INTRODUCTION

Giardia lamblia (syn. Giardia intestinalis, Giardia duodenalis) is a flagellated unicellular eukaryotic microorganism that commonly causes diarrheal disease throughout the world. It is the most common cause of waterborne outbreaks of diarrhea in the United States (18) and is occasionally seen as a cause of food-borne diarrhea (47a, 227). In developing countries, there is a very high prevalence and incidence of infection, and data suggest that long-term growth retardation can result from chronic giardiasis (105). In certain areas of the world, water contaminated with G. lamblia cysts commonly causes travel-related giardiasis in tourists (33).

Giardia species have two major stages in the life cycle. Infection of a host is initiated when the cyst is ingested with contaminated water or, less commonly, food or through direct fecal-oral contact. The cyst is relatively inert, allowing prolonged survival in a variety of environmental conditions. After...
exposure to the acidic environment of the stomach, cysts excyst into trophozoites in the proximal small intestine. The trophozoite is the vegetative form and replicates in the small intestine, where it causes symptoms of diarrhea and malabsorption. After exposure to biliary fluid, some of the trophozoites form cysts in the jejunum and are passed in the feces, allowing completion of the transmission cycle by infecting a new host.

The clinical aspects of giardiasis have been reviewed recently (260), as has the host immune response to giardiasis (95) and epidemiology (106). This review will focus primarily on the biology of the organism and deal relatively little with the clinical disease or the host-parasite interaction. Since a previous review of Giardia species (3), considerable progress has been made in the understanding of the organism, and these new advances will be emphasized. Many of the major current advances have been facilitated by the ongoing progress of the genome project, based at the Marine Biological Laboratories and involving collaborators at the University of Illinois, University of Texas at El Paso, University of California at San Diego, and University of Arizona (210). The sequences as of June 2000 give at least a single-pass read of 85% of the genome, and the results are available at www.mbl.edu/Giardia.

CLASSIFICATION AND EVOLUTION OF GIARDIA SPECIES

History of the Discovery and Species Designation of Giardia

An appropriate classification for Giardia spp. is critical to an understanding of the pathogenesis and epidemiology of infection, as well as the biology of the organism. This process has been difficult for a number of reasons. (i) The (presumed) asexual nature of the organism does not allow mating experiments to allow species designation. For clonal organisms in the same clade, there are no well-defined criteria for species designation; these designations remain controversial. (ii) Many of the earlier descriptions of Giardia spp. assumed a different species for each host and consequently overestimated the number of species. Subsequent species descriptions based on morphologic differences detected by light microscopy have probably underestimated the differences among isolates, strains, or species. (iii) Cross-transmission experiments of Giardia from one host to another have yielded inconsistent results. (iv) The available tools for distinguishing Giardia isolates have been inadequate until the recent introduction of molecular and electron microscopic techniques for classifying Giardia spp. In view of these concerns, a review of the history of the description of Giardia and the designation of Giardia species is warranted.

Giardia was initially described by van Leeuwenhoek in 1681 as he was examining his own diarrheal stools under the microscope (66). The organism was described in greater detail by Lamb in 1859, who thought the organism belonged to the genus Cercomonas and named it Cercomonas intestinalis (172). Thereafter, some have named the genus after him while others have named the species of the human form after him (i.e., G. lamblia). In 1879, Grassi named a rodent organism now known to be a Giardia species, Dimorphus muris, apparently unaware of Lamb’s earlier description. In 1882 and 1883, Kunstler described an organism in tadpoles (G. agilis) that he named Giardia, the first time Giardia was used as a genus name. In 1888, Blanchard suggested the name Lamblia intestinalis (29), which Stiles then changed to G. duodenalis in 1902 (314). Subsequently, Kofoid and Christiansen proposed the names G. lamblia in 1915 (165) and G. enterica in 1920 (166). There continued to be controversy about the number of Giardia species for many years, with some investigators suggesting species names on the basis of host of origin and others focusing on morphology. For example, over 40 species names had been proposed on the basis of host of origin (169). Simon, on the other hand, used morphologic criteria to distinguish between G. lamblia and G. muris and accepted the name G. lamblia for the human form (304). In 1952, Filice published a detailed morphologic description of Giardia and proposed that three species names be used on the basis of the morphology of the median body; G. duodenalis, G. muris, and G. agilis (104). The species name G. lamblia became widely accepted through the 1970s. Since the 1980s, some have encouraged the use of the name G. duodenalis, and in the 1990s, the name G. intestinalis has been encouraged by other investigators (170). At this time there does not appear to be adequate reason to abandon the term G. lamblia, which has been widely accepted in the medical and scientific literature.

G. lamblia is pear shaped and has one or two transverse, claw-shaped median bodies; G. agilis is long and slender (100) and has a teardrop-shaped median body; and the G. muris trophozoite is shorter and rounder and has a small, rounded median body. G. lamblia is found in humans and a variety of other mammals, G. muris is found in rodents, and G. agilis is found in amphibians (Table 1).

For the Giardia isolates grouped with G. lamblia on the basis of morphologic criteria discernible by light microscopy, differences that can be detected by electron microscopy have allowed the description of additional species, G. psittaci from parakeets (78) and G. ardeae from herons (81). Another species, Giardia microti, has been suggested on the basis of host specificity for voles and muskrats, differences in the cyst as assessed by electron micrography (97), and by differences of the 18S rRNA sequences compared with G. lamblia of human origin (331).

Genotypes of G. lamblia

Molecular classification tools have been of great value in understanding the pathogenesis and host range of Giardia isolates obtained from humans and a variety of other mammals. The first study of the molecular differences of G. lamblia isolates (26) was a zymode analysis of five axenized isolates, three from humans, one from a guinea pig, and one from a cat, using six metabolic enzymes. Zymode analysis consists of the typing of organisms based on the migration of a set of enzymes on a starch gel in the presence of an electric field. The migration depends on the size, structure, and isoelectric point of these enzymes. Since these properties are a function of the primary amino acid sequence, differences in the zymodemes should reflect differences in the sequences of the genes encoding these enzymes. In 1985, restriction fragment length polymorphism analysis of 15 isolates was performed using random probes (251). These studies resulted in the description of three groups; group 3 was so different from groups 1 and 2 that the
suggestion of a separate species designation was made. Subsequently, a number of other molecular classification studies have been performed using zymodeme analysis (2, 16, 47, 212, 214, 234, 277, 316) and restriction fragment length polymorphism analysis (65, 85–87, 90, 137). Pulsed-field gel electrophoresis (PFGE) chromosome patterns have also been studied (7, 42, 143, 167, 291) but are of limited value for classification because of the frequent occurrence of chromosome rearrangements (4, 179). Likewise, classification by surface antigens (250) is limited by antigenic variation of the variant-specific proteins (VSPs) (6, 243). These studies have been very useful, but the conclusions that can be drawn from these types of data are limited by the semiquantitative nature of the data. To allow a more quantitative comparison of \textit{Giardia} isolates, sequence comparisons of the small-subunit rRNA, triosephosphate isomerase (\textit{tim}), and glutamate dehydrogenase (GDH) genes have been utilized in a number of subsequent studies (17, 92, 193, 229–231).

These studies have all confirmed the division of \textit{G. lamblia} human isolates into two major genotypes (Table 2). The first consists of Nash groups 1 and 2, Mayrhofer assemblage A, groups 1 and 2, and the Polish isolates, while Nash group 3, Mayrhofer assemblage B, groups 3 and 4, and the Belgian isolates form the other major genotype. For example, the \textit{tim} nucleotide sequences of the group 1 and 2 isolates diverged by 1% in the protein coding region and 2% in the flanking regions, while groups 1 and 3 diverged by 19% and the flanking regions were so dissimilar as to preclude their alignment (193). The small-subunit or 18S rRNA (SS rRNA) sequence shows a 1% divergence between groups 1 and 3 (333), reflecting its more highly conserved sequence. In addition to their marked genetic differences, the two genotypes may have a number of important biologic differences. For example, the GS isolate (group 3) was significantly more pathogenic in infections of human volunteers than was the WB isolate (group 1) (249). Group 3 organisms also appear to grow much more slowly in axenic culture than do genotype 1 organisms (155).

More recently, a number of additional assemblages (genotypes) have been proposed for \textit{Giardia} isolates from a variety of mammals. These isolates are morphologically identical to human \textit{G. lamblia}, but sequences of their protein-coding regions differ. These studies have allowed the identification of a dog isolate that is genetically distinct from human \textit{G. lamblia}. \textit{Giardia} isolates from dogs have been notoriously difficult to

<table>
<thead>
<tr>
<th>TABLE 1. \textit{Giardia} species</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Species name</strong></td>
</tr>
<tr>
<td>-----------------</td>
</tr>
<tr>
<td><strong>Light microscopy</strong></td>
</tr>
<tr>
<td>\textit{G. agelis}</td>
</tr>
<tr>
<td>\textit{G. muris}</td>
</tr>
<tr>
<td>\textit{G. lamblia}</td>
</tr>
<tr>
<td>\textit{G. ardeae}</td>
</tr>
<tr>
<td>\textit{G. psittaci}</td>
</tr>
<tr>
<td>\textit{G. microti}</td>
</tr>
</tbody>
</table>

<sup>a</sup> NA, not available.

<table>
<thead>
<tr>
<th>TABLE 2. Genotypes of \textit{G. lamblia}</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Proposed designation</strong></td>
</tr>
<tr>
<td>-----------------</td>
</tr>
<tr>
<td>Genotype A-1</td>
</tr>
<tr>
<td>Genotype A-2</td>
</tr>
<tr>
<td>Genotype B</td>
</tr>
<tr>
<td>C</td>
</tr>
<tr>
<td>D</td>
</tr>
<tr>
<td>E (or A-livestock)</td>
</tr>
<tr>
<td>F</td>
</tr>
<tr>
<td>G</td>
</tr>
</tbody>
</table>
axenized, in comparison to isolates from humans (213), leading to the proposal that \textit{Giardia} isolates from dogs were different from cat or human isolates. However, a few dog isolates have been axenized and characterized (17). To further evaluate the zoonotic potential of dog \textit{Giardia}, suckling-mouse infections were established from 11 consecutive infected dogs. On the basis of sequence analysis, these were assigned to two assemblages (C and D) that were quite distinct from assemblages A and B (230). A PCR-based study of nine fecal isolates from dogs found that one of the nine was similar to human isolates while the other eight were different (138). These results suggest that most dog isolates are genetically distinct from those found in humans and have little or no potential for zoonotic transmission.

Separate assemblages (E through G) have also been proposed for hoofed livestock (92), cat (229), and rat (229) isolates (Table 2). This same sequence-based study demonstrated that \textit{G. microti} was a member of this compilation of seven assemblages (229). Further studies of \textit{Giardia} obtained from cattle have demonstrated that some of the isolates belong to the livestock assemblage (assemblage E) while others belong to assemblage A (genotype 1) and thus may have the potential for human infection (259). Assemblages C through G have not yet been isolated from humans, suggesting the likelihood that some genotypes of \textit{G. lambia} have a broad range of host specificity that includes humans while others appear to be more restricted in their host range and may not pose a risk of zoonotic transmission. Whether these seven assemblages should be considered separate species should await further data and consensus. To unify the designation of the \textit{G. lambia} genotypes, I propose an approach that takes into account all the currently available data. Priority should go to the first classification suggested by Nash et al. (251) in 1985. However, “group” has been used in different ways by different authors. In addition, it is clear from data published after the initial description that Nash groups 1 and 2 are closely related while group 3 is genetically distinct enough to allow consideration of a separate species or subspecies name as suggested (251). The assignment of species status to clonal organisms is problematic (324) and should be delayed until there is general agreement among investigators, but agreement on genotype designation should facilitate a better understanding and coordination of the literature. Therefore, I propose that genotype A be accepted as the designation for Nash group 1 (A-1), Nash group 2 (A-2), assemblage A, and the Polish isolates. Genotype B would then designate Nash group 3, assemblage B, and the Belgian isolates. It may also be appropriate to use a similar designation for assemblages C through G, since these also reflect significant genetic differences as well as host differences.

**Host-Parasite Coevolution**

Whenever members of a grouping of parasitic organisms parasitize different hosts, it is appropriate to address the question of coevolution of the host and parasite. For \textit{Giardia} spp., the ability to do so has been limited because of the uncertainty about the host specificity of a specific \textit{Giardia} species or genotype as well as the problem of appropriate classification of isolates. The development of good molecular classification tools and a better understanding of host specificity makes it reasonable to address the possibility of coevolution of different species or genotypes of \textit{Giardia} with their hosts. The greater difference between \textit{G. lamblia} and \textit{G. ardeae} than among the \textit{G. lamblia} genotypes supports the idea that the divergence between \textit{G. lamblia} and \textit{G. ardeae} accompanied the divergence between birds and mammals. However, \textit{G. lamblia} and \textit{G. ardeae} are closer to each other than to \textit{G. muris}, the opposite of the expected finding, since mice diverged from other mammals more recently than from birds. Sequence information is not yet available from \textit{G. agilis} to allow a similar comparison of the amphibian and mammalian \textit{Giardia} species. The closer-than-expected relationship between \textit{G. ardeae} and \textit{G. lamblia} could perhaps be explained if transmission of \textit{Giardia} between birds and mammals was followed by divergence of the two lineages with their hosts.

