Vol. 68, No. 1

Glycosylation of Homologous Immunodominant Proteins of *Ehrlichia chaffeensis* and *Ehrlichia canis*

JERE W. McBRIDE, XUE-JIE YU, AND DAVID H. WALKER*

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

The glycoprotein genes of *Ehrlichia chaffeensis* (1,644 bp) and *Ehrlichia canis* (2,064 bp) encode proteins of 548 to 688 amino acids with predicted molecular masses of only 61 and 73 kDa but with electrophoretic mobilities of 120 kDa (P120) and 140 kDa (P140), respectively. The 120-kDa protein gene of *E. chaffeensis* contains four identical 240-bp tandem repeat units, and the 140-kDa protein gene of *E. canis* has 14 nearly identical, tandemly arranged 108-bp repeat units. Conserved serine-rich motifs identified in the repeat units of P120 and P140 were also found in the repeat units of the human granulocytotropic ehrlichiosis agent 130-kDa protein and of the fimbria-associated adhesin protein Fap1 of *Streptococcus parasanguis*. Nearly the entire (99%) *E. chaffeensis* P120 gene (1,616 bp), the 14-repeat region (78%) of the *E. canis* P140 gene (1,620 bp), and a 2-repeat region from the *E. chaffeensis* P120 gene (520 bp) were expressed in *Escherichia coli*. The recombinant proteins exhibited molecular masses ranging from 1.6 to 2 times larger than those predicted by the amino acid sequences. Antibodies against the recombinant proteins reacted with *E. chaffeensis* P120 and *E. canis* P140, respectively. Carbohydrate was detected on the *E. chaffeensis* and *E. canis* recombinant proteins, including the two-repeat polypeptide region of *E. chaffeensis* P120. A carbohydrate compositional analysis identified glucose, galactose, and xylose on the recombinant proteins. The presence of only one site for N-linked (Asn-Xaa-Ser/Thr) glycosylation, a lack of effect of *N*-glycosidase F, the presence of 70 and 126 Ser/Thr glycosylation sites in the repeat regions of P120 and P140, respectively, and a high molar ratio of carbohydrate to protein suggest that the glycans may be O linked.

Ehrlichia chaffeensis and *Ehrlichia canis* are obligate intracellular bacteria that exhibit tropism for monocytes and macrophages and are responsible for the diseases human monocytotropic ehrlichiosis and canine ehrlichiosis, respectively (10). Recently, the 120- and 140-kDa protein genes from *E. chaffeensis* and *E. canis* have been cloned, expressed, and characterized (18, 19). The 120-kDa protein (P120) of *E. chaffeensis* and the 140-kDa protein (P140) of *E. canis* are immunodominant and *E. chaffeensis* P120 appears to be surface exposed (9a). The proteins are homologous, and each has a region of serine-rich tandem repeats. The recombinant *E. canis* P140 and *E. chaffeensis* P120 exhibit molecular masses much larger than those predicted by the amino acid sequences, and antibodies produced against the recombinant proteins recognized native *E. chaffeensis* and *E. canis* proteins of similar sizes (18, 19). Two proteins (P100 and P130) from the human granulocytotropic erhlichiosis (HGE) agent have also been cloned, and the recombinant proteins exhibited higher-than-predicted molecular masses (12).

The existence of glycoproteins in eukaryotic cells has been known for years and was thought to be restricted to these cells. However, more recently, glycoproteins have been identified in eubacteria and archaebacteria (6, 7). The best characterized prokaryotic glycoproteins are surface layer (S-layer) proteins, but other membrane proteins unrelated to the S-layer have also been reported in eubacteria, such as species of *Streptococcus*, *Mycobacterium*, and *Borrelia* (4, 11, 15). Characterization of eubacterial glycoproteins thus far has revealed diverse carbohydrate structures compared to eukaryotes, which have more conserved glycan structures. Many specific functions of glycoproteins in prokaryotes have been reported, including maintenance of cell shape, protein stability, protection against proteolysis, and adherence to target cells (7).

Glycans are usually attached to proteins by an asparagine residue (N glycosylation) or to a serine or threonine (O glycosylation). Both types of glycosylation have been reported for eubacterial S-layer and outer membrane proteins (7). N- and O-linked glycans typically consist of a conserved core region with variable branching sugars, but glycans without a conserved core region have also been reported (8).

The purpose of this study was to determine whether the recombinant and native proteins from *E. canis* and *E. chaffeensis* are glycosylated. The extent and character of the glycans on the ehrlichial proteins were examined by using carbohydrate compositional analysis, glycosidases, and lectins to better understand the glycan structure on recombinant and native ehrlichial glycoproteins.

