Cryptosporidium, a waterborne enteric parasite, is a frequent cause of diarrheal disease outbreaks worldwide. Thus far, the few antigens shown to be important for attachment to and invasion of the host cell by Cryptosporidium are all mucin-like glycoproteins. In order to identify other antigens that could be important for Cryptosporidium host-parasite interactions, the Cryptosporidium genome databases were mined for other mucin-like genes. A single locus of seven small mucin sequences was identified on chromosome 2 (CpMuc1 to -7). Reverse transcriptase PCR analysis demonstrated that all seven CpMucs were expressed throughout intracellular development. CpMuc4 and CpMuc5 were selected for further investigation because of the significant sequence divergence between Cryptosporidium parvum and C. hominis alleles. Rabbit anti-CpMuc5 and -CpMuc4 antibodies identified several polypeptides in C. parvum lysates, suggestive of proteolytic processing of the mucins. All polypeptides were larger than the predicted molecular weight, which is suggestive of posttranslational processing, most likely O-glycosylation. In immunofluorescence assays, both anti-CpMuc4 and -CpMuc5 antibodies reacted with the apical region of sporozoites and revealed surface-exposed epitopes. The antigens were not shed during excystation but did partition into the aqueous phase of Triton X-114 extractions. Consistent with a role in attachment and invasion, CpMuc4 and CpMuc5 could be detected binding to fixed Caco-2A cells, and anti-CpMuc4 peptide antibodies inhibited Cryptosporidium infection in vitro. Sequencing of CpMuc4 and CpMuc5 from C. hominis clinical isolates identified several polymorphic alleles. The data suggest that these antigens are integral for Cryptosporidium infection in vitro and may be potential vaccine candidates.
proteins, and p23, predicted to contain mucin-type O-glycosylation sites, can be purified by Helix pomatia lectin affinity chromatography (A. M. Cevallos and H. Ward, unpublished data), suggesting that it is O-glycosylated. The oligosaccharides decorating these mucins exhibit exposed T [Gal(1-3)-GalNAc]-1-3-Ser/Thr) determinants that are normally cryptic on mammalian cells because of additional carbohydrate decoration (48).

Publication of the Cryptosporidium genome databases permitted identification of genes encoding other mucin-like glycoproteins (1, 51). A text search of the C. parvum genome database for “mucins” identified 31 genes (34). Among these were seven genes clustered on a single locus on chromosome 2, indicative of coordinated expression and/or biological function. The orthologous genes in C. hominis were significantly divergent from the C. parvum genes, which raised the possibility that these loci might be polymorphic among Cryptosporidium isolates and might be targets of selective immune pressure. In this study, we describe a preliminary investigation of the products of the most polymorphic genes on this locus, CpMuc4 and CpMuc5.

**MATERIALS AND METHODS**

Parasites. C. parvum Iowa isolate oocysts were purchased from Bunch Grass Farm (Deary, ID). C. parvum parasite lysates were generated as described previously (8). For collection of “shed” or exocytosed proteins, hypochlorite-treated oocysts were excysted in phosphate-buffered saline (PBS) at 37°C for 1 h, the supernatant was collected, and protease inhibitors (protease inhibitor cocktail set g) were included as a positive control for PCR.

For PCR amplification of the CpMuc4 and CpMuc5 genes from clinical samples, DNA samples were extracted from deidentified stool samples from Bangladeshi (22) and Indian (2, 28) patients infected with Cryptosporidium. To amplify CpMuc4 and CpMuc5, a nested PCR strategy was performed using Platinum Pfu polymerase (Invitrogen) and primers specific for the published CpMuc4 and CpMuc5 sequences (Table 2). Primers specific for the published CpMuc5 sequence (Table 2) were designed to amplify CpMuc5 from C. hominis Ile and Ile clinical isolates.

Cloning and expression of CpMuc4 and -5. The coding sequences of CpMuc4 and CpMuc5 minus the putative signal sequences, were amplified from C. parvum Iowa oocyst DNA and cloned into the pET32XaLIC vector following the manufacturer’s directions for ligation-independent cloning (Novagen, Madison, WI). Fusion proteins expressed from this vector contain an internal S tag, internal and C-terminal His tags, and N-terminal thioridoxin tag sequences. The vectors were transformed into Escherichia coli BL21 cells, and expression was induced with 1 mM isopropyl-β-D-thiogalactopyranoside (IPTG). The fusion proteins were purified by metal affinity chromatography (Talon; Clontech, Mountain View, CA) and resolved in preparative SDS-PAGE gels as described previously (9).

