Giardiasis is a major cause of waterborne diarrheal disease (1–3), contributing greatly to the global burden of malnutrition. Despite its importance, neither the basic biology of *Giardia lamblia* nor the pathophysiology of infection is well understood. *Giardia* infection initiates with ingestion of infectious cysts that can persist for months in cold fresh water (4). Exposure of cysts to gastric acid during passage through the stomach triggers excystation (4, 5), and after entry into the small intestine, the parasite emerges and divides into two equivalent binucleate trophozoites (5, 6). Trophozoites attach to and colonize the human small intestinal epithelium, where they can cause disease. However, if they detach and are carried downstream, they must encyst to survive outside the host. Our laboratory has reproduced the giardial life cycle *in vitro* (7–9).

The “biological goal” of excystation is differentiation into a form that can survive in the environment and infect a new host. The construction of the extracellular cyst wall is therefore essential as it is the single structure that allows the parasite to persist in fresh water, resist disinfectants, pass through the stomach of the new host, and infect in the small intestine. This 300-nm-thick fibrous structure (10–12) is tight enough to exclude small molecules, such as water and gastric acid. On the other hand, it must be capable of transmitting the physiological stimuli that regulate activation and emergence (13). Therefore, the encystation pathway may be thought of as a giardial virulence mechanism necessary for survival outside the host and infection.

The cyst wall is at least 40% protein, and the remainder is carbohydrate (14, 15). The two known cyst wall proteins (CWPs) contain similar domains, including a putative N-terminal signal peptide and five tandem, leucine-rich repeats (LRR) (16, 17), and are reported to be phosphorylated (18). In the regions of similarity, both proteins are acidic, but CWP2 has a very basic 121-residue C-terminal extension. Expression of the two CWPs increases greatly soon after exposure to encystation signals (16, 17, 19, 20). LRR motifs of both prokaryotic and eukaryotic proteins have diverse functions and cellular localizations but are associated with protein-protein interactions (21–23).

Earlier, we discovered a novel regulated secretory pathway that is induced during giardial excystation (19, 24). The appearance of large membrane-bounded encystation secretory vesicles (ESVs) at 6–24 h is the earliest morphological change in giardial encystation (17, 19, 25, 26). CWP1 and -2 are concentrated within the ESVs for transport to the nascent cyst wall (16, 17, 20, 25). CWP1 and -2 also contain 18 and 19 cysteine residues, respectively, and form disulfide-bonded heterodimers and oligomers soon after synthesis (16, 17, 27, 28). However, the mechanism by which CWPs are targeted to the regulated secretory pathway is not well understood (28).

*Giardia* belongs to one of the earliest diverging eukaryotic lineages (29). There is little understanding of the biology of the origins of eukaryotes, particularly of the endomembrane-mediated system of protein modification, sorting, and transport (30).
The Giardia cyst wall is a relatively simple form of eukaryotic extracellular matrix with crucial structural and signaling functions whose assembly requires regulated protein expression and transport. Therefore, formation of the ESVs and cyst wall in this protist is an excellent model for the study of early forms of biosynthesis of extracellular structures in eukaryotes. Because of the evolutionary and functional importance of the cyst wall, we queried the Giardia Genome Project Database (www.mbl.edu/Giardia, Ref. 31) with the coding sequences of CWP1 and CWP2 (16, 17) to identify additional CWP genes. We identified five putative leucine-rich proteins. Only one of these (CWP3) is up-regulated during encystation and is transported by the ESV pathway and incorporated into the cyst wall. In contrast, the genes encoding the other four proteins are actually down-regulated during encystation. Therefore, we characterized the expression and detailed the targeting requirements of cwp3.

Other important parasites, e.g. Cryptosporidium, Toxoplama, Toxoplasma, whose differentiation has not been achieved. Project Database make it a valuable model for other parasites whose differentiation has not been achieved.

EXPERIMENTAL PROCEDURES

Giardia Culture and Encystation—Trophozoites of G. lamblia isolate WB (ATCC 50803), clone C6, were cultured in modified TYI-S-33 medium (33). Encystation was induced by a modification (34) of our earlier procedure (35). Briefly, trophozoites that were grown to late log phase in growth medium were harvested and encysted for the times indicated in TYI-S-33 medium containing 12.5 mg/ml bovine bile at pH 7.8 at a beginning density of 5 × 10^6 cells/ml.