MATERIALS AND METHODS

Ehrlichiae and purification. *E. canis* (Oklahoma isolate) and *E. chaffeensis* (Arkansas strain) were provided by Jacqueline Dawson (Centers for Disease Control and Prevention, Atlanta, Ga.). Propagation of ehrlichiae was performed in DH82 cells with Dulbecco modified Eagle medium supplemented with 10% bovine calf serum and 2 mM L-glutamine at 37°C. The intracellular growth in DH82 cells was monitored by the presence of *E. canis* and *E. chaffeensis* morulae by using general cytologic staining methods. Cells were harvested when 100% of the cells were infected with ehrlichiae and were then pelleted in a centrifuge at 17,000 × *g* for 20 min. Cell pellets were disrupted with a Braun-Sonic 2000 sonicator twice at 40 W for 10 s on ice. The lysate was loaded onto discontinuous gradients of 42, 36, and 30% Renografin (*E. chaffeensis*) or onto a 32% continuous Percoll gradient (*E. canis*) as previously described (14). Ehrlichiae were collected and washed with sucrose-phosphate-glutamate buffer (218 mM sucrose, 3.8 mM KH$_2$PO$_4$, 7.2 mM K$_2$HPO$_4$, 4.9 mM glutamate, pH 7.0) and pelleted by centrifugation.

* 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.

Gene cloning and expression and purification of recombinant proteins. The *E. chaffeensis* P120 gene (1,864 bp) (including a 260-bp noncoding region corresponding to the carboxy terminus), the *E. canis* P140 gene (1,620 bp), and a region consisting of two repeat units (520 bp) from the *E. chaffeensis* P120 gene were cloned into pGEX expression vectors as previously described (17–19). All constructs were transformed into *Escherichia coli* BL21, and the recombinant fusion proteins were expressed after IPTG (isopropyl-β-D-thiogalactopyranoside) induction for 4 h at 37°C. Bacteria were harvested by centrifugation at 5,000 × *g* for 20 min and resuspended in phosphate-buffered saline with protease inhibitors (Roche Molecular Biochemicals, Indianapolis, Ind.). The suspension was sonicated to lyse the bacteria, and the insoluble material was pelleted by centrifugation. The supernatant was removed, and the soluble recombinant proteins were purified by using glutathione-Sepharose (Amersham Pharmacia Biotech, Piscataway, N.J.).

Detection of native *E. chaffeensis* and *E. canis* glycoproteins. Purified *E. chaffeensis* and *E. canis* organisms grown in vitro were subjected to sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE), and the separated proteins were transferred to nitrocellulose. The membrane was incubated in blocking buffer (1% nonfat milk, 0.1 M Tris, pH 7.5) for 1 h. Blots were incubated with homologous and heterologous rabbit (*E. chaffeensis* P120) or mouse (*E. canis* P120) antisera against the respective recombinant proteins at 1:100 for 1 h. Bound antibody was detected with a secondary goat anti-rabbit or goat anti-mouse immunoglobulin G (α-chain) alkaline phosphatase-labeled conjugate (Kirkegaard & Perry Laboratories, Gaithersburg, Md.) and visualized with 5-bromo-4-chloro-3-indolyl phosphate–nitroblue tetrazolium (BCIP-NBT) substrate.

Detection of glycoproteins. Glycoprotein detection was performed with a glycoprotein detection kit according to the protocol of the manufacturer (Bio-Rad Laboratories, Hercules, Calif.). Briefly, recombinant proteins were incubated in solution with periodate, oxidizing adjacent hydroxyl groups in carbohydrates to generate free aldehyde groups. The oxidized aldehydes were subsequently labeled with biotin hydrazide. The labeled proteins were subjected to SDS-PAGE and transferred to nitrocellulose. The membrane-bound biotin-labeled glycoproteins were detected with alkaline phosphatase-labeled streptavidin and visualized with BCIP-NBT substrate.

Monosaccharide analysis. Analysis of the monosaccharide ratio and total carbohydrate from *E. chaffeensis* P120 and *E. canis* P140 recombinant proteins and the glutathione *S*-transferase (GST) protein was performed by gas chromotography. Purified fusion proteins (100 μg) were cleaved with thrombin for 24 h, dialyzed against water, and dried. Samples were hydrolyzed with 1 M methanolic HCl for 16 h at 80°C. The released sugars were derivatized with Tri-Sil, and the samples were analyzed on a gas chromatograph with a Supelco column. *myo*-Inositol (20 μg) was added as a standard. The molar ratio of carbohydrate to protein was determined based on the quantity of recombinant protein submitted for analysis (100 μg), as estimated with stained SDS-polyacrylamide gels. Monosaccharide analysis was performed at the University of Georgia Complex Carbohydrate Research Center, Athens.