Antibody production. Rabbits were immunized as described previously, using recombinant CpMuc4 and CpMuc5 fusion proteins exsiced from SDS-PAGE gels (Harlan Bioproducts for Science, Indianapolis, IN) (9). Immunoglobulin G (IgG) was isolated from pooled preimmune rabbit sera and rabbit anti-CpMuc4 and -CpMuc5 sera by protein A affinity chromatography (Pierce Biosciences, Rockford, IL). Antipeptide antibodies were generated by immuunizing rabbits with the CpMuc4 peptide (NPFPAGVSLSSPRPR)116 coupled to keyhole limpet hemocyanin. The anti-CpMuc4- peptide IgG was affinity purified from the sera on a peptide Sepharose column (Dr. Dragonfly Sciences, Wellesley, MA).

Immunooassays. Parasite lysates were resolved by SDS-PAGE on 12% acrylamide gels and transferred to polyvinylidene difluoride membranes. Western blotting was performed as previously described (9). For immunofluorescence assays, oocysts were excysted in eight-well chamber slides that had been coated with poly-l-lysine, and sporozoites were allowed to glide on the slides for 1 h at 37°C. Slides were rinsed with PBS and fixed for 10 min with freshly made 1% paraformaldehyde. Half of the samples were permeabilized with 0.5% TX-100 in PBS. The slides were blocked with PBS containing 2% normal goat serum in PBS and then warmed to 37°C. Slides were rinsed with PBS and fixed for 15 min at 37°C. The proteins were acetone precipitated overnight at –20°C and resuspended in 1× sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) sample buffer.

**PCRs.** For reverse transcriptase PCR (RT-PCR), Caco-2A epithelial cells were infected with C. parvum oocysts for 6, 12, 24, 48, and 72 h, and total RNA was extracted with RNeasy kits (Qiagen, Hercules, CA). The mucin transcripts were amplified by RT-PCR, using Superscript III RT (Invitrogen), HotstarTaq master mix (Qiagen), and the primers described in Table 1 (GPMuc1 through CpMuc7, forward [F] and reverse [R]), targeted to the putative coding sequence of each mucin. g/p40/15 was amplified as a positive control, using published primers (9). DNA isolated from Iowa oocysts (Gnome kit; QI OGene, Irvine, CA) was included as a positive control for PCR.

**TABLE 1. Primers for RT-PCR**

| Primer | Sequence (5'-3') | Product size (bp) |
|--------|------------------|-------------------|
| GPMuc1-F | AAG TTC AGG TCA CTC CTA ACA ACA | 407 |
| GPMuc1-R | TCA GGT TCT TCC AAT CCT AAC A | 582 |
| GPMuc2-R | TTT CAG TCT TTC TAG TAG TGG T | 701 |
| GPMuc3-R | TCA TTT GAG CAT AAG AAG AAG AGA | 615 |
| GPMuc4-F | GCT GCA AGC GGT GTT GAC GAA | 512 |
| GPMuc4-R | GAA GGT GAA GGA GGA GTT CTG GA | 402 |
| GPMuc5-F | ATC TGC GAT TTG TTA CTT T | 546 |
| GPMuc5-R | CCG CTC TAC CAC CAT TAC CA | 765 |

**TABLE 2. Primers for nested PCR of clinical samples**

| Target | Primer | Primer sequence (5'-3') |
|--------|--------|-------------------------|
| ChMuc4 | Primary F | GAA GCA TGC GAA TTA AAA CTT T |
| Primary R | TGT AGT TAG AAA AAG AGC GAG AAT TG |
| Nested F | CAC ATT CAA TCA GTC TGG CA |
| Nested R | CAT TGA TGG TGG AAA CCA CA |
| ChMuc5 | Primary F | CAA TTC TCG CTT TCT CAA TCT AA |
| Primary R | CTC ATG AAC ATT GTA ACA CAC AAA G |
| Nested F | GAC GAG TAT TAA ACT TGG AGA C |
| Nested R | CAT CTA ACT TTA AAC AAT AGC AGT |
| ChMuc5 | Primary F | CAA TTT TCG CTT TCT CAA CTA |
| Primary R | ACT GCT CTT AAC CTT ACT TTA ACA |
| Nested F | GGT TGT AAA TTA TGA ACT CCT GTC |
| Nested R | GGC AAT CTC ACT AGT ACT TAC ACA |

For PCR amplification of the CpMuc4 and CpMuc5 genes from clinical samples, DNA samples were extracted from deidentified stool samples from Bangladeshi (22) and Indian (2, 28) patients infected with Cryptosporidium. To amplify CpMuc4 and CpMuc5, a nested PCR strategy was performed using Platinum Pfu polymerase (Invitrogen) and primers specific for the published CpMuc4 and CpMuc5 sequences (Table 2). Primers specific for the published CpMuc5 sequence (Table 2) were designed to amplify CpMuc5 from C. hominis Ile and Ile clinical isolates.
in any other apicomplexan genome (3) or in GenBank (7). Alignment of the CpMucs (AlignX and VectorNTI; Invitrogen) did not identify any regions of homology between the members of the locus. Further analysis of the CpMucs with the program BlockMaker (http://blocks.fhcrc.org/blocks/make_blocks.html), which searches for small regions of homology among a group of sequences, did not identify any short regions of homology among the CpMucs.