Isolation and Analysis of the cwp3 Gene—We searched the G. lamblia Genome Project Database (www.mbl.edu/Giardia, Ref. 31) with amino acid sequences of the cwp1 and cwp2 genes (GenBank accession numbers S61925 and S61924, respectively), using the BLAST program (36). We amplified a 1.2-kb sequence containing the cwp3 gene from genomic DNA by PCR using primers cwp3F (5'-CGCGCTAGCTATGCTAAGTTT-3') and cwp3R (5'-GGCGGCTAGCCATGTTTGGG-3') and cwp3AUE (5'-GGCGGAATTCTTGAATGAGTCACTCTCTGTTAGGCGGCTAGCTTGCTTTTGCCGTTAGAAGAG-3') and cwp3AUER generated a 0.7-kb product that was digested with Ncol and EcoRI. Another PCR with primers cwp3NDR (5'-GGCGGCATGTAGCTGCTTGCTTTTGCCGTTAGAAGAG-3') and cwp35NF generated a 0.3-kb PCR product that was digested with Ncol and NheI and cloned into NcoI/EcoRI-digested pNLp02-1 with the 0.7-kb NcoI/EcoRI fragment. The resulting ND2 contains a cwp3 gene lacking the 44 amino acids N-terminal to LRRs (residues 17–60). A similar strategy was used to construct R1D, R2D, R3D, R4D, R1–D, and CD1, which contain cwp3 gene with deletion of first LRR (residues 61–83), second LRR (residues 84–107), third LRR (residues 108–131), fourth LRR (residues 132–155), first to fourth (residues 61–155) LRRs, and the region from residues 156–201. The resulting products were sequenced (Applied Biosystems). For semiquantitative RT-PCR of RT-PCR products were cloned into pGEM-T easy vector (Promega) and sequenced (Applied Biosystems). For semiquantitative RT-PCR of RT-PCR products were cloned into pGEM-T easy vector (Promega) and sequenced (Applied Biosystems).
brane was incubated with anti-gPDI-1 serum (38) and detected with goat anti-rabbit antibody (Pierce, 5000) and enhanced chemiluminescence (Amersham Biosciences).

**Immunofluorescence Assay**—Stably transfected cells were harvested after 24 h in growth or encystation medium under G418 selection, washed in PBS, and attached to glass coverslip (2 × 10⁶ cells/coverslip) and then fixed and stained (41). Water-resistant cysts were harvested after 48 h by centrifugation and five treatments with double distilled water for 5 min and then kept in cold distilled water overnight. Cells were reacted with anti-AU1 or AU5 monoclonal antibody (Covance, Princeton, NJ, 500 in blocking buffer) and anti-mouse ALEXA 568 (Molecular Probes, 500 in blocking buffer) as the detector. The images were evaluated on a Zeiss LSM 510 laser scanning confocal microscope.

**Immunoelectron Microscopy**—Encysting cells were processed as described (25). Briefly, washed parasites were fixed in suspension for 15 min in freshly prepared 4% formaldehyde, in PBS, pH 7.4, and then pelleted and resuspended in 4% formaldehyde in PBS and fixed for an additional 18–24 h at 4 °C. The cells were washed briefly in PBS and resuspended in 1% low temperature gelling agarose. After cooling, the agarose blocks were trimmed into mm³ pieces; cryoprotected with 2.3 M glycerol and 22.6 M DMSO overnight at −70 °C and rapidly frozen in liquid nitrogen. Ultrathin cryosections were cut on a Leica UCT ultramicrotome equipped with an FC-S cryoattachment and collected onto formvar/carbon-coated nickel grids. Sections were washed through several drops of 1× PBS containing 5% fetal calf serum, 10 mM glycine, pH 7.4; blocked in 10% fetal calf serum for 30 min; and incubated overnight in 20 μg/ml anti-AU1 antibody. After washing, the grids were incubated for 2 h in 5 μm AU-conjugated donkey anti-mouse (Jackson Immunoresearch Laboratories). The grids were then washed through several drops of PBS followed by several drops of ddH₂O and embedded in an aqueous solution containing 3.2% agarose blocks were trimmed into mm³ pieces; cryoprotected with 2.3 M glycerol and 22.6 M DMSO overnight at −70 °C and rapidly frozen in liquid nitrogen. Ultrathin cryosections were cut on a Leica UCT ultramicrotome equipped with an FC-S cryoattachment and collected onto formvar/carbon-coated nickel grids. Sections were washed through several drops of 1× PBS containing 5% fetal calf serum, 10 mM glycine, pH 7.4; blocked in 10% fetal calf serum for 30 min; and incubated overnight in 20 μg/ml anti-AU1 antibody. After washing, the grids were incubated for 2 h in 5 μm AU-conjugated donkey anti-mouse (Jackson Immunoresearch Laboratories). The grids were then washed through several drops of PBS followed by several drops of ddH₂O and embedded in an aqueous solution containing 3.2% polyvinyl pyrrolidone (10 K), 0.2% methyl cellulose (400 centipoises), and 0.1% uranyl acetate. The grids were observed and photographed on a Phillips 420 transmission electron microscope at 80 kV.

**Sequence Analysis and Alignment**—The cwp3 and lrp1–4 genes were amplified from genomic DNA. An alignment of the LRRs of Cwp1, Cwp2, and Cwp3 was prepared with ClustalW (42). The cwp3 sequence was most closely deposited in GenBank™ under accession number AF061927. The lrp3 sequence was previously deposited in GenBank™ under accession number AF310726. We have deposited lrp1, lrp2, and lrp4 in GenBank™ under accession numbersAY225413, AY225414, and AY225415.