**\textit{Giardia} and Other Diplomonads as Early-Branching Eukaryotes**

Traditionally, all living organisms have been classified as prokaryotes or eukaryotes, and some still argue for retaining the two major divisions (208). However, the most widely accepted classification now utilizes three major divisions, Archaea (archaeabacteria), Bacteria (eubacteria), and Eukarya (eukaryotes) (353), which can then be divided into kingdoms. With either classification system, \textit{G. lamblia} is clearly a eukaryotic organism and has been considered a member of the protozoa, the more “animal-like” of the unicellular eukaryotes. These protozoan organisms have traditionally been classified by their morphology into flagellates, ciliates, amebae (rhizopods), and sporozoans. Thus, \textit{G. lamblia} was classified with the flagellated protozoans, including the kinetoplastids (e.g., \textit{Leishmania} spp. and \textit{Trypanosoma} spp.), parabasalids (e.g., \textit{Trichomonas vaginalis}), and Dientamoeba (e.g., \textit{Dientamoeba fragilis}) (185). \textit{Giardia} has been placed in the order Diplomonadida (two karyomastigonts, each with four flagella, two nuclei, no mitochondria, and no Golgi complex; cysts are present, and it can be free-living or parasitic) and the family Hexamitidae (six or eight flagella, two nuclei, bilaterally symmetrical, and sometimes axostyles and median or parabasal bodies), along with the mole parasite \textit{Sporonucleus muris} (145) and the free-living organism \textit{Hexamita inflata}. Some of the higher orders of classification do not appear to be phylogenetically valid, such as the placement of all flagellated protozoans together. However, the family Hexamitidae does appear to be a monophyletic group (see below).

One recent classification system with one prokaryotic and five eukaryotic kingdoms has retained Protozoa as one of the six kingdoms (45), but most recent proposals have suggested abandoning the term “protozoa” in favor of the more general but at the same time more precise term “protista” (56, 57, 350). This review will use the term “protista,” but it should be noted that in the clinical context, \textit{G. lamblia} is usually referred to as a protozoan.

Recent classifications of the eukaryotic microbial organisms have depended primarily on molecular comparisons. An ideal molecular classification system would be based on a gene(s) that is required for all life and that is sufficiently highly conserved across all forms of life that accurate alignments allow comparison and classification of all organisms. In many ways,
rRNA has been the most useful gene for molecular comparisons, because RNA sequences are highly conserved across life and because the function of the rRNA is so central to the biology of the organism. Therefore, the most widely accepted classification scheme has been based on SS (18S) rRNA sequences. Based on comparisons of SS rRNA sequences, G. lamblia was proposed as one of the most primitive eukaryotic organisms (308), along with T. vaginalis and the microsporidia. The use of SS rRNA sequences to place Giardia as an early-branching eukaryote has been criticized because of the high G+C content of the SS rRNA of Giardia (75%) and because Giardia spp. are parasitic organisms; artifacts may be introduced by high rates of mutation accompanying host adaptation. An analysis of the early-branching eukaryotes suggested that the basal position of Giardia was an artificial result of long-branch attraction due to a greater evolutionary rate of Giardia (315). Analysis of the large subunit of RNA polymerase II and reanalysis of the eF1 and eF2 sequences also supported the idea of the long-branch artifact (133). However, no such effect was shown for the eRF3 tree (142). In addition, the absence in G. lamblia of the highly conserved N-terminal domain of eRF3 found in other eukaryotes including T. vaginalis suggested the divergence of G. lamblia before the acquisition of the N-terminal domain (142). It should also be noted that the phylogenetic placement of Hexamita, a free-living organism in the same family (as determined by morphologic criteria), avoids artifacts due to high G+C content or parasitism. The G+C content of the Hexamita inflata SS rRNA is 51%, and H. inflata was found to be monophyletic with Giardia (183, 332).

The classification of H. inflata with G. lamblia is also supported by a comparison of the glyceraldehyde-3-phosphate dehydrogenase sequences of G. lamblia, Trepomonas agilis, H. inflata, and Spirotrichonympha sp. (287). A comparison of the SS rRNA sequences of Giardia, Hexamita, Trepomonas, and Spirotrichonympha, all diplomonads, showed that all were phylogenetically related and that the last three comprised one clade while Giardia occupied another clade (46). These sequence-based comparisons have led to the proposed following classification system for Giardia: kingdom Protozoa, subkingdom Archezoa (includes the phyla Metamonada and Microsporidia), subphylum Epharyngia, class Trepomonadacea, subclass Diplzoza, and order Giardiida (includes the families Octomitidae and Giardiidae). In this classification system, the diplomons are considered to comprise the subclass, Diplzoza, and Hexamita, Trepomonas, and Spirotrichonympha are all members of the other diplozoan order, Distomatida (46).

An examination of the different genes used for the phylogenetic classification of Giardia shows that the genes used in transcription and translation generally yield a basal position for Giardia, although artifact due to long-branch attraction has not been ruled out (Table 3). Interestingly, a phylogenetic tree of one of the genes thought to be of mitochondrial origin (cpt660) also suggests that Giardia is an early-branching eukaryote (107, 286). Of the other groups of genes, the metabolic genes do not yield a clear position, but some suggest a eubacterial origin for Giardia. Cytoskeletal genes give mixed results, with actin giving an early divergence but tubulin yielding a divergence that is later than that of Entamoeba histolytica. Among other classes of genes, HSP70 and cathepsin B phylogenies suggest an early divergence for Giardia. Thus, although the answers are not conclusive, most of the data suggest that G. lamblia and the other diplomons are among the most basal of the extant eukaryotes.

Evolution of Eukaryotes

G. lamblia is a typical eukaryotic organism in that it has a distinct nucleus and nuclear membrane, cytoskeleton, and endomembrane system, but it lacks other organelles that are nearly universal in eukaryotes, such as nucleoli and peroxisomes. In addition, G. lamblia is anaerobic, lacking mitochondria or any of the components of oxidative phosphorylation.

The hypothesis that eukaryotes arose through endosymbiotic events, resulting in the origin of plastids and mitochondria and the use of oxidative phosphorylation as the major source of energy production, has become widely accepted (120). One version of this hypothesis is that a common ancestor of the archaeabacteria and eukaryotes evolved a nucleus and cytoskeleton. A descendant then endocytosed a eubacterium, resulting in the development of a mitochondrion. The observation that certain amitochondriate eukaryotes (e.g., G. lamblia and other diplomons as well as Trichomonas vaginalis) were basal on an 18S rRNA tree led to the suggestion that those organisms descended from Archezoa before the plastid symbiotic event (44, 46). However, the subsequent recognition of mitochondrial genes in the genomes of T. vaginalis (39) and G. lamblia (268) has led to the suggestion that the amitochondrial protists have secondarily lost their mitochondria.

An alternative version of the symbiotic hypothesis is that eukaryotes developed through a symbiotic association of an archaeabacterium with a eu bacterium (207). In this proposal, the eu bacterial endosymbiont provided the pathways for organisms depending on the metabolism of pyruvate by mitochondrial respiration as well as those depending on pyruvate metabolism by pyruvate:ferredoxin oxidoreductase (PFOR), either in the cytosol (G. lamblia) or in hydrogenosomes (T. vaginalis). According to this proposal, the organism leading to G. lamblia subsequently lost the enzymatic components of mitochondrial respiration.

BIOLCHEMISTRY AND METABOLISM

Axenic Growth of Trophozoites

G. lamblia trophozoites obtained from a rabbit, chinchilla, and cat were first grown axenically (in the absence of exogenous cells) in 1970 (222). HSP-1 medium, a subsequent modification reported in 1976, contained phytone peptone, glucose, I-cysteine HCl, Hanks solution, and human serum (223). This study reported the first human G. lamblia isolate, Portland-1 (P-1). P-1 and WB, an isolate obtained from a symptomatic human who probably acquired his infection in Afghanistan (307), belong to the same genotype and have been used for many of the studies of G. lamblia. The growth medium has subsequently been modified, and currently the most commonly used medium is modified TYI-S-33 (159) (Table 4). Among the notable requirements for axenic growth are the absolute requirements for a low O2 concentration, a high cysteine concentration, and the requirement for exogenous lipids, which are obtained from the serum component. When kept at 37°C, the trophozoites adhere to the glass wall of the tube in which
they are grown. This adherence is dependent on glycolysis and on contraction of the proteins of the ventral disk (99). To date, the other species of *Giardia* (e.g., *G. muris* and *G. agilis*) have not been grown axenically.

**Carbohydrate Metabolism**

Most eukaryotic organisms depend primarily on aerobic metabolism for their energy production. However, certain eukaryotes, including *Trichomonas* spp., *Entamoeba* spp., and *Giardia* spp., are characterized by their lack of mitochondria and cytochrome-mediated oxidative phosphorylation. They rely on fermentative metabolism (even when oxygen is present) for energy conservation. Glycolysis and its brief extensions generate ATP, with generation dependent only on substrate level phosphorylation. Glucose is not completely oxidized to CO$_2$ and H$_2$O as in aerobic metabolism but is incompletely catabolized to acetate, ethanol, alanine, and CO$_2$. The balance of end product formation is sensitive to the O$_2$ tension and glucose concentration in the medium. The proposed pathways of energy production from glucose and aspartate are shown in Fig. 1 to 3.

The metabolism of trophozoites is markedly affected by small changes in oxygen concentration. Under strictly anaerobic conditions, alanine is the major product of carbohydrate metabolism (72, 263, 265) (Fig. 2). Even with the addition of minimal amounts of O$_2$ (i.e., concentrations of <0.25 μM), ethanol production is stimulated and alanine production is inhibited (263). With further increases in O$_2$ concentration, ethanol and alanine production are inhibited. At O$_2$ concentrations of >46 μM, alanine production is completely inhibited and acetate and CO$_2$ are the predominant products of energy metabolism. These oxygen concentrations are likely to be relevant to the intestinal milieu in which the trophozoites replicate since the oxygen concentration in this environment is estimated to vary between 0 and 60 μM. Thus, the pathway of metabolism of pyruvate appears to be altered for differing anaerobic or microaerophilic environments.

Despite the anaerobic metabolism, trophozoites do produce

<table>
<thead>
<tr>
<th>Table 3. Molecular phylogeny of <em>G. lamblia</em></th>
</tr>
</thead>
<tbody>
<tr>
<td>Gene</td>
</tr>
<tr>
<td>SS (18S) rRNA</td>
</tr>
<tr>
<td>Translational apparatus</td>
</tr>
<tr>
<td>RNA polymerase III</td>
</tr>
<tr>
<td>EF1-alpha</td>
</tr>
<tr>
<td>EF2</td>
</tr>
<tr>
<td>Eukaryotic release factor 1 (eRF1)</td>
</tr>
<tr>
<td>Eukaryotic release factor 3 (eRF3)</td>
</tr>
<tr>
<td>RNA polymerase II large subunit (RPB1)</td>
</tr>
<tr>
<td>Transcription apparatus</td>
</tr>
<tr>
<td>Fibrillarin</td>
</tr>
<tr>
<td>Mitochondrial genes</td>
</tr>
<tr>
<td>cpn60</td>
</tr>
<tr>
<td>Varyl-tRNA synthetase</td>
</tr>
<tr>
<td>Cytoskeleton</td>
</tr>
<tr>
<td>Annexins</td>
</tr>
<tr>
<td>Actin</td>
</tr>
<tr>
<td>β-Tubulin</td>
</tr>
<tr>
<td>Metabolic pathways</td>
</tr>
<tr>
<td>Triosephosphate isomerase</td>
</tr>
<tr>
<td>Glutamate dehydrogenase (NADP dependent)</td>
</tr>
<tr>
<td>Adenylate kinase</td>
</tr>
<tr>
<td>Glyceraldehyde-3-phosphate dehydrogenase</td>
</tr>
<tr>
<td>Pyruvatephosphate dikinase</td>
</tr>
<tr>
<td>Malate dehydrogenase</td>
</tr>
<tr>
<td>Malate dehydrogenase, decarboxylating</td>
</tr>
<tr>
<td>Fructose-1,6-bisphosphate aldolase (class II)</td>
</tr>
<tr>
<td>Acetyl-CoA synthetase</td>
</tr>
<tr>
<td>Other genes</td>
</tr>
<tr>
<td>Cathepsin B (cysteine protease)</td>
</tr>
<tr>
<td>Signal recognition peptide</td>
</tr>
<tr>
<td>HSP70 (cytosolic and Grp78/Bip)</td>
</tr>
<tr>
<td>CDC2 (member of cyclin-dependent kinase family)</td>
</tr>
</tbody>
</table>
TABLE 4. Medium for in vitro growth of *G. lamblia*  

