +263 4 794980. Fax: +263 4 794980.

¹ Department of Infectious Diseases, College of Veterinary Medicine, University of Florida, PO Box 110880, Gainesville, FL 32611-0880, USA

² Department of Veterinary Microbiology and Pathology, Washington State University, Pullman, WA 99163-7043, USA

³ Tick Research Unit, Department of Parasitology and Tropical Veterinary Medicine, University of Utrecht, PO Box 80.165, 3508 TD Utrecht, The Netherlands

⁴ Kenya Agricultural Research Institute, SR-CRSP, PO Box 58137, Nairobi, Kenya

Major immunogenic polypeptides (21, 32, 40, 46, 58, 85 and 160 kDa) of *Cowdria ruminantium* were identified by immunoprecipitation and immunoblotting. A pUC13 library of *C. ruminantium* genomic DNA was screened with hyperimmune sheep serum to identify *Escherichia coli* colonies which expressed genes encoding these immunogenic proteins. A recombinant *E. coli* colony, F5.2, was identified containing plasmid insert DNA of 2773 bp. The cloned DNA insert contained two long open reading frames (ORFs) of 627 bp (complete) and 831 bp (incomplete), both potentially encoding proteins containing an N-terminal signal peptide. Deletion experiments suggested that the hyperimmune sheep serum recognized a protein that was encoded by the 627 bp ORF. The 627 bp ORF was amplified by polymerase chain reaction (PCR), subcloned and expressed to a high level in *E. coli*. A sheep antiserum made to the expressed recombinant fusion protein recognized a 21 kDa protein of all strains of *C. ruminantium* tested, confirming that the 627 bp ORF encodes a native 21 kDa protein in *C. ruminantium*. Similarly, the recombinant protein was recognized by all sera tested from heartwater-infected animals. The antigenic conservation of the 21 kDa protein and its immunogenic nature are reasons for further testing of this recombinant protein in subunit diagnostic tests.

Keywords: *Cowdria ruminantium*, rickettsia, heartwater, immunogenic proteins, serodiagnosis

INTRODUCTION

Heartwater or cowdriosis is an infectious, economically important disease of livestock which is caused by the Gram-negative rickettsia *Cowdria ruminantium* and transmitted by ticks of the genus *Amblyomma* (Uilenberg, 1983). Economic losses due to heartwater are significant due to the high mortality rates, which range from 20 to 90% (Uilenberg, 1983, 1990). Losses due to morbidity are likely but have not been assessed. Heartwater is endemic in Africa south of the Sahara, and has also spread to the Caribbean islands (Burridge, 1985; Perreau *et al.*, 1980). Its presence in the Caribbean region threatens the livestock industries of North, Central and South America, as some of the indigenous *Amblyomma* ticks of these

continents have been experimentally shown to transmit *C. ruminantium* (Barre *et al.*, 1987). To reduce the losses caused by heartwater, it is important to detect its presence in an area and to monitor its spread. Hence, improved diagnostic methods to detect infected and carrier animals (Andrew & Norval, 1989) are essential.

The traditional methods of diagnosis of heartwater include: (a) recognition of clinical signs, which are not unique to heartwater and can be confused with other diseases (Uilenberg, 1983); (b) microscopic examination of brain biopsies (Purchase, 1945), which is impractical for field use; and (c) subinoculation of blood into susceptible animals (Uilenberg, 1983), which is cumbersome and expensive. Improved methods for diagnosis of heartwater by utilization of nucleic-acid-based assays are being investigated (Waghela *et al.*, 1991; Mahan *et al.*, 1992; Yunker *et al.*, 1993). These have been complemented by the development of serological assays (Du Plessis *et al.*, 1987; Jongejan *et al.*, 1991; Mahan *et al.*, 1993; Semu

† **Present address:** University of Florida/USAID/SADC Heartwater Research Project, PO Box CY 551, Causeway, Zimbabwe.

The GenBank accession number for the sequence reported in this paper is L07385 (mnemonic: COW IMM DOM).

et al., 1992), but these assays suffer from poor sensitivity and specificity (Logan *et al.*, 1986; Du Plessis *et al.*, 1987; Jongejan *et al.*, 1991; de Vries *et al.*, 1993; Mahan *et al.*, 1993). Although diagnosis of heartwater by nucleic-acid-based assays would be specific, these assays may still lack the sensitivity that is required to detect carrier animals at all times (Mahan *et al.*, 1992). Hence, the development of a specific and sensitive serological diagnostic assay for heartwater remains an important goal.

Data on the molecular cloning of a *C. ruminantium* gene that encodes a 21 kDa immunogenic protein are presented here.

METHODS

Origin of *C. ruminantium* strains. Five strains of *C. ruminantium* from Zimbabwe were used: Crystal Springs, Nyatsanga, Lemco T3, Palm River and Zwimba (Byrom & Yunker, 1990); two from South Africa, Welgevonden (Du Plessis, 1985) and Ball 3 (Haig, 1952); one from Nigeria (Van Winkelhoff & Uilenberg, 1981); and the Gardel strain from Guadeloupe (Uilenberg *et al.*, 1985). *C. ruminantium* organisms were harvested from cell cultures following complete lysis of endothelial cells, and processed by the methods described below.