**RESULTS**

**Identification and Characterization of the cwp3 Gene**—To identify possible new cyst wall proteins, we queried the G. lamblia Genome Project Database (www.mbl.edu/Giardia, Ref. 31) with amino acid sequences of the cwp1 and cwp2 genes. This search detected five putative open reading frames with E values <10⁻⁴, lrp1, lrp2, lrp3, lrp4, and cwp3. During the course of this study, two of these proteins, LRP3 (“virus receptor protein”) and CWP3 were deposited in GenBank™ by others but not published. We found that lrp1, lrp2, lrp3, and lrp4 were expressed at significantly lower levels during encystation (Fig. 1A). Therefore, we focused on cwp3, which was up-regulated (Fig. 1B). The cwp3 coding region with 262 nt of 5′- and 120 nt of 3′-flanking regions was retrieved and amplified. Comparing genomic and cDNA sequences showed that the cwp3 gene contained no introns. Primer extension analysis revealed a major transcription start site, −13, and three minor transcription start sites, −16, −17, and −18, relative to the translation start site of cwp3 in 24-h encysting cells. No primer extension product was detected in vegetative cells. All transcription start sites are located within a 9-bp AT-rich sequence, −13/−21 (AAAAAAATAA), as has frequently been observed for initiation of giardial transcripts (16, 17, 38, 43–45). A classic polyadenylation signal (AGTTAAC, Refs. 46 and 47) was 67 nt downstream of the stop codon, and the polyadenylation sites shown by 3′RACE were 11 and 21 bases further downstream (data not shown). Short untranslated regions and lack of introns are typical of giardial transcripts (44–45).

The deduced CWP3 protein contains 247 amino acids with a predicted molecular mass of ~27.3 kDa and a pI of 4.45. (Fig. 2). The N-terminal 17 amino acids of CWP3 comprise a predicted signal sequence (Signal P prediction, Ref. 48) with the most likely cleavage site between positions 17 and 18. The 44-amino-acid region N-terminal to the LRR-like repeats (N-flanking region) exhibited 52% identity to that of CWP1 or CWP2. Importantly, all of the 14 cysteines in CWP3 are positionally conserved with the corresponding cysteine residues in CWP1 and CWP2. CWP1 and -2 have several additional cysteine residues, of which two pairs are conserved. Unlike CWP1 and CWP2, CWP3 only contains four complete LRRs, which are located within a 9-bp AT-rich sequence, −13/−21 (AAAAAAATAA), as has frequently been observed for initiation of giardial transcripts (16, 17, 38, 43–45). A classic polyadenylation signal (AGTTAAC, Refs. 46 and 47) was 67 nt downstream of the stop codon, and the polyadenylation sites shown by 3′RACE were 11 and 21 bases further downstream (data not shown). Short untranslated regions and lack of introns are typical of giardial transcripts (44–45).

The deduced CWP3 protein contains 247 amino acids with a predicted molecular mass of ~27.3 kDa and a pI of 4.45. (Fig. 2). The N-terminal 17 amino acids of CWP3 comprise a predicted signal sequence (Signal P prediction, Ref. 48) with the most likely cleavage site between positions 17 and 18. The 44-amino-acid region N-terminal to the LRR-like repeats (N-flanking region) exhibited 52% identity to that of CWP1 or CWP2. Importantly, all of the 14 cysteines in CWP3 are positionally conserved with the corresponding cysteine residues in CWP1 and CWP2. CWP1 and -2 have several additional cysteine residues, of which two pairs are conserved. Unlike CWP1 and CWP2, CWP3 only contains four complete LRRs, which consist of central residues 61–179 (~13 kDa) (Fig. 3). Each LRR is 23 or 24 residues in length with 7 hydrophobic residues, most of which are leucine, although some are valine, isoleucine, methionine, or tryptophan. CWP3 may also have an incomplete fifth LRR with 3 hydrophobic residues either absent or not in close proximity.
Fig. 2. Sequence alignment of CWPs. Amino acids that are identical in all three CWPs are black. Amino acids identical in two CWPs are indicated in gray. The 44 amino acids following the predicted signal peptide and N-terminal to the LRRs have dots underneath. The location of the four LRRs is underlined. The fifth (incomplete) LRR is indicated by a dashed underline. The 14 cysteine residues of CWP3 are all positionally conserved with CWP1 and -2 and are shown with a solid triangle.

Fig. 3. Sequence alignment of the LRRs in CWP1, CWP2, and CWP3. Amino acids that are similar or identical to the consensus (according to Clustal W) are indicated in gray.

