JOURNAL OF CLINICAL MICROBIOLOGY, Jan. 2000, p. 369–374 Vol. 38, No. 1
0095-1137/00/$04.00+0
Copyright © 2000, American Society for Microbiology. All Rights Reserved.

Molecular Cloning and Characterization of the 120-Kilodalton Protein Gene of *Ehrlichia canis* and Application of the Recombinant 120-Kilodalton Protein for Serodiagnosis of Canine Ehrlichiosis

XUE-JIE YU, JERE W. MCBRIDE, C. MARCELA DIAZ, AND DAVID H. WALKER*

Department of Pathology, WHO Collaborating Center for Tropical Diseases, University of Texas Medical Branch, Galveston, Texas 77555-0609

Received 11 December 1998/Returned for modification 2 April 1999/Accepted 24 August 1999

The 120-kDa outer membrane protein (p120) is a potential adhesin of *Ehrlichia chaffeensis*, and recombinant p120 is very useful for serodiagnosis of human monocytotropic ehrlichiosis. The analogous gene of p120 in *Ehrlichia canis* was cloned, sequenced, and expressed. Like the *E. chaffeensis* p120, the *E. canis* p120 contains tandem repeat units. However, neither the repeat number nor the amino acid sequences in the repeats are identical in the two *Ehrlichia* species. The repeat units are hydrophilic and by probability analysis are predicted to be surface exposed in both species. The repeat regions of the p120s of the two species have common amino acid sequences that are predicted to be surface exposed. The overall amino acid sequence of the *E. canis* p120 is 30% homologous to that of *E. chaffeensis* p120. Protein immunoblotting demonstrated that the recombinant *E. canis* p120 reacted with convalescent sera from dogs with canine ehrlichiosis. These results indicate that the recombinant p120 is a potential antigen for the serodiagnosis of canine ehrlichiosis.

Ehrlichia spp. are obligate intracellular gram-negative bacteria which reside in the endosomes of hematopoietic cells and infect various animal hosts including humans, domestic and wild *Canidae*, deer, horses, sheep, cattle, and wild rodents. Each member of the tribe *Ehrlichieae* has its own particular target cell tropism. Most species of *Ehrlichia* are either monocytotropic (*E. canis*, *E. chaffeensis*, *E. sennetsu*, *E. risticii*, and *E. muris*) or granulocytotropic (human granulocytic ehrlichia [HGE], *E. equi*, *E. phagocytophila*, and *E. ewingii*) with the exceptions of *Cowdria ruminantium*, which grows in the endothelial cells of the host, and *Anaplasma marginale*, an erythrocyte parasite. Although ehrlichiae were described in the early part of this century, they were primarily considered pathogens of veterinary importance in the United States until this decade. Two new human *Ehrlichia* pathogens (*E. chaffeensis* and a human *E. phagocytophila*-like organism) were discovered in the United States (4, 6, 10, 17) recently. *E. canis*, the prototype species of the genus, is the etiologic agent of canine ehrlichiosis.

Canine ehrlichiosis is a worldwide disease transmitted by the brown dog tick, *Rhipicephalus sanguineus* (12, 16). *E. canis* causes a mild transient acute febrile illness, which may progress to severe illness and a fatal syndrome (tropical canine pancytopenia) (5, 11, 23). Each year millions of dollars are spent treating companion and working dogs infected with *E. canis* worldwide. Recently *E. canis*, or an antigenically indistinguishable organism, has been isolated from a human (19) and could be considered a public health threat. Understanding the genetic and antigenic composition of *E. canis* is essential for studying the pathogenesis of canine ehrlichiosis and developing an effective vaccine.

Previously we cloned and sequenced the *p120* gene of *E. chaffeensis* (25). Very recently we demonstrated the *E. chaffeensis* p120 to be an outer membrane protein that is preferentially expressed on the dense-core ultrastructural form of *E. chaffeensis* but not on the reticular cell (19a). The p120 appears to be an adhesin of *E. chaffeensis* because a noninvasive, nonadherent strain of *Escherichia coli* expressing the p120 acquired the ability to adhere to and enter cultured mammalian cells (Popov et al., submitted). The p120 is an immunodominant protein of *E. chaffeensis*, and it reacts with sera from most patients with monocytotropic ehrlichiosis (27). *E. canis* and *E. chaffeensis* are genetically and antigenically closely related species (2, 3, 7). The homologies between *E. canis* and *E. chaffeensis* are 98% for the 16S rRNA gene and 89% for the *nadA* gene (26). Since the p120 appears to be important in the attachment and serodiagnosis of *E. chaffeensis*, we hypothesized that an *E. chaffeensis* p120 analogue exists in *E. canis* and possesses similar biological functions. In this study we cloned, sequenced, expressed, and characterized the *p120* gene of *E. canis* and evaluated the recombinant p120 of *E. canis* for serodiagnosis of canine ehrlichiosis by Western blotting.

MATERIALS AND METHODS

Ehrlichia. *E. canis* Oklahoma was kindly provided by Jacqueline Dawson (Centers for Disease Control and Prevention, Atlanta, Ga.). *E. canis* Florida and three North Carolina isolates (Demon, DJ, and Jake) were kindly provided by Edward B. Breitschwerdt (College of Veterinary Medicine, North Carolina State University, Raleigh). *E. canis* Louisiana was kindly provided by R. E. Corstvet (Louisiana State University, Baton Rouge). Ehrlichiae were cultivated in DH82 cells, a canine macrophage-like cell line (9). DH82 cells were harvested with a cell scraper when 100% of the cells were infected with ehrlichiae. The cells were centrifuged at 17,400 × *g* for 20 min. The pellets were disrupted with a Braun-Sonic 2000 sonicator at 40 W for 30 s twice on ice. The cell lysate was loaded onto discontinuous gradients of 42 to 36 to 30% Renografin and then centrifuged at 80,000 × *g* for 60 min. Ehrlichiae in the heavy and light bands were collected (24) and washed by centrifugation with sucrose-phosphate-glutamate buffer (218 mM sucrose, 3.8 mM KH_2PO_4, 7.2 mM K_2HPO_4, 4.9 mM glutamate, pH 7.0).

DNA preparation. *E. canis* genomic DNA was prepared from Renografin density gradient-purified ehrlichiae by using an IsoQuick nucleic acid extraction kit according to the instructions of the manufacturer (ORCA Research Inc., Bothell, Wash.). The genomic DNA was used in Southern blotting and in PCR for detecting the *E. canis p120* gene.

* Corresponding author. Mailing address: Department of Pathology, 301 University Blvd., University of Texas Medical Branch, Galveston, TX 77555-0609. Phone: (409) 772-2856. Fax: (409) 772-2500. E-mail: dwalker@utmb.edu.