Preparation of antisera. Hyperimmune sheep serum was prepared from sheep 175, which had been previously infected with 2 ml of a near-terminal cell culture containing *C. ruminantium* (Crystal Springs strain). Subsequently, this sheep recovered and was resistant when rechallenged twice at 4 week intervals with 10 ml homologous strain infected sheep blood stabilate. At the time of rechallenge, a susceptible sheep was inoculated with the same stabilate to show that the challenge was virulent. This sheep exhibited clinical signs of heartwater (Uilenberg, 1983) and was positive on brain biopsy (Purchase, 1945). Two weeks after the last inoculation, sheep 175 was bled and serum was stored at -20°C . Antisera were also collected 2–6 weeks after a single or repeated homologous *C. ruminantium* infection of sheep (infected with Welgevonden, Crystal Springs or Mbizi strains), goats (infected with Crystal Springs, Mbizi, Plumtree or Towla strains), and cattle (infected with Ball 3, Palm River or Welgevonden strains). Antisera were also collected from cattle located in a heartwater endemic area, at the Mbizi Quarantine Station in the southern lowveld of Zimbabwe.

Antisera against the expressed protein product of the 627 bp ORF were prepared by inoculating sheep with 180 μg of recombinant protein mixed with Freund's complete adjuvant for the first inoculation followed by another four inoculations 3 weeks apart in incomplete Freund's adjuvant. Control sheep were inoculated similarly except that PBS was substituted for the recombinant protein. Two weeks after the final inoculation the sheep were bled and the sera prepared and stored at -20°C .

Analysis of *C. ruminantium* antigens by immunoblotting and immunoprecipitation. *C. ruminantium* organisms were harvested from terminally infected endothelial cell cultures (Bezuidenhout *et al.*, 1985; Byrom & Yunker, 1990). Culture supernatants were centrifuged at 400 *g* for 10 min to remove host cell debris, and organisms in the supernatant were washed twice in PBS by centrifugation at 30 000 *g* for 30 min and then resuspended in a small volume of PBS. *C. ruminantium* proteins (20–60 μg per lane) were separated by SDS-PAGE on a 12% (w/v) gel and electroblotted to nitrocellulose membranes for immunoblotting (Mahan *et al.*, 1993). Rainbow molecular mass standards (Amersham) were run on each gel for estimation of molecular size.

Immunoprecipitation of biosynthetic labelled proteins of *C. ruminantium* and of uninfected endothelial cell cultures was conducted using Protein G beads (Pharmacia), by the method of Hines *et al.* (1989). Each well of a 24-well plate was seeded with 10^6 cells of either infected (30% of cells infected) or uninfected cells in methionine-free Dulbecco's minimum essential media and 60 μCi (2.2 MBq) [^{35}S]methionine. The infected cultures were harvested after 48 h, pooled, and the organisms or endothelial cells were then lysed and prepared for immunoprecipitation (Hines *et al.*, 1989). The immunoprecipitated antigens were detected by 12% (w/v) SDS-PAGE followed by fluorography.

Preparation of *C. ruminantium* DNA. *C. ruminantium* organisms were harvested and resuspended in saline EDTA (0.15 M NaCl, 0.1 M Na_2EDTA), frozen and thawed twice, and digested with lysozyme at 5 mg ml^{-1} , for 30 min at 37°C . Thereafter, the DNA was treated with RNase A (100 $\mu\text{g ml}^{-1}$) at 37°C for 1 h; followed by proteinase K (100 $\mu\text{g ml}^{-1}$) and SDS at a final concentration of 2% (w/v), for 1 h at 37°C . The DNA was extracted once with phenol, phenol/chloroform, chloroform and precipitated by ethanol. The precipitated DNA was washed once with 70% (v/v) ethanol, dried and resuspended in TE buffer (10 mM Tris pH 7.5, 0.1 mM Na_2EDTA).

Preparation of a *C. ruminantium* genomic DNA library. Genomic *C. ruminantium* DNA (Crystal Springs strain) was sheared by sonication, filled in (Sambrook *et al.*, 1989), and ligated to *EcoRI/NotI* linkers (Invitrogen). The linker-ligated DNA was phosphorylated (Sambrook *et al.*, 1989) and size-selected through S-400 spin columns (Pharmacia) to select for fragments larger than 600 bp. Ligation of the size-selected DNA to *EcoRI*-digested pUC13 was accomplished at 8°C in the presence of T4 DNA ligase at an insert:vector molar ratio of 2:1. The recombinant plasmid DNA was used to transform Max Efficiency *E. coli* DH5 α competent cells [GIBCO BRL, Genotype: F^{-} $\phi 80\text{d}lacZ\Delta\text{M15 } \Delta(lacZYA-argF)\text{U169 } deoR \text{ } recA1 \text{ } endA1 \text{ } hsdR17 (\text{r}_K^-, \text{m}_K^+) \text{ } supE44 \text{ } \lambda^- \text{ } thi-1 \text{ } gyrA96 \text{ } relA1$], to ampicillin resistance in the presence of X-Gal and IPTG (Sambrook *et al.*, 1989).

Screening of the genomic DNA library. *E. coli* colony blots on nitrocellulose filters were lysed in chloroform vapour for 15 min and washed 3 times in TBS (20 mM Tris, 500 mM NaCl pH 7.5). Bacterial debris was removed by rolling the filters face down on Whatman chromatography paper with a pipette. The filters were blocked in 1% (w/v) gelatin in TBS for 1 h on a rocking platform and reacted overnight at room temperature with sheep 175 hyperimmune serum diluted 1:10 000 in TBS, 1% (w/v) gelatin. The filters were washed 3 times, 15 min each wash in TBS, 1% (w/v) gelatin and 0.05% (v/v) Tween 20, and then reacted with ^{125}I -labelled protein G (2.5 μCi) in TBS, 1% (w/v) gelatin, 0.05% (v/v) Tween 20 for 1–2 h at room temperature. After another three washes (15 min each wash) in TBS, 1% gelatin, 0.25% (v/v) Tween 20, the filters were dried and the signal detected by autoradiography. All screening assays were conducted in duplicate.

