Polymorphic Mucin Antigens CpMuc4 and CpMuc5 Are Integral to Cryptosporidium parvum Infection In Vitro

Roberta M. O’Connor,1* Patrick B. Burns,1 Tin Ha-Ngoc,1 Kristen Scarpato,1 Wasif Khan,2 Gagandeep Kang,3 and Honorine Ward1

Division of Geographic Medicine and Infectious Diseases, Tufts Medical Center, Boston, Massachusetts1; International Center for Diarrheal Disease Research, Dhaka, Bangladesh2; and Department of Gastrointestinal Sciences, Christian Medical College, Vellore, India3

Received 11 September 2008/Accepted 20 November 2008

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 investigate 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 CpMuc genes were expressed throughout intra-cellular 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.

Cryptosporidium spp. are apicomplexan parasites that are responsible for diarrheal disease outbreaks worldwide. With the AIDS epidemic, Cryptosporidium gained recognition as a significant opportunistic pathogen, as the parasite causes a chronic, debilitating, and often fatal diarrheal disease in the immunocompromised (11, 19, 25, 26). In areas of endemicity in developing countries, cryptosporidiosis is associated with growth and developmental delays (14, 20, 38) and has been identified as a separate risk factor for childhood death in malnourished children (27). To date, there is no vaccine to prevent this disease, and there is no treatment that is effective in all patients (15).

To identify candidate vaccine antigens, many studies have focused on identifying the Cryptosporidium sporozoite antigens that are involved in invasion of host epithelial cells. In general, this work is hampered by the lack of in vitro propagation methods and systems for genetic manipulation. To date, very few sporozoite antigens have been characterized. Additionally, two species of Cryptosporidium cause most human disease: these are Cryptosporidium hominis and C. parvum (30). C. hominis is the species found in most human infections; however, because this species does not naturally infect other hosts, most experimentation has been done using the zoonotic species C. parvum. In some cases, C. hominis antigen genes are significantly divergent from the C. parvum orthologues (1, 16, 51). The Cpgp40/15 locus (8, 9, 42, 47), which encodes surface proteins important for attachment of and invasion into host cells by sporozoites, not only exhibits sequence divergence between the two species but is also highly polymorphic among different C. hominis isolates, suggesting that the gene products are targets of selective immune pressure.

Characterization of the attachment and invasion process of Cryptosporidium zootes suggests that this parasite may follow the paradigm described for other apicomplexan parasites (10, 45, 46). Upon contact with host cells, the specialized apical complex organelles (rhoptries, dense granules, and micronemes) release antigens that participate in recognition of and attachment to the host cell, invasion, and formation of the parasitophorous vacuole in which the parasite replicates (13). In contrast to other apicomplexan parasites, Cryptosporidium appears to employ mucin-like glycoproteins in these processes. Mucins are glycoproteins that have an amino acid composition consisting of 20 to 55% serine, threonine, and proline residues, with such extensive O-linked glycosylation of the serine and threonine residues that 40 to 80% of the molecular weight is attributable to O-linked carbohydrate (44). To date, the four Cryptosporidium zootes antigens that have been identified as integral to attachment and invasion are all glycoproteins, including CSL (23, 36), gp900 (5, 32), gp40/15 (8, 9, 33, 42, 47), and p23 (31). Both gp900 and gp40/15 are mucin-like glyco-
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 oloosaccharides decorating these mucins exhibit exposed T [Gal(b1-3)GalNAc] and/or Tn (GalNAc-α1-3-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 reverse transcriptase PCR (RT-PCR), Caco-2A epithelial cells. When the cells reached 80 to 90% confluence, the cells were infected with C. parvum infected Caco-2A cells. The cells were fixed with methanol and stained with DAPI (4',6-diamidino-2-phenylindole) (Vector Laboratories, Burlingame, CA), and examined by differential interference contrast and fluorescence microscopy using a Zeiss AxiosImager Z1.1 microscope (Carl Zeiss Microscopy, Jena, Germany). Images were captured with an IEEE1394 digital camera (Hamamatsu, Hamamatsu, Japan). Colocalization of the fluorescent labels was done using Velocity software (Improvision Inc., Lexington, MA). To improve the resolution of some samples, 2 stacks with 0.1-μm spacing were collected and deconvolved using Velocity’s iterative deconvolution program.