Southern blotting. The *E. chaffeensis p120* gene was amplified by PCR with primer pair pxcf3b (CAG CAA GAG CAA GAA GAT GAC) and pxar4 (ACA TAA CAT TCC ACT TTC AAA). The 1.2-kb PCR product lacked 138 nucleotides at the beginning, and 192 nucleotides at the end, of the structural gene of the *E. chaffeensis* p120. DNA was labeled during PCR by incorporating digoxigenin-dUTP with the PCR DIG probe synthesis kit (Roche Molecular Biochemicals, Indianapolis, Ind.) and was used as a probe to detect the homologous gene in *E. canis* by Southern blotting. DNA hybridization was performed at 42°C overnight with the Dig Easy hybridization buffer, and the digoxigenin-labeled DNA bound to the *E. canis* genomic DNA was detected with nitroblue tetrazolium and BCIP (5-bromo-4-chloro-3-indolylphosphate) by following the instructions of the manufacturer (Roche Molecular Biochemicals). The quality and quantity of the *E. canis* genomic DNA were monitored with a probe of the *E. canis p120* gene. The *E. canis* probe was amplified by PCR with primers 515f (GAA ATC CAT CAA GTG AAG TT) and 356r (TGA AGG CAT AGG ATT TAA TAA AGG) and labeled with digoxigenin. The *E. canis* probe spanned 1,620 nucleotides from 174 to 1795 in the structural gene of the *E. canis* p120.

PCR amplification of the *E. canis p120* gene. Primers were designed based on the DNA sequence of the *E. chaffeensis p120* gene (Fig. 2) (25). The *E. canis p120* gene was amplified by PCR with 30 cycles of 94°C for 30 s, 52°C for 1 min, and 72°C for 2 min. The PCR product was purified by using a QIAquick PCR purification kit (Qiagen Inc., Santa Clarita, Calif.) and was cloned into pCR2.1 TA cloning vector (Invitrogen, Carlsbad, Calif.). The resultant recombinant plasmid was designated pCR120.

DNA sequencing. DNA was sequenced with an ABI Prism 377 DNA sequencer (Perkin-Elmer Applied Biosystems, Foster City, Calif.). Both DNA strands of the *E. canis p120* gene were sequenced. The nonrepeat regions were sequenced by primer extension. The repeat region was sequenced by unidirectional deletion.

Unidirectional deletion of the *E. canis* gene of p120. The repeat region was deleted from the 5′ end of the *E. canis p120* gene by using restriction endonuclease *Spe*I partial digestion. Plasmid pCR120 was first completely digested with *Xba*I, which had a unique cleavage site on the plasmid sequence near the 5′ end of the *E. canis* gene of p120. Then pCR120 was partially digested with *Spe*I. *Spe*I had a unique cleavage site in each repeat of the *E. canis p120* gene but had no cutting site outside the repeat region, including the plasmid vector sequence. To ensure an appropriately representative partial digestion, an aliquot was removed from the digestion mixture every 5 min. The digestion was stopped by adding EDTA to a final concentration of 50 mM and by heating at 70°C for 10 min. After *Xba*I and *Spe*I digestion, various numbers of repeat units between *Xba*I and each *Spe*I cleavage site were removed (deleted) to generate deleted plasmid DNAs with noncompatible ends (*Xba*I at the 3′ end and *Spe*I at the 5′ end). The restriction enzyme-digested mixture was treated with Klenow fragment to fill in the ends. The restriction mixture was then separated by electrophoresis on a 1% agarose gel to remove the plasmids from the internal repeats because their molecular sizes differed significantly. The mixture of the deleted plasmids was extracted from the gel by using a QIAquick gel extraction kit (Qiagen Inc.) and self-ligated by using T4 ligase. The deleted plasmids were transformed into *E. coli* DH5α and selected for sequencing according to their sizes.

Alternatively, the repeat region of the *E. canis p120* gene was unidirectionally deleted from the 3′ end by using Exonuclease III with the Erase-a-Base system according to the instructions of the manufacturer (Promega, Madison, Wis.).

Determining the number of repeats in the *E. canis p120* gene. The recombinant plasmid pCA120 was digested completely with *Eco*RI. There is an *Eco*RI cleavage site on both sides of the vector DNA sequences that flank the insert, and there is no *Eco*RI cleavage site in the insert. The insert was separated from the vector by agarose gel electrophoresis. The insert DNA was excised from the agarose gel and purified by using the QIAquick gel extraction kit (Qiagen Inc.). The DNA insert was digested partially with *Spe*I as described above. The digestion mixtures were separated in a 1% agarose gel and vacuum transferred onto a nylon membrane. The DNA bands in the nylon membrane were hybridized with an oligonucleotide probe (CGC AAG ATA AAG TGG GAA TTT) which was derived from the sequence upstream of the repeat region of the *E. canis p120* gene. The DNA probes were labeled by using digoxigenin-11-dUTP with a DIG oligonucleotide tailing kit according to the manufacturer's protocol (Boehringer Mannheim Co., Indianapolis, Ind.).

Gene analysis. The DNA and deduced amino acid sequences were analyzed with the Wisconsin GCG software package (Genetics Computer Group, Inc., Madison, Wis.) and DNASTAR software (DNASTAR, Inc., Madison, Wis.). The deduced protein was analyzed by using the PSORT program (World Wide Web site: http://psort.nibb.ac.jp), which predicts the presence of signal sequences by the methods of McGeoch (18) and von Heijne (22) and detects potential transmembrane domains by the method of Klein et al. (15).

Expression of the *E. canis p120* gene in *E. coli*. Directly cloning the *E. canis p120* gene into the pGEX expression vector (Amersham Pharmacia Biotech, Piscataway, N.J.) was prevented by the absence of matched restriction endonuclease cleavage sites between DNA sequences of the *p120* gene and the multiple cloning site of the pGEX vector. The coding region of the *E. canis p120* gene was amplified with 515f and 356r primers. The PCR-amplified DNA corresponds to amino acids 58 to 598 of the *E. canis* p120 leaving out the DNA sequences encoding 57 and 90 amino acids at the beginning and the end, respectively, of the protein. The PCR-amplified DNA was cloned into pCR2.1 TA cloning vector (Invitrogen) to obtain the *Eco*RI cleavage sites on both ends of the insert. The insert in a recombinant plasmid was cut by *Eco*RI and separated from the plasmid DNA in an agarose gel. The insert was extracted from the agarose gel by using a QIAquick gel extraction kit (Qiagen Inc.) and cloned into *Eco*RI-digested pGEX vector. The *E. canis* protein was expressed in *E. coli* BL21 as a glutathione *S*-transferase (GST) fusion protein. The GST fusion protein was affinity purified by using glutathione Sepharose 4B beads (Amersham Pharmacia Biotech). The *E. canis* recombinant p120 was cleaved from the GST fusion protein with thrombin.

