Available online at www.sciencedirect.com
SCIENCE DIRECT®
International Journal for Parasitology 33 (2003) 235–255
www.parasitology-online.com

Invited review

Molecular biology of the amitochondriate parasites, *Giardia intestinalis*, *Entamoeba histolytica* and *Trichomonas vaginalis*

Stepanka Vanacova[a,b], David R. Liston[a], Jan Tachezy[b], Patricia J. Johnson[a,*]

[a]*Department of Microbiology, Immunology and Molecular Genetics, David Geffen School of Medicine, University of California, Los Angeles, CA 90095, USA*
[b]*Department of Parasitology, Charles University, Prague, Czech Republic*

Received 23 October 2002; received in revised form 3 December 2002; accepted 3 December 2002

Abstract

The amitochondriates are an assembly of unicellular protists that lack mitochondria, and often other typical eukaryotic organelles, such as peroxisomes. Relatively little research has been conducted on amitochondriates, even though marine waters are likely to be teeming with such organisms, representing yet unknown protist lineages. The three amitochondriates that have been studied in any detail are the three mucosal human parasites: *Giardia*, *Trichomonas* and *Entamoeba*. These pathogens have worldwide distribution and are the most commonly encountered parasites in North America and Europe. Despite significant differences in their lifecycles and pathogenic properties, *Giardia*, *Trichomonas* and *Entamoeba* are customarily grouped together based on their being microaerophilic, their anaerobic carbohydrate metabolism, their lack of mitochondria, and their placement on deep-branching lineages in eukaryotic phylogenetic trees. During the last decade, the development of functional tools has allowed molecular analyses of gene expression to be initiated on these divergent eukaryotes. The resulting data indicate significant differences between the organisation of genetic information and mechanisms of gene regulation in amitochondriates and other organisms. Promoter architecture, as well as the regulatory transcription factors required to mediate promoter activity, lacks the conservation observed for the transcriptional apparatuses of metazoa and even appear to be unique within the amitochondriates. Our knowledge of the molecular biology and gene expression in amitochondriates is still in its infancy and a discussion of the current status of research in this area is presented.

Keywords: *Giardia*; *Entamoeba*; *Trichomonas*; Gene expression; Parasite; Genome

1. Introduction

Studies on gene expression in eukaryotes have concentrated mainly on animals, plants and fungi, while a rather sparse amount of information is available for parasitic protozoa. Several intriguing discoveries, such as RNA editing and *trans*-splicing, mechanisms of gene regulation observed in parasitic kinetoplastids (African trypanosomes and *Leishmania* (Boothroyd and Cross, 1982; Benne et al., 1986; Sutton and Boothroyd, 1986; Miller and Wirth, 1988; Tessier et al., 1991; Hajduk et al., 1992; Benne, 1994) have stimulated interest in gene expression in protozoa. These studies have focussed on kinetoplastids (eg. trypanosomes and *Leishmania*), apicomplexa (*Plasmodium* and *Toxoplasma)* and ciliates (*Tetrahymena* and *Paramecium*), while considerably less attention has been paid to amitochondriate anaerobic protozoa.

The amitochondriate organisms that have received the greatest attention are the human-infective parasites, *Entamoeba*, *Giardia* and *Trichomonas*. Renewed interest in these parasites was provoked by phylogenetic analyses of small subunit (SSU) RNA genes (Sogin et al., 1996), and translation elongation factors EF-1a and EF-2 genes (Shirakura et al., 1994; Hashimoto et al., 1994; Hashimoto and Hasegawa, 1996) that place them near the base of the eukaryotic tree. Although the exact phylogenetic positions of the amitochondriates are currently uncertain, considerable evidence supports early divergence of *Trichomonas* and *Giardia*, whereas *Entamoeba* most likely branches higher in the tree. The deep branching of these organisms and their lack of mitochondria were originally interpreted to represent a relic of a premitochondrial phase of eukaryote evolution. Although mitochondria are absent in trichomo-

* Corresponding author. Tel.: +1-310-825-4870; fax: +1-310-206-5231.
E-mail address: johnsonp@ucla.edu (P.J. Johnson).

0020-7519/03/$30.00
doi:10.1016/S0020-7519(02)00267-9

nads, they possess another type of ATP-generating organelle, the hydrogenosome, with considerable biochemical and biogenetic similarity with mitochondria and the two organelles appear to share a common ancestry (reviewed in Dyall and Johnson, 2000; Lang et al., 1999). Whilst the fermentative anaerobic biochemistry of these parasites has been intensively studied, only rudimentary information exists about their molecular biology. The aim of the following text is to summarise the current data on the molecular mechanisms of the gene expression in these three amitochondriate eukaryotes.

2. Biology of amitochondriates

The three amitochondriates, *Entamoeba histolytica*, *Giardia intestinalis* (commonly referred to as *Giardia lamblia*) and *Trichomonas vaginalis* are highly prevalent human-infective parasites with a worldwide distribution. *Entamoeba* and *Giardia* cause severe diarrhoea and *Trichomonas* causes vaginitis, and occasionally urethritis in males. *Entamoeba histolytica* is estimated to cause severe disease in 48 million people, killing 70,000 each year (WHO/PAHO/UNESCO, 1997). Although *G. intestinalis* rarely causes lethal disease, it is the most common intestinal parasite of humans in developed countries. In Asia, Africa and Latin America, about 200 million people have symptomatic giardiasis, with about 500,000 new cases reported each year (WHO, 1996). *Trichomonas vaginalis* is likewise a very prevalent parasite. Trichomoniasis is the most common non-viral sexually transmitted infection worldwide with an estimated 170 million new cases occurring each year (WHO, 2001).