As expected for mucin-like glycoproteins, all are rich in proline, serine, and threonine and have predicted mucin-type O-glycosylation sites (Table 3) (18). CpMuc5 has one predicted N-glycosylation site. CpMuc1, -3, and -6 are predicted to have signal anchors instead of signal peptides, but the Signal-P program that was used for these predictions (6) may not accurately predict cleavage sites on C. parvum proteins. In gp40, the actual signal peptide cleavage site was 11 amino acids downstream of the predicted site (9). None of the CpMucs are predicted to have glycosylphosphatidylinositol anchor sites or transmembrane domains. The same locus is present in the C. hominis genome (51), and the predicted amino acid sequences of the ChMuc1-7 genes exhibit significant sequence divergence from their C. parvum homologues, particularly the Muc4 and Muc5 sequences (Table 3).

**CpMuc1-7 are expressed throughout intracellular development.** To confirm that the CpMuc1-7 mucin genes were expressed during infection, RT-PCR was performed on RNA extracted from Caco-2A cells infected with C. parvum for 12 to 72 h. All seven mucins were expressed throughout intracellular development (Fig. 1), suggesting that these genes do encode parasite proteins and are not pseudogenes. Background bands for the reactions lacking RT indicate genomic DNA (gDNA) contamination of some of the RNA samples (Fig. 1, lanes 6 to 9).

Anti-CpMuc4 and anti-CpMuc5 IgGs recognize multiple bands in sporozoite lysates that are larger than the predicted molecular sizes of the proteins. Because the Muc4 and Muc5 genes exhibited the greatest degrees of sequence divergence between C. hominis and C. parvum, these antigens were selected for further analysis. Rabbit IgGs raised against E. coli recombinant CpMuc4 and -5 and the Muc4 peptide 102PNPF AGVSLSSPRPR116 were used to probe Western blots of excysted oocyst lysates (Fig. 2). The native antigens were present in very low abundance in excysted oocyst lysates; a total of 4 × 10⁶ oocysts/lane was needed to visualize CpMuc4 by Western blotting, and a total of 2 × 10⁶ oocysts/lane was needed for

### Table 3. In silico analysis of CpMuc1-7 gene products

| Gene product | CryptoDB gene ID | No. of amino acids | Predicted molecular mass (Da) | Amino acid composition (%) | No. of O-glycosylation sites | No. of N-glycosylation sites | Signal peptide or anchor | % Identity to ChMuc |
|--------------|-----------------|--------------------|-------------------------------|---------------------------|-----------------------------|-----------------------------|-------------------------|---------------------|
| CpMuc1      | cgD2_390        | 155                | 17,437                        | 7.74                      | 10.32                       | 7.1                         | 3                      | 0                   | Anchor             |
| CpMuc2      | cgD2_400        | 194                | 21,040                        | 7.73                      | 16.5                        | 10.31                       | 30                     | 0                   | Peptide            |
| CpMuc3      | cgD2_410        | 261                | 28,170                        | 7.28                      | 27.59                       | 3.45                        | 50                     | 0                   | Anchor             |
| CpMuc4      | cgD2_420        | 205                | 22,008                        | 9.76                      | 14.63                       | 3.9                         | 9                      | 0                   | Peptide            |
| CpMuc5      | cgD2_430        | 206                | 21,176                        | 14.56                     | 19.9                        | 4.85                        | 36                     | 1                   | Peptide            |
| CpMuc6      | cgD2_440        | 134                | 15,314                        | 10.45                     | 11.94                       | 3                           | 1                      | 0                   | Anchor             |
| CpMuc7      | cgD2_450        | 182                | 19,345                        | 6.04                      | 17.03                       | 5.5                         | 10                     | 0                   | Peptide            |

* As predicted by the NetO Glyc (18) and NetNGlyc (http://www.cbs.dtu.dk/services/NetNGlyc) programs.

* As predicted by the SignalP 3.0 program (6).
CpMuc5 (compared to 10^6 oocysts/lane for gp40/15). Preimmune IgG did not recognize any Cryptosporidium proteins (Fig. 2, lane 1). The anti-CpMuc5 IgG revealed two major bands, of 55 and 25 kDa, and two minor bands, of 75 and 32 kDa, for C. parvum oocyst lysates (Fig. 2, lane 2). The rabbit anti-CpMuc4 IgG consistently revealed a major band at 30 kDa and sometimes showed two or three minor bands, at 98, 64, and/or 25 kDa (Fig. 2, lane 3), for C. parvum lysates. The anti-CpMuc4-peptide IgG recognized the 30-kDa band (Fig. 2, lane 4), and recognition of this band could be inhibited with an excess of peptide (Fig. 2, lane 5).