<table>
<thead>
<tr>
<th>Component</th>
<th>Quantity/liter</th>
<th>Final concn</th>
</tr>
</thead>
<tbody>
<tr>
<td>K&lt;sub&gt;2&lt;/sub&gt;HPO&lt;sub&gt;4&lt;/sub&gt;</td>
<td>1 g</td>
<td>4.4 mM</td>
</tr>
<tr>
<td>H&lt;sub&gt;2&lt;/sub&gt;O</td>
<td>600 m</td>
<td>4.4 mM</td>
</tr>
<tr>
<td>Trypsinase (casein peptone)</td>
<td>2 g</td>
<td>34 mM</td>
</tr>
<tr>
<td>Yeast extract</td>
<td>2 g</td>
<td>16.5 mM</td>
</tr>
<tr>
<td>α-Glucose</td>
<td>10 g</td>
<td>56 mM</td>
</tr>
<tr>
<td>NaCl</td>
<td>20 g</td>
<td>1.1 mM</td>
</tr>
<tr>
<td>Cysteine-HCl</td>
<td>20 g</td>
<td>22.8 mg</td>
</tr>
<tr>
<td>Ascorbic acid</td>
<td>200 mg</td>
<td>10%</td>
</tr>
<tr>
<td>Ferric NH&lt;sub&gt;4&lt;/sub&gt;-citrate (Sigma F-5879)</td>
<td>40 ml of 6.5% solution</td>
<td>40 ml of 10 M solution</td>
</tr>
<tr>
<td>Bovine bile (Sigma B-8381)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sodium</td>
<td>10%</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Component</th>
<th>Quantity/liter</th>
<th>Final concn</th>
</tr>
</thead>
<tbody>
<tr>
<td>Serum (bovine or fetal calf)</td>
<td>100 ml</td>
<td>10%</td>
</tr>
</tbody>
</table>

<sup>a</sup> Formula of TYI-S-33 (159), the most commonly used medium for growth of *G. lamblia* trophozoites. The ferric-NH<sub>4</sub> citrate and bovine bile are generally kept as solutions at 4°C. All components except the serum are dissolved in water, brought to 200 ml, sterilized on a 0.22- or 0.45-μm-pore-size filter, and added to 700 ml of distilled water. Heat-inactivated (56°C for 20 min) serum is then added. During in vitro growth, the organisms are grown in sealed glass containers nearly filled with medium. When this is impossible, such as during cloning by limiting dilution in 96-well plates (243), a sealed bag containing an anaerobic generator is used. At 4°C, the shelf life is limited to 5 to 7 days, primarily because of the degradation of the cysteine.

Glucose supplies the major source of energy derived from carbohydrates (147). Glucose is converted to pyruvate by the Embden-Meyerhof-Parnas and hexose monophosphate shunt pathways (Fig. 1). For most eukaryotic and prokaryotic organisms, the conversion of fructose-6-phosphate to fructose-1,6-bisphosphate is an irreversible and regulated step that is catalyzed by an ATP-dependent phosphofructokinase. However, for *Giardia* (219), as well as for *T. vaginalis* (221) and *E. histolytica* (279), this reaction is catalyzed by a pyrophosphate-dependent phosphofructokinase. In contrast to the ATP-dependent enzyme, this enzyme catalyzes a reversible reaction and is not a regulated enzyme (220). The pyrophosphate-dependent phosphofructokinase from *Giardia* has been cloned and characterized (274, 289). Triosephosphate isomerase, an enzyme that catalyzes a reversible conversion between dihydroxyacetone phosphate and δ-glyceraldehyde-3-phosphate, was cloned by complementation in *Escherichia coli* and characterized (238).

Two enzymes that convert phosphoenolpyruvate into pyruvate, ATP-dependent pyruvate kinase (271), and pyrophosphate-dependent pyruvate phosphate dikinase (PPDK) have been identified (37, 132, 141). There is a potential energy advantage in the reaction mediated by PPDK, since two molecules of ATP can be generated by a coordinated reaction involving PPDK and adenylate kinase. Adenylate kinase (288) converts two ADP molecules into ATP + ADP by a reaction that is essentially energy neutral, and PPDK converts phosphoenolpyruvate plus AMP into pyruvate + ATP, resulting in some oxygen free radicals, and a mechanism for detoxification of oxygen is necessary. Detoxification of oxygen free radicals in aerobic organisms is typically accomplished by an enzymatic pathway beginning with superoxide dismutase. Superoxide dismutase activity has been detected by some (188), but other investigators did not detect superoxide dismutase, catalase, or peroxidase activities in *G. lamblia* (35) and proposed an H<sub>2</sub>O<sub>2</sub>-producing NADH oxidase as the major route of oxygen detoxification in *G. lamblia* (36).

Routine *Giardia* medium contains 50 mM glucose, and many of the metabolic studies of trophozoites have been performed with that medium. When the glucose concentration is reduced to 10 mM, there is little effect on trophozoite growth (299). At glucose concentrations below 10 mM, replication rates are reduced by nearly 50% and ethanol production is markedly reduced, alanine production is moderately reduced, and acetate production is unaffected. Thus, glucose promotes trophozoite growth but is not absolutely essential. The growth enhancement by glucose was not provided by other sugars.

In *T. vaginalis*, glycolytic enzymes (glucose to pyruvate) are found in the cytosol but those involved in pyruvate oxidation are localized in a separate organelle, the hydrogenosome (152). In *Giardia*, there is no metabolic compartmentalization; rather, all the reactions occur in the cytosol or on the cytosolic surfaces of membranes for PFOR-mediated reactions (188).

Glucose Metabolism

FIG. 1. Metabolism of glucose to phosphoenolpyruvate. Many of the enzymes have been documented in terms of enzymatic activity, isolation of the enzyme, or cloning of the gene encoding the enzyme. The enzymes are labeled as follows: 1, hexokinase (188); 2, glucose phosphate isomerase (proposed); 3, pyrophosphate-dependent phosphofructokinase (186, 219, 274, 289); 4, fructose bisphosphate aldolase (128, 188); 5, triosephosphate isomerase (238); 6, glyceraldehyde-3-phosphate dehydrogenase (287); 7, phosphoglycerate kinase (proposed); 8, phosphoglyceromutase (proposed); 9, enolase (proposed). Certain enzymes (for steps 2, 7, 8, and 9) are suggested on the basis of pathways in other organisms but have not yet been proved for *Giardia*. This figure has been adapted from material presented in prior reviews (55, 146, 150) and updated from more recent literature as cited for individual enzymes.
the net generation of two ATP molecules during the conversion of phosphoenolpyruvate to pyruvate, rather than the one ATP produced by the pyruvate kinase reaction. Their relative roles in glycolysis have not yet been determined, but the higher specific activity of pyruvate kinase suggests that it may play a major role in glycolysis (271).

The conversion of pyruvate to acetyl coenzyme A (acetyl-CoA) (Fig. 3) is catalyzed by PFOR, which utilizes ferredoxin rather than NAD as the electron acceptor (188, 326, 327), in place of the pyruvate dehydrogenase complex found in aerobic eubacteria and eukaryotes. PFOR is also found in *E. histolytica* (280) and *T. vaginalis* (352).

Metronidazole is an antimicrobial agent with a broad spectrum of activity against anaerobic bacteria and protozoa. It is activated when its 5-nitro group is reduced by ferredoxin that has in turn been reduced by PFOR, generating a toxic nitro radical. Both metronidazole and sodium nitrite, another respiratory inhibitor thought to act by destroying the iron-sulfur center of PFOR, were toxic to *G. muris* trophozoites, but only sodium nitrite was toxic to cysts. It was proposed that they had different effects because of the inability of metronidazole to enter the cyst, but this proposal was not directly tested (262).

In a study of different *G. lamblia* isolates, some of which were fully susceptible to metronidazole and some of which were relatively resistant, decreased ferredoxin levels were associated with drug resistance (192).

Acetyl-CoA can be converted directly to acetate by ADP-forming acetyl-CoA synthetase (293), resulting in the production of ATP from ADP as acetyl-CoA is converted to acetate. Alternatively, acetyl-CoA is converted to ethanol, using acet-
aldehyde as an intermediate, by the bifunctional enzyme alcohol dehydrogenase E (59, 292, 295). Alcohol dehydrogenase E has an acetaldehyde dehydrogenase activity in the amino terminus that catalyzes the conversion of acetyl-CoA to acetaldehyde and an alcohol dehydrogenase activity in the carboxyl terminus that converts the acetaldehyde to ethanol.

**Amino Acid Metabolism**

Amino acids are becoming increasingly recognized as important components of the energy metabolism of *G. lamblia*. The uptake of aspartate, alanine, and arginine from the extracellular medium, as well as the documentation of glucose-independent metabolism, suggests the potential importance of amino acid metabolism for energy production in *Giardia* (215, 297, 299).

The arginine dihydrolase pathway is one potential source of energy (74, 297, 300) (Fig 4). This pathway is present in a number of prokaryotic organisms, but among eukaryotes it has been documented only in *T. vaginalis* (191) and *G. lamblia*. In the arginine dihydrolase pathway, arginine is converted to ornithine and ammonia with the generation of ATP from ADP by substrate-level phosphorylation. Ornithine is subsequently exported in exchange for extracellular arginine by a transporter mechanism (298).

Aspartate is another potential source of energy. It is converted to oxaloacetate by aspartate transaminase, entering the intermediary pathway, where it is converted to pyruvate via a malate intermediate (215) (Fig. 2).

Alanine also appears to play an important role in allowing trophozoites to adapt to hypoxic challenge. With an isosmotic extracellular environment, the intracellular alanine concentration is 50 mM (161, 270), and with a hypoxic challenge, the concentration of alanine rapidly decreases by an active transport mechanism. (In addition to alanine, potassium appears to play a role in osmoregulation of trophozoites [204].)

The secretion of alanine occurs via an alanine transporter that also transports L-serine, glycine, and L-threonine, L-glutamine, and L-asparagine (73, 258). This transporter acts as an antiporter, exchanging intracellular alanine for these other amino acids from the extracellular environment (298).

Aside from the synthesis of alanine as a by-product of energy metabolism, the only other amino acid for which de novo synthesis has been documented is valine. Thus, *Giardia* lacks synthesis of most amino acids and depends on scavenging them from the intestinal milieu in which the trophozoite replicates.

One of the notable requirements for axenic growth of *G. lamblia* trophozoites is the absolute requirement for a relatively high concentration of cysteine (16 mM). Cysteine also provides a partial protection from the toxicity of oxygen that is not seen with other reducing agents, including cystine, and therefore appears to be a specific effect of cysteine (109, 110, 113). Cysteine is not synthesized de novo and is not synthesized from cystine (202). It appears to be imported into the cell by passive diffusion, although active transport may account for some of the acquisition of cysteine (202). The importance of free thiol groups on the surfaces of trophozoites was demonstrated by the toxicity of thiol-blocking agents that are unable to penetrate intact cells (116). This toxicity suggests that these agents are reacting with thiol groups on the trophozoite surfaces, killing the trophozoites. Cysteine appears to be the major thiol group present (34). When trophozoites are metabolically labeled with radiolabeled cysteine, most of the label is incorporated into the VSPs (6, 11), suggesting that these surface proteins may play a role in protection of the trophozoite from oxygen toxicity (see “Structure and biochemistry of the VSPs” below).

**Lipid Metabolism**

The growth of trophozoites predominantly in the duodenum and jejunum initially suggested the possible importance of bile in the growth of *Giardia* trophozoites. Short-term axenic growth in the absence of serum can be supported by bile (111). The biliary lipids cholesterol and phosphatidylcholine and the bile salts glycocholate and glycodeoxycholate will also support growth in the absence of serum (111). Serum is required for longer-term axenic growth, but it has been shown that the Cohn IV-1 fraction of bovine serum (enriched in alpha globulins, lipoproteins, and growth factors) can substitute for whole serum (194). In fact, insulin-like growth factor II, which is present in fraction IV-1, stimulates trophozoite growth and cysteine uptake. The same fraction from a number of other mammals was also effective in supporting growth, although in some cases antibody depletion was required.

*G. lamblia* trophozoites do not have the capacity of de novo synthesis of fatty acids (149), with the possible exception of certain minor fatty acids (75). However, free fatty acids are toxic to trophozoites, demonstrating 50% lethal doses of 2 to 12 μM (283). In fact, the toxicity of human milk for *G. lamblia* trophozoites appears to be mediated through products of milk lipolysis (129, 283). The trophozoites appear to satisfy their lipid requirements by obtaining cholesterol and phosphatidylcholine from the external environment (94, 111, 200). The cholesterol and phospholipids are supplied by lipoproteins, β-cyclodextrins, and bile salts, with transfer of lipids to the

---

**FIG. 4. Arginine dihydrolase pathway (74, 297, 300). The enzymes are labeled as follows: 1, arginine deiminase (163); 2, ornithine transcarbamoylase (297, 300); 3, carbamate kinase (226).**

Arginine $\rightarrow$ Citrulline $\rightarrow$ Carbamyl Phosphate $\rightarrow$ CO$_2$ + NH$_4$ + ADP $\rightarrow$ ATP

Ornithine (exported in exchange for extracellular arginine)
Parasite surface being facilitated by bile salts (200). It has also been suggested that a low level of endocytosis of lipids occurs (200). Conjugated bile acids appear to be taken up by a carrier-mediated mechanism that includes different carriers for cholate and cholyglycine (60).