Encystation-induced Expression of the cwp3 Gene—A 1.2-kb transcript (Fig. 1B) that likely represents cwp3, whose cDNA is 0.9 kb (Fig. 2), was detected in Northern blots at very low levels in vegetative and early encysting cells. cwp3 message increased ∼100-fold and was maximal at 24 h, remaining high at 48 h, relative to loading controls (Fig. 1B, lower panel). In contrast, the genes encoding the four other LRR proteins were expressed at significantly lower levels at 24 h of encystation (Fig. 1A). To determine the expression of CWP3 protein, we prepared a pNC3 construct in which the cwp3 gene is controlled by its own promoter and contains an AU1 epitope tag at its C terminus (see Fig. 5). When the samples are reduced, a 0.8-kDa band, consistent with the predicted molecular mass of CWP3 with the AU1 tag (0.8 kDa), was detected with anti-AU1 antibody (Fig. 1D). Its levels increased very small (0.1%) and increased (to ∼30%) during encystation. The 27.3-kDa molecular mass of CWP3 with the AU1 tag (∼0.8 kDa) was detected with anti-AU1 antibody (Fig. 1C). Its levels increased gradually during encystation and reached a maximum at 24 h encystation relative to the constitutively expressed loading control (49), suggesting that expression was similar at the steady-state mRNA and protein levels. The expression of epitope-tagged CWP3 protein did not influence the growth or encystation of these cells, indicating that the transfected CWP3 is not highly expressed or that it is not toxic if overexpressed. When the samples are not reduced, the small amount of CWP3 expressed in trophozoites or early in encystation migrated as a 43-kDa band. As encystation progressed, increasing amounts of the 43-kDa species of CWP3 were observed. Strikingly, much of the protein appeared in larger heterodisperse complexes (Fig. 1D), indicating the increasing formation of intermolecular disulfide linkages. CWP1 and CWP2 were expressed with similar kinetics and formed similar polydisperse disulfide-bonded complexes (Fig. 1D) (16, 17, 27, 28).

Localization of the CWP3 Protein and Regions Needed for Targeting—Immunoelectron microscopy showed that the AU1-tagged CWP3 in pNC3 transfectants (see Fig. 5) was localized within typical membrane-bounded ESVs (Fig. 4A) and in the wall of water-resistant cysts (Fig. 4B). No staining was observed in non-transfected controls. CWP3 localized uniformly in the interior, but not the membrane of the ESVs and throughout the cyst wall, as shown previously for CWP1 and CWP2 (16, 17, 20). Indirect immunofluorescence assays confirmed that the AU1-tagged wild-type CWP3 localized to the ESVs in encysting cells (see Fig. 6A) and to the cyst wall in water-resistant cysts (see Fig. 6B). ESVs are identified in confocal microscopy by their typical large size and shape (Ref. 27 and data not shown). The number of positively stained cells detected in vegetative trophozoites was very small (0.1%) and increased (to ∼30%) during encystation. To determine the expression of CWP3 protein, we prepared a pNC3 construct in which the cwp3 gene is controlled by its own promoter and contains an AU1 epitope tag at its C terminus (see Fig. 5). When the samples are reduced, a 0.8-kDa band, consistent with the predicted molecular mass of CWP3 with the AU1 tag (0.8 kDa), was detected with anti-AU1 antibody (Fig. 1C). Its levels increased gradually during encystation and reached a maximum at 24 h encystation relative to the constitutively expressed loading control (49), suggesting that expression was similar at the steady-state mRNA and protein levels. The expression of epitope-tagged CWP3 protein did not influence the growth or encystation of these cells, indicating that the transfected CWP3 is not highly expressed or that it is not toxic if overexpressed. When the samples are not reduced, the small amount of CWP3 expressed in trophozoites or early in encystation migrated as a 43-kDa band. As encystation progressed, increasing amounts of the 43-kDa species of CWP3 were observed. Strikingly, much of the protein appeared in larger heterodisperse complexes (Fig. 1D), indicating the increasing formation of intermolecular disulfide linkages. CWP1 and CWP2 were expressed with similar kinetics and formed similar polydisperse disulfide-bonded complexes (Fig. 1D) (16, 17, 27, 28).