DNA hybridization. *C. ruminantium* genomic DNA samples from different strains were blotted onto nylon membranes (Genescreen Plus, Dupont), in amounts of 100, 10, and 1 ng using the Hybridot blot apparatus (BRL). These blots were probed with the 2125 bp *EcoRI* fragment of pF5.2 DNA isolated from a 1% low-melting-temperature agarose gel (Sambrook *et al.*, 1989) and labelled by the random primer extension method (Boehringer Mannheim) using [^{32}P]dCTP (Mahan *et al.*, 1992).

As a specificity control, a blot of genomic DNA from *C. ruminantium*, *Anaplasma marginale*, *Babesia bigemina*, *Babesia bovis*,

Fig. 1. (a) Immunogenic proteins of *C. ruminantium* (Crystal Springs) revealed by immunoblotting. Lanes 1–8 represent sheep 175 hyperimmune antiserum reactions at dilutions of 1/100, 1/500, 1/1000, 1/2000, 1/4000, 1/8000, 1/10000 and 1/12000 respectively. The position of molecular mass markers is shown on the left (kDa). (b) Immunogenic proteins of *C. ruminantium* revealed by immunoprecipitation. [35 S]methionine-labelled *C. ruminantium* or uninfected bovine endothelial cell proteins were reacted with pre- or post-immune sheep 175 serum. Lanes 1 and 2, pre-immune serum (1/10 dilution) with uninfected bovine endothelial cell antigen and *C. ruminantium* antigen respectively. Lane 3, post-immune serum (1/100 dilution) with uninfected bovine cell antigen. Lanes 4, 5 and 6, post-immune serum (1/10, 1/100 and 1/1000 dilutions respectively) with *C. ruminantium* antigen. The position of molecular mass markers is shown on the left (kDa).

Trypanosoma brucei parasites, and bovine, caprine and ovine DNA from peripheral blood leucocytes was also probed with the 32 P-labelled 2125 bp *Eco*RI fragment. The blot contained 10, 1 and 0.1 ng of *C. ruminantium* DNA (Crystal Springs strain) and 400, 40 and 4 ng of the other DNA samples.

DNA sequencing. The DNA insert in pF5.2 was sequenced as double-stranded DNA using Sequenase (United States Biochemical), as recommended by the manufacturer. Plasmid-specific primers were used in initial reactions and then new oligonucleotides were synthesized, based on the sequences obtained ('primer walking'). Insert DNA was completely sequenced on both strands. The DNA sequence obtained was analysed using PCGENE (Intelligenetics) and the GCG (University of Wisconsin) programs. Database searches were performed through the Interdisciplinary Center for Biotechnology Research (ICBR), University of Florida at the National Center for Biotechnology Information, Bethesda, MD, USA, using the GENINFO(R) Experimental BLAST Network Service and the algorithm described by Altschul *et al.* (1990).

Subcloning of the 627 bp ORF into pFLAG and purification of the expressed recombinant protein. Primers incorporating either *Hind*III (5' primer, CAGGGAAGCTTGCAATTTT-TTGGGATATTC) or *Sal*I (3' primer, CAGCGGTCGACT-TCTATATAAAAGTCTGTAC) restriction sites were prepared from the sequence of the 627 bp ORF and used to amplify this ORF from *Cowdria* genomic DNA of the Crystal Springs strain. The PCR-amplified DNA was directionally subcloned into the expression vector pFLAG (International Biotechnologies) doubly digested with *Hind*III and *Sal*I. Transformed *E. coli* colonies were analysed for expression of recombinant protein by 7.5–15% (w/v) SDS-PAGE and silver staining. An *E. coli* colony expressing recombinant protein (no. 10) was

selected and expanded in 500 ml cultures of L-broth induced with 0.1 mM IPTG and the protein purified by monoclonal antibody affinity chromatography, as described by the manufacturer of the pFLAG vector system.

ELISA. ELISA plates were coated with 50 μ l per well of recombinant *C. ruminantium* protein (4 μ g ml $^{-1}$) in 0.14 M NaCl, 0.02 M Na $_2$ HPO $_4$, 0.003 M KH $_2$ PO $_4$, pH 7.2 containing 0.02% (w/v) sodium azide (PBS/azide) and incubated overnight at 4 °C. After five washes with buffer containing PBS/azide and 0.05% (v/v) Tween 20, the wells were blocked for 60 min at room temperature with 1% (w/v) bovine serum albumin in PBS/azide. The plates were washed five times and incubated for 60 min with 50 μ l per well preimmune or immune serum diluted in PBS/azide. The plates were washed (five times) and incubated at room temperature for a further 60 min in the presence of rabbit anti-sheep IgG conjugated to alkaline phosphatase (1:1000; added at 50 μ l per well). After another five washes, the substrate, *p*-nitrophenylphosphate (1 mg ml $^{-1}$; Sigma), in 0.16 M NaHCO $_3$, 0.14 M Na $_2$ CO $_3$, 0.02 M MgCl $_2$, pH 9.6, was added at 50 μ l per well to each well and incubated for 30 min at room temperature. Absorbance was read at 405 nm.