The major fatty acids found in axenically grown trophozoites are palmitic acid, stearic acid, and oleic acid (75). Fatty acid desaturase activity, including desaturation of oleate to linoleate and linolenate, has been documented (75). Arachidonic acid is incorporated into neutral lipids, phospholipids, and a wide variety of cellular lipids (108), while palmitic acid, myristic acid, and oleic acid are transesterified primarily into phospholipids (28), including cellular phospholipids, (e.g., phosphadiyglycerol, phosphatidycholine, phosphatidylethanolamine, and phosphatidylinositol) (149, 154, 228, 313). Interesterification also occurs with incorporation of conjugated fatty acids into phosphatidylglycerol (108). The toxicity of certain analogs of phosphatidylglycerol for trophozoites has been documented, although the mechanism of this toxicity has not been determined (108).

Isoprenoids are lipids derived from mevalonate that are commonly found in eukaryotic cells. The most notable end product is cholesterol, but isoprenoids are also incorporated into proteins such as the GTP-binding proteins by posttranslational modification. Isoprenylation of proteins has been demonstrated by the incorporation of radiolabeled mevalonate into trophozoite proteins (197). Incorporation of mevalonate and cell growth were inhibited in a reversible manner by competitive inhibitors of 3-hydroxy-3-methylglutaryl (HMG)-CoA reductase. Inhibitors of later steps of isoprenylation permanently inhibited cell growth.

**Purine and Pyrimidine Salvage**

Many of the pathogenic protozoa, including *G. lamblia* (339), depend on salvage pathways for obtaining purine nucleosides. In addition, *G. lamblia* (13), as well as *T. vaginalis* (340) and *Trichomonas foetus* (341), lack pyrimidine synthetic pathways and depends on salvage pathways for obtaining both purine and pyrimidine nucleosides.

Studies of purine metabolism using radiolabeled precursors have demonstrated the incorporation of adenine, adenosine, guanine, and guanosine into nucleotides but no incorporation of components of the de novo synthetic pathway, such as formate, glycine, hypoxanthine, inosine, or xanthine (23, 224, 339). The likely scenario is that the purine nucleosides (adenosine and guanosine) are imported by a transporter with broad specificity for nucleosides as well as deoxyribonucleosides (19, 63). The purine nucleosides are then broken down to the bases by their respective hydrolases (Fig. 5). Phosphoribosyl 1-pyrophosphate is synthesized by phosphoribosyl 1-pyrophosphate synthetase (171) and reacts with the salvaged purine bases to produce the nucleoside-5'-monophosphate in a reaction catalyzed by the respective monophosphate phosphoribosyltransferase (PRTase). The PRTases from most eukaryotes utilize hypoxanthine or xanthine as substrate, but the *G. lamblia* PRTase is highly specific for guanine (261, 311), indicating that guanine is the only source for guanine nucleotides. The amino acid sequence of the PRTase enzyme shows less than 20% identity to the human enzyme (311). The crystallographic structure of the *G. lamblia* PRTase has suggested possible reasons for the unique substrate of the *G. lamblia* enzyme, such as an aspartate substitution for leucine at position 181 (303).

*G. lamblia* trophozoites also depend on the salvage of exogenous thymidine, cytidine, and uridine for the synthesis of the pyrimidine nucleotides (13, 190) (Fig. 6). Uridine is probably imported by the thymidine transporter (62, 89) and broken down into uracil by a hydrolase (13) or phosphorylase (181). Cytidine and uridine can also be imported by a broad-specificity transporter (63), but most cytidine probably enters the synthetic pathway for UTP synthesis rather than being converted more directly to CTP (Fig. 6). The majority of CTP is probably generated from UTP by CTP synthetase (151). A low-affinity nucleobase transporter has also been documented (89). Although the transporter’s function has not been determined, it may be involved in export of unusable (e.g., thymine) or excessive quantities of nucleobases.

DNA synthesis also depends on the salvage of exogenous deoxynucleotides. *G. lamblia* trophozoites lack ribonucleotide reductase and do not incorporate purine bases or nucleosides into DNA. Rather, they depend on the salvage of exogenous purine deoxynucleosides (20) (Fig. 7). A single purine deoxynucleoside kinase for the synthesis of deoxynucleotides from their respective deoxynucleosides (deoxyadenosine and deoxyguanosine) and a pyrimidine deoxynucleoside kinase catalyze the production of the pyrimidine deoxynucleotides from the deoxynucleosides thymidine and cytosine (176).

**FIG. 5.** Purine ribonucleoside salvage pathways. The enzymes are labeled as follows: 1, adenosine hydrolase (339); 2, adenine phosphoribosyltransferase (APRTase); 3, guanosine hydrolase (339); 4, guanine phosphoribosyltransferase (GPTase) (339).

**FIG. 6.** Pyrimidine ribonucleoside salvage pathways. The enzymes are labeled as follows: 1, uracil phosphoribosyltransferase (UPRTase) (58); 2, uridine/thymine phosphorylase (181) (a uridine hydrolase [13] not confirmed); 3, uridine phosphoribosyltransferase (kinase) (13) (not confirmed [336]); 4, UMP kinase (151); 5, UDP kinase (151); 6, CTP synthetase (151, 187); 7, CDP kinase (151); 8, cytosine phosphoribosyltransferase (CPRTase) (13) (low level of activity [151]); 9, cytidine hydrolase (13); 10, cytidine deaminase (336). The initial pyrimidine salvage pathway was described in (reference) and has been updated from more recent literature (see the references cited for individual enzymes).
CELL BIOLOGY

Trophozoite Structure

The *G. lamblia* trophozoites are pear-shaped and are approximately 12 to 15 µm long and 5 to 9 µm wide. The cytoskeleton includes a median body, four pairs of flagella (anterior, posterior, caudal, and ventral), and a ventral disk (Fig. 8 and 9). Trophozoites have two nuclei without nucleoli that are located anteriorly and are symmetric with respect to the long axis (Fig. 9). Lysosomal vacuoles, as well as ribosomal and glycogen granules, are found in the cytoplasm. Golgi complexes become visible in encysting trophozoites but have not been confirmed to be present in vegetative trophozoites (117). However, stacked membranes suggestive of Golgi complexes have been demonstrated (174, 309).

**FIG. 8.** Trophozoite coronal section. A coronal view of a trophozoite demonstrates the nuclei (N), endoplasmic reticulum (ER), flagella (F), and vacuoles (V). A mechanical suction is formed when the ventral disk (VD) attaches to an intestinal or glass surface. Components of the ventral disk include the bare area (BA), lateral crest (LC), and ventrolateral flange (VLF). A magnified view of the ventral disk is shown in Fig. 10.

---

G. *lamblia* trophozoites have two nuclei that are nearly identical in appearance (Fig. 8 and 9). They replicate at approximately the same time (351) and are both transcriptionally active (153) as determined by uridine incorporation into nuclear RNA. Both nuclei have approximately equal numbers of rDNA genes as determined by in situ hybridization using the rDNA probe (153). Both have approximately equal amounts of DNA as determined by the intensity of nuclear staining with 4′,6-diamidino-2-phenylindole (DAPI) (153) or propidium iodide (25), although the possession of equal amounts of DNA by the two nuclei has been questioned (M. Frisardi and J. Samuelson, Woods Hole Molecular Parasitology Meeting, abstr. 263A, 1999, and abstr. 202C, 2000). It is generally assumed that the two nuclei have the same complement of genes and chromosomes, and results from our laboratory using fluorescence in situ hybridization with single-copy genes support this assumption (L. Yu, C. W. Birky, and R. D. Adam, unpublished observations). These features all differ from those of the ciliated protista that are binucleate, such as *Tetrahymena* and *Paramecium* (276). These organisms have a smaller micronucleus that contains the genomic DNA but is not transcriptionally active. DNA from the micronucleus is amplified into multiple copies, forming the macronucleus, from which transcription occurs. *Giardia* spp. and the other diplomonads appear to be unique in their possession of two nuclei that are identical by the above parameters.

The nuclei of higher eukaryotes have readily visible nucleoli,
which are the sites of rRNA transcription. Fibrillarin is a major component of nucleoli and is required for pre-rRNA processing as well as for viability in Saccharomyces cerevisiae (144, 296, 325). Nucleoli have not been identified in Giardia lamblia nuclei, and antibody to Giardia lamblia fibrillarin diffusely stains the nuclei of Giardia lamblia, suggesting that rRNA transcription and processing are not localized to certain regions of the nuclei (240).

Axenic cultures have not been synchronized, and the division time is very brief, so studies of cellular and nuclear division have been limited to observations of trophozoite replication based on static microscopy. On the basis of these observations, it has been proposed that during cellular division, the nuclei move laterally followed by nuclear replication, resulting in trophozoites with four nuclei. The trophozoites then divide in the longitudinal plane in such a way as to maintain the left-right asymmetry (48, 104, 153). Electron micrographs of excysting organisms undergoing cytokinesis have shown cells that appear to be joined laterally or with the ventral disks opposing each other (130). Laterally joined cells would be consistent with the above mechanism of cytokinesis, but replication that results in opposing ventral disks would invert the left-right asymmetry with each division. Thus, it is not yet possible to be certain of the morphology of cytokinesis.

A number of drugs that arrest the cell cycle of mammalian cells have been tested for their effect on trophozoites. Colchicine (mammalian G₂+M) and gamma irradiation (mammalian G₂+M) had no effect, while hydroxyurea (mammalian G₂+M) and razoxane (mammalian G₂/S, sometimes G₂+M), arrested the cell cycle in the G₂+M phase (140). To date, axenic cultures of Giardia lamblia have not been synchronized.

Cytoskeleton and Motility

Trophozoites colonize the small intestine of their host, predominating in the mid-jejunum. They attach by their concave ventral surfaces (ventral adhesive disk) to the intestinal wall, where they obtain the necessary nutrients and avoid transport beyond the jejunum. The ventral disk mediates a mechanical attachment not only to the intestinal wall but also to the surface of the container used for axenic growth. Neither cellular invasion nor receptor-mediated attachment has been documented for Giardia spp. Therefore, the cytoskeleton and especially the ventral disk play a key role in the survival of the organism in the intestine of the host.

The ventral disk is a unique and important component of the Giardia lamblia cytoskeleton. Grossly, it appears as a concave structure with a maximum depth of 0.4 μm covering the entire ventral surface. The edge narrows into a lateral crest, and a more flexible ventrolateral flange (Fig. 8) surrounds the disk (82). The disk contains the contractile proteins actinin, α-actinin, myosin, and tropomyosin (103) as the biochemical basis for the contraction of the disk involved in adherence.

FIG. 9. Trophozoite cross section. A cross-sectional view of a trophozoite demonstrates the nuclei (N), flagella (F), vacuoles (V), and endoplasmic reticulum (ER).
ment depends on active metabolism and is inhibited by temperatures below 37°C, increased oxygen levels, or reduced cysteine concentrations (109, 113). Ultrastructurally, the ventral disk includes a set of microtubules containing 13 protofilaments and linked to the ventral membrane. These microtubules form the base of the microribbons (dorsal ribbons) that extend nearly perpendicular from the membrane (Fig. 8 and 10). The protein constituents of the dorsal ribbons (microribbons) include a set of giardins, which are alpha-coiled-helix proteins approximately 29 to 38 kDa in size. The giardins line the edges of the microribbons but are not found in the microtubules (273). Although 23 forms of these proteins have been separated by two-dimensional electrophoresis, the N-terminal sequences of some variants are identical, suggesting that post-translational modification accounts for some of the forms (273). Several giardins, α1-giardin (273), α2-giardin (15), β-giardin (10, 134), and γ-giardin (257), have been cloned, and their sequences confirm the alpha-helical structure of these proteins. The α1-giardin and α2-giardin sequences demonstrate approximately 80% identity at the nucleotide and amino acid levels (15), while the other giardins have no significant sequence similarity. Of the two sequences published for the β-giardin gene, the one found in reference (134) is the correct one except that the start codon is 13 amino acids upstream from the reported cDNA sequence (2 amino acids upstream from the reported genomic sequence) (10), yielding a slightly higher predicted molecular mass (R. Adam, unpublished results). A protein with immunologic cross-reactivity with β-giardin, but significantly larger, has also been identified as a cytoskeletal protein (head stalk protein; HPSR2) (206). This 183-kDa protein has a long coiled-coil stalk and an N-terminal hydrolytic domain. Whether it has the same cellular distribution as the giardins has not been reported.

Tubulins are not found in the microribbons but are found in the microtubules (310). The microtubules of the ventral disk, as well as those of the flagella, are presumably composed of αβ-tubulin. Posttranslational modifications of the G. lamblia tubulins include acetylation and polyglycylation (310, 345, 346). Benzimidazoles are compounds used primarily as antihelminthic agents, but they also have significant in vitro activity against G. lamblia as well as in vivo efficacy for giardiasis (49, 68, 70, 122, 233). Their effect is thought to be mediated by their interaction with β-tubulin. G. lamblia trophozoites exposed to albendazole (or other benzimidazoles) lose their ability to adhere despite normal flagellar movement (70). This suggests a difference between the tubulin found in the ventral disk and that found in the flagella. This, as well as the observation that adherence can occur in the absence of flagellar function (103), suggest that the ventral disk is important for adherence but the flagella are not. With more prolonged exposure (e.g. 24 h) to albendazole, the ventral disk becomes fragmented, with dislocation of the microribbons and microtubules of the ventral disk (49). Substantial electron-dense precipitates can be found in the microtubules and microribbons and to a lesser extent in the other tubulin-containing structures, such as the median body and flagella (49). The pronounced effect of albendazole on the microribbons, despite the apparent lack of tubulin in the microribbons, raises the possibility that the benzimidazoles are reacting with proteins other than tubulins, such as the giardins.