Lectin analysis. Lectin-binding analysis was performed with recombinant proteins and native *E. canis* and *E. chaffeensis* proteins by using a glycan differentiation kit (Roche Molecular Biochemicals) according to the manufacturer's protocol. The following digoxigenin-labeled lectins were used: *Galanthus nivalis* agglutinin recognizes terminal mannose linked α(1-3), α(1-6), or α(1-2); *Sambucus nigra* agglutinin recognizes sialic acid linked α(2-6) to galactose; *Maackia amurensis* agglutinin recognizes sialic acid linked α(2-3) to galactose; peanut agglutinin recognizes core disaccharide galactose β(1-3) *N*-acetylgalactosamine; and *Datura stramonium* agglutinin recognizes Gal α-(1-4) GlcNAc. Each was incubated with *E. canis*, *E. chaffeensis*, and uninfected DH82 cells separated by SDS-PAGE and with purified recombinant proteins. Lectins were detected with an antidigoxigenin alkaline phosphatase-labeled antibody and visualized with BCIP-NBT substrate.

Enzymatic deglycosylation. Analysis of glycan linkages of *E. chaffeensis* P120 and *E. canis* P140 was performed with a panel of glycosidase enzymes. Recombinant proteins were incubated with endoglycosidases (*N*-glycosidase F and *O*-glycosidase) and exoglycosidases (neuraminidase, β-galactosidase, and glucosaminidase) supplied in a enzymatic deglycosylation kit under denaturing conditions according to the protocol of the manufacturer (Bio-Rad Laboratories). The recombinant proteins were incubated independently with the enzymes to examine specific glycan linkages.

Protein analysis. Amino acid motifs were identified by using a Clustal alignment, and *Emini* surface probability plots were performed with protein sequence analysis software (Protean; DNAStar, Madison, Wis.).

RESULTS

Expression of recombinant proteins in *E. coli*. The *E. chaffeensis* P120 gene (1,864 bp; partial), the *E. canis* P140 gene (1,620 bp; partial), and the *E. chaffeensis* repeat region (520 bp) were cloned into pGEX expression vectors (Fig. 1). The recombinant proteins expressed in *E. coli* BL21 cells exhibited
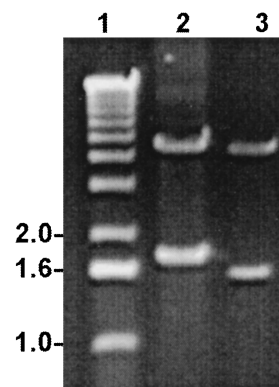

FIG. 1. Agarose gel electrophoresis of *E. chaffeensis* P120 gene and *E. canis* P140 gene pGEX constructs digested to release inserts containing incomplete *E. chaffeensis* P120 gene (1,864 bp; 99%) (lane 2) and *E. canis* P120 gene (1,620 bp; 78%) (lane 3). Lane 1, molecular size markers (1-kb ladder).

molecular masses 1.6 to 2 times larger than those predicted by the amino acid sequence for each gene. The predicted molecular mass of the *E. chaffeensis* P120 recombinant protein was 67 kDa, that of the *E. canis* P140 protein was 57 kDa, and that of the *E. chaffeensis* two-repeat region was 20 kDa. However, the 100-kDa molecular mass of recombinant *E. chaffeensis* P120 observed on SDS-PAGE was approximately 1.6 times larger than the protein size predicted by the amino acid sequence, and the recombinant two-repeat region of *E. chaffeensis* P120 (60 kDa) was also 1.6 times larger than predicted by the amino acid sequence (Fig. 2). The 14 tandem repeats of *E. canis* P140 exhibited a molecular mass of 112 kDa, or approximately double the size predicted by the amino acid sequence (Fig. 2). All molecular masses were determined by subtracting the mass of 28-kDa GST fusion protein from the observed molecular mass of the recombinant fusion protein.

Identification of native glycoproteins. Native proteins with molecular masses of 120 and 140 kDa from *E. chaffeensis* and *E. canis*, respectively, reacted with antisera against the corresponding homologous recombinant glycoproteins (Fig. 3). Antiserum against each respective recombinant protein did not cross-react with the heterologous native protein (not shown).

Glycoprotein detection. Recombinant proteins representing *E. chaffeensis* P120, *E. canis* P140, and the two-repeat region of *E. chaffeensis* P120 were analyzed for carbohydrate content. Recombinant *E. canis* P140, *E. chaffeensis* P120, *E. chaffeensis* P120 repeat, and GST proteins were labeled with biotin to detect carbohydrate residues (Fig. 4). The GST fusion protein (28 kDa) was not labeled (not shown).