RESULTS

Identification of immunogenic proteins of *C. ruminantium*

An antiserum prepared from hyperimmune sheep 175 recognized several major proteins of *C. ruminantium* in immunoblots and immunoprecipitation including 21, 32, 40, 46, 58, 85 and 160 kDa proteins (Fig. 1a, b). In the immunoblotting assay, a 30–32 kDa protein was most

```

1  AAATAATATAGAACTACAGCACAGGTAATTAGTAATCTTATATCACAAAGCTTGCCAATATATAACTTCCGTAAAAATA 80
81  TCAAAATATCAGATCCAACCTTGATATGATACTATCTAATCTATAGTTTGTATAACTTATAAACTCTTACAAAAATACCAA 160
161  TAGGTAACAATAAATTAATTTAAAGCTTTAATGCCCTATATAATATAAAAGCAAGTAAAGTAATTTTATAGTTATTCATT 240
241  AATACATTATGCAATTTCAAACTTACCATTATGTTACATACCTATATTAATTATTAAGTTTAGTAATTAATTATATGAA 320
321  TTATAATTACTATATAAATATGTAATATTAACTTAATAAGTTCTTTCATATTACATAAATATTGATAGATAAAAAA 400
401  AATTTAAGATTATTACGTAGTAAAGCATCTATAAACATACCTTAAACAAAGTGAAGTATCTCAAGTATCATACCTATC 480
481  TAACCTCTACATAAATATCAACCAATATATTAACAAATTCATTACAAAATCCTATAATGCAATATATTAACAACATGCA 560
561  TTTAAGAAAAAGGTATAAGCTATCTCATAGTGATTATTAGTATTACAATTATAAATAAATACATAGTACAAATTAAC 640
641  TTTCAAGAAATTCATACAAAGTACCAACATCTATATTAATAATTACTTGAAGAAAAATTTATAAGTACATACCTATTCT 720
721  ACATAACATTAGTATAAAAAACATCAATGCATATTCAGAAATCTACTACAAAATCAAAATACATAAATAAGACAATTA 800
801  AAACATATTATCTACTAATATATAAATAAAGCTATTTAATTAATTTATAAATAACGAACTCCTCTACAGCATTTTAA 880
881  GAATAATACCTACTACACTAGGAATTATTTACTTGCTATATTTTATGCTAGTCTGTAATAATACATATACAAATAGATT 960
961  CATAACTCAACTCTGAATGTAACTACTACACTAAAAATCATTTTTTATAGCGTAAGTATATACAATTTTACTACA 1040
1041  TAACATAAAGAAAAACATAGAAGTTTAAATTTTATTTAACTAAACAAATACTAACCAAAATATAGAAATTTTAGATTTCAC 1120
1121  AATTTGCATATTTAAATACTACAAAGATAAAACATACTATAAAATTTTGAAGTACATACCAACATCTGTTTAT 1200
1201  GCATATATGATAGTGTGTTTTTATAAAAACTCTAATTTTATCAATGGAGCACATCATGAAGCTTCAAGTTTATACTT 1280
                                     M E H I M K A I K F I L
1281  AATCTATGTTTACTATTTGCAGCAATTTTTTGGGATATTCTTACATAACAAAACAGGTATATTCCAACCAAATTACA 1360
N L C L L F A I F L G Y S Y I T K Q G I F P K L H
1361  CGACTCTCCTGATGTTAATATATCGAACAAAGCGGATATAAATACTAGCTTTAGCTTAATTAATCAGGATGGTATTACGA 1440
D S P D V N I S N K A D I N T S F S L I N Q D G I T I
1441  TATCTAGTAAAGACTTCCTTGGAAAACATATGTTAGTCCTTTTGGGTTTTCTTCTGTAAAACATTTGCCCCATGGAA 1520
S S K D F L G K H M L V L F G F S S C K T I C P M E
1521  CTAGGGTTAGCATCCCAATCTAGATCAACTTGGCAACGAATCTGACAAGTTACAAGTAGTCTTTTAACTATTGATCC 1600
L G L A S T I L D Q L N E S D K L Q V V F I T I D P
1601  AACAAAGATACTGTAGAACTAAAGAGTTTCACAAAATTTTGACTCAGGATTCAAATGTTAAACGAAACATTG 1680
T K D T V E T L K E F H K N F D S R I Q M L T G N I E
1681  AAGCTTATTAATCAAAAGTACAAAGGTACAAAGTATGTAGGTACGCCAGACATGATAACCAATTAACCATTTCTGGA 1760
A I N Q I V Q G Y K V Y V G Q P D N D N Q I N H S G
1761  ATAATGTATGTAGACAAGAAAGGAGAATATTTAACCAATTTTGTACCAGATTTAAAGTCAAAAGCGCTCAAGTGGA 1840
I M Y I V T D K K G E Y L T H F V P D L K S K E P Q V D
1841  TAAATTAATTTCTTTAATTAAGCAGTATCTTAAATTAATACTTAATTAATAGAAATAGTACAGACTTTTATATAGAATC 1920
K L L S L I K Q Y L *
1921  TAACCTTTAGGATATATCTAATGAAGAAGTTTCTTTTAGTCTTATTTTTACTGCTTGTAAATGTTACCCAAAGATTCTA 2000
M K K F L L V L F L L L V M L P K D S N
2001  ATGCGGAACATATACATGTTGTTGGATCATCTACAGCATTTCCATTTATCGCAGCAATAGCAGAAGAAATTTGGGAGGTTT 2080
A E H I H V V G S S T A F P F I A A I A E E F G R F
2081  TCAGATTATGGAACACCTATAATAGAGTCTGTGGGAGTGGTATGGGTTTTAGTATGTTTTGTCAAAGTGTAGAAAACAG 2160
S D Y G T P I I E S V G S G M G F S M F C Q S V E N S
2161  TACGCCAGATATAGCTATGTCATCTCGTAAGATAAAGGATGCAGAGGTAGAATTATGTAAGGTAATGACGTTTCATGACA 2240
T P D I A M S S R K I K D A E V E L C K S N D V H D I
2241  TTATTGAAATCATTAGGATATGATGGTATTGTTATTGCAAACTCTAACAAATAGCAATAAGCTTGATTTTACAAAAA 2320
I E I I I G Y D G I V I A N S N N S N K L D F T K K
2321  GATCTATTCAAAGCTTTAAGCAAGTATGCAACGTGAGAAGATATACATAGTATACCAGTAAATGATTTTAAGTATTG 2400
D L F K A L S K Y A T S E E Y T H S I P V N D F K Y W
2401  GTCAGAAATTAATAAGGTTCCCAATATTGATATTGAAGTTTACGGACCATACAAAAACACAGGTACTTATAATATAC 2480
S E I N N R F P N I D I E V Y G P Y K N T G T Y N I L
2481  TAATCGAAGAAATATGCAGGATCTTGTATGAATCATAAAAATTTTATTGAAGTATACCCAGACTTAAAAAAGACAA 2560
I E I M Q D S C M N H K N F F I E V Y P D L K K R Q
2561  CACGCATGCAGTATGATCCGCAATGATGGCAAGTACATTGAAGTTGCAGCTAATGAAAACATTATTATACAAAAAATTGC 2640
H A C S M I R N D G K Y I E V A A N E N I I I Q K I A
2641  AAAAAAATGCTGCTTTTGGTATTTTATGCTTTTAAATACAGAAATCAAGATAAAATACATGGAATAAAATG 2720
K N N A A F G I F S F S F L I Q N Q D K I H G N K I A
2721  CAGGTGTGGAACCTACATATGAACTATTTCTCTGAAAAATATATTTATCA 2773
G V E P T Y E T I S S G K Y I L S