Vinculin is a protein that binds α-actinin and mediates the attachment of actin filaments to membrane sites. Its location in attachment regions of the ventral disk has led to the suggestion that it may be involved in the attachment process (241). Further functional studies are required to confirm or refute this proposal.

The trophozoite has four pairs of flagella that begin at two sets of basal bodies that are near the midline and anteroventral to the nucleus. They emerge from the anterior, posterior, caudal, and ventral regions of the trophozoite. Paraflagellar rods extend along one side of the two ventral flagella (102, 135). Nine pairs of microtubules encircle two microtubules to form the flagella (Fig. 8). The flagella appear to be important for motility but not for attachment. In addition, their early emergence through the cyst wall during the process of excystation suggests their importance in excystation (38) (see “Excystation” below).

The median body is a component of the cytoskeleton that is located in the midline and dorsal to the caudal flagella and consists of a group of microtubules in a tight bundle. It is unique to Giardia spp., and its morphology helps define the morphologic characteristics of the different Giardia species (Table 1). G. lamblia trophozoites typically have two median bodies that are shaped like claw hammers. The median body has been proposed as the assembly site for microtubules to be incorporated into the ventral disk (216). Caltractin/centrin, a calcium binding protein that is responsible for basal-body ori-
entation, has been identified in the basal bodies as well as the paraflagellar rods and median bodies (21, 216). A 101-kDa coiled-coil protein has also been localized to the median body by immunofluorescence microscopy (205).

**Protein Transport and Degradation**

**Endomembrane protein transport system.** The endoplasmic reticulum (ER) and Golgi complex are part of a eukaryotic endomembrane system involved in protein folding and translocation. In higher eukaryotes, proteins destined for secretion have a signal sequence that directs them to the ER as are translated in the ribosomes. The signal sequence binds to the signal recognition particle (SRP). This complex then binds to the SRP receptor (SR) on the cytoplasmic portion of the ER. The SR is a dimeric protein consisting of the membrane-anchored SRQ and the GTPase, SRα. After their translocation to the ER, chaperonins such as BiP (the HSP70 homologue found in the ER) aid in folding and further transport. Although structures consistent with ER had previously been identified by electron microscopy (EM), there was some doubt until recently as to the existence of the ER in *G. lamblia*. However, the cloning and characterization of SRQα, as well the identification of an extensive membrane system labeled with antibody to BiP (309), has clearly demonstrated the existence of the ER. One of the important aspects of protein folding includes the correct formation of disulfide bonds, which is accomplished in the lumen of the ER by protein disulfide isomerase. Three protein disulfide isomerase genes from *G. lamblia* have been cloned and characterized, and their products have been localized to the ER (162).

Golgi complexes have not been detected in trophozoites by standard microscopic techniques but have been demonstrated in encysting organisms (117, 282) (see “Encystation” below). More recently, transmission and freeze fracture EM of nitrobenzoxadiazole (NBD) ceramide-labeled log phase trophozoites demonstrated heavy staining in the perinuclear region in a pattern similar to that of the Golgi complex from other organisms (174). ADP-ribosylation factor (ARF) is a GTP binding protein required for the budding of clathrin-coated and COP I-coated vesicles from the Golgi complex. ARF is expressed in *G. lamblia* trophozoites and has been used to complement ARF function in *S. cerevisiae* (182, 239). Antibody to ARF and to β-COP (from the COP I coat) label perinuclear vesicles in a manner that is inhibited by brefeldin A (195), also suggesting that *G. lamblia* trophozoites may have a form of Golgi complex. However, further investigation is required to further confirm the presence of a Golgi complex in trophozoites and to determine its characteristics.

Specific importation of proteins into the nuclei has also been documented by using the simian virus 40 nuclear localization signal to direct green fluorescent protein (GFP) to the nuclei (76). Ran is a nuclear GTP binding protein that is involved in the import of proteins in the nuclei of *Xenopus* oocytes and participates in cell cycle progression in yeast (*spp1*). It has been cloned and characterized in *G. lamblia*, but its cellular location and function have not yet been determined (50).

**Endosome-lysosome vacuoles.** Most eukaryotes have a system of endosomes and lysosomes that degrade and recycle endogenous proteins or those acquired by endocytosis or phagocytosis from the extracellular space. Early endosomes internalize endocytosed proteins to allow for their subsequent return to the cell membrane or transport to late endosomes (or, alternatively, maturation of early into late endosomes) followed by transport to and degradation by the lysosomes. Both are acidic, with an endosome pH of <6 and a lysosome pH of 5. Trophozoites have numerous vacuoles encompassing the periphery of the cell (Fig. 8 and 9), which fulfill at least some criteria of endosomes and lysosomes. These vacuoles are acidic, as shown by their uptake of acidine orange (101, 156). They concentrate exogenous ferritin and lucifer yellow, suggesting their potential role in endocytosis (30, 173). *G. lamblia* virus particles appear to be concentrated into the vacuoles by an endocytotic mechanism (523) (see “Giardia lamblia virus” below). Pulse-chase labeling with horseradish peroxidase showed early and persistent labeling of vacuoles, suggesting that there is no distinction between early and late endocytic vesicles as is found in higher eukaryotes (173). Labeling of a smaller portion of the vacuoles with chemicals that label the ER, such as glucose-6-phosphatase and zinc iodide-osmium tetroxide, and three-dimensional reconstructions (173), as well as EM using anti-BiP antibody (309), have suggested a continuity of these vacuoles with the ER. The vacuoles also contain a variety of hydrolase activities, such as acid phosphatase, proteinase, and RNase, indicating their lysosomal characteristics (98, 189). Thus, the vacuoles appear to function as early and late endosomes and lysosomes and may be functionally associated with the ER (173).

Cysteine proteinase activity occurs in the endosome-lysosome vacuoles (189). The major cysteine proteinase activity of *G. lamblia* trophozoites has been found in proteins with molecular masses of 40 and 105 kDa (123), 35 and 95 kDa (348), and 38-kDa thiol (269). It appears likely that there are two major sizes of cysteine proteinases, one in the 35- to 40-kDa size range and another in the 95- to 105-kDa range. A family of three cysteine protease genes (CP1, CP2, and CP3) have been found to be members of the cathepsin B subgroup of the peptidase family C1 (344). These cysteine proteinases are approximately 30 kDa in size, perhaps corresponding to the smaller of the above-noted proteinase activities. Only CP2 and CP3 are expressed; CP2 is found in the vacuole and is involved in excystation (see “Excystation” below).

**Proteasomes.** While the lysosome degrades endocytosed proteins, the proteasome is a large complex that degrades improperly folded endogenous cytosolic proteins. The 20S proteasome complex is about 700 kDa and contains the protease activity. The complex can be converted into a 26S complex by the addition of proteins that add regulatory activity and bring the size to 2,000 kDa. Archaea have the simplest proteasome, consisting of two alpha and two beta heptamers, while the eukaryotic 20S proteasome contains seven different alpha and seven different beta subunits organized in a manner similar to the Archaea. Since *G. lamblia* has certain similarities to the Archaea but other eukaryotic features, it is of interest that the 20S proteasome of *Giardia* demonstrates a distinctly eukaryotic pattern with 14 subunits (77). Proteins are targeted to the proteasomes by ubiquitin, which appears to be present in *Giardia* as a single-copy gene, in contrast to the multiple gene copies found in other eukaryotes (168).
Cyst Structure

Encystation occurs after organisms have undergone nuclear replication but before cytokinesis; therefore, cysts contain four nuclei. They are approximately 5 by 7 to 10 μm in diameter and are covered by a wall that is 0.3 to 0.5 μm thick and composed of an outer filamentous layer and an inner membranous layer with two membranes. The outer portion of the cyst wall is covered by a web of 7- to 20-nm filaments (79, 80). Four major proteins have been identified in the outer cyst wall, 29, 75, 88, and 102 kDa in size (80). The sugar component of the outer portion is predominantly galactosamine in the form of N-acetylgalactosamine (GalNAc) (148). Earlier claims that the cyst wall is composed of chitin (N-acetylgalactosamine) have been refuted (3).

As the environmentally stable form of the life cycle, cysts have a metabolic rate only 10 to 20% of that found in trophozoites (262). The respiration of both cysts and trophozoites is stimulated by ethanol, while trophozoites are stimulated only by glucose.

Encystation

Promoting and inhibiting factors. As trophozoites replicate and colonize the intestinal surface, some encyst in the jejunum after exposure to biliary secretions. Encystation has been performed in vitro by exposing trophozoites to an environment that mimics that of the jejunum (114, 115, 301). Specific conditions that promoted encystation included a mildly alkalotic pH of 7.8 and conjugated bile salts plus fatty acids (114). Other investigators reported that expression of cyst wall protein, a marker of the encystation process, was detected 90 min after exposure to medium with lipoprotein-deficient serum (196). Encystation was abolished when cholesterol was added, leading to the proposal that encystation results from cholesterol starvation. The relative importance of cholesterol starvation and addition of bile salts and fatty acids for the induction of encystation remains controversial.

Events of encystation. Encystation can be divided into two phases, early and late. The timetables determined by various studies have differed somewhat, probably because of the different approaches used. After the initiation of encystation conditions, the organisms do not all enter the encystation process at the same time. Therefore, morphologic studies that examine the time course of the first encysting organisms tend to give earlier times for the encystation events than studies that depend on pooled organisms. These morphologic studies indicate that the early phase is complete within about 10 h and the late phase is complete by 16 h (83). For this discussion, we will assume that the early and late events from the various studies are comparable.

The early phase consists of the intracellular synthesis and transport of cyst wall components. A Golgi-like stack of membranes become readily visible by EM and are probably the means of transporting cyst-specific molecules to the encystation-specific vesicles (ESV) (282). The Golgi-like membrane labels with NBD-ceramide and protein secretion is inhibited by brefeldin A, as expected for the Golgi complex (195). Expression of BiP, the molecular chaperone found in the ER, is increased during encystation (199). The ESV become visible by light microscopy (96) and are probably involved in the transport of building blocks to the nascent cyst wall. A set of proteins of 21 to 39 kDa are expressed in the early phase (281). A monoclonal antibody to several bands of 26 to 44 kDa initially labels the ESV and subsequently labels the cyst wall, indicating transport of these molecules to the cyst wall (211, 342). The genes encoding two of the proteins incorporated into the cyst wall have been cloned and characterized, a 26-kDa CWP1 (236) and a 39-kDa CWP2 (198). They are related leucine-rich proteins and form a stable complex that is transported to the cyst wall by the ESV (127, 198). Transfection studies performed with fusions of modified GFP and the CWP1 gene have demonstrated the importance of the N terminus for directing the protein to the ESV (127). Localization to the cyst wall was dependent on the leucine-rich middle portion. A 110-bp 5′-flanking region controlled encystation-specific expression, while the 3′ flanking region down-regulated the level of steady RNA produced. Presumably, there is significant overlap between the 21- to 39-kDa and 26- to 44-kDa proteins and CWP1 and CWP2, respectively, but the exact relationship has not been determined.

The late phase of encystation consists of the appearance on the trophozoite plasmalemma of sites for initiation of the assembly of cyst wall filaments followed by the assembly of the filamentous portion of the cyst wall. A set of larger (66-, 78-, 92-, and 103-kDa) proteins are expressed in addition to the lower-molecular-mass proteins (281). Perhaps this includes the 75-, 88-, and 102-kDa proteins found in the cyst wall (80), but these larger proteins have not yet been characterized. One encystation-specific protein, enc6, has not been sequenced in its entirety, but its transcript size of 4.4 kb would be compatible with the 102/103-kDa protein (278). By the time encystation is completed, motility disappears. The outer portion becomes rounded and filamentous, and the organisms are no longer attached to a surface. The ESV disappear, and the internal portion includes two trophozoites with four nuclei that have not yet completed cytokinesis.

Enzymatic pathways. The major cyst wall sugar, N-acetylgalactosamine, is produced by an enzymatic pathway that is induced during encystation. The activities of each of the enzymes in the proposed pathway have been documented (203) (Fig. 11), and two of the enzymes have been cloned and characterized. In the first step, fructose-6-phosphate from the glucose metabolism pathway (Fig. 1) is converted to glucosamine-6-phosphate by glucosamine-6-phosphate-isomerase (312, 334). One form of this enzyme (gpi2, or Gln6PI-A) is expressed constitutively at low levels during the entire life cycle. Transcription of the other form of the of this enzyme (gpi1, or Gln6PI-B) is greatly increased during encystation (164, 312, 334). This up-regulated gene has two transcripts with different initiation sites. The shorter transcript is expressed constitutively at low levels, while expression of the longer transcript with the 147 nucleotide 5′ untranslated region (UTR) is markedly increased during encystation (334). Another enzyme in this pathway, UDP-N-acetylgalactosamine pyrophosphorylase (40, 41), is allosterically regulated (41), with a sixfold increase in activity from activation by glucosamine-6-phosphate.

During encystation, the uptake of oxygen and glucose decreases substantially so that glucose uptake is undetectable within 16 h of exposure to encystation conditions (264). However, aspartate uptake is unchanged, resulting in the suggestion...
that gluconeogenesis may result from the importation of amino acids.