All three organisms are lumenal, extracellular parasites. *Entamoeba histolytica* and *G. intestinalis* are enteric parasites of the small and large intestines, respectively and *T. vaginalis* resides in the urogenital tract. The life cycle of *Entamoeba* and *Giardia* consists of two stages – the dormant tetranucleated cyst and the vegetative trophozoite. The cyst is excreted in the environment and is the infectious stage. The motile trophozoite, which is released by the cyst upon transmission to a new host, divides and colonises the intestine. Unlike *Giardia* and *Entamoeba*, which are primarily transmitted by oral ingestion of cysts, *T. vaginalis* is sexually transmitted by direct human contact. It exists entirely as a flagellated trophozoite and has no cyst stage. As a result the parasite cannot survive outside the host; although in response to some stress conditions, pseudocyst formation has been observed (Abonyi, 1995).

The trophozoites of all three organisms divide asexually by binary cell fission, however, the molecular mechanisms underlying cell division have not been investigated. Characterisation of mitosis in *T. vaginalis* has demonstrated the presence of all five mitotic phases: prophase, metaphase, anaphase, telophase, and cytokinesis (Gómez-Conde et al., 2000). The chromatin of *T. vaginalis* undergoes different degrees of condensation, from heterochromatic chromosomes in metaphase to decondensed chromatin at anaphase. Interestingly, these cells form an extranuclear spindle during mitotic division with the nuclear membrane remaining intact (Brugerolle, 1975). Studies on the *Giardia* cell cycle have been complicated as common drugs (e.g. colchicin) that typically arrest eukaryotic cell cycle do not function in this parasite (Hoyne et al., 1989). The only evidence for distinct mitotic stages so far is based on the comparative analysis of DNA content of *Giardia* and *Escherichia coli* (Bernander et al., 2001). *Entamoeba* chromosomes do not appear to condense at any stage during mitosis (Dvorak et al., 1995; Willhoeft and Tannich, 2000), and little is known about the organisation of DNA during the cell cycle of this organism.

3. DNA content and structure

In *Giardia*, chromatin analysis by transmission electron microscopy (TEM) and digestion with micrococcal nuclease reveals nucleosomal conformations typically found in eukaryotes (Triana et al., 2001). Genes encoding histones H1, H2a, H3 and H4 have all been identified in *Giardia* (Wu et al., 2000; Triana et al., 2001) and *Entamoeba* (Födinger et al., 1993; Sanchez et al., 1994; Binder et al., 1995; Scharfetter et al., 1997). In contrast, limited information on histone genes in *T. vaginalis* is available (Marinets et al., 1996). H3 and H4 histone genes have been characterised in this organism and the amino acid sequence of these proteins display significant divergence from other eukaryotic homologues (Marinets et al., 1996). Interestingly, *Entamoeba* histone genes are more divergent than those identified in *Giardia* and *Trichomonas* (Binder et al., 1995; Marinets et al., 1996), although the latter two organisms branch deeper in eukaryotic trees.

Genome sequencing projects are well underway for both *G. intestinalis* (http://www.mbl.edu/Giardia/) and *E. histolytica* (http://www.tigr.org/tdb/e2kl/eha1) and one is in the planning stage for *T.vaginalis*. With the genome sequence almost entirely available for *Giardia* strain WB, the haploid genome size has been assessed at approximately 12 Mbp. The karyotype is also better studied for *Giardia* than the other two parasites, however its ploidy and chromosome number, which appears to vary between strains (Upcroft and Upcroft, 1999), remains controversial. *Giardia* trophozoites are unusual as they contain two morphologically indistinguishable nuclei of almost identical size. The two nuclei appear to contain the same amount of DNA, both are transcriptionally active (Adam, 2000), and encode equal copies of rRNA genes (Kabnick and Peattie, 1990). Unlike those of many other eukaryotes, the nuclei of *Giardia* lack discernible nucleoli. *Giardia* trophozoites of strains WB and GS appear to be tetraploid with their DNA arranged into five chromosomes (Adam et al., 1988; Le Blancq and Adam, 1998) that are flanked by telomeres composed of tandem

arrays of [TAGGG] repeats, similar to those of other eukaryotes (Adam et al., 1991; Le Blancq et al., 1991b). Reassociation kinetics studies and quantitative analysis of genes accessible in the gene databases and *Giardia* genome project estimated the G + C content of the *Giardia* genome to be 42–49% (Boothroyd et al., 1987; Smith et al., 1998; McArthur et al., 2000).

A comparative analysis of DNA content in different life cycle stages, using *E. coli* DNA as a control, suggests that during the trophozoite vegetative growth, both nuclei switch from a diploid to tetraploid genome content (Bernander et al., 2001), with cells in the G(2) phase having a ploidy of 8N (two nuclei, each with a 4N ploidy). However, other studies have estimated the ploidy of certain strains to be 10–12N (Adam, 2001) and considerable variation has been observed between strains (Upcroft et al., 1989; Korman et al., 1992; Upcroft et al., 1996). Further studies are required to determine the ploidy at different stages and whether the ploidy of all chromosomes is equal or whether aneuploidy is a feature of this parasite.