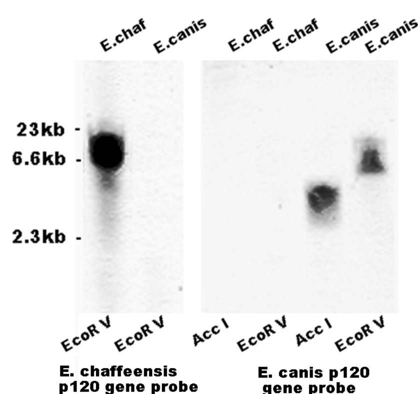


FIG. 1. Southern blot. Shown is the hybridization of the *E. chaffeensis* and *E. canis p120* gene probes with restriction enzyme *Acc*I- and/or *Eco*RV-digested *E. chaffeensis* (E. chaf) and *E. canis* genomic DNA.

Immunization of mice. BALB/c mice were immunized with recombinant *E. canis* p120-GST fusion protein. The recombinant protein was mixed with an equal volume of Freund's complete adjuvant for the first injection and with Freund's incomplete adjuvant for the subsequent injections. Mice were immunized intraperitoneally or subcutaneously with 50 μg of the recombinant p120-GST fusion protein four times at 1-week intervals.

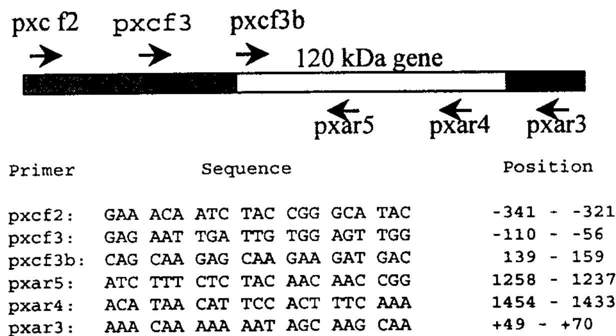
Protein immunoblotting. Ehrlichial recombinant proteins were separated on 10% Tris-HCl Ready Gel with a preparative comb (Bio-Rad Laboratories, Hercules, Calif.). The protein was electrotransferred onto a nitrocellulose membrane by using a Trans-Blot SD semidry transfer cell (Bio-Rad Laboratories). The protein on the membrane was incubated with canine sera by using a Mini-Protean II multiscreen apparatus (Bio-Rad Laboratories). Nine convalescent dog serum samples and five normal dog sera were obtained from the Louisiana Veterinary Medical Diagnostic Laboratory (Baton Rouge). These samples were positive for *E. canis* by an immunofluorescence procedure. Sera were diluted 1:100 for protein immunoblotting.

Nucleotide sequence accession number. The DNA sequence of the *E. canis p120* gene was assigned GenBank accession no. AF112369.

RESULTS

Cloning the *E. canis p120* gene. Southern blotting demonstrated that the *E. chaffeensis p120* gene probe failed to hybridize with restriction enzyme-digested *E. canis* genomic DNA under conditions in which the probe gave strong hybridization with *E. chaffeensis* genomic DNA (Fig. 1). The control probe from the *E. canis p120* gene hybridized with *E. canis* DNA but not *E. chaffeensis* DNA. These results indicated that the *E. canis p120* gene differed substantially from the homologous *E. chaffeensis p120* gene.

We further attempted to amplify the homologous *p120* gene in *E. canis* Oklahoma by PCR. Primers derived from the *E. chaffeensis p120* gene and sequences flanking the gene had been used previously for sequencing the *E. chaffeensis p120* gene (Fig. 2). Three forward primers were paired with three reverse primers to form nine pairs of primers. A 2.5-kb DNA fragment was amplified from *E. canis* genomic DNA by the primer pair pxcf2 and pxar3, derived from the noncoding DNA sequences flanking the *E. chaffeensis p120* gene (Fig. 2). No DNA was amplified by using primers derived from the coding region of the *E. chaffeensis* p120 gene. The 2.5-kb PCR product was cloned into pCR2.1 TA cloning vector to generate the recombinant plasmid pCA120.



Primer	Sequence	Position
pxcf2:	GAA ACA ATC TAC CGG GCA TAC	-341 - -321
pxcf3:	GAG AAT TGA TTG TGG AGT TGG	-110 - -56
pxcf3b:	CAG CAA GAG CAA GAA GAT GAC	139 - 159
pxar5:	ATC TTT CTC TAC AAC AAC CGG	1258 - 1237
pxar4:	ACA TAA CAT TCC ACT TTC AAA	1454 - 1433
pxar3:	AAA CAA AAA AAT AGC AAG CAA	+49 - +70

FIG. 2. DNA sequences and positions of oligonucleotide primers derived from the *E. chaffeensis p120* gene (open box) and the DNA sequences flanking the gene (shaded boxes). The positions of primers are indicated as minus and plus for DNA sequences upstream and downstream of the *p120* gene, respectively. Nine pairs of primers were formed by combining each forward primer with each reverse primer and were used to amplify the *E. canis p120* gene by PCR.

DNA sequence analysis of the *E. canis p120* gene. Preliminary sequencing data indicated that the 2.5-kb PCR product of *E. canis* contained tandem repeats with 108 nucleotides each. The presence of the repeats made the sequencing difficult to accomplish by primer walking. Restriction enzyme analysis of the DNA sequences demonstrated that each repeat has a unique *Spe*I endonuclease cleavage site. Therefore, the number of repeats was determined by *Spe*I partial digestion and Southern blotting. Southern blotting demonstrated that there were 14 repeats in the *E. canis p120* gene (Fig. 3). The repeat region of the *E. canis p120* gene was sequenced by unidirectional deletion of the DNA fragment in pCA120.

DNA sequencing demonstrated that the DNA insert contained an open reading frame (ORF) of 2,064 nucleotides which encoded 688 amino acids (Fig. 4). This ORF was designated the *E. canis p120* gene. There were no consensus DNA sequences of the *E. coli* promoter near the 5′ end of the gene. The N terminus of the deduced amino acids did not share consensus sequence with *E. coli* signal peptides. DNA sequencing confirmed that there were 14 tandem repeats in the *E. canis p120* gene (Fig. 4). At the amino acid level, the homology of all repeats was greater than 94% (Fig. 5). Preceding the first repeat there is an incomplete repeat that has a seven-amino-acid deletion (Fig. 5) and that is 70% homologous to the other repeats.