Analysis of protein glycosylation sites. Analysis of predicted glycosylation sites on the *E. canis* P140 recombinant protein revealed only 2 possible N-linked glycosylation sites (amino acids 7 and 32; Asn-Xaa-Ser/Thr) and 147 possible O-linked glycosylation sites (Ser or Thr), and analysis of predicted glycosylation sites on *E. chaffeensis* P120 revealed 91 O-linkage sites and two N-linkage sites. The *E. chaffeensis* P120 and *E. canis* P140 proteins have a high serine-and-threonine content that is particularly concentrated in the repeat regions of the proteins (Table 1; Fig. 5). *E. canis* P140 has 110 serine residues, of which 98 are located in the repeat region, and *E. chaffeensis* P120 has 67 serine residues, with 52 located in the repeat region. Threonine, another potential O-linked glycosylation site, was present at many sites in the repeat regions of each protein. *E. chaffeensis* P120 has 18 threonine residues in the entire repeat region, and *E. canis* P140 has 28 threonine
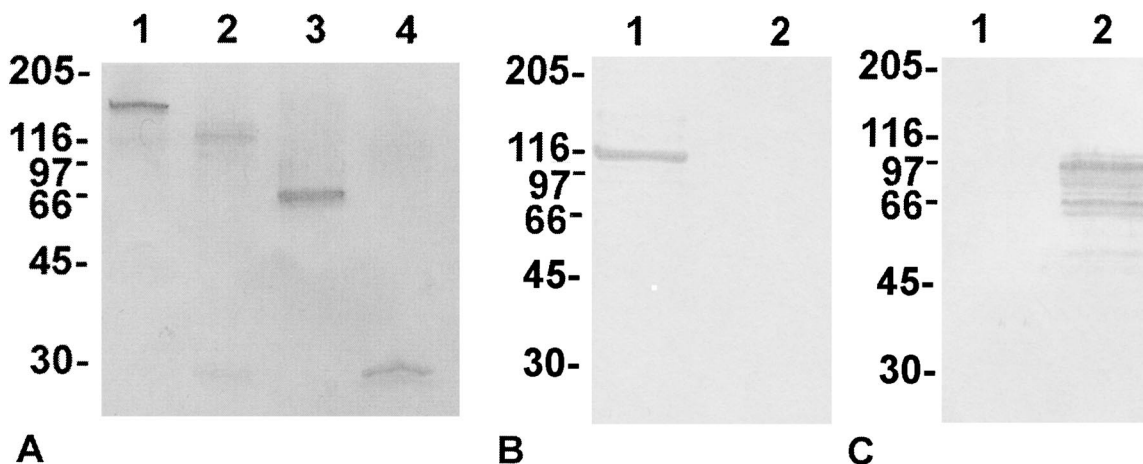
FIG. 2. (A) SDS-PAGE of recombinant ehrlichial GST fusion proteins expressed in *E. coli* BL21 cells. Lane 1, *E. canis* P140; lane 2, *E. chaffeensis* P120; lane 3, *E. chaffeensis* P120 two-repeat units; lane 4, GST protein (28 kDa). (B and C) Specificity of anti-recombinant *E. canis* P140 (B) and anti-recombinant *E. chaffeensis* P120 (C) against recombinant *E. canis* P140 (lanes 1) and *E. chaffeensis* P120 (lanes 2) proteins after removal of GST by thrombin protease. Numbers on the left of each panel are molecular masses in kilodaltons.

residues in the repeat region. A total of 16 amino acids (20%) of 80 were identified as potential O-linked glycosylation sites on each *E. chaffeensis* P120 repeat, while *E. canis* P140 had 9 of 36 (25%) in each repeat. Although the conserved N-linked glycosylation motif Asn-Xaa-Ser/Thr was present in two locations on each protein, only one of these sites was present in each of the expressed recombinant proteins. Homologous serine- and threonine-rich motifs were identified in the repeat regions of *E. chaffeensis* P120, *E. canis* P140, HGE agent P130, and *Streptococcus parasangius* Fap1 (Fig. 6).

Glycosidase analysis. Incubation of *E. canis* and *E. chaffeensis* P120 recombinant proteins with *N*-glycosidase F (N-linked glycans), *O*-glycosidase (Galβ1→3GalNAc), neuraminidase (sialic acid residues), β-galactosidase (β-linked 1-4 terminal galactose), or glucosaminidase (β-linked *N*-acetylglucosamine) did not result in deglycosylation as determined by a reduction in the electrophoretic mobilities of the recombinant proteins as evaluated by SDS-PAGE.