Cytological techniques have suggested the presence of four to eight major chromosomes in each *Giardia* nucleus (Kabnick and Peattie, 1990; Erlandsen and Rasch, 1994). In parallel, pulsed field gel electrophoretic analyses revealed a presence of four to five major bands ranging in size from 1.6–4 Mbp (Adam et al., 1988; Le Blancq and Adam, 1998) with evidence for minor chromosomes less than 1 Mbp (Adam et al., 1988). Hybridisation analyses revealed the presence of five chromosomal linkage groups (Adam et al., 1988; Korman et al., 1992; Le Blancq and Adam, 1998). However, a single definitive picture of *Giardia* chromosome content has not emerged, as different karyotypes are seen in different clinical isolates, as mentioned above. The heterogeneity has been reported to be mainly the result of changes in hypervariable subtelomeric regions (Le Blancq and Adam, 1998). The variable components of the subtelomeric regions are composed of short arrays of repeats such as rDNA operons that can move among homologous as well as non-homologous chromosomes (Le Blancq et al., 1991a; Hou et al., 1995; Le Blancq and Adam, 1998). Upcroft and colleagues have suggested that this high variability in *Giardia* genome ploidy may be a dynamic response to different stress conditions such as environmental or drug pressure (Upcroft et al., 1992, 1996; Upcroft and Upcroft, 1999).

The nucleus *of E. histolytica* is large and appears to contain little condensed chromatin. The interphase nuclei are round in shape with chromosomal DNA concentrated in the centre of the nucleus and episomal rDNA on the periphery (Willhoeft and Tannich, 2000). The genome has a low G + C content (about 22.4%) (Tannich and Horstmann, 1992); however the G + C content of coding regions is approximately 33%. Early studies reported a ploidy of up to 14N (Byers, 1986). As *E. histolytica* chromosomes do not condense at any stage of the cell cycle, standard cytogenetic tools cannot be used to observe individual chromosomes. The situation has been further complicated by the observation that *E. histolytica* cultures are made up of a heterogeneous population of cells that contain varying amounts of DNA (Gangopadhyay et al., 1997b). Moreover, there appears to be an extensive chromosome length variability among different *Entamoeba* isolates caused by substantial size variations of homologous chromosomes (Petter et al., 1993; Willhoeft and Tannich, 1999). Recently, 68 independent cDNAs were used to probe pulse field gel electrophoresis (PFGE) separated genomic DNA from different *Entamoeba* isolates and revealed the presence of 14 independent linkage groups (Willhoeft and Tannich, 1999, 2000). This analysis suggests there are only four separate chromosomes. Based on these studies, the size of the haploid *Entamoeba* genome was estimated to be up to 20 Mbp and functional ploidy of a single cell would be at least 4N. The ongoing *E. histolytica* genome sequencing project will eventually allow accurate determination of chromosome number and ploidy.

The trichomonad genome has been estimated by reassociation kinetics CoT analyses to contain more than 50% of repetitive sequences and approximately 2.5×10^7 bp of unique DNA. The overall G + C content was also estimated to be about 36% (Wang and Wang, 1985b). *Trichomonas vaginalis* has six monocentric chromosomes (Drmota and Král, 1997; Yuh et al., 1997, 1998) that can be visualised by cytological chromosome spreads. These chromosomes appear to be haploid except in a small percentage of the cells (0.1%) where diploid metaphases and four nuclei were observed suggesting that *T. vaginalis* may undergo meiosis (Drmota and Král, 1997). The ability to disrupt single-copy genes by homologous recombination also indicates that *T. vaginalis* is haploid (our unpublished results). In contrast to *Giardia* and *Entamoeba*, no chromosomal size heterogeneity was reported when 15 different *T. vaginalis* isolates were compared (Lehker and Alderete, 1999). However, PFGE analyses, using the same conditions on additional strains, reveal different results from that reported and indicate considerable heterogeneity (J. Upcroft, P. Upcroft published results).

Unlike in most eukaryotic cells, no extranuclear mitochondrial DNA is present in *Entamoeba*, *Giardia* or *Trichomonas*, as they lack this organelle. Since mitochondria and hydrogenosomes seem to share a common ancestry (Dyall and Johnson, 2000), there have been several attempts made to detect extrachromosomal DNA in the hydrogenosomes of trichomonads, all of which have been unsuccessful (Turner and Müller, 1983; Wang and Wang, 1985a; Clemens and Johnson, 2000). The crypton organelle (mitosome) of *E. histolytica*, that is believed to be derived from an ancestral mitochondria, also appears to contain no DNA (Tovar et al., 1999; Ghosh et al., 2000a).

A double stranded RNA (dsRNA) virus has been detected in some strains of both *Trichomonas* and *Giardia* (Wang and Wang, 1985a, 1986; Flegr et al., 1987, 1988). These dsRNA viruses contain two overlapping open reading

frames encoding the viral capsid protein and the RNA-dependent RNA polymerase (White and Wang, 1990; Wang et al., 1993; Khoshnan et al., 1994; Khoshnan and Alderete, 1995; Su and Tai, 1996). Unlike *Giardia* and *Trichomonas*, *E. histolytica* has not been reported to contain a dsRNA virus, but it does contain an intriguing nuclear-localised plasmid that has been characterised in considerable detail (Bhattacharya et al., 1989; Dhar et al., 1995; Lioutas et al., 1995). This plasmid contains varying numbers of ribosomal DNA (rDNA) genes (Lioutas et al., 1995). The most abundant have 15–26 kb of rDNA and are present at about 200 copies per cell (Bhattacharya et al., 1989). No genes encoding ribosomal RNAs have been reported on *Entamoeba* chromosomes, thus all rRNA appears to be plasmid encoded. The rDNA circles consist of one or more rRNA transcription units arranged as inverted repeats. The transcription unit contains coding sequences for small subunit, 5.8S and large subunit (LSU) rRNAs (Ramachandran et al., 1993). The regions upstream and downstream of these units contain several families of short tandem repeats, which are similar to intergenic spacers separating rDNA units in a variety of organisms (Mittal et al., 1991, 1992; Sehgal et al., 1993, 1994). It is not known whether the rDNA circle encodes any transcripts other than rRNAs. The rDNA circle EhR1 in strain HM-1:IMSS encodes a 0.7 kb polyadenylated transcript detectable on Northern blots, which lacks any significant open reading frame (ORF) (Burch et al., 1991; Sehgal et al., 1993, 1994). It is not yet known which RNA polymerase is responsible for transcription of the episomal rDNA. In the HK-9 strain, the putative promoter sequence [AGTGAAAATATACTATACAGGAGAAG] and the sequence of the 5′ flanking region are very different from known consensus promoters (Michel et al., 1995) for RNA polymerase I, the polymerase responsible for rDNA transcription in other eukaryotic cells. Interestingly, Jansson et al. (1994) have reported a correlation between the presence of the EhR1 rDNA plasmid and the virulence of *E. histolytica* strains.