```

Fig. 2. DNA sequence and predicted protein sequence of pF5.2 plasmid insert DNA. The underlined areas show recognition sites for *EcoRI* (648 bp) and *HincII* (1666 bp). ORF1 of 627 bp (1245 bp to 1871 bp) and ORF2 of 831 bp (from 1943 bp to 2773 bp) are shown. The vertical bold bars indicate predicted sites of cleavage after the N-terminal signal peptides of ORF1 or ORF2.

immunodominant, reacting with antiserum diluted down to 1/12000. The relative recognition of this protein compared to other *C. ruminantium* proteins (in particular the 58 kDa) was assay-dependent and not as evident using immunoprecipitation of *C. ruminantium* proteins metabolically labelled with [³⁵S]methionine (Fig. 1b).

Screening of *C. ruminantium* genomic DNA library

A genomic library of *C. ruminantium* DNA from the Crystal Springs (Zimbabwe) strain was screened with sheep 175 hyperimmune antiserum. Screening of approximately 3000 recombinant *E. coli* colonies resulted in the identification of a positive colony, F5.2, which reacted

with the serum from hyperimmune sheep 175 but not with preinfection serum, or with ¹²⁵I-protein G alone. Isolated individual F5.2 colonies also reacted with hyperimmune serum but not with preinfection serum. Furthermore, F5.2 expressed the protein in the absence of induction by IPTG.

Characterization of the DNA insert of pF5.2

The DNA insert in plasmid pF5.2 was sequenced and determined to be 2773 bp (Fig. 2), and of high overall A+T content (74%). The sequence contained two long open reading frames (ORFs): one complete, of 627 bp (ORF 1), and the other incomplete, of 831 bp

C. ruminantium: 5 MKAIKFILNLCLLFAAIFLGYSYITKQIF 34
M+ K + NL L AA+FLGYSY+ K+GIF
A. marginale: 1 MRIFKIVSNLLLFVAAVFLGYSYVKKGIF 30

C. ruminantium: 57 SLINQDGITISSKDFLGKHMVLVFGFSSCKTICPMELGLASTILDQLGNESDKLQVVFIT 116
+L+N +G+T+SS DF GKHMLV+FGFS+CK CP ELG+AS +L +LG+ +DKLQVVFIT
A. marginale: 54 NLVNHEGVTVSSGDFGGKHMVLVFGFSACKYTCPTELGMASQLLSKLGHDHAKLQVVFIT 113

C. ruminantium: 117 IDPTKDTVETLKEFHKNFDSRIQMLTGNIEAINQIVQGYKVVVG 160
+DP DTV LKE+HK+FD+RIQMLTG I +V+ YKVYVG
A. marginale: 114 VDPKNDTVAKLKEYHKSFDARIQMLTGEEADIKSVVENYKVVVG 157

C. ruminantium: 164 NDNQINHSIMYIVDKKGEYLTHFVPLKSKEPQVDKLLSLIKQYL 209
+D I+HS MY+++ KG Y+ HF PD + E Q ++L + ++
A. marginale: 162 SDGDIDHSTFMYLINGKGRYVGHFAPDFNASEGQGEELFKFVSGHM 207

Fig. 3. Similarity between the amino-acid sequence of *C. ruminantium* ORF1 and that of MSP-5, the conserved outer-membrane protein of *A. marginale*. The centre line reveals regions of identity (letter) or conservative substitutions (+).