Lectin binding. Recombinant and native *E. chaffeensis* and *E. canis* proteins did not react with the lectins specific for conserved glycan core structures that were tested.

Monosaccharide composition and carbohydrate/protein ratio. Monosaccharides attached to the recombinant glycoproteins were identified by using gas chromatography. The molar percentages of the monosaccharides and the estimated molar ratios of carbohydrate to protein represented on *E. chaffeensis* P120 and *E. canis* P140 are shown in Table 2. Carbohydrate was not detected on the GST protein.
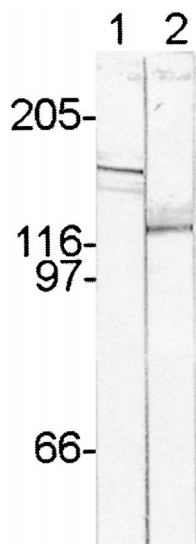

FIG. 3. Immunoblot showing electrophoretically separated proteins of *E. canis* and *E. chaffeensis* organisms, demonstrating the reaction of *E. canis* P140 (lane 1) and *E. chaffeensis* P120 (lane 2) proteins with antisera produced against the corresponding recombinant proteins. The recognized ehrlichial proteins are approximately twice as large as those predicted by the amino acid sequences encoded by the genes. Numbers on the left are molecular masses in kilodaltons.
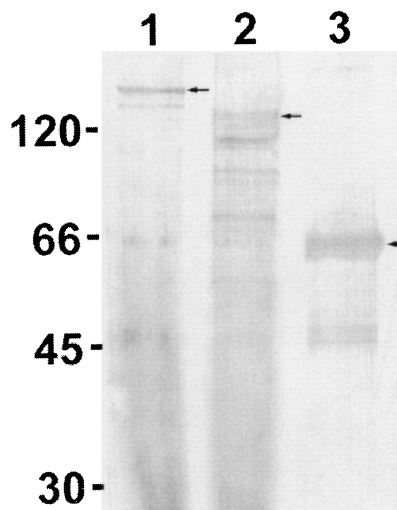

FIG. 4. Detection of carbohydrate on the expressed recombinant fusion proteins after biotin labeling of oxidized sugars. Lane 1, *E. canis* P140 (arrow); lane 2, *E. chaffeensis* P120 and proteolytic degradation products (arrow); lane 3, *E. chaffeensis* P120 two-repeat units (arrow). Numbers on the left are molecular masses in kilodaltons.

TABLE 1. Predicted and observed molecular masses and serine-
and-threonine contents of ehrlichial glycoproteins and repeat regions

Protein	Molecular mass (kDa)		Ser/Thr content (n/%)	
	Predicted	Observed[a]	Entire open reading frame	Repeat region
E. chaffeensis P120	61	100/120	91/16	70/20
E. canis P140	73	112[b]/140	147/21	126/25
HGE agent P100	61	100/ND[c]	65/11	30/11
HGE agent P130	66	130/ND	156/25	80/30

[a] The first value is for the recombinant expressed as detected by immunoblotting; the second value is for the native protein.
[b] Fourteen-repeat region (78% of the open reading frame).
[c] ND, no data.

DISCUSSION

The existence of prokaryotic glycoproteins has been demonstrated in previous studies which characterized S-layer glycoproteins and other membrane-associated glycoproteins in prokaryotes (6, 7). The larger-than-predicted molecular masses of recombinant and native *E. canis* P140 and *E. chaffeensis* P120 appear to be a result of an unusual posttranslational O-linked glycosylation. The two-repeat region that was expressed from the *E. chaffeensis* P120 gene did not contain any N-linked glycosylation sites and still exhibited 1.6 times the predicted molecular mass of the protein. There was one N-linkage site for glycosylation in the other larger P120 and P140 recombinant proteins, but glycans could not be removed by *N*-glycosidase F, an enzyme specific for N linkages. Together, this evidence supports the attachment of glycans to P120 and P140 via O linkages.

The recombinant proteins expressed in *E. coli* were found to have three carbohydrate residues, i.e., glucose, galactose, and xylose. The compositional analysis did not reveal any conserved core sugars such as *N*-acetylglucosamine, *N*-acetylgalactosamine, or mannose. The sugar composition observed in the recombinant proteins suggests an unusual glycan modification. Others have reported O-linked glucose moieties on proteins

FIG. 6. Clustal alignment of shared homologous serine-rich motifs within the repeat units of *S. parasangius* Fap1, *E. canis* P140, *E. chaffeensis* P120, and HGE agent P130. Shared identical amino acids are boxed, and potential glycosylation sites in each motif are shown in boldface.

such as epidermal growth factor, tissue plasminogen activator, and coagulation factors VII, IX, and XII. Xylose-glucose O linkages to serine residues have been described previously and appear to be similar to the composition and linkage observed on the ehrlichial recombinant proteins (8). A compositional analysis of the sugars found on recombinant P120 and P140 indicated that glucose was the primary residue on the recombinant proteins. There were differences in the molar percentage of sugars on each recombinant protein, a manifestation of the microheterogeneity that exists between the two glycoproteins. These findings also suggest that glucose may be the primary attachment via O linkage to serine or threonine. A more detailed structural analysis would elucidate the exact nature of the sugar chains or monosaccharide linkages.