Replication of the rDNA plasmids initiates from multiple dispersed locations throughout the rDNA molecules rather than from a single site. The frequency of initiation is two times higher within the rDNA transcription units than in the intergenic spacers (Dhar et al., 1996). In addition, replication may terminate at many locations in the plasmid (Dhar et al., 1998). The absence of a fixed replication origin is a novel finding, previously unobserved for other plasmids.

4. Methods of introducing foreign DNA

4.1. Transfection systems, gene silencing and inducible gene expression

During the past decade, techniques for transient and selectable transfection have been developed for *Giardia*, *Entamoeba* and *Trichomonas*. Transfection techniques have been primarily based on the introduction of closed-circular plasmids containing exogenous reporter genes flanked by 5′and 3′ untranslated regions of parasite genes using electroporation. Chloramphenicol acetyltransferase (CAT) (Nickel and Tannich, 1994; Delgadillo et al., 1997), firefly luciferase (Purdy et al., 1994; Yee and Nash, 1995; Delgadillo et al., 1997) and green fluorescent protein (GFP) have been used for transient expression. Neomycin phosphotransferase (Hamann et al., 1995; Vines et al., 1995; Delgadillo et al., 1997; Sun et al., 1998), puromycin *N*-acetyltransferase (Singer et al., 1998; Dan et al., 2000; Dan and Wang, 2000) and hygromycin phosphotransferase (Hamann et al., 1997; Ramakrishnan et al., 1997; Dan et al., 2000; Dan and Wang, 2000) have been used for selectable transfection. In addition to using plasmid vectors and electroporation to transfect *Giardia*, Wang and colleagues have elegantly engineered the *Giardia* double-stranded RNA virus as a means to introduce and express both exogenous and endogenous genes (Yu and Wang, 1996; Yu et al., 1996a,b).

To extend the use of DNA transfection in *G. intestinalis* and *E. histolytica*, tetracycline-inducible transfection systems were developed by integrating the bacterial *tet* operator–repressor elements into episomal DNA transfection vectors (Hamann et al., 1997; Ramakrishnan et al., 1997; Sun and Tai, 2000). In *Giardia*, a vector containing two copies of the tetracycline operator (*tetO*) inserted between the 32-bp long Ras-like nuclear protein (*ran*) gene minimal promoter was used for regulated expression of the reporter luciferase gene (*luc* +) (Sun and Tai, 2000). This construct also contains the selectable marker *neo*, driven by the *ran* promoter, and the tetracycline repressor (TetR) that is controlled by the *Giardia* gene promoters. Without tetracycline, binding of TetR to the *tetO*-modified promoter site results in the repression of *luc* + gene transcription. Luciferase activity is inducible by increasing tetracycline concentrations from 0.1 to 10 μg/ml. In *Entamoeba*, two different tetracycline-inducible systems have been developed. One utilises a single episomal DNA plasmid construct bearing the TetR gene under the control of *E. histolytica* lectin (hlg5) (Buß et al., 1995; Purdy et al., 1996) gene promoter, the selectable hygromycin resistance gene under the control of actin gene promoter, and *tetO* in front of a CAT reporter gene (Hamann et al., 1997). The second system contains two constructs: a plasmid containing the tet-repressor gene under the control of ferredoxin untranslated regions and a hygromycin resistance gene and another episome with a *tetO*-controlled luciferase gene together with the neomycin resistance gene (Ramakrishnan et al., 1997; Sun and Tai, 2000). In both systems, the *tetO* is inserted between the TATA-box and the ATTCA initiator element of *E. histolytica hgl5* lectin gene promoter (the promoter structure of this gene is discussed in detail subsequently). The repression of expression of the downstream reporter gene is diminished in the presence of 1 μg/ml of tetracycline. These inducible gene expression systems

should allow essential genes to be disrupted without a lethal phenotype.

Homologous gene replacement (i.e. gene knock-out) has been reported for *Giardia* (Singer et al., 1998) and has been achieved in *Trichomonas* (our unpublished results; Land et al., manuscript in preparation) by introducing selectable markers flanked by long regions of parasite DNA on linearised plasmids. Unfortunately, similar attempts to knock out genes in *Entamoeba* have been unsuccessful. An alternative approach of disrupting genes of interest is the use of antisense RNA. Such an RNA-based repression/silencing system has been developed in *Giardia* using its dsRNA virus (Yu et al., 1996a,b; Yu and Wang, 1996). A transfection vector that consists of the target gene sequence in the reverse orientation flanked by 5′ and 3′ untranslated regions of the viral genome and a puromycin selectable marker gene has been utilised. This cassette is transcribed in vitro, and the recombinant transcript is then transfected into *Giardia* trophozoites via electroporation (Yu et al., 1995). The transcript replicates inside the transfected cells in the presence of the wild type dsRNA virus. The efficiency of a functional gene knock-out can be made more efficient by the inclusion of a hammerhead ribozyme inserted within the antisense target gene sequence (Dan et al., 2000; Dan and Wang, 2000).