CpMuc4 and CpMuc5 are soluble antigens that are not shed during excystation. Neither CpMuc4 nor CpMuc5 was shed from sporozoites into the excystation medium (Fig. 3, lanes 2). In TX-114 extraction and phase separation procedures, CpMuc4 and CpMuc5 were found in the aqueous phase (Fig. 3, lanes 3), as would be expected for secreted proteins with neither glycosylphosphatidylinositol anchors nor transmembrane domains. However, the 64-kDa CpMuc4 band did partition equally between the detergent and aqueous phases (Fig. 3, left panel, lanes 3 and 4), suggesting that some CpMuc4 may be associated with the membrane.

CpMuc4 and CpMuc5 localize to the apical surface of sporozoites and to an unidentified compartment within intracellular merozoites. In IFAs of paraformaldehyde-fixed sporozoites, CpMuc5 (Fig. 4, top panels, red fluorescence) and CpMuc4 (not shown) both localized to the apical region of sporozoites and were not shed during gliding. This is clearly illustrated by colocalization with gp15 (Fig. 4, top panels, green fluorescence), which is found on the sporozoite membrane and is shed in trails during gliding (9, 17, 42). In intracellular-stage organisms, CpMuc4 (Fig. 4, lower panels) and CpMuc5 (not shown) displayed a punctate reactivity within the merozoites. In both sporozoites and meronts, the localization of CpMuc4 and CpMuc5 was indistinguishable. Preimmune rabbit IgG did not react with sporozoites or intracellular-stage organisms (not shown).

Anti-CpMuc4 and anti-CpMuc5 antibodies reacted with paraformaldehyde-fixed, nonpermeabilized sporozoites, suggesting that these mucin antigens may have surface-exposed epitopes. To confirm this observation, anti-CpMuc4 and anti-CpMuc5 were colocalized with the anti-gp900 MAb 4G12 (45). gp900 is a microneme glycoprotein that does not have surface-exposed epitopes in intact sporozoites (35) and thus can serve...
as a control for surface reactivity. Permeabilized and nonpermeabilized sporozoites were probed with MAb 4G12 (Fig. 5, green fluorescence) and either anti-CpMuc4-peptide IgG or anti-CpMuc5 IgG (Fig. 5, red fluorescence). Images were deconvolved to improve resolution. Anti-CpMuc reagents reacted with the apical region of both nonpermeabilized and permeabilized sporozoites, in a punctate pattern (Fig. 5). In contrast, 4G12 reacted only with permeabilized sporozoites, in both apical and posterior regions, and did not colocalize with CpMuc4 or CpMuc5, suggesting that the mucin antigens are not present in micronemes (Fig. 5, lower panels).

Antibodies raised to a CpMuc4 peptide inhibit *Cryptosporidium* infection in vitro. Affinity-purified anti-CpMuc4-peptide antibodies were tested for the ability to inhibit *C. parvum*...
FIG. 6. Anti-Muc4-peptide antibody and F(ab) fragments inhibit *C. parvum* infection in vitro. (A) Data pooled from three independent experiments and analyzed by two-way analysis of variance (***, *P* < 0.001; **, *P* < 0.01; *, *P* < 0.05). (B) F(ab) fragments were added to the assay at 35 mg/ml. The experiment was performed once, and data were analyzed by two-tailed *t* test (**, *P* = 0.0014).

**FIG. 7.** CpMuc4 and CpMuc5 bind to Caco-2A epithelial cells in a dose-dependent and saturable manner. Data were pooled from two independent experiments. Error bars indicate standard errors of the means.

infection in vitro. Protein A-purified IgG from normal rabbits was run in parallel as a control. Neither anti-CpMuc4-peptide IgG nor control IgG inhibited oocyst excystation. *C. parvum* infection of intestinal epithelial cells was almost completely ablated in the presence of 20 μg/ml of antibody (Fig. 6) (*P* < 0.01) compared to the case with control IgG, and inhibition diminished with lower concentrations of antibody. *C. parvum* infection was also inhibited by F(ab) fragments prepared from the antipeptide IgG (*P* = 0.0014).

**CpMuc4 and CpMuc5 bind to fixed Caco-2A cells.** CpMuc4 and CpMuc5 from TX-114-extracted aqueous-phase proteins were detected binding to Caco-2A cells in a dose-dependent manner (Fig. 7). The control glycoprotein ovalbumin bound to cells at very low levels, although it was added to the cells at a higher concentration than that of the mucins. β-Galactosidase, another control glycoprotein, was not detected binding to Caco-2A cells at any concentration (not shown).