<table>
<thead>
<tr>
<th>Primer</th>
<th>Sequence (5′–3′)</th>
<th>Product size (bp)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CpMuc1-F</td>
<td>TGA AGT CAC CTA ATG ACA ACA</td>
<td>407</td>
</tr>
<tr>
<td>CpMuc1-R</td>
<td>TCA GTG TTT CCC ATG TTT AA</td>
<td>582</td>
</tr>
<tr>
<td>CpMuc2-R</td>
<td>TAT TCA ACT TTA AAC ACT TAC ACA</td>
<td>615</td>
</tr>
<tr>
<td>CpMuc3-F</td>
<td>GTA GAT CAA AAG GTA GTA GTA GTA</td>
<td>512</td>
</tr>
<tr>
<td>CpMuc4-F</td>
<td>GCA GGA GTG GAG GAG GAG</td>
<td>402</td>
</tr>
<tr>
<td>CpMuc4-R</td>
<td>AGT AAC AGG GGA GTT TTT</td>
<td>546</td>
</tr>
</tbody>
</table>

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 ChMuc4 and ChMuc5, a nested PCR strategy was performed using Platinum Pfu polymerase (Invitrogen) and primers specific for the published ChMuc4 and ChMuc5 sequences (Table 2). Primers specific for the published CpMuc5 sequence (Table 2) were designed to amplify ChMuc5 from C. hominis IIe and IIk 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 isolate 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 St tag, internal and C-terminal His tags, and N-terminal thioredoxin 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 previously described (9). Antiglobulin production. Rabbits were immunized as described previously, using recombinant CpMuc4 and CpMuc5 fusion proteins exsized 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 immunizing rabbits with the CpMuc4 peptide 107-125PPFAGVSLSSPRPRP116 coupled to keyhole limpet hemocyanin. The anti-CpMuc4-4 peptide IgG was affinity purified from the sera on a peptide Sepharose column (Dragonfly Sciences, Wellesley, MA).

Immunoblotting. Parasite lysates were resolved by SDS-PAGE on 12% acrylamide gels and transferred to nitrocellulose filters. Immunoblotting 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% or 4% paraformaldehyde. Half of the samples were permeabilized with 0.5% TX-100 in PBS. The slides were blocked with 2% normal goat serum in PBS and then probed with anti-CpMuc4 and anti-CpMuc5 antibodies. Reactivity was detected with Alexa Fluor 594 conjugated goat anti-rabbit IgG. CpMuc4 and CpMuc5 antibodies were colocalized with either the anti-gp90 monoclonal antibody (MAb) 4G12 (45), detected with Alexa Fluor 488 conjugated goat anti-mouse IgG, or with the anti-gp15 MAb CrA2 (9), detected with fluorescein isothiocyanate-conjugated goat anti-mouse IgA. For localization of CpMuc4 and CpMuc5 in intracellular-stage organisms, eight-well chamber slides were seeded with Caco-2A cells. When the cells reached 80 to 90% confluence, the cells were infected with C. parvum oocysts for 18 h. Cells were fixed with methanol and probed as described above. The slides were mounted with Vectashield mounting medium containing 4',6-diamidino-2-phenylindole (DAPI) (Vector Laboratories, Burlingame, CA) and examined by differential interference contrast and fluorescence microscopy using a Zeiss AxiosImager Z1.1 microscope (Carl Zeiss Microscopy, Jena, Germany). Images were captured with an IEEE1394 digital camera (Hamamatsu, Hamamatsu, Japan). Colocalization of the fluorescent labels was done using Velocity software (Improvision Inc., Lexington, MA). To improve the resolution of some samples, 2 stacks with 0.1-μm spacing were collected and deconvolved using Velocity’s iterative deconvolution program.

<table>
<thead>
<tr>
<th>Target</th>
<th>Primer</th>
<th>Primer sequence (5′–3′)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ChMuc4</td>
<td>Primary F</td>
<td>GAA GCA TGC GAA TTA AAA CT T</td>
</tr>
<tr>
<td>Primary R</td>
<td>TCG ATG TAG AAA AAC AGC GAG AAT TG</td>
<td></td>
</tr>
<tr>
<td>Nested F</td>
<td>CAC ATT CAA TCA GCT TGG CA</td>
<td></td>
</tr>
<tr>
<td>Nested R</td>
<td>CAT TGA TGT TGG AAA CCA CAA</td>
<td></td>
</tr>
<tr>
<td>ChMuc5</td>
<td>Primary F</td>
<td>CAA TCC TCG TTT TCA TCT AA</td>
</tr>
<tr>
<td>Primary R</td>
<td>CTC ATG AAC ATT GTA ACA CAC AAA G</td>
<td></td>
</tr>
<tr>
<td>Nested F</td>
<td>GAC GAG TAT AAA ACT TGG AGA AGA C</td>
<td></td>
</tr>
<tr>
<td>Nested R</td>
<td>CAT TCA ACT TTA AAC AAT AGC AGC</td>
<td></td>
</tr>
<tr>
<td>CpMuc5</td>
<td>Primary F</td>
<td>CAA TTT TCG TTT TCA TCT GA</td>
</tr>
<tr>
<td>Primary R</td>
<td>GAA ACC TCA AAC ACT AAC ACT ACA</td>
<td></td>
</tr>
<tr>
<td>Nested F</td>
<td>GGT TGG AAA TTA TGA ATG CCT GTC</td>
<td></td>
</tr>
<tr>
<td>Nested R</td>
<td>CAC AAT CTC ACT AGT ACT TAC ACA</td>
<td></td>
</tr>
</tbody>
</table>
Generation of F(ab) fragments. F(ab) fragments of anti-Muc4-peptide IgG and preimmune IgG were produced using immobilized papain (Pierce Chemical Company) following the manufacturer’s directions. In brief, the antibodies were resuspended in 20 mM sodium phosphate, pH 7.0, 0.1 mM EDTA, and 20 mM cysteine HCl, combined with the papain-agarose, and incubated overnight at 37°C. After incubation, the papain-agarose was removed by centrifugation, and the Fc fragments were removed by incubation with protein A agarose (Kirkegaard & Perry Laboratories, Gaithersburg, MD). The protein concentration of F(ab) fragments was estimated by measuring the A280.