Localization of the CWP3 Protein and Regions Needed for Targeting—Immunoelectron microscopy showed that the AU1-tagged CWP3 in pNC3 transfectants (see Fig. 5) was localized within typical membrane-bounded ESVs (Fig. 4A) and in the wall of water-resistant cysts (Fig. 4B). No staining was observed in non-transfected controls. CWP3 localized uniformly in the interior, but not the membrane of the ESVs and throughout the cyst wall, as shown previously for CWP1 and CWP2 (16, 17, 20). Indirect immunofluorescence assays confirmed that the AU1-tagged wild-type CWP3 localized to the ESVs in encysting cells (see Fig. 6A) and to the cyst wall in water-resistant cysts (see Fig. 6B). ESVs are identified in confocal microscopy by their typical large size and shape (Ref. 27 and data not shown). The number of positively stained cells detected in vegetative trophozoites was very small (0.1%) and increased (to ∼30%) during encystation. A series of deletions was designed to identify the region(s) of CWP3 that are needed for accurate transport to the cyst wall via the ESVs in the regulated secretory pathway (summarized in Fig. 5). Deletion of the predicted N-terminal signal peptide (residues 1–16, construct ND, Fig. 6, C and D) or of any single LRR (residues 61–83, construct R1D; residues 84–107, construct R2D; residues 108–131, construct R3D; residues 132–155, construct R4D, Fig. 6, E and F, and data not shown) or all LRRs (residues 61–155, construct R1–4D) completely prevented targeting to ESVs and cyst wall. Instead, these CWP3 mutants localized unevenly in the cytoplasm (Fig. 6, C–F, phenotype of Fig. 5). It is not clear whether the mistargeted deletion proteins were within membrane-bounded vesicles that were smaller than ESVs or in small clumps in the cytosol. We never observed staining of cyst wall in constructs that did not localize to the ESVs. However, with the C-terminal deletion (residues 202–247, construct CD2), ESV localization was normal, but we did not detect any staining at all in water-resistant cysts (deletion CD2, NS phenotype, Fig. 5). Finally, deletion of either the highly conserved 44 amino acids N-terminal to the LRRs (residues 17–60, construct ND2) or residues 156–201 (construct CD1), including the position corresponding to the incomplete fifth LRR, decreased targeting to ESVs and cyst wall. In approximately half the staining cells, the mutant pro-
tein was within ESVs, whereas in the remaining cells, it was in smaller puncti (± phenotype).

Some of the deletion derivatives had fewer expressing cells with lower levels of total staining in IFA. Nonetheless, their phenotype (ability to traffic to the ESVs or cyst wall) was evident. In the case of the ND deletion, both the transcript and protein levels were decreased (data not shown), suggesting the possibility of a regulatory element in the N-terminal region. For all other deletions, the steady-state transcript levels were similar to wild-type transfectants. However, the steady-state protein levels of the ND2, R2D, and CD2 constructs were lower, indicating that these deleted proteins may be unstable or easily degraded. Interestingly, of the individual LRR deletions, only R2D had a lower protein level, and the largest deletion R1–4D was highly abundant.

Because Northern and RT-PCR analyses showed decreased levels of steady-state mRNA of the ND N-terminal deletion (data not shown), we examined the role of the N-terminal open reading frame sequence in regulating gene expression. We prepared a construct pNC35 in which luciferase reporter expression is driven only by the 262-nt 5’ flanking sequence of cwp3. The induction ratio (luciferase expression in 21-h encysting cells relative to that in trophozoites) of this construct was only ~4.5 (data not shown). This is much lower than that of the gMyb2 transcription factor gene (~75-fold, ref. 34) or the glucosamine-6-phosphate isomerase gene (induction ratio ~250-fold, Ref. 38) in similar constructs. Addition of the N-terminal 47 nt of the coding sequence upstream of luciferase (pNC35+) increased the induction ratio of the CWP3 luciferase fusion by ~6-fold, to ~28.8, suggesting some role for this region in up-regulating expression of the cwp3 gene in encystation.

Cleavage of the basic C-terminal extension of CWP2 was reported to be necessary for this protein to be incorporated into the cyst wall (17, 50). However, in the course of these studies, we found CWP2 with a C-terminal AU1 epitope tag (pNC2) in the ESVs in encysting cells (Fig. 7A) and the wall of water-resistant cysts (Fig. 7B). This suggests that C-terminal cleavage of CWP2 may not be required for incorporation into the wall. In addition to the 39-kDa precursor, an ~26 kDa fragment similar in size to the mature CWP2 retained its C-terminal epitope tag (Fig. 7C), suggesting that the cleavage of CWP2 may not be C-terminal. A CWP1 construct (pNC1) with an AU5 epitope tag at its C terminus also targeted to ESVs and the cyst wall (not shown).

**DISCUSSION**

The giardial cyst wall, which enables it to survive in the environment and is an important model for extracellular matrix assembly and function in an early diverging protist. Two cyst wall proteins and a homopolymer of GalNAc have been described to date (15–17). Additional components may exist; however, our mining of the Genome Project Database suggests that any other encystation proteins are not closely related to CWP1, -2, or -3. Although these components of the cyst wall have been characterized, neither the early steps nor the late steps in the formation of the extracellular CW fibrils are well understood. It is likely that the CWPs are targeted to the secretory pathway by their typical N-terminal signal peptides. However, it is not known how they are sorted or targeted to the regulated secretory pathway (24, 28).