**ChMuc4 and ChMuc5 genes sequenced from clinical isolates exhibit extensive polymorphisms.** Because the Muc4 and Muc5 coding sequences were significantly different between the *C. hominis* and *C. parvum* loci (56.3 and 71% identity, respectively) (Table 3), we sequenced these genes from *Cryptosporidium* clinical samples to determine whether there was diversity at these loci between isolates. As part of another project, the species and subgenotypes of *Cryptosporidium* isolates from Bangladeshi patients were determined by restriction fragment length polymorphism (RFLP) at the small-subunit (SSU) rRNA locus (21) and sequencing or RFLP of the fragment length polymorphism (RFLP) at the small-subunit loci (24, 42, 49; K. Hira et al, unpublished data). However, the Muc4 genes from two IIk and three Iic isolates were significantly divergent from either the *C. parvum* or *C. hominis* sequence (see Fig. S1B in the supplemental material). The Muc4 Iic/IIk sequence was 63.4% homologous to the ChMuc4 sequence and 49.2% homologous to the CpMuc4 sequence, but all five samples were identical to each other at the nucleotide sequence level, except for Bang7, which had a silent single nucleotide polymorphism (SNP) (see Fig. S1 in the supplemental material).

The ChMuc5 gene sequences also exhibited polymorphisms among clinical samples (see Fig. S2 in the supplemental material). Three samples had one SNP but were otherwise identical to the published ChMuc5 sequence; this group included Ia, Ib, and Iik gp40/15 subgenotypes. The second group was also identical to the ChMuc5 sequence but lacked the SNP of the first group and had a single amino acid polymorphism (SAAP) at position 127 (P to Q). This group included five Bangladeshi subtypes (Ib, Id, Ie, If, and Iik). The third group consisted of two Bangladeshi Iik samples and four Indian Iic samples that had sequences that differed from CpMuc5 by six SAAPs (see Fig. S2B in the supplemental material). Finally, there was one Bangladeshi gp40/15 subgenotype Ie sample that shared 71% identity with CpMuc5 and 83% identity with ChMuc5 (see Fig. S2B in the supplemental material).

**DISCUSSION**

Despite the public health importance of *Cryptosporidium*, very few studies have investigated the host-parasite interactions of this pathogen, and very few parasite antigens have been characterized. The antigens that have been well investigated and implicated in the process of host cell invasion have been shown or predicted to have mucin-type O-glycosylation. The observation that the *Cryptosporidium* genomes contain 31 ORFs encoding putative mucin antigens serves to further highlight the importance of this class of antigens to host-parasite interactions. Of these, the CpMuc1-7 locus was particularly interesting for several reasons. The location of the genes to-
gether in a single locus suggested that their expression is regulated in a coordinated fashion and/or that these mucins may be integrated in the same biological process. Although coordinated function of multiple apical complex proteins has been reported for other apicomplexan proteins (4, 40), this phenomenon has not yet been observed in Cryptosporidium. Alignment of the C. parvum and C. hominis (ChMuc1-7) Muc1-7 loci identified extensive polymorphisms (Table 3), especially at the Muc4 (56.3% identity between C. hominis and C. parvum orthologs) and Muc5 (71% identity between orthologs) loci, suggesting that the gene products might be important virulence determinants subject to immune pressure. Recent identification of Toxoplasma virulence genes by quantitative trait loci mapping also identified highly polymorphic loci encoding rhoptry proteins (37, 43).

Antibodies to CpMuc4 and CpMuc5 identified several bands in oocyst lysates. The sizes of the bands are suggestive of proteolytic processing. For example, the 75-kDa Muc5 polypeptide may be the precursor to the 55- and 25-kDa bands, and the 98-kDa Muc4 band may be processed into the 64- and 30-kDa bands. The observation that the anti-Muc4-peptide antibody reveals only the 30-kDa Muc4 band further supports this hypothesis, although this may be just a reflection of the low abundance of the higher-molecular-weight polypeptides (Fig. 2). However, all of these polypeptide bands are significantly larger than the predicted molecular sizes of the CpMuc4 (22 kDa) and CpMuc5 (21 kDa) deduced amino acid sequences, strongly suggesting that the polypeptides may be modified by O-linked glycosylation, as predicted (Table 3). Thus, an alternative explanation is that the multiple bands represent differentially glycosylated forms of the antigens. This issue will have to be resolved by glycan analysis, pulse-chase analysis, and proteomic approaches in future experiments.

The combined data reported here strongly support a role for CpMuc4 and CpMuc5 in sporozoite attachment to and invasion of host cells. Both antigens appear to have surface-exposed epitopes and localize to the apical region of the parasite. Apical complex antigens in other apicomplexans are known to facilitate attachment to the cell surface, gliding motility, and subsequent invasion and intracellular development of the parasite (39, 41). The lack of colocalization with the microneme glycoprotein gp90 suggests that CpMuc4 and CpMuc5 are located in different apical compartments (rhoptries or dense granules) or in a different subgroup of micronemes. Currently, there are no markers for Cryptosporidium rhoptries and dense granules; identification of the CpMuc4 and CpMuc5 compartment will require localization by immunoelectron microscopy.