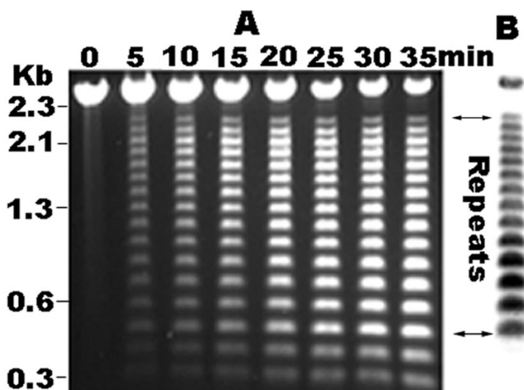


FIG. 3. (A) Agarose gel electrophoresis of the *E. canis p120* gene partially digested with *Spe*I at various time points. (B) Southern blotting determination of the number of repeats. DNA digested for 35 min with *Spe*I from the gel in panel A was transferred to a nylon membrane and hybridized with a digoxigenin-labeled oligonucleotide probe which anneals to the DNA sequences upstream of the repeat region of the *E. canis p120* gene.

```
-340 AAACAATCTACCGGGCATACTTCAACACAATCAGTATATTTGCATCTTATGCACTTATCG
-280 GTAACGAAGTGTGTCATTACAGAGTTATTAATAATAAAGTAACCATTTTTATTGTAATGT
-220 TTTTTCTTGCCAAGTTCAATTAATTTATTGTTTACATAAGGTATAAATGCGGATTATGGT
-160 TAAATTATGCATGTCGTAAGTATAAAATAAGTTGATAAGTGTTTTGTTATATCCTAATAG
-100 ATATAGGAGGCATTGGTTCTATATAAATGTTATTTTATGATAAATAATTAATTTTTAACA
 -40 GGATGAATTTGTGCAATGTATTTAAATTAAGAGGATTTTTATGGATATTGATAACAATAA 20
                                                  M  D  I  D  N  N  N
     TGTGACTACATCAAGTACGCAAGATAAAAGTGGGAATTTAATGGAAGTGATTATGCGTAT 80
      V  T  T  S  S  T  Q  D  K  S  G  N  L  M  E  V  I  M  R  I
     ATTAAATTTTGGTAATAATTCAGATGAGAAAGTAAGCAATGAAGACACTAAAGTTCTTGT 140
      L  N  F  G  N  N  S  D  E  K  V  S  N  E  D  T  K  V  L  V
     AGAGAGTTTACAACCTGCTGTGAATGACAATGTAGGAAATCCATCAAGTGAAGTTGGTAA 200
                                          Primer 515f
      E  S  L  Q  P  A  V  N  D  N  V  G  N  P  S  S  E  V  G  K
     AGAAGAAAATGCTCCTGAAGTTAAAGCGGAAGATTTGCAACCTGCTGTAGATGGTAGTGT 260
      E  E  N  A  P  E  V  K  A  E  D  L  Q  P  A  V  D  G  S  V
     AGAACATTCATCAAGTGAAGTTGGGAAAAAAGTATCTGAAACTAGTAAAGAGGAAAGTAC 320
      E  H  S  S  S  E  V  G  K  K  V  S  E  T  S  K  E  E  S  T
     TCCTGAAGTTAAAGCAGAAGATTTGCAACCTGCTGTAGATGGTAGTATAGAACATTCATC 380
      P  E  V  K  A  E  D  L  Q  P  A  V  D  G  S  I  E  H  S  S
     AAGTGAAGTTGGAGAAAAAGTATCTAAAACTAGTAAAGAGGAAAGTACTCCTGAAGTTAA 440
      S  E  V  G  E  K  V  S  K  T  S  K  E  E  S  T  P  E  V  K
     AGCAGAAGATTTGCAACCTGCTGTAGATGATAGTGTGGAACATTCATCAAGTGAAGTTGG 500
      A  E  D  L  Q  P  A  V  D  D  S  V  E  H  S  S  S  E  V  G
     AGAAAAAGTATCTGAAACTAGTAAAGAGGAAAATACTCCTGAAGTTAAAGCAGAAGATTT 560
      E  K  V  S  E  T  S  K  E  E  N  T  P  E  V  K  A  E  D  L
     GCAACCTGCTGTAGATGGTAGTATAGAACATTCATCAAGTGAAGTTGGAGAAAAAGTATC 620
      Q  P  A  V  D  G  S  I  E  H  S  S  S  E  V  G  E  K  V  S
     TAAAACTAGTAAAGAAGAAAGTACTCCTGAAGTTAAAGCAGAAGATTTGCAACCTGCTGT 680
      K  T  S  K  E  E  S  T  P  E  V  K  A  E  D  L  Q  P  A  V
     AGATGATAGTGTGGAACATTCATCAAGTGAAGTTGGAGAAAAAGTATCTGAAACTAGTAA 740
      D  D  S  V  E  H  S  S  S  E  V  G  E  K  V  S  E  T  S  K
     AGAAGAAAATACTCCTGAAGTTAAAGCGGAAGATTTGCAACCTGCTGTAGATGGTAGTGT 800
      E  E  N  T  P  E  V  K  A  E  D  L  Q  P  A  V  D  G  S  V
     AGAACATTCATCAAGTGAAGTTGGGAAAAAAGTATCTGAAACTAGTAAAGAGGAAAGTAC 860
      E  H  S  S  S  E  V  G  K  K  V  S  E  T  S  K  E  E  S  T
     TCCTGAAGTTAAAGCAGAAGATTTGCAACCTGCTGTAGATGATAGTGTGGAACATTCATC 920
      P  E  V  K  A  E  D  L  Q  P  A  V  D  D  S  V  E  H  S  S
     AAGTGAAGTTGGAGAAAAAGTATCTGAAACTAGTAAAGAGGAAAATACTCCTGAAGTTAG 980
      S  E  V  G  E  K  V  S  E  T  S  K  E  E  N  T  P  E  V  R
     AGCAGAAGATTTGCAACCTGCTGTAGATGGTAGTGTAGAACATTCATCAAGTGAAGTTGG 1040
      A  E  D  L  Q  P  A  V  D  G  S  V  E  H  S  S  S  E  V  G
     AGAAAAAGTATCTGAAACTAGTAAAGAGGAAAGTACTCCTGAAGTTAAAGCAGAAGATTT 1100
      E  K  V  S  E  T  S  K  E  E  S  T  P  E  V  K  A  E  D  L
     GCAACCTGCTGTAGATAGTAGTATAGAACATTCATCAAGTGAAGTTGGGAAAAAAGTATC 1160
      Q  P  A  V  D  S  S  I  E  H  S  S  S  E  V  G  K  K  V  S
     TGAAACTAGTAAAGAGGAAAGTACTCCTGAAGTTAAAGCAGAAGATTTGCAACCTGCTGT 1220
      E  T  S  K  E  E  S  T  P  E  V  K  A  E  D  L  Q  P  A  V
     AGATGGTAGTGTAGAACATTCATCAAGTGAAGTTGGAGAAAAAGTATCTGAAACTAGTAA 1280
      D  G  S  V  E  H  S  S  S  E  V  G  E  K  V  S  E  T  S  K
     AGAGGAAAATACTCCTGAAGTTAAAGCAGAAGATTTGCAACCTGCTGTAGATGGTAGTGT 1340
      E  E  N  T  P  E  V  K  A  E  D  L  Q  P  A  V  D  G  S  V
     AGAACATTCATCAAGTGAAGTTGGAGAAAAAGTATCTGAAACTAGTAAAGAGGAAAATAC 1400
      E  H  S  S  S  E  V  G  E  K  V  S  E  T  S  K  E  E  N  T
     TCCTGAAGTTAAAGCGGAAGATTTGCAACCTGCTGTAGATGGTAGTGTAGAACATTCATC 1460
      P  E  V  K  A  E  D  L  Q  P  A  V  D  G  S  V  E  H  S  S
     AAGTGAAGTTGGAGAAAAAGTATCTGAAACTAGTAAGGAAGAAAGTACTCCTGAAGTTAA 1520
      S  E  V  G  E  K  V  S  E  T  S  K  E  E  S  T  P  E  V  K
     AGCGGAAGATTTGCAACCTGCTGTAGATGGTAGTGTGGAACATTCATCAAGTGAAGTTGG 1580
      A  E  D  L  Q  P  A  V  D  G  S  V  E  H  S  S  S  E  V  G

     AGAAAAAGTATCTGAGACTAGTAAAGAAGAAAGTACTCCTGAAGTTAAAGCGGAAGATTT 1640
      E  K  V  S  E  T  S  K  E  E  S  T  P  E  V  K  A  E  D  L
     GCAACCTGCTGTAGATGGTAGTGTGGAACATTCATCAAGTGAAGTTGGAGAAAAAGTATC 1700
      Q  P  A  V  D  G  S  V  E  H  S  S  S  E  V  G  E  K  V  S
     TGAGACTAGTAAAGAGGAAAGTACTCCTGAAGTTAAAGCGGAAGTACAGCCTGTTGCAGA 1760
      E  T  S  K  E  E  S  T  P  E  V  K  A  E  V  Q  P  V  A  D
     TGGTAATCCTGTTGCTTTAAATCCTATGCCTTCAATTGATAATATTGATACTAATATAAT 1820
                     Primer 356r
      G  N  P  V  P  L  N  P  M  P  S  I  D  N  I  D  T  N  I  I
     ATTCCATTACCATAAAGACTGTAAAAAAGGTTCAGCTGTAGGAACAGATGAAATGTGTTG 1880
      F  H  Y  H  K  D  C  K  K  G  S  A  V  G  T  D  E  M  C  C
     TCCTGTATCAGAATTAATGGCTGGGGAACATGTTCATATGTATGGAATTTATGTCTATAG 1940
      P  V  S  E  L  M  A  G  E  H  V  H  M  Y  G  I  Y  V  Y  R
     AGTTCAATCAGTAAAGGATTTAAGTGGTGTATTTAATATAGATCATTCTACATGTGATTG 2000
      V  Q  S  V  K  D  L  S  G  V  F  N  I  D  H  S  T  C  D  C
     TAATTTAGATGTTTATTTTGTAGGATACAATTCTTTTACTAACAAAGAAACAGTTGATTT 2060
      N  L  D  V  Y  F  V  G  Y  N  S  F  T  N  K  E  T  V  D  L
     AATATAATATTGTAGTACGTAAGCTTTATAAAATTGTATATTGAATAGCAAGTAATGCTA 2120
      I  *
     ATGCAGTATTGCTTGCTATTTTTTTGTTT                                2149
```