A successful inhibition of gene expression in *E. histolytica* has also been reported using RNA antisense methods (Moshitch-Moshkovitch et al., 1998). The constructs utilised here are based on the finding that the 5′ untranslated regions of a specific ribosomal protein L21 gene does not bind to polyribosomes in this organism (Moshitch-Moshkovitch et al., 1997). Therefore, target genes that are placed in the antisense direction downstream of this 5′ untranslated regions can be transcribed to produce large quantities of free RNA, to subsequently reduce the endogenous gene's expression (Moshitch-Moshkovitch et al., 1998; Bracha et al., 2000). This technique has been successfully applied to significantly decrease expression of various genes in *E. histolytica*, thus allowing functional studies of a spectrum of *Entamoeba* proteins (Ankri et al., 1998, 1999a,b; Bracha et al., 1999, 2000).

5. Transcription

5.1. Transcription of protein-encoding genes by an α-amanitin-resistant RNA polymerase

Transcription in eukaryotes is carried out by three distinct DNA-dependent RNA polymerases. Typically, RNA polymerase I transcribes ribosomal RNAs, RNA polymerase II is responsible for the transcription of protein-encoding genes, and RNA polymerase III handles the synthesis of small RNAs, such as tRNAs. These three RNA polymerases display differential sensitivity to the fungal toxin α-amanitin. In metazoa, RNA polymerase I is resistant to concentrations of α-amanitin up to 1 mg/ml, while RNA polymerase III is moderately sensitive with 50% inhibition at 50–150 μg/ml α-amanitin. RNA polymerase II is the most sensitive, displaying 50% inhibition at 2–20 μg/ml α-amanitin (Lindell et al., 1970). Surprisingly, transcription of protein-encoding genes has been found to be resistant to high levels of α-amanitin in *E. histolytica*, *G. intestinalis*, and *T. vaginalis* (Lioutas and Tannich, 1995; Quon et al., 1996; Yee et al., 2000). In *T. vaginalis* 50% inhibition of protein-encoding gene transcription was seen with 250 μg/ml of α-amanitin, while synthesis of protein-encoding genes was resistant to 1 mg/ml α-amanitin in both *E. histolytica* and *G. intestinalis*. This raises the question of which RNA polymerase is responsible for the transcription of protein-encoding genes in these protists. There is precedent for transcription of protein-encoding genes by RNA polymerase I in parasitic protozoans. In the African trypanosome, *Trypanosoma brucei*, transcription of the variant surface glycoprotein and procyclin genes is resistant to 1 mg/ml α-amanitin (Kooter and Borst, 1984; Rudenko et al., 1989). To date all available evidence is consistent with RNA polymerase I transcription of these genes, unlike all other protein-encoding genes in this organism which appear to be transcribed by an α-amanitin-sensitive RNA polymerase II (reviewed in Lee and Van der Ploeg, 1997).

In the case of *T. vaginalis*, the available data favour RNA polymerase II transcription of protein-encoding genes (Quon et al., 1996; Liston and Johnson, 1999; Vanacova et al., 2001a). Experiments examining the sensitivity of transcription to both α-amanitin and tagetitoxin in various trichomonad species provide evidence of transcription by three distinct RNA polymerases (Vanacova et al., 2001a). In *T. vaginalis*, while transcription of rRNAs and protein-encoding genes is resistant to α-amanitin, synthesis of tRNAs is inhibited by tagetitoxin, typical of an RNA polymerase III-like enzyme. However, there is a great deal of diversity in the sensitivity of transcription to α-amanitin among the trichomonad species examined (Vanacova et al., 2001a). For example unlike *T. vaginalis*, transcription of protein-encoding genes is highly sensitive to α-amanitin in *Tritrichomonas foetus*. Quon et al. isolated and characterised the gene encoding the LSU of RNA polymerase II (RBP1) from *T. vaginalis* (Quon et al., 1996). A comparison of the amino acid sequences of the putative α-amanitin binding site in RBP1s from *T. vaginalis* and higher eukaryotes, revealed that several residues that were conserved among the α-amanitin-sensitive RPB1s, were not conserved in the *T. vaginalis* enzyme (Quon et al., 1996). These differences could allow α-amanitin-resistant transcription of protein-encoding genes by the *T. vaginalis* RNA polymerase II. The amino acid sequence of the *T. vaginalis* RBP1 was found to be 41% identical to the murine RBP1 and 60% similar to both the murine and *Plasmodium falciparum* RBP1s. Although the *T. vaginalis* RBP1 shares a high degree of similarity with its metazoan counterparts, one unique feature is the composition of its C-terminal

domain. In higher eukaryotes, the C-terminal domain consists of a heptapeptide, YSPTSPS, which is repeated 26 times in yeast and 52 times in mammalian RBP1s (Dahmus, 1995). The *T. vaginalis* RBP1 C-terminal domain, however, consists of a dipeptide, SP, which is repeated 19 times (Quon et al., 1996). Even though it lacks the heptapeptide repeat, due to the large number of serine residues it is likely that the *T. vaginalis* RBP1 C-terminal domain is phosphorylated during transcription initiation like its higher eukaryotic counterparts. As RNA polymerase subunits from *E. histolytica* or *G. intestinalis* have yet to be characterised, it is unknown whether they share any of the unique features of the *T. vaginalis* enzyme.