Fig. 4. Subcloning and expression of bp 1302–1918 (Fig. 3) in pFLAG/*E. coli*: 7.5–17.5% SDS-PAGE analysis and silver staining of total proteins from individual *E. coli* colonies, induced with 0.1 mM IPTG. Lane 1, molecular mass standards (kDa); lanes 2–5, recombinant *E. coli* colonies 4, 5, 7 and 10 respectively; lane 6, *E. coli* containing non-recombinant pFLAG plasmid vector, overloaded to show lack of expression of 23 kDa recombinant polypeptide.

Fig. 5. Immunoblot reactions of sheep 177 serum prepared against the purified recombinant 23 kDa protein with various strains of *C. ruminantium*. The immunoblot shows that the post-immunization sheep serum (1:5000 dilution; a) recognized a 21 kDa protein of *C. ruminantium* which was conserved in antigenicity and molecular size between various strains of *C. ruminantium*. There was no reaction with the pre-immunization serum (b). CS, Crystal Springs; WV, Welgevonden; B, Ball 3; L, Lemco; G, Gardel; N, Nigeria; Z, Zwimba.

(ORF 2). ORF 1 potentially encoded a 23.6 kDa protein with a predicted (von Heijne, 1986) N-terminal signal peptide. With the signal peptide removed (probable cleavage site is between amino acid residues 19 and 20),

translation of a mature 21.4 kDa protein is predicted. ORF 2 was interrupted by vector sequence at the 3' end and also contained a predicted N-terminal signal peptide with a probable cleavage site between amino acid residues 21 and 22. This ORF potentially encoded, therefore, an incomplete polypeptide of 31.2 kDa or 28.8 kDa with signal peptide removed. Deletion and subcloning experiments indicated that ORF 1 encoded the protein recognized by sheep 175 serum (data not shown). Database searches using the 'BLAST' program (Altschul *et al.*, 1990) revealed significant similarity between the protein encoded by ORF 1 and the MSP-5 outer-membrane protein of *A. marginale* (Fig. 3). The GAP Program (GCG Sequence Analysis Software, University of Wisconsin) revealed

Fig. 6. (a) Immunoblot reactions of sera from heartwater-infected sheep, cattle and goats with Crystal Springs strain *C. ruminantium* antigen. Antisera were taken after single or multiple infections of sheep, with Welgevonden (strips 4–17), Mbizi (strip 18), and Crystal Springs (strips 19–22); of cattle with Palm River (strips 23–29), Ball 3 (strips 30–33) and Welgevonden (strips 34 and 35); and of goats with Mbizi (strip 36), Plumtree (strip 37), Crystal Springs (strip 38) and Towla (strip 39) strains of *C. ruminantium*. The 21 kDa protein is consistently recognized by all sera (arrow) although with varying intensity. Molecular mass markers are shown to the left; negative control sheep serum is shown as lane 1, sheep 177 serum (see Fig. 5) is included as a positive control for recognition of the 21 kDa protein (lane 2), and lane 3 contains the reaction of sheep 175 serum. (b) Immunoblot reactions of sera from cattle (strips 4–21) from heartwater endemic area of Zimbabwe with the Crystal Springs strain *C. ruminantium* antigen. These cattle sera consistently recognize the 21 kDa protein (arrow) although the intensity of the reaction varies. Molecular mass markers are shown to the left; negative control sheep serum is shown as lane 1, sheep 177 serum (see Fig. 5) is included as a positive control for recognition of the 21 kDa protein (lane 2), and lane 3 contains the reaction of sheep 175 serum.

55.5% identity between the ORF 1 and the coding DNA sequence for MSP-5 (data not shown). No significant similarities were found in DNA or protein sequence databases to ORF 2.

Confirming the specificity of the DNA insert in pF5.2, a 2125 bp *EcoRI* fragment encompassing both ORFs (*EcoRI* site at position 648 within the 2773 bp insert; Fig. 2) hybridized at high stringency to DNA from all different heartwater isolates tested but not to DNA from *A. marginale*, *B. bigemina*, *B. bovis*, *T. brucei*, or cattle, sheep or goats (data not shown).

The 627 bp ORF encodes the immunogenic 21 kDa protein of *C. ruminantium*

Immune recognition of the polypeptide product of ORF 1 was confirmed by amplifying this ORF by PCR and subcloning into pFLAG for high-level expression of its encoded polypeptide. Individual recombinant *E. coli* colonies were analysed by SDS-PAGE and the expression of a novel 23 kDa protein was detected in colonies 4, 5 and 10 after IPTG induction, but not in colony 7 or in

control colony-containing nonrecombinant pFLAG vector (Fig. 4). The molecular size of the recombinant protein was greater than that predicted by the ORF 1 sequence because of the fusion of nine amino acids encoded by the pFLAG vector sequences. The junction regions of insert/plasmid DNA from colonies 4, 5 and 7 and the complete insert DNA in colony 10 was sequenced. These data showed that DNA was correctly inserted for expression in colonies 4, 5 and 10 and incorrectly inserted in colony 7. Furthermore, the inserted DNA sequence in colony 10 agreed with that of the ORF 1 (Fig. 2), indicating no amplification errors.

Subsequently, recombinant protein was purified from colony 10 after IPTG induction, employing the calcium-dependent binding of monoclonal antibody M1 specific for a four-amino-acid portion of the nine-amino-acid N-terminal fusion peptide. Sera of sheep immunized with the recombinant protein specifically recognized a 21 kDa native protein of seven *C. ruminantium* strains in immunoblots (Fig. 5). Similarly, the recombinant protein was specifically recognized by post-immune serum from sheep

175 and by other sera from heartwater immune cattle and sheep in an ELISA and immunoblotting assay (data not shown). These reactions demonstrate the conservation of molecular size and antigenicity of the 21 kDa protein, and confirm the assignment of the 627 bp ORF 1 to the immunogenic 21 kDa protein of *C. ruminantium*.