C. parvum in vitro infection assay. The effects of anti-CpMuc4-peptide IgG and F(ab) fragments on C. parvum infection of Caco-2A cells were determined using an in vitro assay as described previously (8). In brief, 96-well plates were seeded with Caco-2A cells, and the cells were grown to confluence. Oocysts were removed by incubation with protein A agarose (Kirkegaard & Perry Laboratories, Gaithersburg, MD). The percent excystation of the sequence 5’-H11032 was estimated by measuring the number of oocysts remaining after incubation/total number of oocysts, suggesting that the genes are expressed. Analysis of the sequence 5’ of this locus identified another ORF encoding a putative mucin. We named the mucins CpMuc and numbered them CpMuc1 through -7 (5’ to 3’). The proteins are unique to Cryptosporidium, as no homologues were found 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/blcks/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 SignalP 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 CpMuc1-7 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 Caco2A 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 102-PNPF AGVSLSSPFRPR116 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>
<thead>
<tr>
<th>Gene product</th>
<th>CryptoDB gene ID</th>
<th>No. of amino acids</th>
<th>Predicted molecular mass (Da)</th>
<th>Amino acid composition (%)</th>
<th>No. of O-glycosylation sites</th>
<th>No. of N-glycosylation sites</th>
<th>Signal peptide or signal anchor</th>
<th>% Identity to ChMuc</th>
</tr>
</thead>
<tbody>
<tr>
<td>CpMuc1</td>
<td>cgd2_390</td>
<td>155</td>
<td>17,437</td>
<td>7.74</td>
<td>10.32</td>
<td>7.1</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>CpMuc2</td>
<td>cgd2_400</td>
<td>194</td>
<td>21,040</td>
<td>7.73</td>
<td>16.5</td>
<td>10.31</td>
<td>30</td>
<td>0</td>
</tr>
<tr>
<td>CpMuc3</td>
<td>cgd2_410</td>
<td>261</td>
<td>28,170</td>
<td>7.28</td>
<td>27.27</td>
<td>27.59</td>
<td>3.45</td>
<td>56</td>
</tr>
<tr>
<td>CpMuc4</td>
<td>cgd2_420</td>
<td>205</td>
<td>22,008</td>
<td>9.76</td>
<td>14.63</td>
<td>3.9</td>
<td>9</td>
<td>0</td>
</tr>
<tr>
<td>CpMuc5</td>
<td>cgd2_430</td>
<td>206</td>
<td>21,176</td>
<td>14.56</td>
<td>19.9</td>
<td>4.85</td>
<td>36</td>
<td>1</td>
</tr>
<tr>
<td>CpMuc6</td>
<td>cgd2_440</td>
<td>134</td>
<td>15,314</td>
<td>10.45</td>
<td>11.94</td>
<td>3</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>CpMuc7</td>
<td>cgd2_450</td>
<td>182</td>
<td>19,345</td>
<td>6.04</td>
<td>17.03</td>
<td>5.5</td>
<td>10</td>
<td>0</td>
</tr>
</tbody>
</table>

---

Amino acid composition is shown as percentage of total mass. Predicted molecular masses were calculated using the SignalP 3.0 program (6). As predicted by the NetOGlyc (18) and NetNGlyc (http://www.cbs.dtu.dk/services/NetNGlyc) programs.

**RESULTS**

In silico analysis of CpMuc1 to -7. A text search for “mucin” in the CryptoDB database (34) yielded 31 open reading frames (ORFs) encoding putative mucin-like glycoproteins. Among these were genes for six small-molecular-weight mucins clustered on a single locus on chromosome 2 which were also in the EST database, suggesting that the genes are expressed. Analysis of the sequence 5’ of this locus identified another ORF encoding a putative mucin. We named the mucins CpMuc and numbered them CpMuc1 through -7 (5’ to 3’). The proteins are unique to Cryptosporidium, as no homologues were found in any other apicomplexan genome (3) or in GenBank (7).
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* infection in vitro. 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).
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.

**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.

**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).

### 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 not previously described in 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...
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 a different apical compartment (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 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 Cpgp40/15 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 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 IIj (50). The confusing aspect of this nomenclature system is that none of the methods used to identify Cryptosporidium species distinguish between anthroponotic and zoonotic isolates, resulting in “subgenotypes” such as IIC and IIC (and probably IIk) that have been found only in humans and have been termed “anthroponotic 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 Cpgp40/15 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 Cpgp40/15 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-


ization of a major sporozoite surface glycoprotein of *Cryptosporidium parvum*. Funct. Integr. Genomics **1**:207–217.