We attempted to further characterize the glycan structure by utilizing lectins and glycosidases. Lectins and glycosidases recognize very specific motifs, and the negative results that we obtained with both approaches confirmed that the glycosylation on the native and recombinant proteins does not contain common core structures often found in N- or O-linked eukary-
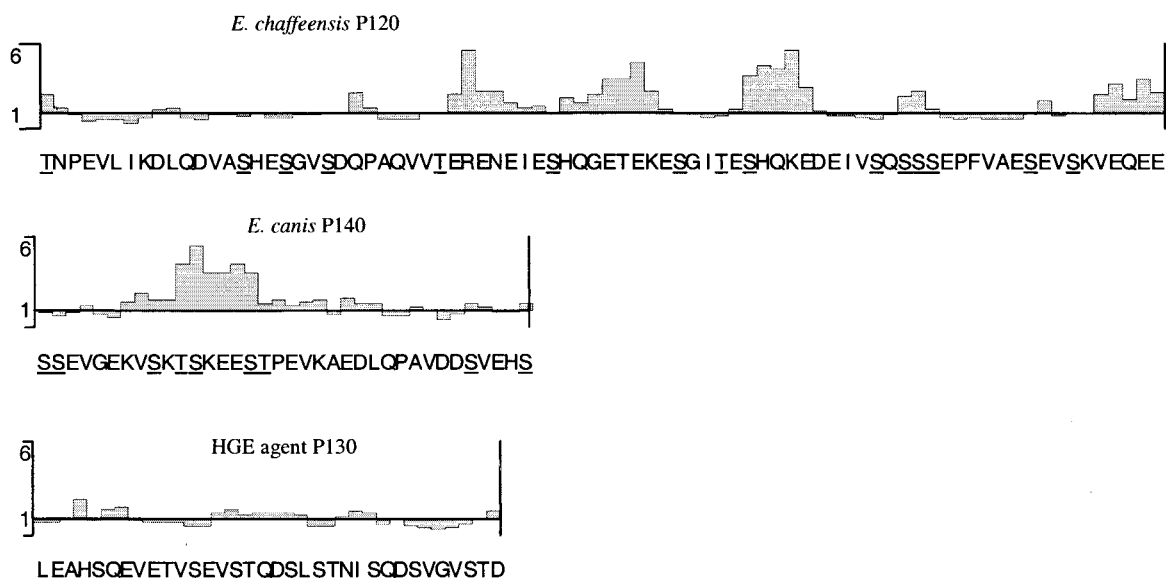

FIG. 5. Schematic representation of predicted surface-exposed regions as determined by an *Emini* surface probability plot of *E. chaffeensis* P120, *E. canis* P140, and HGE agent P130 repeat regions. Potential O-linked glycosylation sites within each repeat unit are underlined. A value of >1 represents an increased probability of surface location.

TABLE 2. Carbohydrate compositional analysis of recombinant
E. chaffeensis P120 and *E. canis* P140 and estimated
carbohydrate-to-protein molar ratio

Monosaccharide	mol % in:	
	E. chaffeensis P120	*E. canis* P140
Glucose	64.5	92.4
Galactose	34	4.5
Xylose	1.5	3.1
Molar ratio (carbohydrate/protein)	18:1	17:1

otic glycans. Lectins that recognize *N*-acetylglucosamine or *N*-acetylgalactosamine failed to react with the recombinant or native *E. chaffeensis* P120 and *E. canis* P140. Lectins and glycosidases are most commonly used to characterize N- or O-linked glycoproteins, which contain conserved core structures. Characterization with lectins becomes more difficult when glycans lack these core structures. Our findings and those of others suggest that some prokaryotic glycoproteins exhibit unconventional glycan structures.