Recognition of the 21 kDa protein of *C. ruminantium* by sera from infected animals and animals from a heartwater endemic area

Sera from sheep, goats and cattle infected with different strains of *C. ruminantium*, and sera from cattle in a heartwater endemic area of Zimbabwe, consistently recognized the 21 kDa protein (although with varying intensity; Fig. 6a, b) by immunoblotting. These reactions justify further investigation of this recombinant protein for development of serological diagnostic assays for heartwater.

DISCUSSION

The cloned *C. ruminantium* DNA insert in plasmid pF5.2 had a high (74%) overall A + T content, which was also observed in two previously sequenced segments of *C. ruminantium* DNA (70% A + T: Waghela *et al.*, 1991). The ORFs identified in this study were enriched in G + C compared to the remainder of the insert DNA, which may be one useful criterion for identification of *C. ruminantium* genes. Expression of the recombinant 21 kDa protein in F5.2 *E. coli* was independent of orientation in the pUC13 vector and not inducible by IPTG, suggesting that *C. ruminantium* regulatory sequences were functioning in expression of this gene. This is not unexpected given the close taxonomic relationship between this prokaryote and *E. coli* (Dame *et al.*, 1992).

Database searches (Altschul *et al.*, 1990) revealed significant similarity between the 21 kDa *C. ruminantium* protein and MSP-5, the conserved 19 kDa outer-membrane protein of *A. marginale* (Fig. 3). MSP-5 has been proposed as an antigen suitable for serological diagnosis of bovine anaplasmosis since it was consistently recognized by *A. marginale*-infected bovine sera (Visser *et al.*, 1992). *A. marginale* and *C. ruminantium* are closely related rickettsiae, based on 16S ribosomal RNA analysis (Dame *et al.*, 1992), and it may be possible to use the equivalent 21 kDa polypeptide of *C. ruminantium* for diagnosis of heartwater.

The 21 kDa *C. ruminantium* polypeptide is conserved in size and antigenicity in seven heartwater strains from Africa and the Caribbean, and the native and recombinant 21 kDa protein are consistently recognized by sera from infected livestock and from cattle from a heartwater endemic area. We propose further characterization of this protein for use in serological diagnosis (and perhaps vaccines, once its location on the organisms is established). In the light of the close phylogenetic relationship of *C. ruminantium*, *A. marginale* and *Ehrlichia* sp. (Dame *et al.*, 1992; van Vliet *et al.*, 1992), and the likelihood of cross-reactions between immunogenic proteins of these

organisms (Mahan *et al.*, 1993), development of a 21 kDa protein specific serological assay for heartwater diagnosis will require identification of epitopes unique to *C. ruminantium*. Together with the use of diagnostic nucleic acid probes (Mahan *et al.*, 1992; Waghela *et al.*, 1991), these molecular approaches should significantly improve our understanding of the epidemiology and pathogenesis of heartwater and lead to its effective control.

ACKNOWLEDGEMENTS

We acknowledge the excellent technical assistance of Anna Lundgren, Annie Moreland, Sondra Kamper and Teresa Harkins and synthesis of oligonucleotides by core facilities of the Interdisciplinary Center for Biotechnology Research, University of Florida. This work received support from the United States Agency of International Development (USAID) Co-operative Agreement no. AFR-0435-A-00-9084-00, USAID grants DAN 1328-6-SS-4093-00 and DPE-5542-6-SS-7073-00 and the United States Department of Agriculture/CRGO no. 91-34135-6178.

REFERENCES

- Altschul, S. F., Gish, W., Miller, W., Myers, E. & Lipman, D. J. (1990). A basic local alignment search tool. *J Mol Biol* **215**, 403–410.
- Andrew, H. R. & Norval, R. A. I. (1989). The carrier status of sheep, cattle and African buffalo recovered from heartwater. *Vet Parasitol* **34**, 261–266.
- Barre, N., Uilenberg, G., Morel, P. C. & Camus, E. (1987). Danger of introducing heartwater on to the American mainland: potential role of indigenous and exotic *Amblyomma* ticks. *Onderstepoort J Vet Res* **54**, 405–417.
- Bezuidenhout, J. D., Paterson, C. L. & Barnard, B. J. M. (1985). In-vitro cultivation of *Cowdria ruminantium*. *Onderstepoort J Vet Res* **52**, 113–120.
- Burridge, M. J. (1985). Heartwater invades the Caribbean. *Parasitol Today* **1**, 1–3.
- Byrom, B. & Yunker, C. E. (1990). Improved culture conditions for *Cowdria ruminantium* (Rickettsiales), the agent of heartwater disease of domestic ruminants. *Cytotechnology* **4**, 285–290.
- Dame, J. B., Mahan, S. M. & Yowell, C. A. (1992). Phylogenetic relationship of *Cowdria ruminantium*, agent of heartwater, to *Anaplasma marginale* and other members of the order Rickettsiales determined on the basis of 16S rRNA sequence. *Int J Syst Bacteriol* **42**, 270–274.
- Du Plessis, J. L. (1985). A method for determining the *Cowdria ruminantium* infection rate of *Amblyomma hebraeum*: effects in mice injected with tick homogenates. *Onderstepoort J Vet Res* **52**, 55–61.
- Du Plessis, J. L., Camus, E., Oberem, P. T. & Malan, L. (1987). Heartwater serology: some problems with the interpretation of results. *Onderstepoort J Vet Res* **54**, 327–329.
- Haig, D. A. (1952). Note on the use of the white mouse for the transport of strains of heartwater. *J South African Vet Med Assoc* **23**, 167–170.
- von Heijne, G. (1986). A new method for predicting signal peptide cleavage sites. *Nucleic Acids Res* **14**, 4683–4690.
- Hines, S. A., McElwain, T. F., Buening, G. M. & Palmer, G. H. (1989). Molecular characterization of *Babesia bovis* merozoite surface proteins bearing epitopes immunodominant in protected cattle. *Mol Biochem Parasitol* **37**, 1–10.
- Jongejan, F., Thielemans, M. J. C., de Groot, M., van Kooten, J. S. & van der Zeyst, B. A. M. (1991). Competitive enzyme-linked