Additional evidence that transcription of protein-encoding genes is likely carried out by an RNA polymerase II in *T. vaginalis* comes from the characterisation of their promoters (discussed in detail subsequently). These promoters have a bipartite structure with a conserved core promoter and upstream regulatory elements, which is consistent with RNA polymerase II promoters in other eukaryotes (Quon et al., 1994; Liston et al., 1999; Liston and Johnson, 1999; Tsai et al., 2001). A conserved core promoter element, the initiator, has been identified surrounding the transcription start-site of every protein-encoding gene examined so far (Quon et al., 1994; Liston and Johnson, 1999) (discussed in detail subsequently). This initiator is structurally and functionally similar to its metazoan counterpart, which is recognised by transcription factors associated with RNA polymerase II transcription (Smale, 1997; Liston and Johnson, 1999).

In addition, the structure of protein-encoding gene promoters in *E. histolytica* resembles that of RNA polymerase II promoters in other eukaryotes (Purdy et al., 1996; Singh et al., 1997) (discussed in detail subsequently). In contrast, promoters of protein-encoding genes in *G. intestinalis* have a unique structure as compared to the RNA polymerase II promoters of other eukaryotes (Sun and Tai, 1999; Yee et al., 2000; Elmendorf et al., 2001b) (discussed further subsequently). Therefore, while the available evidence for *T. vaginalis* and *E. histolytica* supports RNA polymerase II transcription of protein encoding genes, more information is needed to ascertain the nature of the RNA polymerases responsible for protein-encoding gene transcription in *G. intestinalis*.

Analyses of the structure and expression of a number of genes from all three amitochondriates have provided evidence that genes in these protists are arranged in single transcription units. For example, in *Giardia* and *E. histolytica*, coding regions tend to be closely linked unidirectionally within the genome (Petter et al., 1992; Bruchhaus et al., 1993; McArthur et al., 2000). Certain multi-copy genes in *E. histolytica* were found to be unlinked and distributed on distinct chromosomes (Bruchhaus et al., 1993; Ramakrishnan et al., 1996; Willhoeft and Tannich, 1999). In all three parasites, the 5′ sequences of cDNAs are co-linear with genomic copies of the corresponding genes, indicating the absence of a *trans*-spliced 5′ splice-leader RNA. In contrast, *cis*-splicing introns have been found in *Entamoeba* and *Giardia*. However, this may be relatively rare as introns have been reported in only a few *E. histolytica* genes so far (Willhoeft et al., 2001) (Lohia and Samuelson, 1993; Plaimauer et al., 1994; Urban et al., 1996) and in only one *G. intestinalis* gene (Nixon et al., 2002). *Entamoeba histolytica* introns are relatively short, ranging from 46 to 115 bp, and possess all the conserved elements found in the introns of higher eukaryotes. The dinucleotides GU and AG are located at the donor and acceptor splice sites and the branchpoint sequence, UUCUUAAU, matches the eukaryotic consensus, PyNPyPyPuAPy (Plaimauer et al., 1994; Urban et al., 1996). The single analysed *Giardia* intron also contains the typical eukaryotic consensus sequences involved in the splicing reaction, with one notable exception: the dinucleotide located at the 5′ donor site is CU instead of GU (Johnson, 2002). Despite the loss of conservation of this dinucleotide, there is remarkable sequence identity between the first 10 nucleotides of the donor site of the *Giardia*intron (CUAUGUUGAG) and the single *cis*-spliced intron characterised in trypanosomes (GUAUGGUGAG) (Mair et al., 2000). Moreover, these sequences are similar to the consensus intron donor sequences of yeast (GUAUG), mammals (GUAAG) (Burge et al., 1998) and *E. histolytica* (GUUUG) (Willhoeft et al., 2001) implying that the same mechanism may be used to recognise the donor site of protist, yeast and mammalian introns (Johnson, 2002).

To date, no introns have been identified in any trichomonad genes. However, a single copy gene encoding the essential spliceosomal protein, PRP8, has been detected in *T. vaginalis* (Fast and Doolittle, 1999). The 7 kb open reading frame of this gene is preceded by a typical *T. vaginalis* initiator and analysis of the PRP8 amino acid sequence shows a high degree of homology to other eukaryotic PRP8s, especially in regions directly involved in *cis*-splicing (Fast and Doolittle, 1999). This finding suggests that trichomonads possess functional spliceosomes and therefore some of their genes may have introns. Data obtained from a *Trichomonas* genome project should reveal these, if indeed they are present.

5.2. Promoter structure

The development of transfection methods has greatly enhanced our understanding of transcription in these parasites (Nickel and Tannich, 1994; Purdy et al., 1994; Gilchrist et al., 1995; Hamann et al., 1995; Vines et al., 1995; Yee and Nash, 1995; Delgadillo et al., 1997; Singer et al., 1998; Sun et al., 1998). These techniques have allowed the use of reporter gene assays (described earlier) to map the location of DNA elements required for transcription and to then analyse the sequence requirements for their function. This has given us a much clearer picture of what constitutes a promoter for protein-encoding genes in these