Our finding that the recombinant proteins were glycosylated by *E. coli* was contrary to previous reports. It has generally been accepted and demonstrated that *E. coli* is unable to glycosylate various proteins that are glycoproteins in a native state (1). However, a glycosylated pilus of *E. coli* has been reported and provides evidence that the organism is capable of protein glycosylation to some extent (13). Our observations of larger-than-predicted sizes of recombinant proteins in *E. coli* were also similar to those when recombinant proteins from the HGE agent were expressed (12). This evidence supports the concept that specific glycosylation modifications are performed by *E. coli*. The recent sequencing of the entire *E. coli* genome has also led to the identification of glycotransferase enzymes that may be involved in protein glycosylation (2). The inability of *E. coli* to glycosylate many eukaryotic recombinant proteins points to possible differences in biosynthetic pathways that may exist between eukaryotes and prokaryotes, or perhaps only proteins transported to the plasma membrane are glycosylated. It is clear that *E. coli* can glycosylate some proteins and may preferentially glycosylate certain proteins of prokaryotic origin.

The repeat regions of the *E. chaffeensis* and *E. canis* proteins appear to be targeted for glycosylation. Other large proteins (P100 and P130) with repeat regions cloned from the HGE agent also demonstrate larger-than-predicted molecular masses (12) and seem likely to be glycosylated. Carbohydrate has also been detected on the Fap1 fimbria-associated adhesin protein of *S. parasanguis* (15). We identified several similar serine-rich motifs in the repeat regions of HGE agent P130, *E. chaffeensis* P120, *E. canis* P140, and *S. parasanguis* Fap1. The conservation of these serine/threonine motifs among the four repeat regions suggests that they may be the primary targets for glycosylation. The number of serine and threonine residues in each protein correlated with the magnitude of the increased molecular sizes of P130 and P140 compared to P120 of *E. canis* and P100 of the HGE agent. *E. canis* P140 and HGE agent P130 exhibit a higher ratio of observed-to-predicted molecular mass than *E. chaffeensis* P120, and these proteins have 5 to 14% more serine and threonine residues, respectively. O-linked glycosylation most commonly occurs on serine and threonine residues, although tyrosine is occasionally glycosylated (1). The high serine content found in the repeat regions of these proteins is most likely where the glycans are attached.

Expression of the two-repeat region from *E. chaffeensis* P120 resulted in a protein 1.6 times larger than that predicted by the amino acid sequence. Extrapolation from the mass of the two-repeat region to the entire repeat region completely accounted for the increased molecular mass of the recombinant *E. chaffeensis* P120. This calculation suggests that glycosylation of the recombinant protein occurs in the serine-rich repeat region of the protein. The recombinant *E. canis* P140 protein exhibited twice the molecular mass predicted by the amino acid sequence. This difference was also more exaggerated than the increased molecular mass of *E. chaffeensis* P120. *E. canis* P140 does have more potential glycosylation sites in the repeat region; in addition, the repeat region of *E. canis* P140 is much larger than the repeat region of *E. chaffeensis* P120. These differences may account for the larger increase in the molecular mass of *E. canis* P140. The recombinant *E. chaffeensis* P120 was slightly smaller than the native protein recognized by antiserum raised against the recombinant protein. This suggests that glycosylation of the native glycoproteins may be slightly different in the number of sugars or character of the sugar residues.

The similarities among P120, P140, HGE agent P130, and the fimbria-associated adhesin protein Fap1 of *S. parasangius* suggest that these proteins may have similar locations and functions. The majority of glycosylated eubacterial proteins described thus far are S-layer proteins or other proteins associated with the outer membrane (7). Glycosylation of P120 and P140 further supports previous data that P120 is an *E. chaffeensis* surface protein. The function of *E. chaffeensis* P120 and *E. canis* P140 has not been determined, but *E. chaffeensis* P120 has been localized to the surface of *E. chaffeensis* and the extracellular fibrillar matrix in the morula by immunoelectron microscopy (9a). This observation is consistent with the case for other prokaryotic glycoproteins which are surface or membrane associated (6). In addition, it appears that in prokaryotes glycosylation occurs at the outer side of the plasma membrane (5) where surface proteins would be transported prior to their final destination. An S-layer morphology has been observed on the surface of *Rickettsia prowazekii* (9), and a protein believed to be an S-layer protein has been described (3). Most S-layer proteins described thus far contain a signal sequence and a high proportion of acidic amino acids; however, the S-layer protein identified in *R. prowazekii* does not contain a signal sequence, nor does *E. chaffeensis* P120 or *E. canis* P140, but these proteins contain a high proportion of acidic amino acids as often found in S-layer proteins. These similarities suggest that P120 and P140 could be analogous proteins in *E. chaffeensis* and *E. canis*.