In higher eukaryotic cells, there is disagreement on how proteins are sorted to the matrix of regulated secretory granules (51). In general, discrete targeting sequences have not been identified, and aggregation of secretory proteins within the trans-Golgi network is thought to be central (52). Understanding the regulated secretory pathway in *Giardia* is important because it is an early diverging eukaryote lacking certain typical eukaryotic organelles (2). It has frequently been hypothesized that the giardial Golgi apparatus is assembled during encystation (24, 53) and that the ESVs originate from the trans-Golgi network. Protein sorting does occur in the giardial regulated secretory pathway because constitutive proteins destined for the plasmalemma or the lysosomal compartments do not appear in the ESVs (24). On the other hand, recent data suggest that in *G. lamblia*, certain functions of the endoplasmic reticulum and Golgi co-localize or overlap spatially and temporally in a novel regulated secretory pathway (54).

*Giardia* has conserved signals for targeting constitutive proteins to the endoplasmic reticulum and lysosomal compartments, suggesting that they appeared early in eukaryotic evolution (53, 55). However, we hypothesize that the ability of the three CWPs to form intermolecular complexes is required for their concentration within the regulated secretory granules and incorporation into the nascent cyst wall. The three CWPs have similar kinetics of expression and form heterodisperse complexes.
higher molecular weight disulfide-bonded complexes. All are evenly distributed in the ESV matrix and cyst wall. The CWPs are cysteine-rich and characterized by LRR-like repeats, but the relative importance of disulfide bonding versus LRR interactions is not clear.

It is striking that all of the 14-cysteine residues of CWP3 are positionally conserved with CWP1 and -2, although the latter have several additional cysteines. This suggests possible structural specificity in disulfide cross-linking despite the heterodisperse nature of the CWP complexes. Formation of correct disulfide bonds, which is crucial for protein folding and maturation, occurs co-translationally in the endoplasmic reticulum (56–58). Exposure of live yeast or mammalian cells to DTT prevents formation of disulfide bonds in nascent proteins in the endoplasmic reticulum and can keep them in reduced form, detectable in non-reducing SDS-PAGE, preventing further modification and transport to the cell surface. Importantly, after proteins have folded to their mature conformation, they are largely resistant to DTT in situ. Removal of DTT permits normal refolding of the reduced proteins.

We recently used this powerful paradigm to probe the maturation of giardial CWPs within the secretory pathway. We found that the hetero-oligomer formation between CWPs was critical for formation and stability of ESVs. Treatment of live encysting cells with DTT prevented ESV formation (27). Moreover, later in encystation, exposure to DTT caused rapid and reversible disassembly of the ESVs, concomitant with reduction of the CWPs to monomers (27). This suggests that formation and stability of giardial regulated secretory granules may depend upon the accumulation of disulfide-bonded CWP cargo. The prolonged susceptibility of CWPs to reduction is highly unusual among secretory proteins. It suggests that despite the extensive disulfide bonding, the CWPs remain incompletely folded and in a relatively soluble or fluid form within the ESVs. We propose that this delayed maturation may be an important strategy to prevent the CWPs from forming fibrils before secretion.

Our current deletion analyses also support the importance of the LRR-like repeats. These detailed studies indicate that presence of each of the four LRRs is essential for targeting of CWP3 to the regulated secretory pathway. Deletion of any one of the four putative LRRs completely prevented CWP3 from forming hetero-oligomeric complexes not distinguishable from those of wild type (data not shown).

The regions outside the LRRs are also important for targeting. Deletion of the N-terminal signal peptide prevented targeting of CWP3 to the ESV pathway (as expected), probably because it could not enter the secretory pathway. Of the constructs that did target to the ESVs, only CD2 lacking the C-terminal 46 amino acids was not seen in either the cytoplasm or wall of mature cysts. This may be because ESV-containing CD2 proteins were not able to complete their traffic to the cyst wall. Alternatively, cells expressing CD2 protein in their ESVs did not go on to form water-resistant cysts. Perhaps this is

![Image](http://www.jbc.org/)

**Fig. 6.** Localization of CWP3 and its deletion derivatives. The pNC3 (A and B), ND (C and D), and R3D (E and F) transfectants were cultured in encystation medium for 24 h (A, C, and E) or 48 h (B, D, and F) and then subjected to immunofluorescence analysis using anti-AU1 antibody for detection. A and B show that the product of wild-type pNC3 localizes to the ESVs and cyst wall in encysting trophozoites and cysts, respectively. In A, the appearance of discrete ESVs and seemingly more diffuse staining is due to the differences in the focal plane of the optical section for each cell and to the brightness of the stain. C shows that the product of the N-terminal deletion (amino acids 1–16), ND, was in a smaller, more punctate pattern that did not correspond to ESVs. D shows that the product of ND did not localize to the cyst wall in water-resistant cysts. E and F show that the localization of the third LRR deletion (amino acids 132–155) R3D was indistinguishable from C and D.

**Fig. 7.** Incorporation of the C terminus of CWP2 into the cyst wall. CWP2 with a C-terminal AU1 epitope tag in the pNC2 transfectants localized to the ESVs in encysting cells (A) and to the wall of water-resistant cysts (B). C, Western analyses of CWP1 or CWP2 with C-terminal epitope tag. The pNC2 (CWP2) transfectants and pPC1 (CWP1) were cultured in encystation medium and harvested at 24 h. AU1-tagged CWP2 protein and AU5-tagged CWP1 were detected by anti-AU1 and anti-AU5 antibody, respectively.