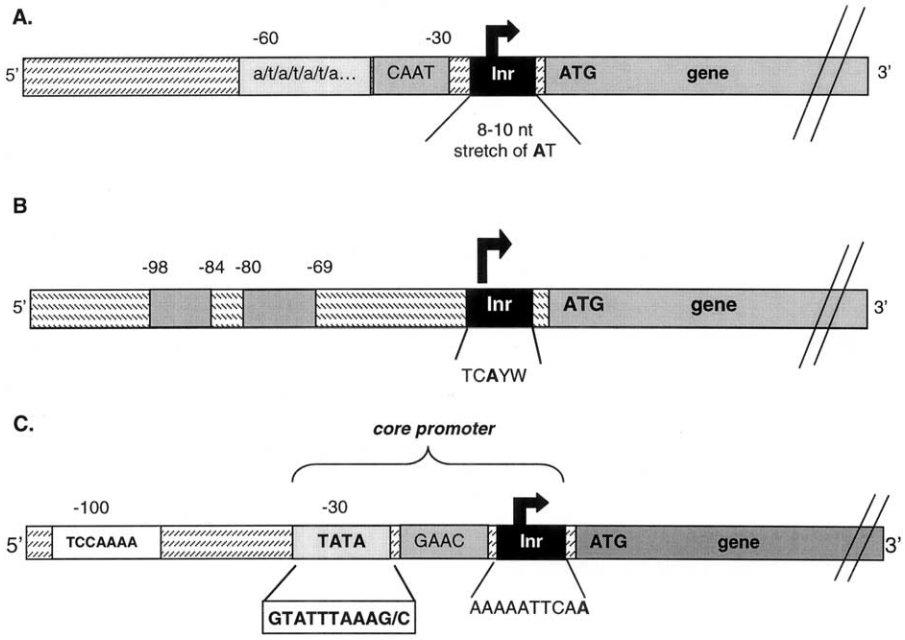

Fig. 1. A schematic diagram of promoter structure for *G. intestinalis* (A), *T. vaginalis* (B), and *E. histolytica* (C).

organisms. In higher eukaryotes, promoters of protein-encoding genes consist of complex arrays of specific DNA sequences that combine diverse gene-specific enhancer elements with more commonly shared core promoter elements, such as the TATA-box and initiator (reviewed in Lemon and Tjian, 2000). The core promoter regions of higher eukaryotic genes are quite heterogeneous (reviewed in Smale, 1997). Some contain both a TATA-box and an initiator, while others contain only one of these elements. In some genes, however, the core promoter elements are undefined, as they lack either of these sequences. Promoters in *E. histolytica* and *T. vaginalis* appear to have a similar bipartite structure as other eukaryotes with a conserved core promoter region and upstream regulatory elements (Purdy et al., 1996; Singh et al., 1997; Liston et al., 1999; Liston and Johnson, 1999; Tsai et al., 2001). Interestingly, the core promoters of both parasites do not display the same heterogeneity seen in higher eukaryotes, each has a highly conserved core promoter structure that appears to be shared by nearly all examined genes (Quon et al., 1994; Purdy et al., 1996; Liston and Johnson, 1999). Promoter regions in *G. intestinalis*, unlike those of either *E. histolytica*, *T. vaginalis*, or higher eukaryotes, are extremely compact, as only approximately 50 bp are required for expression of a reporter gene in transfected cells (Sun and Tai, 1999; Yee et al., 2000; Elmendorf et al., 2001b). While the promoters examined so far have some common sequence elements, such as upstream AT-rich and CAAT-like sequences and a Py-poly A sequence surrounding the transcription start-site, these elements are not very conserved among the various *G. intestinalis* promoters examined so far, suggesting that the sequences required for promoter activity are highly degenerate in this protist. A schematic diagram of promoter structure for all three organisms is shown in Fig. 1.

5.2.1. Promoters in E. histolytica

The first identification of potential promoter elements in *E. histolytica* came from the examination of the upstream sequences of several genes (Bruchhaus et al., 1993). Typical eukaryotic promoter elements, such as a consensus TATA-box or classic initiator element, were not found in these upstream sequences. However, those studies identified two elements: the first was the sequence ATTCA or ATCA surrounding the transcription start-site, and the second element was an unusual TATA-like motif, TATTTAAA, located approximately 30 bp further upstream (Bruchhaus et al., 1993).

One of the first *E. histolytica* promoters to be analysed, the promoter of the gene encoding the 170 kDa LSU of the GalNAc-specific lectin *hgl2*, was found to be restricted to a fragment containing the 550 bp upstream of the transcription start-site (Buß et al., 1995). Within this region, two elements were identified by sequence comparison with another lectin heavy chain gene: (1) an element 100 bp upstream of the transcription start-site which was similar to the CCAAT-box sequence found in some higher eukaryotic promoters and (2) a 15 bp sequence at −520 from the transcription start-site (Buß et al., 1995). Mutation of these two sequences, the TATA-like motif, or the ATTCA element at the transcription start-site reduced promoter activity to between 20 and 52% of wild-type, demonstrating their importance for full activity of the *hgl2* promoter.

The best characterised promoter in *E. histolytica* is from the *hgl5* gene, which like *hgl2* encodes the GalNAc-specific lectin. Purdy et al. found that full promoter activity was

confined to a fragment containing 272 bp upstream of the transcription start-site (Purdy et al., 1996). Linker-scanning analysis of this region identified five upstream regulatory elements (UREs) critical for *hgl5* expression. Four of these elements were positive regulatory elements: URE5 at −219 to −200, URE4 at −189 to −160, URE2 at −69 to −60, and URE1 at −49 to −40, while one negative regulatory sequence, URE3, was identified at −89 to −80 (Purdy et al., 1996). Two of these upstream elements have been studied in more detail. URE4 consists of two nine-base pair repeats of the sequence AAAAATGAA (Schaenman et al., 1998) and recently a 28 kDa protein, EhEBP1, and a 18 kDa protein, EhEBP2, both of which recognise this sequence have been isolated by DNA-affinity chromatography (Schaenman et al., 2001). While analysis of these proteins did not reveal any motifs common to DNA-binding proteins, both proteins contain sequences homologous to the RNA recognition motif, a domain which has been found in a large number of RNA-binding proteins and several sequence-specific DNA binding proteins (Nagai et al., 1995). The larger EhEBP1 has two copies of the RNA recognition motif, while EhEBP2 has one. Both proteins also contain a 54 or 58 amino acid regions, respectively, rich in acidic and basic residues (Schaenman et al., 2001). Interestingly, over-expression of EhEBP1 led to a seven-fold drop in expression of a reporter gene under the control of the *hgl5* promoter (Schaenman et al., 2001). This suggests that either EhEBP1 is actually a transcriptional repressor, and other proteins function through URE4 to activate transcription, or that the function of EhEBP1 requires a cofactor, present in limiting amounts, which was sequestered by the overexpression of EhEBP1. Further experiments with EhEBP1 and EhEBP2 will be required to elucidate the exact mechanisms of URE4-regulated transcription.