Identification of ehrlichial P120, P140, P130, and P100 and steptococcal Fap1 based on their immunoreactivity clearly demonstrates that these glycoproteins are immunodominant (12, 15, 17, 18). Immunoreactivity of the native and recombinant proteins is strong, and further investigation is needed to determine the role that carbohydrate has in immune recognition of these proteins. The recombinant *E. chaffeensis* P120 and *E. canis* P140 have demonstrated usefulness as serodiagnostic antigens (17, 19). Consistent detection of antibodies in preconvalescent-phase sera by using recombinant *E. chaffeensis* P120 (16) suggests that the attached glycans may elicit a T cell-independent immune response resulting in early production of antibody directed at these proteins.

REFERENCES

1. **Beguin, P., J. Miller, and J.-P. Aubert.** 1993. Cellulose degradation by *Clostridium thermocellum*: from manure to molecular biology. FEMS Microbiol. Lett. **100:**523–528.
2. **Blattner, F. R., G. Plunkett III, C. A. Bloch, N. T. Perna, V. Burland, M. Riley, et al.** 1997. The complete genome sequence of *Escherichia coli* K12. Science **277:**1453–1462.
3. **Carl, M., M. E. Dobson, W.-M. Ching, and G. A. Dasch.** 1990. Characterization of the gene encoding the protective paracrystalline-surface-layer protein of *Rickettsia prowazekii*: presence of a truncated identical homolog in *Rickettsia typhi*. Proc. Natl. Acad. Sci. USA **87:**8237–8241.
4. **Dobos, K. M., K. Khoo, K. Swiderek, R. Brennan, and J. T. Belisle.** 1996. Definition of the full extent of glycosylation of the 45-kilodalton glycoprotein of *Mycobacterium tuberculosis*. J. Bacteriol. **178:**2498–2506.
5. **Lechner, J.** 1989. Structure and biosynthesis of prokaryotic glycoproteins. Annu. Rev. Biochem. **58:**173–194.
6. **Messner, P.** 1997. Bacterial glycoproteins. Glycoconj. J. **14:**3–11.
7. **Moens, S., and J. Vanderleyden.** 1997. Glycoproteins in prokaryotes. Arch. Microbiol. **168:**169–175.
8. **Nishimura, H., S. Kawabata, W. Kisiel, S. Hase, T. Ikenaka, T. Takao, Y. Shimonishi, and S. Iwanaga.** 1989. Identification of a disaccharide (Xyl-Glc) and a trisaccharide (Xyl$_2$-Glc) O-glycosidically linked to a serine residue in the first epidermal growth factor-like domain of human factors VII and IX and protein Z and bovine protein Z. J. Biol. Chem. **264:**20320–20325.
9. **Palmer, E. L., L. P. Mallavia, T. Tzianabos, and J. F. Obijeski.** 1974. Electron microscopy of the cell wall of *Rickettsia prowazekii*. J. Bacteriol. **118:**1158–1166.
9a.**Popov, V. L., X. J. Yu, and D. H. Walker.** The 120-kDa outer membrane protein of *Ehrlichia chaffeensis*: preferential expression on dense-core cells and gene expression in *Escherichia coli* associated with attachment and entry. Microb. Pathog., in press.
10. **Rikihisa, Y.** 1991. The tribe *Ehrlichieae* and ehrlichial diseases. Clin. Microbiol. Rev. **4:**286–308.
11. **Sambi, V., C. Stefanelli, and R. Cevenini.** 1992. Detection of glycoproteins in *Borrelia burgdorferi*. Arch. Microbiol. **157:**205–208.
12. **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.
13. **Tomoeda, M., M. Inuzuka, and T. Date.** 1975. Bacterial sex pili. Prog. Biophys. Mol. Biol. **30:**23–56.
14. **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.
15. **Wu, H., P. Mintz, M. Ladha, and P. M. Fives-Taylor.** 1998. Isolation and characterization of Fap1, a fimbriae-associated adhesin of *Streptococcus parasanguis* FW213. Mol. Microbiol. **28:**487–500.
16. **Yu, X. J., P. Crocquet-Valdes, L. C. Cullman, V. L. Popov, and D. H. Walker.** 1999. Comparison of *Ehrlichia chaffeensis* recombinant proteins for diagnosis of human monocytotropic erhlichiosis. J. Clin. Microbiol. **37:**2568–2575.
17. **Yu, X. J., P. 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.
18. **Yu, X. J., P. Crocquet-Valdes, and D. H. Walker.** 1997. Cloning and sequencing of the gene for a 120-kDa immunodominant protein of *Ehrlichia chaffeensis*. Gene **184:**149–154.
19. **Yu, X. J., J. W. McBride, C. M. Diaz, and D. H. Walker.** Molecular cloning and characterization of the 120-kilodalton protein gene of *Ehrlichia canis* and serodiagnostic application for canine ehrlichiosis. J. Clin. Microbiol., in press.