immunosorbent assay for heartwater using monoclonal antibodies to a *Cowdria ruminantium* specific 32 kilodalton protein. *Vet Microbiol* **28**, 199–211.

Logan, L. L., Holland, C. J., Mebus, C. A. & Ristic, M. (1986). Serological relationship between *Cowdria ruminantium* and certain *Ehrlichia* spp. *Vet Rec* **119**, 458–459.

Mahan, S. M., Waghela, S. D., McGuire, T. C., Rurangirwa, F. R., Wassink, L. A. & Barbet, A. F. (1992). A cloned DNA probe for *Cowdria ruminantium* hybridizes with eight heartwater strains and detects infected sheep. *J Clin Microbiol* **30**, 981–986.

Mahan, S. M., Tebele, N., Mukwedey, D., Semu, S., Nyathi, C. B., Wassink, L. A., Kelly, P. J., Peter, T. & Barbet, A. F. (1993). An immunoblotting diagnostic assay for heartwater based on the immunodominant 32 kDa protein of *Cowdria ruminantium* detects false positives in field sera. *J Clin Microbiol* **31**, 2729–2737.

Perreau, P., Morel, P. C., Barre, N. & Durrand, P. (1980). Existence de la cowdriose (heartwater) a *Cowdria ruminantium* chez les ruminants des Antilles françaises (La Guadeloupe) et des Mascareignes (La Réunion et Ile Maurice). *Rev Elev Med Vet Pays Trop* **33**, 21–22.

Purchase, H. S. (1945). A simple and rapid method for demonstrating *Rickettsia ruminantium* (Cowdry, 1925) in heartwater brains. *Vet Rec* **57**, 413–414.

Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989). *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.

Semu, S., Mahan, S. M., Yunker, C. E. & Burridge, M. J. (1992). Development and persistence of *Cowdria ruminantium* specific antibodies following experimental infection of cattle, as detected by the indirect fluorescent antibody test. *Vet Immunol Immunopathol* **33**, 339–352.

Uilenberg, G. (1983). Heartwater (*Cowdria ruminantium* Infection): Current Status. *Adv Vet Sci Comp Med* **27**, 427–480.

Uilenberg, G. (1990). Present and future possibilities for the control of cowdriosis and anaplasmosis. *Vet Q* **12**, 39–45.

Uilenberg, G., Camus, E. & Barre, N. (1985). A strain of *C. ruminantium* isolated in Guadeloupe (French West Indies). *Rev Elev Med Vet Pays Trop* **38**, 34–42.

Visser, E. S., McGuire, T. C., Palmer, G. H., Davis, W. C., Shkap, V., Pipano, E. & Knowles, D. (1992). The *Anaplasma marginale* MSP-5 gene encodes a 19-kilodalton protein conserved in all recognized *Anaplasma* species. *Infect Immun* **60**, 5139–5144.

van Vliet, A. H. M., Jongejan, F. & van der Zeijst, B. A. M. (1992). Phylogenetic position of *Cowdria ruminantium* (Rickettsiales) determined by analysis of amplified 16S Ribosomal DNA sequences. *Int J Syst Bacteriol* **42**, 494–498.

de Vries, N., Mahan, S. M., Ushewokunze-Obatolu, Norval, R. A. I. & Jongejan, F. (1993). Correlation between antibodies to *Cowdria ruminantium* (Rickettsiales) in cattle and the distribution of *Amblyomma* vector ticks in Zimbabwe. *Exp Appl Acarol* **17**, 799–810.

Van Winkelhoff, A. J. & Uilenberg, G. (1981). Heartwater: cross-immunity studies with strains of *Cowdria ruminantium* isolated in West and South Africa. *Trop Anim Health Prod* **13**, 160–164.

Waghela, S. D., Rurangirwa, F. R., Mahan, S. M., Yunker, C. E., Crawford, T. B., Barbet, A. F., Burridge, M. J. & McGuire, T. C. (1991). A cloned DNA probe identifies *Cowdria ruminantium* in *Amblyomma variegatum* ticks. *J Clin Microbiol* **29**, 2571–2577.

Yunker, C. E., Mahan, S. M., Waghela, S. D., McGuire, T. C., Rurangirwa, F. R., Barbet, A. F. & Wassink, L. A. (1993). Detection of *Cowdria ruminantium* by means of a DNA probe, pCS20 in infected bont ticks, *Amblyomma hebraeum* the major vector of heartwater in Southern Africa. *Epidemiol Infect* **110**, 95–104.

Received 15 December 1993; revised 9 March 1994; accepted 23 March 1994.