---

**A New Cyst Wall Protein in Giardia**
A New Cyst Wall Protein in Giardia

21707

lateral in encystation. These data suggest that targeting of CWP3 to the ESV pathway is governed by complex aspects of overall protein conformation that can be perturbed by deletion of many segments of this protein, leading to subtle defects in interaction with other CWPs.

A number of LRR-containing proteins have now been crystallized (59), and certain patterns have emerged. Each LRR repeat unit consists of a β-strand and an α-helix, resulting in a horseshoe-shaped region with hydrophilic residues in the β-sheets facing inward. The LRR repeats range from 22 to 25 amino acids, and most LRR proteins contain 4 or more repeats. The repeats generally contain 7 leucine or isoleucine residues and an asparagine involved in hydrogen bonding, arranged in patterns that may be conserved among families of LRR proteins but vary between different families. Valine, alanine, phenylalanine, methionine, or tryptophan can be substituted for the Leu or Ile. The putative LRRs in the three CWPs generally conform to this pattern, although CWP3 appears more divergent. The CWPs have conserved Leu or Ile residues in positions 2, 4, 9, 13, 17, 20, and 23 and Asn in position 7. Blast searches and especially Hidden Markov Model analyses suggest similarity with LRR structures in proteins from all three kingdoms. However, the CWPs differ from these LRR-containing proteins in having many cysteine residues. Their interchain disulfide bonding makes structural predictions problematic. In addition, it is likely that the CWPs interact with the polyGalNAc cyst wall fibril chains, although to date, there is no information on the nature of these interactions or whether they form within the cell or after secretion (15).

Hehl et al. (28) expressed CWP1-GFP chimeras to define the protein domains needed for CWP heterodimer formation and targeting to the cyst wall. Their work demonstrated that: 1) stably transfected CWP1 is expressed in encystation with similar kinetics to the native protein; 2) CWP1 with a large C-terminal GFP epitope tag can be incorporated into the cyst wall; and 3) CWP1–2 hetero-oligomer formation is important and involves cysteine residues throughout the proteins. However, the methods required to achieve GFP fluorescence may have affected the secretory pathway. Exposure of live cells to reduced temperatures is commonly employed to produce specific blocks in protein transport (60). GFP fluorescence requires exposure to oxygen (61), which Giardia does not tolerate well (62). Transfected encysting Giardia were exposed “for several hours” to air or oxygen in the cold. Apparently, this did not affect CWP1 that had already been incorporated into the cyst wall. However, Hehl et al. (28) reported the “apparent condensation of GFP-labeled vesicular structures” into “sometimes apparently only one major ESV localized primarily to the anterior dorsal part of the cell,” after exposure of the cells to cold and oxygen during encystation. No other group has reported ESV similar to this. Moreover, it is not consistent with the many sites of cyst wall fiber formation seen over the entire encysting cell (12). Therefore, the interpretation regarding this aspect of the encystation secretory pathway (28, 54) should be viewed with caution.

Using four GFP chimeras with progressively longer segments of CWP1, Hehl et al. (28) reported that a chimera containing only the signal peptide and region N-terminal to the LRRs “localized to distinct vesicular structures in the cytoplasm” but not to the cyst wall (28). This is in contrast to our finding (above) that constructs lacking any one of the LRRs did not target to the ESVs. Their deletion constructs were visualized prior to maximal ESV formation (5-h encystation), and the cells were exposed to cold and oxygen. For these reasons and because they did not perform ultrastructural analyses, the nature of these vesicles is not clear. Localization of two longer constructs that included all five LRRs to the cyst wall was reported not to be affected by absence of the C terminus. In contrast, we found that the most C-terminal region of CWP3 was also needed for incorporation into the cyst wall. The reasons for the differences between our findings and those of Hehl et al. with CWP1 are not clear. It is possible that CWP1 and -3 have distinct structural requirements for targeting to the regulated secretory pathway. Alternatively, the small (6-amino-acid) AU1 epitope tag may be less likely to affect trafficking than GFP in Giardia.

Taken together, our data suggest that sequences throughout the CWP3 molecule are important for efficient targeting to the regulated secretory pathway and incorporation into the cyst wall. This may be because most domains are required for correct overall protein conformation and transport. The LRR deletion constructs localized irregularly in the cytoplasm in punctate patterns that were smaller than ESVs. We do not know whether these deletion proteins are within membrane-bounded vesicles. However, the staining pattern resembles that of the lysosome-like peripheral vesicles in size and distribution. If the deletions were transported to these vesicles, they might be destined for destruction. However, of the LRR deletions, only R2D seemed to have significantly decreased protein stability.