FIG. 4. *E. canis p120* gene sequence and the deduced amino acids. The nucleic acids of repeats 1, 3, 5, 7, 9, 11, and 13 are underlined. Arrows indicate the sequences and directions of primers that were used to amplify the DNA fragment to express the gene.

Sequence homology of the p120s of *E. canis* and *E. chaffeensis*. Searching the SwissProt database by using the FastA program revealed that the amino acid sequence of the *E. canis* p120 is most closely related to that of the *E. chaffeensis* p120. The amino acid identity of p120s of *E. canis* and *E. chaffeensis* is 30%. A comparison of the amino acid sequences of *E. chaffeensis* and *E. canis* showed that they are more conserved on the N terminus and in the repeat region of p120. The amino

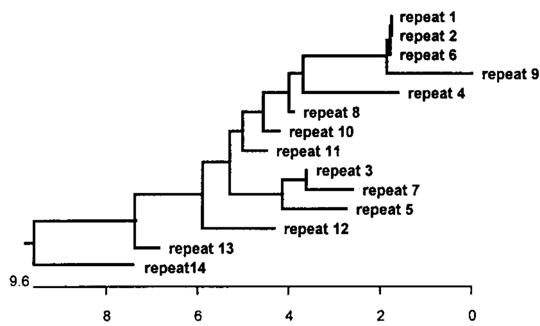


FIG. 5. Phylogenetic relationships of the repeat units of the *E. canis* p120. The scale represents the percent difference in amino acid sequence.

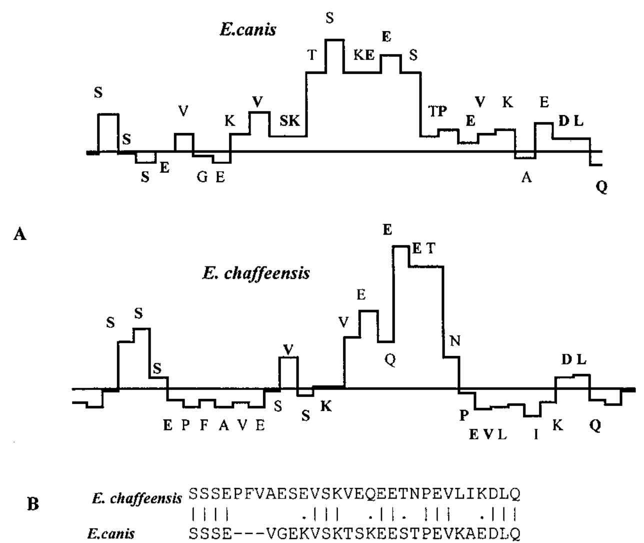


FIG. 7. Comparison of surface-exposed amino acids in repeat units of the p120s of *E. canis* and *E. chaffeensis*. (A) Surface probabilities of amino acids. Boldface letters indicate the amino acids conserved between *E. canis* and *E. chaffeensis*. (B) Alignment of the amino acid sequences shown in panel A. Lines represent identical amino acids. Dots represent conserved replacements. Dashes indicate gaps that were introduced for optimal alignment of the amino acid sequences.

acid identity is 50% for the first 32 amino acids of the N termini of the 120-kDa proteins of *E. canis* and *E. chaffeensis*.