Both antigens bind to intestinal epithelial cells in a dose-dependent manner that is suggestive of a ligand-receptor interaction. However, the binding assays were done with TX-114 aqueous-phase sporozoite extract, which contains many proteins and possibly even the other CpMuc antigens. It is therefore possible that CpMuc4 and CpMuc5 do not interact directly with a host receptor(s) but bind indirectly via other sporozoite proteins. The observation that anti-CpMuc4-peptide antibodies inhibit infection in vitro also supports an important role for this antigen. Since F(ab) fragments derived from these antibodies also inhibit infection, this effect is not due to nonspecific aggregation of sporozoites. It remains unclear how soluble antigens such as these would mediate an interaction between the sporozoite and the host cell. It is possible that like the gp40/gp15 antigen complex (29), these antigens associate with other membrane-bound antigens. An association of CpMuc4 and CpMuc5 with membrane components was not seen in the TX-114 phase separations, but this may have been due to a weak or transitory interaction.

The discovery of different CpMuc4 and CpMuc5 alleles in C. hominis clinical samples is also indicative of the importance of the gene products to host-parasite interactions. To date, the Cppg40/15 locus (also known as Gp60) is the only other Cryptosporidium locus to exhibit significant polymorphisms that translate into extensive amino acid changes that could change immune recognition of the antigen. A common method of categorizing Cryptosporidium isolates from humans and animals, and the method used to categorize the Indian and Bangladeshi samples (K. Hira and H. Ward, unpublished observations), is to determine the species by RFLP at the SSU rRNA locus (21) and then to identify the gp40/15 allele by sequencing or RFLP analysis (24, 42, 49). This has led to the identification of at least eight C. hominis gp40/15 alleles, Ia through Ig, and nine C. parvum gp40/15 alleles, Ia through Ij (50). The confusing aspect of this nomenclature system is that none of the methods used to identify Cryptosporidium species distinguish between anthropo- notic and zoonotic isolates, resulting in “subgenotypes” such as Iic and Ile (and probably IIk) that have been found only in humans and have been termed “anthropo- notic C. parvum” (50). Interestingly, the Muc4 and Muc5 loci in these isolates exhibited the most diversity. In all Iic samples and two of three IIk samples, the Muc4 allele was unique, being significantly different from either the C. parvum or C. hominis allele. In contrast, the Muc5 alleles in these samples were similar to the C. parvum allele, with only six SAAPs and the elongation of the polyserine domain by two serines. The third IIk sample (Bang9) carried C. hominis alleles at both Muc4 and Muc5 (one SAAP) loci. A fourth IIk sample, from which only Muc5 was sequenced, also carried the C. hominis allele but was in the group that lacked the SAAP. A unique Muc5 allele was found in one of two Ie samples (see Fig. S2B in the supplemental material). Although it appears that the CpMuc4 and CpMuc5 alleles segregate differently from the Cppg40/15 alleles, sequencing of the complete CpMuc1-7 loci from geographically diverse isolates will be required to determine the relationship among these loci and between these and the Cppg40/15 locus.

In summary, we have begun an investigation of two polymorphic mucin genes that appear to be integral to Cryptosporidium infection. These data add to the body of evidence suggesting that this pathogen relies on mucin antigens for attachment to and invasion of host cells. Further investigation of these and other mucin genes identified in the Cryptosporidium genome databases are essential for understanding the biology of this parasite and for the development of vaccines to prevent this disease.

ACKNOWLEDGMENTS

This research was funded in part by a 2004 developmental grant to R.M.O. from the Life Span/Tufts/Brown Center for AIDS Research (grant P30AI042853) from the National Institute of Allergy and Infectious Diseases and by Public Health Service grant R21AI070037 (R.M.O.) from the National Institute of Allergy and Infectious Dis-
REFERENCES

1. Abrahamsen, M. S., T. J. Templeton, S. Enomoto, J. E. Abrahante, G. Zhu, C. A. Lanco, M. D. Deng, C. Liu, G. Wu, S. L. Tsipori, G. A. Buck, P. Xu, A. T. Bankier, P. B. Dear, B. A. Konfortov, H. F. Spriggs, L. Iyer, V. Anantharaman, L. Aravind, and V. Kapur. 2004. Complete genome sequence of the apicomplexan, Cryptosporidium parvum. Science 304:441–445.

2. Ajajhamp, S. P., P. Gladstone, D. Selvadurain, J. P. Muluil, H. Ward, and G. Kang. 2007. Molecular and spatial epidemiology of cryptosporidiosis in children in a semirural community in South India. J. Clin. Microbiol. 45:915–920.