**Excystation**

Promoting and inhibiting factors. In the mammalian host, excystation occurs with exposure to the contents of the proximal small intestine after passage through the acidic environment of the stomach. Excystation was first induced in vitro by exposure of animal- and human-derived cysts to an acidic pH (27). Optimum excystation occurred following exposure to a pH of 1.3 to 2.7. Subsequently, excystation of *G. muris* cysts has been performed at pH 7.5 in phosphate buffer with bicarbonate (101), indicating that an acidic pH was not required for excystation. The inhibition of excystation by 4-4'-diisothiocyanatosulfobenzene-2,2'-disulfonic acid (DIDS) suggested that vacuolar acidification was required. Subsequent studies of cysts produced in vitro found an optimal pH of 4.0. Excystation was facilitated by pancreatic proteases and inhibited by a trypsin inhibitor, suggesting the importance of proteases in excystation. In addition to external proteases, a *G. lamblia* cysteine protease (CP2) of the cathespin B family is required for excystation (344). Cysteine protease activity was found in the endosome-lysosome-like vacuoles. Inhibitors of cysteine proteases prevented excystation but did not affect trophozoite growth or replication. Certain calmodulin antagonists (TFP and W7, but not W5) inhibited excystation, suggesting that calmodulin may be involved in excystation (24). Excystation is inhibited by antibody to cyst wall and by wheat germ agglutinin (WGA), presumably by its reaction with one or more of the glycoproteins found in the cyst wall (218).

Morphology of excystation. After exposure to conditions that promote excystation, the process of excystation is rapid, being completed within 10 min for *G. lamblia* (38) or for *G. muris* (53, 54), and appears to be similar for both species. After the initiation of excystation, one or two pairs of cytoplasmic protrusions (preventral flanges) appear to develop into the ventral flange (130). The peritrophic space and the preventral flange enlarge as the emerging trophozoite separates from the cyst wall. Externally, the emergence of flagella through the opening cyst wall is followed by the entire trophozoite (38). Karyokinesis occurs during encystation, so that the emerging trophozoite has four nuclei but only the eight flagella of a single trophozoite. The oval trophozoite becomes more rounded and then undergoes cytokinesis within 15 to 30 min after the onset of excystation, so that two trophozoites are formed from one cyst.

**GENETICS AND MOLECULAR BIOLOGY**

**Genome Structure**

The genome of *G. lamblia* has the features expected of eukaryotic cells, including linear chromosomes flanked by telomeres that are similar in sequence to those of other eukaryotes (TAGGG) (8, 178). Chromosomal DNA of eukaryotes forms chromatin by associating with four core histones (H2a, H2b, H3, and H4) and a linker histone (H1) to form nucleosomes. *G. lamblia* has all four of the core histones, and these histones are very similar to those of other eukaryotes and have no particular similarity to the histone-like proteins of the Archaea (354). The GC-rich rDNA genes (31, 69), as well as certain protein-coding genes that were sequenced in some of the initial molecular work on *G. lamblia*, suggested a GC-rich genome (3). However, the G+C content of the contigs that have been assembled as part of the genome project have yielded a GC content of 49%, excluding the rDNA genes (210).

PFGE separations of chromosomal DNA have been used to demonstrate that *G. lamblia* trophozoites have five chromosome linkage groups ranging in size from approximately 1.6 to 3.8 Mb (7, 177) (Fig 12). The initial report of PFGE separations of *G. lamblia* chromosomes identified four intensely staining ethidium bromide-stained chromosomal bands for most isolates (including ISR and WB) and five bands for JH (7). Probes specific for each of the five chromosomal bands of the JH isolate were used to demonstrate that the smallest of the four bands of the ISB and WB isolates represented two comigrating chromosomes.

Bands that stained more faintly with ethidium bromide were shown to be size variants of chromosome 1, first by using several chromosome-specific probes (7) and then by more detailed mapping of the various size variants (4, 139). These studies showed that chromosome 1 varied in size from approximately 1.1 to 1.9 Mb and that the size variation occurred in both subtelomeric regions (4, 139). Previous studies had shown that tandem repeats of the rDNA (ribosomal DNA) unit were adjacent to the telomeric repeat region (TAGGG) (8, 179, 180) and were frequently involved in subtelomeric rearrangements. The rDNA repeat is found on different chromosomes among isolates of the same genotype (8). The rDNA repeats are found on one end of chromosome 1 of the ISR isolate, and about 30% of the size difference is accounted for by the difference in rDNA copy number (4). The remainder of the size difference...
have not yet been assembled into the contigs. The reason for
G. lamblia DNA (93). These values of approximately 12 Mb are
2000, the sum of the lengths of the contigs was approximately
lamblia genome data (www.mbl.edu/Giardia). As of December
mide-stained bands on PFGE separations (167, 328), but these
appears to be due to differences in repetitive DNA near the
other telomere. These studies of chromosome 1 have identified
at least three (4) or four (139) size variants of chromosome 1
within cloned isolates, leading to the suggestion that tropho-
zoites are tetraploid at least for chromosome 1. Even though
the central regions of the Giardia chromosomes appear to be
fairly constant (Adam, unpublished), there is substantial vari-
ability at the telomeres.

Other studies have shown greater numbers of ethidium bro-
mide-stained bands on PFGE separations (167, 328), but these
probably represent size variations of the five chromosomes.
One report suggesting that Giardia contained eight chromo-
somes and that a chromosome of approximately 650 to 800 kb
(called chromosome 3) was duplicated (329) most probably
represents a description of the size variants of chromosome 1.

The sum of the sizes of the chromosomes (3.8, 3.0, 2.3, 1.6,
and 1.6 Mb) (7) yields an estimated haploid genome size of
12.3 Mb, similar to the size estimate of 10.6 to 11.9 Mb ob-
tained by densitometric analysis of the NorI fragments of G. lamblia DNA (93). These values of approximately 12 Mb are
compatible with the estimates obtained by analysis of the G. lamblia genome data (www.mbl.edu/Giardia). As of December
2000, the sum of the lengths of the contigs was approximately
11 Mb, not including the rDNA repeats or the vsp genes, which
have not yet been assembled into the contigs. The reason for
the discrepant values of $3.0 \times 10^7$ (251) and $8.0 \times 10^7$ (31)
obtained by C$_r$T analysis is not clear.

In addition to the identification of size variants of chromo-
some 1, other evidence for polyploidy includes the identifica-
tion of multiple alleles of repeat-containing vsp genes. The
vsp46 gene is located on chromosome 4, and at least three
different alleles have been identified (356); these three alleles
vary in the copy number of a 195-bp repeat but have identical
flanking regions and map to a single chromosomal location,
indicating the presence of at least three copies of chromosome
4. Similarly, three and four alleles of the vspC5 gene have been
identified in different lines of the WB isolate (Fig. 13) (355),
indicating a ploidy of at least 3 or 4 for chromosome 5. Thus,
the qualitative data suggest a ploidy of at least 4 (or aneuploi-
dy). In contrast, some of the quantitative data have suggested

FIG. 12. PFGE separation of JH and ISR. A PFGE separation of
JH (genotype A-2) chromosomes demonstrates five distinct bands,
while a PFGE separation of ISR (genotype A-1) chromosomes dem-
onstrates four intensely staining and two more faintly staining bands. A
series of chromosome-specific probes have been used to identify five
distinct linkage groups corresponding to the five bands of the JH
isolate (7, 177). The fainter bands of the ISR isolate represent size
variants of chromosome 1, as demonstrated by a set of chromosome-
specific probes (7) as well as by detailed mapping of chromosome 1
(139). One of the size variants comigrates with chromosome 2; hence
that band is labeled 1b2. Part of the size variation of chromosome 1 is due
to variation in the number of telomeric rDNA repeats (4), while much of
the remainder is due to variation near the other telomere (139).

FIG. 13. Southern blot demonstrating allelic heterozygosity of the
vspC5 gene as well as members of the family of genes related to vspC5.
Reproduced from reference 355 with permission of Elsevier Science.
Genomic DNA from WBA6 (expressing VSPA6), WB1269 (an anti-
genic variant derived from WB6 and expressing VSPA6 [CRP72]),
and WBC5 (an antigenic variant derived from WB1269 and expressing
VSPC5) was digested with BamHI, blotted, and probed. (A) Results of
moderate-stringency washing after hybridization with BS176, a probe
extending from $-93$ to $+83$ of vspC5. (B) High-stringency washing. (C)
Results after hybridization with the vspC5 105-bp repeat. The bands
colored C.1 to C.4 indicate four alleles of the vspC5 gene. Cosmids
were cloned for each of the four alleles and were all identical in the
regions flanking the coding region. These cosmids were mapped to a
single chromosomal location by using PFGE and rare-cutting restric-
tion enzymes, demonstrating that these genes were truly allelic in
nature. The largest of the four bands probably represents the ex-
pressed allele as assessed by the size of the transcript on a Northern
blot (358). Note that the smallest of the four alleles is deleted from the
WBC5 genome. The mechanism for this deletion has not been deter-
mined. The bands labeled S1 to S4 represent four different vsp genes
(vspC5-S1 to vspC5-S4) that are similar to vspC5 in the 5’ region.
Cosmids for each of these genes were cloned; each was different in the
flanking regions, and each mapped to a different chromosomal loca-
tion by PFGE. The sequences of vspC5-S1 and vspC5-S2 were $>96\%$
identical to that of vspC5 in the region of the 176-bp probe. The
disappearance of the vspC5-S3 band after high-stringency washing (B)
indicates that it has less similarity to vspC5. The bands for vspC5-S1
to vspC5-S4 are not seen with the repeat probe (C), consistent with the
lack of the 105-bp repeat in these genes.
a higher ploidy. Semiquantitative data based on densitometric scanning of PFGE separations suggested a total of 30 to 50 chromosomal DNA molecules (6 to 10 copies of each chromosome) (7). Similarly, the total DNA content of 1.34 × 10^8 bp (31, 84) divided by the genome complexity of 1.2 × 10^7 bp yields an estimated ploidy of 10 to 12. A possible clarification of the ploidy of *G. lamblia* has been provided by fluorescence-activated cell sorter analysis of vegetatively growing and stationary-phase trophozoites as well as cysts, using *E. coli* chromosomal DNA as a control (25). The results of the fluorescence-activated cell sorter analysis indicate that trophozoites alternate between 4N and 8N (N is the haploid genome), with most of the stationary trophozoites in the 4N state. These results suggest the possibility that trophozoites each contain two diploid nuclei and that DNA replication occurs shortly after nuclear division so that most of the trophozoites are in the G_2_ phase of DNA replication. If this interpretation is correct, trophozoites have an effective ploidy of 4, but quantitative measurements of stationary trophozoites would yield a ploidy of 8.

For any organism with a ploidy greater than 1, a certain degree of allelic heterozygosity is expected. The degree of heterozygosity is generally low in sexual organisms because of meiotic recombination but can become very high in asexual organisms, because of the general lack of introns. Also, in contrast to other eukaryotes, most *Giardia* spp. are assumed to be ancient sexual organisms, as has been demonstrated for the rotifers (347). Since *Giardia* spp. are assumed to be ancient sexual organisms, one might expect a very high degree of allelic heterozygosity. However, the degree of allelic heterozygosity in *G. lamblia* is actually quite low. Although heterozygosity of repeat copy number for the *vpg* genes has been shown (356, 358), as well as eight nucleotide differences between two different alleles of the *vspA6* gene, the degree of heterozygosity identified in 12 kb of sequence of the triosephosphate isomerase (*tim*) gene from multiple isolates was less than 0.02% (17). The degree of allelic heterozygosity identified in the genome project is also very low (A. McArthur and G. Olson, personal communication). The reason for this low degree of heterozygosity has not been determined. Potential reasons could be unrecognized sex in *Giardia* or intermittent loss of a nucleus.

**Transfection System**

The development of transient- and stable-transfection systems has contributed substantially to our understanding of the genetics of *G. lamblia*. The initial description of a transfection system consisted of transient expression of luciferase flanked at the 5’ end by a short region of the GDH gene and at the 3’ end by the putative polyadenylation signal (361). In this and in subsequent transfection systems, DNA was introduced by electroporation. Subsequently, stable episomal transfection was performed using the paromomycin-resistance *pac* gene (305) or the neomycin resistance *neo* gene (318) as selectable markers. Apparently, the bacterial plasmids used as the framework for the transfection vectors contained sequences that functioned as origins of replication in *G. lamblia*.

An interesting difference between the WB isolate (genotype A) and the GS isolate (genotype B) in the disposition of transfected DNA was identified (305). DNA introduced as a supercoiled plasmid into WB replicated as an episomal vector, while linearized vector was integrated into the WB genome by homologous recombination (305). In contrast, transfection of GS with linearized or circular vector DNA resulted in homologous integration, and stable episomal replication of the vector did not occur. Homologous introduction of DNA into the genome did not result in knockouts with total elimination of the wild-type alleles, presumably because of the polyploid nature of the genome. In addition to the expression of drug resistance markers, modified GFP was expressed (305) and has subsequently been used to study protein transport in trophozoites (76, 127).

High levels of luciferase expression have been obtained by using the *G. lamblia* virus as a vector for introduction of foreign DNA (364, 365). Use of the virus vector resulted in very high levels of luciferase expression for at least 30 days in the absence of drug selection.