The amino acid sequences of the *E. canis* and *E. chaffeensis* p120s, especially the repeats, were similar in hydrophobicity, surface probability, and antigenicity. All repeat units in both proteins are predicted to be hydrophilic, surface exposed, and highly antigenic (Fig. 6). The surface-exposed regions of the repeats have common amino acids in both the ehrlichial species (Fig. 7).

Homologous genes in other strains of *E. canis*. A 2.5-kb DNA fragment from each strain of *E. canis* examined, including strains Florida, Louisiana, and the three North Carolina canine isolates (Demon, DJ, and Jake), was amplified with primers pxcf2 and pxar3. The segments of the *p120* genes of all *E. canis* strains were sequenced on both the 5′ and 3′ ends. DNA sequence analysis demonstrated that the DNA sequences both up- and downstream of the repeat region were identical among all strains of *E. canis*. We did not attempt to sequence the complete repeat region for all *E. canis* strains because of the difficulty of sequencing the DNA repeats. We

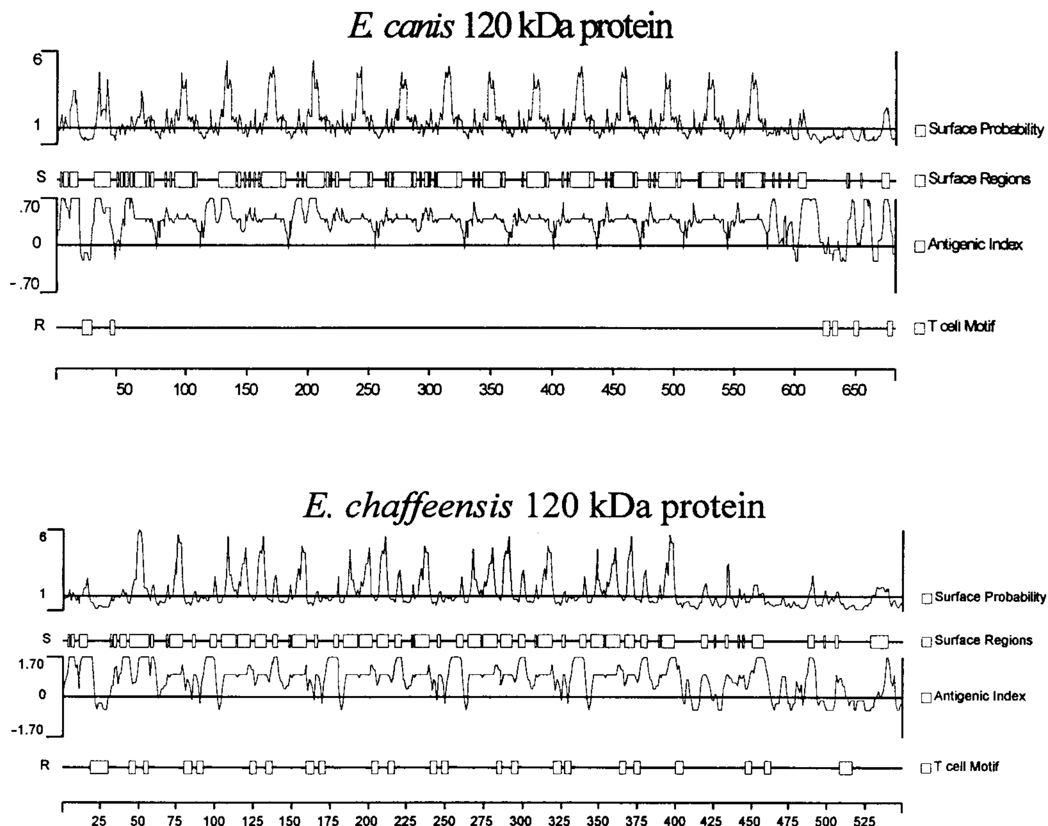


FIG. 6. Surface probabilities, antigenic indices, and T-cell motifs of the p120s of *E. canis* and *E. chaffeensis*.

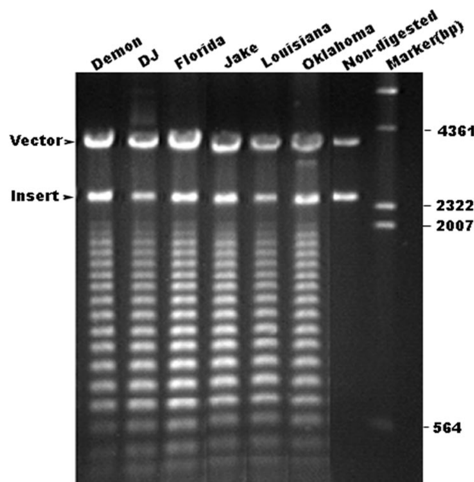


FIG. 8. Agarose gel electrophoresis of the *E. canis p120* genes from six strains of *E. canis* partially digested with *Spe*I. The recombinant pCR2.1 plasmids were first digested with *Eco*RI to release the insert from the vector and then digested partially with *Spe*I. Nondigested, Oklahoma strain *p120* gene DNA was digested with *Eco*RI but not with *Spe*I to show the size of the insert.

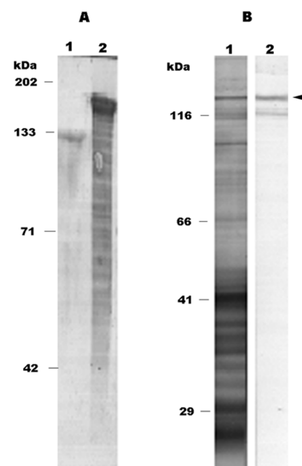


FIG. 9. (A) SDS-PAGE of *E. coli*-expressed *E. canis* p120. Lane 1, *E. canis* recombinant p120 cleaved from the GST fusion protein by thrombin; lane 2, GST fusion protein. (B) Western immunoblot of mouse anti-*E. canis* recombinant p120 sera reacted with *E. canis* antigen (lane 1) and recombinant p120 (lane 2; arrowhead).

sequenced the last repeats of all strains and the first repeat of DJ strain. The sequences of the first repeats of DJ and Oklahoma strains were identical. The sequences of the last repeats were identical among all strains. Homology of the p120 genes from all *E. canis* strains was further demonstrated by their identical *Spe*I restriction physical maps (Fig. 8).