3. Alpersrochea, C., M. Heiges, H. Wang, Z. Wang, S. Fischer, P. Rhodes, J. Miller, E. Kraemer, C. J. Stoeckert, Jr., D. S. Roos, and J. C. Kissinger. 2007. ApIDB: integrated resources for the apicomplexan bioinformatics resource center. Nucleic Acids Res. 35:D427–D430.

4. Babu, D. L., K. T. Andrews, R. F. Waller, D. S. Roos, R. S. Howard, B. S. Crabb, and A. F. Cowman. 2000. RAIP controls rhoptry targeting of RAIP2 in the malaria parasite Plasmodium falciparum. EMBO J. 19:2435–2443.

5. Barnes, D. A., A. Bonnin, J. X. Huang, L. Gousset, J. Wu, J. Gut, P. Doyle, J. F. Duheater, J. T. Hume, and D. M. Bishop. 1998. A novel multi-domain mucin-like glycoprotein of Cryptosporidium parvum mediates invasion. Mol. Biochem. Parasitol. 96:93–110.

6. Bendtsen, J. D., H. Nielsen, G. von Heijne, and S. Brunak. 2004. Improved prediction of signal peptides: SignalP 3.0. J. Mol. Biol. 340:783–795.

7. Benson, D. A., I. Karsch-Mizrachi, D. J. Lipman, J. Ostell, and D. L. Wheeler. 2005. GenBank. Nucleic Acids Res. 33:D34–D38.

8. Ceavallos, A. M., N. Bhat, R. Verdon, D. H. Hamer, B. Stein, S. Tzipori, M. E. Jerome, and H. D. Ward. 2008. Protein coding gene diversity at the highly polymorphic D427–D430. J. Mol. Biol. 380:558–580.

9. Ceavallos, A. M., X. Zhang, M. K. Waldor, S. Jaison, X. Zhou, S. Tzipori, M. R. Neutra, and H. D. Ward. 2000. Molecular cloning and expression of a gene encoded by mucus glycoprotein of Cryptosporidium parvum gp40 and gp15. Infect. Immun. 68:4018–4016.

10. Chen, X. M., S. P. O’Hara, B. Q. Huang, J. B. Nelson, J. J. Lin, G. Zhu, H. D. Ward, and N. F. La Russo. 2004. Apical organelle discharge by Cryptosporidium parvum is temperature, cytoskeleton, and intracellular calcium dependent and required for host cell invasion. Infect. Immun. 72:6806–6816.

11. Colford, J. M., Jr., I. B. Tager, A. M. Hirozawa, G. F. Lemp, T. Aragon, and C. Petersen. 1996. Cryptosporidiosis among patients infected with human immunodeficiency virus. A factor related to symptomatic infection and survival. Am. J. Epidemiol. 144:807–816.

12. Doering, T. L., T. P. Englend, and G. W. Hart. 2003. Detection of glycoporphin A antibodies on proteins. In J. E. Coligan, B. M. Dunn, D. W. Speyer, and J. M. Weng (ed.). Current protocols in protein science. John Wiley and Sons, Inc., Hoboken, N.J.

13. Dubremetz, J. F., N. Garcia-Reguet, V. Conseil, and M. N. Fourmaux. 2008. Protein coding gene sequence diversity at the highly polymorphic Cryptosporidium parvum spv. isolates from South Africa. Infect. Immun. 76:3881–3890.

14. Kambo, L., T. G. Clarke, R. D. Naure, J. A. Lumadue, H. R. Dahlman, P. C. Belitosis, R. E. Chaisson, and C. L. Sears. 1998. Cryptosporidiosis in patients with AIDS: correlates of disease and survival. Clin. Infect. Dis. 27:536–542.

15. McGowan, L. S., A. Hawkins, and I. V. Weller. 1993. The natural history of cryptosporidial diarrhoea in HIV-infected patients. AIDS 7:349–354.

16. Molbak, K., N. Hojlyng, A. Gottschau, J. C. Sa, L. Ingholt, A. P. da Silva, and P. Aaby. 1995. Cryptosporidiosis in infancy and childhood mortality in Guinea Bissau, West Africa. Br. Med. J. 307:417–420.

17. Muthusamy, D., S. S. Rao, S. Ramani, B. Monica, I. Barjeree, O. C. Abrah- man, D. C. Mathai, B. Primrose, J. A. Alai, C. A. Wanke, H. D. Ward, and G. Kang. 2006. Multicus genotyping of Cryptosporidium sp. isolates of human immunodeficiency virus-infected individuals in South India. J. Clin. Microbiol. 44:632–634.

18. O’Connor, R., M. J. Wanyiri, A. M. Cevallos, J. W. Priest, and H. D. Ward. 2007. Cryptosporidium parvum glycoprotein gp40 localizes to the sporozoite surface by association with gp15. Mol. Biochem. Parasitol. 158:867–83.