**Transcription and Translation**

Transcription in *G. lamblia* is distinctly eukaryotic in nature; nonetheless, it has a number of features that are more characteristic of procaryotes. As in all eukaryotes, the transcript is produced in the nucleus and transported to the cytoplasm for translation. The polyadenylation of transcripts is typical for eukaryotes, but the short 5’ UTRs and general lack of introns are more characteristic of procaryotes (although introns may be relatively infrequent in the unicellular eukaryotes). Also, in contrast to other eukaryotes, most *G. lamblia* transcripts do not appear to be capped at their 5’ ends.

Analysis of the transcript of the *α* and *β*-tubulin genes revealed that the 5’ UTR was only 6 nucleotides in length for these genes (160). Subsequent analyses of other *G. lamblia* transcripts performed by primer extension analysis, S1 nuclease protection, and 5’ rapid amplification of cDNA ends have revealed that other transcripts also have short 5’ UTRs, ranging in size from 0 to 14 nucleotides (5). The one exception to date has been the glucosamine-6-phosphate isomerase B gene, which has two transcripts, a constitutive transcript with a 5’ UTR of 3 to 4 nucleotides and a transcript with a 146-nucleotide 5’ UTR that is induced during encystation (164).

The short 5’ UTRs suggested the possibility that *G. lamblia* promoters might be located near the beginning of the open reading frames. This possibility was supported by the observation of a set of highly conserved motifs in the 5’-flanking regions of a number of cytoskeletal genes (136). The consensus sequence AATTTAAAAA was identified at the site of transcription initiation with the actual initiation site located at the TA (or CA). A second region 20 to 35 nucleotides upstream from the transcription initiation site, CAATAAA(A/T)(T/C)AGA(G/T)TC(C/T)GAA was proposed as a promoter region, and a third consensus sequence, CAATTT, was found at −40 to −70. When the upstream region was tested in a transfection assay, a 44-bp sequence of the GDH gene provided maximal transcriptional activity (305), confirming that the GDH promoter was short and located near the coding region. A deletion and mutational analysis of the GDH promoter region confirmed the importance of the AT-rich region at the initiation site as well as a CAAA region 34 bp upstream (360). Specific binding of the immediate 51-bp upstream region to a 68 kDa nuclear extract protein was demonstrated by band shift analysis and by UV cross-linking (360). A deletion and mutational analysis of the *ran* promoter demonstrated that maximal promoter activity...
was present in the region from -51 to -2 and that regions further upstream did not contribute to promoter function (319). The most important component for promoter activity was the -51 to -20 region; smaller portions of that region gave reduced promoter activity.

In addition to the promoter immediately upstream of the initiation codon, a downstream region of 8 to 13 nucleotides has been suggested as a potential E. coli-like Shine-Dalgarno sequence. The identification of a 15-base sequence in the G. lamblia SS rRNA (GUCCCGGCCTCGAG) led to a search of the database for complementary sequences in the 5’ regions of protein-coding sequences. Most of the 59 sequences evaluated had a complementary sequence within 300 bp of the initiation codon. The consensus (e.g., CCGGGGGGGGCUU) markedly increased the translation efficiency (up to 5,000-fold) without significantly affecting the rate of transcription when included in a transfection assay (363).

Transcripts have been analyzed to determine whether they have 7-methylguanosine or other caps at the 5’ ends. Total polyadenylated RNA was analyzed to determine if the 5’ ends were susceptible to T4 RNA ligase (363). They were resistant to 5’ phosphorylation unless pretreated with calf intestinal alkaline phosphatase to remove the 5’ phosphates. These results suggested that the RNA was phosphorylated but not capped. Treatment with the decapping enzyme tobacco acid pyrophosphatase did not increase phosphorylation, also indicating a lack of 5’ capping. These results suggest that most polyadenylated RNA does not have a 7’ methylguanosine cap. However, since the studies were done using total rather than transcript-specific mRNA, it is still possible that individual transcripts could be capped. In fact, for the differentially processed transcript of the glucosamine-6-phosphate isomerase B gene, the constitutive transcript with a short 5’ UTR is not capped while the transcript with the longer (146-nucleotide) 5’ UTR expressed during encystation does have a cap, as demonstrated by RNA ligation after treatment with tobacco acid pyrophosphatase (164). Whether this was a 7’-methylguanosine or another type of cap was not determined.

Transcripts have relatively short 3’ UTRs with short poly(A) tails typically beginning approximately 10 to 30 nucleotides beyond the stop codon. The sequence, AGT PuAAPy, precedes the poly(A) tail by approximately 10 nucleotides and has been proposed as a polyadenylation signal (3, 273).

Small nuclear RNA molecules are involved in the splicing of nuclear pre-mRNA (snRNA U1, U2, U4, U5, and U6) and pre-rRNA (snRNA U3, U8, U14, snR10, and snR30) to produce the mature RNA (256). Many of these molecules from other eukaryotes have a trimethylguanosine cap at the 5’ end. A number of candidate snRNAs from G. lamblia were immunoprecipitated with anti-trimethylguanosine antiserum; caps were confirmed by the susceptibility of the antibody reactivity to the decapping effect of tobacco acid pyrophosphatase. The exact roles and identities of the candidate snRNAs have not yet been determined. It will be of especial interest to determine whether splicing of mRNA occurs since introns have not been identified in the G. lamblia genes characterized to date.

The initial characterization of the rRNA and rDNA genes revealed that the rRNA molecules were much smaller than the typical eukaryotic I8S (SS rRNA) or 28S or large-subunit rRNA (LS rRNA) molecules and in fact were even slightly smaller than the 16S and 23S E. coli rRNA molecules (31, 69). The rDNA repeats were correspondingly small, but had an organization and sequence consistent with that of eukaryotes (see “Giardia and other diplomonads as early-branching eukaryotes” above). In addition, other components of the translation apparatus, including elongation factor 1α, elongation factor 2α, eRF1, eRF3, RPβ1, and RNA polymerase III (Table 3), are distinctly eukaryotic.

**Giardia lamblia Virus**

A double-stranded RNA virus was identified in G. lamblia trophozoites following the observation of a 6.2-kb double-stranded RNA contaminating nucleic acid preparations (64, 337). The virus is nonenveloped and icosahedral, with a diameter of 33 nm, and it infects the nuclei with about 200 copies per nucleus (337). G. lamblia virus (GLV) has a 100-kDa major capsid protein which depends on a cysteine protease for cleavage into the mature protein (362). A second open reading frame encodes a 190-kDa RNA-dependent RNA polymerase (338, 349). The majority of axenic isolates either harbor or are susceptible to GLV infection (including isolates from genotypes A and B), while a minority of isolates are highly resistant to infection (225). GLV infection occurs by endocytosis (323), and susceptibility to infection depends on a specific receptor found of the cell membrane surface (302). A second virus has also been identified; it also has a 6.2-kb double-stranded RNA genome but encodes a slightly smaller capsid protein (95 kDa), which differs significantly from the 100-kDa capsid protein of GLV (322).

**SURFACE ANTIGENS AND ANTIGENIC VARIATION**

**Occurrence and Biological Significance of Antigenic Variation**

G. lamblia trophozoites undergo antigenic variation of a family of immunodominant cysteine-rich surface antigens in vitro and in vivo (6, 12, 243). The initial studies of the surface antigens of G. lamblia showed differences among strains by crossed immunoelectrophoresis and enzyme-linked immunosorbent assay (306) and marked differences in the molecular masses of “excretory-secretory products” from different surface-iodinated G. lamblia isolates (247, 250). These surface antigens varied in number and in size from approximately 50 to 200 kDa in a study of 19 isolates (250). A monoclonal antibody (MAb 6E7) for a 170-kDa surface antigen (initially called CRP170 but now called VSPA6) from the WB isolate was cytotoxic for WB trophozoites but not for isolates expressing other surface antigens (242). WB trophozoites were doubly cloned by limiting dilution and incubated with MAb 6E7; some organisms survived and were totally resistant to the cytotoxic effect of MAb 6E7 (243). One of these cloned lines of organisms (WB1269) had a 68-kDa surface-labeled antigen (initially called CRP68 but now called VSP1269), while another (WB1267) had a 64-kDa surface antigen (VSP1267). A monoclonal antibody (MAb 5C1) reactive with the surface antigen of WB1267 was cytotoxic for WB1267 trophozoites. These trophozoites were incubated with MAb 5C1, and organisms surviving the cytotoxic antibody reacted with neither of the initial MAbs and expressed surface antigens in the 60- to 100-kDa range.
Antigenic variation has subsequently been confirmed in animal models (12, 118, 119) and in infected human volunteers (248). Cloned WB trophozoites expressing the VSPA6 were inoculated into gerbils, and trophozoites collected from their intestines 7 days later demonstrated populations of trophozoites expressing multiple VSP types ranging in size from 50 to 170 kDa (12). Trophozoites collected 28 days after infection were similar to those collected at 7 days. The lack of apparent change from 7 to 28 days argued against a role for acquired immunity in selecting antigenic variants. However, the results were somewhat different in a mouse model of G. lamblia infection. G. lamblia trophozoites in athymic nude mice and in heterozygous nu/nu mice changed VSP type coincident with the development of a humoral anti-VSP antibody response (119). In contrast, scid mice did not develop an antibody response and trophozoites did not undergo antigenic variation. Thus, in the mouse model, the antibody response to the VSP may have been the selective force for antigenic variants.

When human volunteers were inoculated with cloned GS trophozoites by duodenal intubation, four of four became infected with trophozoites expressing VSPH7 (72 kDa). Only 1 of 13 became infected with organisms expressing VSPB6 (200 kDa), suggesting that the ability to survive in the human intestine is greater with one VSP than with another. Whether this is due to selection against one VSP or positive selection for the other has not been determined. The subsequent pattern of change during 22 days after infection showed gradual disappearance of VSPH7 followed by the appearance of other VSPs with different immunoreactivity. The loss of VSPH7 was accompanied by the development of serum antibodies to VSPH7, suggesting the possibility that the immune response led to the antigenic variation.

An alternative or additional possible biological reason for antigenic variation might be adaptation to different intestinal environments. Some evidence for this possibility is provided by the difference in protease susceptibility for different VSP antigen types (252). Trypsin and chymotrypsin were toxic to WB trophozoites expressing VSPA6 (reactive with MAb 6E7). Some organisms survived after exposure to these proteases, and when they were subsequently grown in the presence of trypsin and chymotrypsin, they were no longer susceptible to the cytotoxic effects and expressed alternative VSPs that were not reactive with MAb 6E7.

Antigenic variation has also been documented during the cycle of encystation and excystation (217, 320). In vitro encystation followed by excystation of trophozoites expressing the VSP TSA417 resulted in the loss of TSA417 from the plasma membrane followed by its appearance in the lysosome-like peripheral vacuoles (211, 320). During excystation, the TSA417 transcript was replaced by a variety of other vsp transcripts. This occurred too rapidly to be the result of selection of antigenic variants surviving excystation and may have represented an induced antigenic shift. This antigenic variation during excystation could promote survival by evasion of an intestinal immune response or by adaptation to other important intestinal factors.

Thus, the two major hypotheses regarding the purpose of antigenic variation are (i) evasion of the host immune defense and (ii) enabling of the organisms to survive in different intestinal environments. It should be emphasized that these hypotheses are not mutually exclusive. A better understanding of the biological reasons for antigenic variation would be facilitated if the role of the VSPs was known. It would also be informative to know if the free-living diplomonads such as Hexamita have proteins homologous to the VSPs and if they undergo antigenic variation.

**Structure and Biochemistry of the VSPs**

The initial biochemical studies of the VSPs (then called ES products) indicated that they were protease susceptible. They did not bind to a series of lectins, suggesting that they were not glycosylated (247). A portion of the gene for one of these surface antigens was subsequently cloned from an expression library using MAb 6E7. Sequence analysis revealed a cysteine-rich (12%) protein with frequent CXXC motifs (6). Subsequent sequences of a number of vsp genes have shown a number of characteristics common to all vsp genes. They are all cysteine rich (12%), and most cysteines are present in CXXC motifs. When trophozoites are metabolically labeled with radiolabeled cysteine, most of the label is incorporated into the VSPs (6, 11). It has been suggested that these cysteines are present in the form of disulfide bonds, since free thiol groups have not been detected in purified VSPs (14, 266). However, the presence of thiol groups on the trophozoite surfaces suggests that the VSP cysteines may not all be disulfide bonded (116). Whether the level of disulfide bonding changes in vivo with different oxygenation levels of the surrounding environment has not been determined.

After synthesis, the VSPs are transported through the ER (211) to the membrane, where they diffusely coat the membrane (112, 275). Smaller amounts of VSP can also be detected in the lysosome-like peripheral vacuoles, suggesting the possibility that these vacuoles are involved in recycling of the VSPs (211). The VSPs have an approximately 14- to 17-amino-acid signal peptide that presumably represents a signal peptide for its transport through the ER. The signal peptide is cleaved, 14 amino acids from VSPH7 of the GS isolate (201) and 17 amino acids from TSA417 of the WB isolate (14).