Protein immunoblotting. The *E. canis p120* gene was expressed in *E. coli*. The recombinant protein encoded by a 1,620-bp DNA fragment including all the repeats of the *p120* gene was expressed as a GST fusion protein. The estimated molecular size of the fusion protein on sodium dodecyl sulfate (SDS) gel was approximately 140 kDa, which is much larger than the predicted molecular mass of the entire *E. canis* p120, which is only 73.6 kDa based on the amino acid sequence deduced from the DNA sequence (Fig. 9). Mouse antibodies to the recombinant p120 reacted with a p120 of *E. canis* (Fig. 9). The recombinant *E. canis* p120 reacted with all nine canine convalescent sera but with none of the normal dog sera (Fig. 10).

DISCUSSION

The homology of the amino acid sequences of the p120s of *E. canis* and *E. chaffeensis* is 30%. The DNA sequence homology of the *p120* genes between the two species is 58%. It is surprising that the noncoding sequences flanking the *p120* genes are more conserved than the coding sequences of the *p120* genes of *E. canis* and *E. chaffeensis*. A comparison of 340 nucleotides upstream of the *p120* gene revealed that the noncoding regions adjacent to the *p120* genes of the two species of *Ehrlichia* have 84% homology. From an evolutionary point of view, the coding sequence that is under selection pressure would be expected to be more conserved than the noncoding sequence in which mutation would not be expected to affect the survival of the organism. We believe that the *E. canis p120* gene is the homologue of the *E. chaffeensis p120* considering that they are located in similar positions in the respective genomes, that they are 30% homologous, and especially that they have common motifs in the repeat region. The repeats in both proteins are hydrophilic and are predicted to be surface exposed. Even the total numbers of surface-exposed regions in the repeats of the two proteins are very close in spite of the difference in the numbers of repeat units (the *E. chaffeensis p120* gene has three or four repeats, depending on the strain) (8, 25). The repeat units of both proteins have a common motif consisting of identical amino acids that are hydrophilic and that form the core of the surface-exposed regions of these proteins. These results indicated that the *E. canis* p120 is an outer membrane protein. The repeat units of both proteins are rich in glutamic acid and serine. Glutamic acid and serine each comprise 19% of the amino acids of the *E. canis* repeat unit. Glutamic acid and serine comprise 22 and 15% of the amino acids of the *E. chaffeensis* repeat units, respectively. Like that of the *E. chaffeensis* p120, the predicted molecular mass of the *E. canis* p120 is much smaller than the molecular size estimated on the basis of the electrophoretic mobility of the protein as determined by SDS-polyacrylamide gel electrophoresis (PAGE). The same phenomenon has been reported for other proteins containing repeat domains, including those of *A. marginale* (1), *Plasmodium* spp. (14), and *Staphylococcus aureus* (13, 20) and the HGE 100- and 130-kDa proteins (21). The repeat units of the HGE 100- and 130-kDa proteins have sequences in common with those of the *E. chaffeensis* p120 (21). The aberrant migration of the p120s of *E. canis* and *E. chaffeensis* is caused by glycosylation of the proteins (17a). Since the p120 of *E. chaffeensis* was differentially expressed in different ultrastructural forms of *E. chaffeensis*, this protein may play a role in the pathogenesis of *E. chaffeensis* infection. Whether or not the *E. canis* p120 is preferentially expressed in

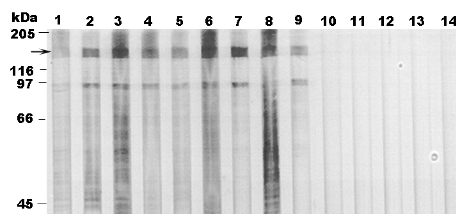


FIG. 10. Western blotting of nine canine convalescent sera (lanes 1 to 9) and five normal canine sera (lanes 10 to 14) reacted with recombinant p120 of *E. canis*. The p120-GST fusion protein is indicated by an arrow.

the dense-core cell of *E. canis* is under investigation. The p120 gene appears to be conserved among all strains of *E. canis* since the known sequences, including the nonrepeat regions as well as the last repeats, are identical among strains of *E. canis* and since all *E. canis* strains have same number of repeats. The high degree of homology of DNA sequences and the identical numbers of repeats of the *p120* genes among the strains of *E. canis* indicated that *E. canis* strains are genetically less diverse than those of *E. chaffeensis*, in which the number of repeats of the *p120* gene differs among strains. p120 is immunodominant in both *E. canis* and *E. chaffeensis* because the recombinant p120s of both species react strongly with either human patient sera (27) or canine sera. Protein immunoblotting demonstrated that rabbit antisera to the *E. chaffeensis* p120 does not cross-react with *E. canis* and that mouse anti-*E. canis* p120 serum does not react with *E. chaffeensis* (data not shown). Therefore, the p120s of *E. canis* and *E. chaffeensis* may be useful for serodiagnosis of canine and human ehrlichiosis, respectively, for which they are both sensitive and specific.

ACKNOWLEDGMENTS

We thank Josie Ramirez-Kim for her assistance in the preparation of this manuscript.

This research was supported by a grant from the Clayton Foundation.