These studies demonstrate the synergy of genomic and biological analyses as only one of the five proteins identified by genomic screening was a functional cyst wall protein. This work may also lead to greater understanding of cyst walls of other parasites and of the evolution of regulated secretory pathways of extracellular matrix secretion and assembly.

Acknowledgments—We are grateful to the G. lamblia Genome Project, Marine Biological Laboratory at Woods Hole, funded by the NIAID, National Institutes of Health under cooperative agreement AI043273 to Dr. Mitchell L. Sogin for preliminary genome data and to Dr. Barbara J. Davids for help with imaging, and to Dr. Andrew McArthur for help with contig analyses.

REFERENCES

1. Marshall, M. M., Naumovitz, D., Ortega, Y., and Sterling, C. R. (1997) Clin. Microbiol. Rev. 10, 67–85
2. Adam, R. D. (1991) Microbiol. Rev. 55, 706–722
3. Pichichero, M. J. G. (1985) Infections of the Gastrointestinal Tract (Blaser, M. J., Smith, P. D., and Radvin, J. L., eds.) pp. 1081–1105, Raven Press, New York
4. Bligh, A. K., and Meyer, E. A. (1979) Nature 277, 301–302
5. Rice, E. W., and Schafer, W. F. III (1981) J. Clin. Microbiol. 14, 709–710
6. Boucher, S.-E., and Gillin, F. D. (1990) Infect. Immun. 58, 3516–3522
7. Gerwig, G. J., van Kuik, J. A., Leeflang, B. R., Kamerling, J. P., Vliegenthart, F. J., and Sauch, J. F. (1978) Science 205, 1040–1043
8. Gillin, F. D., and Reiner, D. S. (1996) Annu. Rev. Microbiol. 50, 679–705
9. Eichinger, D. (2001) Curr. Opin. Microbiol. 4, 421–426
10. Erlandsen, S. L., Bemrick, W. J., and Pawley, J. P. (1989) J. Parasitol. 75, 787–797
11. Erlandsen, S. L., Bemrick, W. J., Schupp, D. E., Shields, J. M., Jarroll, E. L., Sauch, J. F., and Pawley, J. B. (1990) J. Histochim. Cytochem. 38, 625–632
12. Erlandsen, S. L., Macechko, P. T., van Keulen, H., and Jarroll, E. L. (1990) J. Eukaryot. Microbiol. 37, 416–429
13. Hetsko, M. L., McCaffery, J. M., Meng, T.-C., Que, X., and Gillin, F. D. (1998) Exp. Parasitol. 88, 172–183
14. Jarroll, E. L., Macechko, P. T., Steinme, P. A., Bulik, D., Karr, C. D., van Keulen, H., Paget, T. A., Gerwig, G., Kamerling, J., Vliegenthart, J., and Erlandsen, S. J. (2001) J. Eukaryot. Microbiol. 48, 22–26
15. Gerwig, G. J., van Kuik, J. A., Leeflang, B. R., Kamerling, J. P., Vliegenthart, J. F., Carr, C. D., and Jarroll, E. L. (2000) Glycobiology 12, 499–505
16. Mowatt, M. R., Lujan, H. D., Cotten, D. B., Bowers, Y., Yee, J., Nash, T. E., and Stibb, H. H. (1995) Mol. Microbiol. 15, 965–963
17. Lujan, H. D., Mowatt, M. R., Conrad, J. T., Bowers, B., and Nash, T. E. (1995) J. Biol. Chem. 270, 28207–28213
18. Slavin, I., Saura, A., Carranza, P. G., Teuz, M. C., Nores, M. J., and Lujan, H. D. (2002) Mol. Biochem. Parasitol. 122, 95–98
19. Reiner, D. S., Douglas, D., and Gillin, F. D. (1989) Infect. Immun. 57, 963–968
20. McCaffery, J. M., Faubert, G. M., and Gillin, F. D. (1994) Exp. Parasitol. 79, 236–249
21. Mucino, M., Braun, L., Coussart, P., and Ghosh, P. (2000) Proc. Natl. Acad. Sci. U. S. A. 97, 8784–8788
22. Reiner, D. S., Douglas, D., and Gillin, F. D. (1989) Infect. Immun. 57, 702–709
23. Reiner, D. S., McCaffery, M., and Gillin, F. D. (1990) Eur. J. Cell Biol. 53, 142–153
24. McCaffery, J. M., and Gillin, F. D. (1994) Exp. Parasitol. 79, 220–235
Mining the *Giardia lamblia* Genome for New Cyst Wall Proteins
Chin-Hung Sun, J. Michael McCaffery, David S. Reiner and Frances D. Gillin

J. Biol. Chem. 2003, 278:21701-21708.
doi: 10.1074/jbc.M302023200 originally published online April 9, 2003

Access the most updated version of this article at doi: 10.1074/jbc.M302023200

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

This article cites 56 references, 20 of which can be accessed free at http://www.jbc.org/content/278/24/21701.full.html#ref-list-1