The 38 C-terminal amino acids of the VSPs are >90% conserved at the amino acid and nucleotide levels. Each VSP reported to date ends with the amino acid motif CRGKA in both genotype A and B organisms. The conserved C terminus may represent a membrane-anchoring domain. Evidence for this proposal has been reported in work showing a difference in the C termini of membrane-bound and secreted VSP (267). A 3,500-kDa fragment is cleaved from the C terminus of the membrane-associated VSP prior to secretion. The suggested cleavage site is between the K and S of the highly conserved NKSGLS motif (267), which comprises the amino acids 33 to 38 from the C terminus (235). This NKSGLS motif is generally conserved in the genotype A organisms (88) but is not found in the vspH7 gene from GS, a genotype B organism (253).
### TABLE 5. vsp gene families of genotype A-1 G. lamblia

<table>
<thead>
<tr>
<th>Related vsp gene</th>
<th>Comparison to prototype</th>
<th>Chromosomal location</th>
<th>Comments</th>
<th>Reference(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>vspA6</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>vspA6-S1</td>
<td>&gt;90% sequence identity throughout the gene</td>
<td>4</td>
<td>About 21 copies of a 195-bp repeat</td>
<td>9</td>
</tr>
<tr>
<td>vspA6-S2</td>
<td>75–91% identity</td>
<td>4</td>
<td>1+ copies of same 195-bp repeat</td>
<td>237, 357</td>
</tr>
<tr>
<td></td>
<td></td>
<td>5</td>
<td>Four copies of a 201-bp repeat identical to the 195-bp repeat</td>
<td>357</td>
</tr>
<tr>
<td>vsp1267</td>
<td></td>
<td>5</td>
<td>Two identical convergent copies 3 kb apart</td>
<td>235</td>
</tr>
<tr>
<td>vspC5-S1</td>
<td>&gt;96% identity at the 5′ region followed by abrupt divergence</td>
<td>5</td>
<td>26 copies of a 105-bp repeat</td>
<td>355, 358</td>
</tr>
<tr>
<td>vspC5-S2</td>
<td>&gt;96% identity at the 5′ region followed by abrupt divergence</td>
<td>4</td>
<td>Partial copy of one repeat</td>
<td>355</td>
</tr>
<tr>
<td>vspC5-S3</td>
<td>Cross-hybridization to 5′ region</td>
<td>5</td>
<td>Partial copy of one repeat</td>
<td>355</td>
</tr>
<tr>
<td>vspC5-S4</td>
<td>Cross-hybridization to 5′ region</td>
<td>5</td>
<td></td>
<td>355</td>
</tr>
<tr>
<td>TSA417 (vsp417-1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>vsp417-2 (Tsp 11)</td>
<td>Regions of 55–90% identity over much of the gene</td>
<td>96% identity at the 5′ region</td>
<td></td>
<td></td>
</tr>
<tr>
<td>vsp417-4</td>
<td>57–58% identity to vsp417-1 and vsp417-2</td>
<td>5</td>
<td></td>
<td>91</td>
</tr>
<tr>
<td>CRP136</td>
<td></td>
<td>1/2</td>
<td>May be telomeric in location; 23 copies of a 120-bp repeat</td>
<td>51, 330</td>
</tr>
<tr>
<td>CRP65</td>
<td>86–96% identity in nonrepeat region</td>
<td>1/2</td>
<td>Four copies of a 228-bp repeat</td>
<td>52</td>
</tr>
</tbody>
</table>

* a The prototype vsp genes are in bold and the related vsp genes are indented.

b Reported as being on chromosome 3, but probably on chromosome 1 or 2; see the section on genome structure.

The vsp 3′ UTRs have the putative polyadenylation signal common to all G. lamblia strains, AGTPuAAPy (3, 273), immediately preceded by the sequence ACTPyAGPuT, which sometimes begins within the termination codon (88, 320) and (Y. M. Yang and R. Adam, unpublished observations).

On the basis of sequences in vsp genes similar to those shown in zinc binding by nuclear proteins, it has been proposed that the VSP genes may encode zinc binding proteins. The ability to bind zinc was shown for VSPH7 and other VSPs but was also found in regions of the VSP that did not contain the “zinc binding” motifs (254). In addition to zinc, VSPH7 was able to bind other metals including iron. Another laboratory, using VSP4A1 (CRISP90) from a genotype A isolate, demonstrated the ability of reduced but not oxidized VSP to bind zinc. However, the degree of binding was substantially lower than an equimolar zinc:protein amount and was thought to be biologically insignificant (266).

All of the VPSs reported have potential N-linked glycosylation sites (201); biochemical analysis of VSP4A1 demonstrated O-linked GlcNAc rather than the more commonly found N-linked glycosylation (268). VSP4A1 was also palmitoylated in the membrane-anchored portion of the C-terminal region (267, 268). Subsequent work has extended the finding of glycosylation and palmitoylation to other VSPs, including VSPA6 and VSPH7 (131).

**Genomic Organization of the vsp Genes and Potential Mechanisms of Antigenic Variation**

In addition to their marked 3′ similarity, the vsp genes demonstrate different degrees of identity in other regions. In some cases, there is a high degree of similarity throughout the entire gene. For example, there are two identical convergent copies of the vsp1267 gene approximately 3 kb apart (235) and there is greater than 90% identity in the nonrepeat regions of vspA6 and vspA6-S1 (357) as well as CRP136 and CRP65 (52). These examples suggest that the vsp gene repertoire has been expanded by gene duplication followed by divergence. The sequences for one of these vsp genes (vspA6-S1 or vspG3M from two geographically separated isolates (WB from Afghanistan and G3M from Peru) were identical over the regions of comparison (237, 357), suggesting that this divergence is not an extremely rapid process. A number of vsp genes related to TSA417 (vsp417) have been identified from genotype A-1 and A-2 isolates. Comparisons have shown that the differences between genotypes are smaller than the differences between members of this vsp gene family, suggesting that the divergence among the various members of the vsp417 gene family occurred before the divergence between genotypes A-1 and A-2. The vsp genes from the genotype B isolate GS also appear to exist in families (246, 253) but demonstrate similarity to the vsp genes of genotype A only in the conserved 38-amino-acid C terminus and in the CXXC motif.

There are also examples of near identity followed by abrupt divergence, suggesting that in some cases, the vsp repertoire has been expanded by recombination among the vsp genes. For example, vspC5 demonstrates near identity to vspC5-S2 from approximately −150 to +94 and to vspC5-S1 from −150 to +130 (355). A more complete understanding of the vsp repertoire and gene organization will be available when the sequences obtained as part of the Giardia genome project have been assembled.

The vsp genes are present on most if not all the chromosomes and are reasonably dispersed throughout the genome. There does appear to be some clustering of genes as demonstrated by the linkage of some vsp genes (e.g., vspJ267) (Table 5) and by the proximity of vsp genes on some cosmid clones (Yang and Adam, unpublished).
Antigenic variation in African trypanosomes is frequently associated with genome rearrangements, such as duplicative transposition to a telomeric location (32, 290). Therefore, it is of interest to know whether the \textit{vsp} genes are telomeric and whether there are genomic rearrangements associated with antigenic variation. Current data indicate that most of the \textit{vsp} genes are not telomere associated (355–357). The \textit{vspA6} gene is not found in the 150- and 240-kb \textit{Nosi} telomeric fragments of chromosome 4 (356), and the \textit{vsp} conserved region does not hybridize to these regions (Adam, unpublished). Genome rearrangements are not directly associated with antigenic variation of \textit{vspA6} (356) or \textit{vspC5} (358).

\textit{Giardia} trophozoites have a ploidy of at least 4, so there are at least four alleles of each gene. Since the usual definitions of allele cannot be used in a presumably asexual organism, genes have been considered to be allelic when they map to the same genomic location. For most genes, the alleles are identical or nearly identical to each other, sometimes containing heterozygosities at one or two sites (17). However, in the case of the repeat-containing \textit{vsp} genes, such as \textit{vspA6} and \textit{vspC5}, the different alleles can be distinguished by different repeat copy numbers (356, 358) (Fig. 13). Three distinct alleles of the \textit{vspA6} gene have been documented which differ in the copy number (i.e., 8, 9, or 23) of the 195-bp repeat (356). In addition to the differences in repeat copy number, the allele with the greatest repeat copy number contains 8 nucleotide substitutions within the coding region. Some of these substitutions change the amino acid composition, but none change the reading frame. The flanking regions of the three alleles are identical.

Despite the open reading frames and identical flanking regions, a steady-state transcript is expressed only from the allele with the largest repeat copy number as documented by the correct size of band on a Northern blot in addition to direct DNA sequencing, which identified the nucleotide substitutions specific to the allele with the largest repeat copy number. Thus, a transcript is detectable from only one of the alleles despite the absence of apparent sequence alterations that might explain the difference in expression.

The \textit{vspC5} gene is another repeat-containing gene, containing 17 to 26 copies of a 105-bp repeat. In this case, the WBA6 genome contains four distinct alleles of the \textit{vspC5} gene (17, 20, 21, or 26 repeats) while the WBC5 genome (derived from WBA6) contains only three distinct alleles (26, 21, or 20 repeats). There are no sequence differences among the different alleles, either in the coding or in the flanking regions. However, Northern analysis has identified only a single band that is most consistent with a transcript from the allele containing 26 repeats. (Because of the short 5' and 3' UTRs, the transcript size closely matches that of the coding region.)

Initial studies suggested significant recombination involving the \textit{vspA6} (CRP170) gene (6) and deletion of an expression-linked copy of \textit{vspA6} in an isolate no longer expressing \textit{vspA6} (9). However, subsequent data have shown that these recombination events can be explained by changes in repeat copy number for the repeat-containing \textit{vsp} genes (355–358) Other than changes in repeat copy number, chromosome-mapping studies of the \textit{vspA6} and \textit{vspC5} genes have revealed no changes in genomic location and organization in isolates where a particular \textit{vsp} gene is expressed or not expressed (356, 358). This indicates that antigenic variation does not occur by moving the \textit{vsp} genes into expression sites.

The allele-specific expression of the \textit{vsp} genes and the absence of sequence alteration or DNA rearrangements associated with antigenic variation suggest the possibility of an epigenetic form of control for \textit{vsp} genes expression. There is no methylated or “J” DNA in \textit{Giardia} (335), arguing against altered DNA as a means of controlling \textit{vsp} gene expression. Whether changes in chromosome structure or histone acetylation status are involved in the control of \textit{vsp} gene expression has not been studied.

**Other Surface Antigens**

Taglin is a surface antigen that migrates as a 28- and 30-kDa doublet on Western blots (343) and has protease-induced lectin activity (184). After trypsin treatment it binds to mannose-6-phosphate. This antigen is nonvariable, and expression is constant throughout encystation and excystation (320).

\textit{G. lamblia} trophozoites also have a 49-kDa antigen (GP49) that is attached to the membrane surface by a glycosylphosphatidylinositol (GPI) anchor (61). During synthesis of the GPI anchor, exogenous phosphatidylinositol is incorporated, but myo-inositol is converted to phosphatidylinositol prior to incorporation (317). The GPI anchor is not subject to cleavage by phospholipase C, in contrast to a variety of other GPI-anchored surface molecules, including the trypanosome variant surface glycoprotein (43). GP49 is constant among different isolates and does not demonstrate variation within single isolates (61).

**CONCLUSIONS**

Significant progress has been made in the understanding of the phylogenetic relationship of \textit{G. lamblia} to other diplomonads and other eukaryotes. Resolution of the controversies regarding the phylogenetic position of the diplomonads and the relationships among the various \textit{Giardia} species and \textit{G. lamblia} genotypes will depend on further sequence comparisons and biological studies. It is not clear why \textit{G. lamblia} maintains a polyploid genome with two apparently identical nuclei. An understanding of how nuclear symmetry is maintained and the selective advantage conferred by the two nuclei should facilitate our understanding not only of \textit{G. lamblia} but also of general biological principles. The previous decade has contributed substantially to our understanding of protein processing and transport in vegetative trophozoites and encysting organisms. Further investigation promises to aid our understanding not only of the similarities to other eukaryotes but also of the key differences that may provide targets for therapeutic intervention.

The VSPs that coat the surfaces of the trophozoites apparently play an apparent role in the biology of the organism, but this role is not yet understood. A further understanding of the structure and biochemistry of the VSPs may yield insight into the biological role of the VSPs and the reason(s) for antigenic variation. An equally important question is the mechanism by which \textit{G. lamblia} controls the expression of the \textit{vsp} genes. Expression of the \textit{vsp} genes is changed during the process of encystation and excystation as well as during vegetative growth,
but no DNA alterations have been associated with antigenic variation. Therefore, epigenetic mechanisms are the most promising candidates at this time.

The next decade will see the application of the current genomic advances to addressing these interesting and important questions. The genome project will allow the identification of the enzymatic components of the various biochemical pathways and will suggest new avenues of research as well as novel therapeutic strategies.

ACKNOWLEDGMENTS

I thank numerous colleagues for helpful suggestions, including Lidya Sánchez, C. W. Birky, and Ted Nash. I express my gratitude to Lynda Schurig for assistance with the biochemical pathway figures. Thanks to Claire Payne for assistance in obtaining the electron micrographs.

REFERENCES

1. Reference deleted.


190. Linstedt, D., and M. A. Cranshaw. 1983. The pathway of arginine catabo-