REFERENCES

1. **Allred, D. R., T. C. McGuire, G. H. Palmer, S. R. Leib, T. M. Harkins, T. F. McElwain, and A. F. Barbet.** 1990. Molecular basis for surface antigen size polymorphisms and conservation of a neutralization-sensitive epitope in *Anaplasma marginale*. Proc. Natl. Acad. Sci. USA **87:**3220–3224.
2. **Anderson, B. E., J. E. Dawson, D. C. Jones, and K. H. Wilson.** 1991. *Ehrlichia chaffeensis*, a new species associated with human ehrlichiosis. J. Clin. Microbiol. **29:**2838–2842.
3. **Anderson, B. E., C. E. Greene, D. C. Jones, and J. E. Dawson.** 1992. *Ehrlichia ewingii* sp. nov., the etiologic agent of canine granulocytic ehrlichiosis. Int. J. Syst. Bacteriol. **42:**299–302.
4. **Bakken, J. S., J. S. Dumler, S. M. Chen, M. R. Eckman, L. L. Van Etta, and D. H. Walker.** 1994. Human granulocytic ehrlichiosis in the upper Midwest United States. A new species emerging? JAMA **272:**212–218.
5. **Buhles, W. C., Jr., D. L. Huxsoll, and M. Ristic.** 1974. Tropical canine pancytopenia: clinical, hematologic, and serologic response of dogs to *Ehrlichia canis* infection, tetracycline therapy, and challenge inoculation. J. Infect. Dis. **130:**357–367.
6. **Chen, S. M., J. S. Dumler, J. S. Bakken, and D. H. Walker.** 1994. Identification of a granulocytotropic *Ehrlichia* species as the etiologic agent of human disease. J. Clin. Microbiol. **32:**589–595.
7. **Chen, S. M., J. S. Dumler, H. M. Feng, and D. H. Walker.** 1994. Identification of the antigenic constituents of *Ehrlichia chaffeensis*. Am. J. Trop. Med. Hyg. **50:**52–58.
8. **Chen, S. M., X. J. Yu, V. L. Popov, E. L. Westerman, F. G. Hamilton, and D. H. Walker.** 1997. Genetic and antigenic diversity of *Ehrlichia chaffeensis*: comparative analysis of a novel human strain from Oklahoma and previously isolated strains. J. Infect. Dis. **175:**856–863.
9. **Dawson, J. E., Y. Rikihisa, S. A. Ewing, and D. B. Fishbein.** 1991. Serologic diagnosis of human ehrlichiosis using two *Ehrlichia canis* isolates. J. Infect. Dis. **163:**564–567.
10. **Fishbein, D. B., L. A. Sawyer, C. J. Holland, E. B. Hayes, W. Okoroanyanwu, D. Williams, K. Sikes, M. Ristic, and J. E. McDade.** 1987. Unexplained febrile illnesses after exposure to ticks. Infection with an *Ehrlichia*? JAMA **257:**3100–3104.
11. **Greene, C. E., and J. W. Harvey.** 1984. Canine ehrlichiosis, p. 545–561. *In* C. E. Greene (ed), Clinical microbiology and infectious diseases of the dog and cat. The W. B. Saunders Co., Philadelphia, Pa.
12. **Groves, M. G., G. L. Dennis, H. L. Amyx, and D. L. Huxsoll.** 1975. Transmission of *Ehrlichia canis* to dogs by ticks (*Rhipicephalus sanguineus*). Am. J. Vet. Res. **36:**937–940.
13. **Hollingshead, S. K., V. A. Fischetti, and J. R. Scott.** 1986. Complete nucleotide sequence of type 6 M protein of the group A *Streptococcus*. Repetitive structure and membrane anchor. J. Biol. Chem. **261:**1677–1686.
14. **Kemp, D. J., R. L. Coppel, and R. F. Anders.** 1987. Repetitive proteins and genes of malaria. Annu. Rev. Microbiol. **41:**181–208.
15. **Klein, P., M. Kanehisa, and C. DeLisi.** 1985. The detection and classification of membrane-spanning proteins. Biochim. Biophys. Acta **815:**468–476.
16. **Lewis, G. E., Jr., M. Ristic, R. D. Smith, T. Lincoln, and E. H. Stephenson.** 1977. The brown dog tick *Rhipicephalus sanguineus* and the dog as experimental hosts of *Ehrlichia canis*. Am. J. Vet. Res. **38:**1953–1955.
17. **Maeda, K., N. Markowitz, R. C. Hawley, M. Ristic, D. Cox, and J. E. McDade.** 1987. Human infection with *Ehrlichia canis*, a leukocytic rickettsia. N. Engl. J. Med. **316:**853–856.

17a. **McBride, J. W., X. Yu, and D. H. Walker.** Glycosylation of homologous immunodominant proteins of *Ehrlichia chaffeensis* and *Ehrlichia canis*. Infect. Immun., in press.

18. **McGeoch, D. J.** 1985. On the predictive recognition of signal peptide sequences. Virus Res. **3:**271–286.
19. **Perez, M., Y. Rikihisa, and B. Wen.** 1996. *Ehrlichia canis*-like agent isolated from a man in Venezuela: antigenic and genetic characterization. J. Clin. Microbiol. **34:**2133–2139.

19a. **Popov, V. L., X. Yu, and D. H. Walker.** The 120 kDa outer membrane protein of *Ehrlichia chaffeensis*: preferential expression of dense-core cells and gene expression in *Escherichia coli* associated with attachment and entry. Microb. Pathog., in press.

20. **Signas, C., G. Raucci, K. Jonsson, P. E. Lindgren, G. M. Anantharamaiah, M. Hook, and M. Lindberg.** 1989. Nucleotide sequence of the gene for a fibronectin-binding protein from *Staphylococcus aureus*: use of this peptide sequence in the synthesis of biologically active peptides. Proc. Natl. Acad. Sci. USA **86:**699–703.
21. **Storey, J. R., L. A. Doros-Richert, C. Gingrich-Baker, K. Munroe, T. N. Mather, R. T. Coughlin, G. A. Beltz, and C. I. Murphy.** 1998. Molecular cloning and sequencing of three granulocytic *Ehrlichia* genes encoding high-molecular-weight immunoreactive proteins. Infect. Immun. **66:**1356–1363.
22. **von Heijne, G.** 1986. A new method for predicting signal sequence cleavage sites. Nucleic Acids Res. **14:**4683–4690.
23. **Walker, J. S., J. D. Rundquist, R. Taylor, B. L. Wilson, M. R. Andrews, J. Barck, A. L. Hogge, Jr., D. L. Huxsoll, P. K. Hildebrandt, and R. M. Nims.** 1970. Clinical and clinicopathologic findings in tropical canine pancytopenia. J. Am. Vet. Med. Assoc. **157:**43–55.
24. **Weiss, E., J. C. Coolbaugh, and J. C. Williams.** 1975. Separation of viable *Rickettsia typhi* from yolk sac and L cell host components by renografin density gradient centrifugation. Appl. Microbiol. **30:**456–463.
25. **Yu, X.-J., P. A. Crocquet-Valdes, and D. H. Walker.** 1996. Cloning and sequencing of the gene for a 120-kDa immunodominant protein of *Ehrlichia chaffeensis*. Gene **184:**149–154.
26. **Yu, X.-J., and D. H. Walker.** 1997. Sequence and characterization of an *Ehrlichia chaffeensis* gene encoding 314 amino acids highly homologous to the NAD A enzyme. FEMS Microbiol. Lett. **154:**53–58.
27. **Yu, X.-J., P. A. Crocquet-Valdes, L. C. Cullman, and D. H. Walker.** 1996. The recombinant 120-kilodalton protein of *Ehrlichia chaffeensis*, a potential diagnostic tool. J. Clin. Microbiol. **34:**2853–2855.