19. Parida, M. L., Xiao, H. F. Spriggs, L. Iyer, V. Anantharaman, L. Aravind, C. A. Weilman, C. S. Ong, W. R. Mac Kenzie, A. A. Lal, and C. B. Beard. 1997. Genetic polymorphism among Cryptosporidium parvum isolates: evidence of two distinct human transmission cycles. Emerg. Infect. Dis. 3:567–573.

20. Perrym, L. E., D. P. Jasmer, M. W. Rigs, S. G. Bohet, T. C. McGuire, and M. J. Arrowood. 1995. A cloned gene of Cryptosporidium parvum encodes neutralization-sensitive epitopes. Mol. Biochem. Parasitol. 80:137–147.

21. Petersen, C., J. Gut, P. S. Doyle, J. H. Crabb, R. G. Nelson, and J. H. Leech. 1992. Characterization of a ~900,000-M(D) Cryptosporidium parvum sporozoite glycoprotein recognized by protective immune bovine colostral immunoglobulin. Infect. Immun. 60:5132–5138.

22. Priest, J. W., J. P. Kwon, M. J. Arrowood, and P. J. Lammie. 2000. Cloning of the immunodominant 17-kDa antigen from Cryptosporidium parvum. Mol. Biochem. Parasitol. 106:261–271.

23. Puina, D., S. Enomoto, G. A. Buck, M. S. Abrahamsen, and J. C. Kissinger. 2004. Cryptodb: the Cryptosporidium genome resource. Nucleic Acids Res. 32:D329–D333.

24. Putignani, L., A. Possenti, S. Cherchi, E. Pozio, A. Crisanti, and F. Spano. 2008. The thiombsoporin-related protein Cpm1C (CpSP8) belongs to the repertoire of micronemal proteins of Cryptosporidium parvum. Mol. Biochem. Parasitol. 157:98–101.

25. Rigs, M. W., L. A. Stone, P. A. Yount, R. C. Langer, M. J. Arrowood, and D. L. Bentley. 1997. Protective monoclonal antibody defines a circumsporozoite-like glycoprotein exoantigen of Cryptosporidium parvum sporozoites and merozoites. J. Immunol. 158:1787–1795.

26. Salo, A. M., P. C. Goodall, D. C. Mathai, B. Primrose, J. M. W. White, and J. C. Boothroyd. 2007. Toxoplasma cp-coe hosts gene expression by injection of a polyomavirus kinase homologue. Nature 445:324–327.

27. Sallion, S., R. El Showwa, M. El Masri, M. Khalil, N. Blundell, and C. A. Hart. 1991. Cryptosporidiosis in children in Gaza. Ann. Trop. Paediatr. 11:277–281.

28. Sibley, L. D. 2004. Intracellular parasite invasion strategies. Science 304:248–253.

29. Soloff, D., J. F. Dubremetz, and M. Lebrun. 2001. Micromere proteins: structural and functional requirements to promote adhesion and invasion by the apicomplexan parasite Toxoplasma gondii. Int. J. Parasitol. 31:1293–1302.

30. Soloff, D., B. J. Foth, and A. F. Cowman. 2004. Molecular and functional aspects of parasite invasion. Trends Parasitol. 20:567–574.

31. Soder, W. B., J. Gut, J. F. Freeman, M. J. Arrowood, and A. A. Escalante. 2005. Cryptosporidium parvum species diversity in children in a semi-rural community in South India. J. Clin. Microbiol. 44:632–634.

32. Funko, E. A. Gooley, K. L. Williams, and M. B. Slade. 2000. Character-
ization of a major sporozoite surface glycoprotein of *Cryptosporidium parvum*. Funct. Integr. Genomics 1:207–217.

48. Wu, A. M., and S. J. Sugii. 1988. Differential binding properties of Ga1NAc and/or Ga1 specific lectins. Adv. Exp. Med. Biol. 228:205–263.

49. Wu, Z., I. Nagano, T. Boommars, T. Nakada, and Y. Takahashi. 2003. Intraspecies polymorphism of *Cryptosporidium parvum* revealed by PCR-restriction fragment length polymorphism (RFLP) and RFLP-single-strand conformational polymorphism analyses. Appl. Environ. Microbiol. 69:4720–4726.

50. Xiao, L., and U. Ryan. 2008. Molecular epidemiology, p. 119–163. In R. Fayer and L. Xiao (ed.), *Cryptosporidium* and cryptosporidiosis. CRC Press, Boca Raton, FL.

51. Xu, P., G. Widmer, Y. Wang, L. S. Ozaki, J. M. Alves, M. G. Serrano, D. Puin, P. Manque, D. Akiyoshi, A. J. Mackey, W. R. Pearson, P. H. Dear, A. T. Bankier, D. L. Peterson, M. S. Abrahamsen, V. Kapur, S. Tzipori, and G. A. Buck. 2004. The genome of *Cryptosporidium hominis*. Nature 431:1107–1112.