URE3, in contrast to its negative regulatory role in the expression of *hgl5*, was found to be a positive regulator in the ferredoxin gene promoter (Gilchrist et al., 1998). Using a yeast one-hybrid screen, a 22.6 kDa protein, URE3-BP, which binds to the URE3 sequence has been identified (Gilchrist et al., 2001). As with EhEBP1 and EhEBP2, analysis of the URE3-BP amino acid sequence did not identify known DNA-binding domains (Gilchrist et al., 2001). The analysis did reveal the presence of two EF-hand motifs, which is the most common calcium-binding motif found in proteins (Lewit-Bentley and Rety, 2000). The presence of the two EF-hand motifs suggests that the activity of URE3-BP may be regulated by calcium. Gilchrist and colleagues found that relatively high concentrations of calcium, from 100 to 500 μM, inhibited DNA binding by URE3-BP (Gilchrist et al., 2001). Interestingly, they also showed that URE3-BP was localised not only in the nucleus and cytoplasm, but also in the cell membrane (Gilchrist et al., 2001). This suggests that one mechanism used to regulate its function may be its location within the cell. Future experiments should reveal the mechanism of URE3-BP function and how its activity is regulated.

Importantly, these URE sequences appear to be specific for *E. histolytica* promoters and do not resemble known *cis*-acting regulatory elements of other eukaryotes. Several studies, however, have noted sequences similar to metazoan-like regulatory elements upstream of some *E. histolytica* genes. Sequences similar to the C/EBP, GATA-1, OCT, HOX, and AP-1 binding sites were identified by sequence homology in the promoter regions of the *E. histolytica* EhPgp1 and EhPgp5 genes (Gomez et al., 1998; Perez et al., 1998). Several of these elements were able to compete for DNA binding in electrophoretic mobility shift assays (Gomez et al., 1998; Perez et al., 1998). However, as a mutational analysis of these elements was not performed, it is unclear whether the sequence requirements for DNA binding of these elements is the same in both *E. histolytica* and metazoans. In addition, the contribution of these elements to the activity of the EhPgp1 and EhPgp2 promoters has not been determined using a functional assay. Therefore, it is uncertain whether the presence of these metazoan-like elements is simply coincidental or whether they are required for the expression of these genes. Actin expression in *E histolytica* has been shown to be upregulated by exposure to cyclic adenosine monophosphate (cAMP) and pyridylmercuric acetate (PMA) (Manning-Cela and Meza, 1997). In higher eukaryotes these signals function through cAMP response elements and serum response elements, respectively, to regulate gene expression. Ortiz et al. examined an *E. histolytica* actin gene promoter and identified several serum response elements and cAMP response element-like sequences (Ortiz et al., 2000). Interestingly, mutation of these elements abolished upregulation of the actin promoter activity by treatment with cAMP and PMA (Ortiz et al., 2000). Further analysis will be required to determine whether the sequence requirements for the function of the *E. histolytica* serum response elements and cAMP response element-like elements match those of their metazoan counterparts. These results do suggest, however, that *E. histolytica* may contain some homologues of higher eukaryotic *cis*-acting regulatory elements.

The structure of the core promoter region in *E. histolytica* genes appears to be unusual compared to other eukaryotes. Purdy et al. examined the core promoter region of 37 *E. histolytica* genes and found three conserved sequences: (1) The previously identified TATA-like motif with the consensus sequence GTATTTAAA(G/C), (2) a putative initiator-like sequence, AAAAATTCA, surrounding the transcription start-site (this element corresponds to the ATTCA/ATCA sequence described earlier), and (3) a novel element named GAAC, with the consensus sequence GAACT, which has a variable location between the TATA-like and initiator-like motifs (Purdy et al., 1996). Mutational analyses of the *E. histolytica* TATA-like and initiator-like elements indicate that they function in an analogous manner to their higher eukaryotic counterparts in directing transcription start-site locations.

Using primer extension analysis, Singh and colleagues mapped the transcription start-sites from constructs containing various mutations in these core promoter elements (Singh et al., 1997). These results demonstrated that the TATA-like element is dominant for transcription start-site selection in *E. histolytica* (Singh et al., 1997). Transcription initiated 30–31 bp downstream of the TATA-like sequence regardless of whether a functional initiator-like element was present. In addition, if the TATA-like element was shifted 10 or 40 bp upstream, the transcription initiation sites were also shifted by 10 or 40 bp, respectively. While the TATA-like element is dominant for start-site selection, the *E. histolytica* initiator-like element was able to direct accurate transcription initiation when the TATA-like and/or GAAC elements were mutated, indicating that it is able to function independent of the other core promoter sequences (Singh et al., 1997). An *E. histolytica* homologue of the TATA-box binding protein has been cloned by degenerate polymerase chain reaction with primers based on the sequence of the *Acanthamoeba castellanii* TATA-box binding protein (Luna-Arias et al., 1999). The C terminal domain of the 26 kDa *E. histolytica* TATA-box binding protein has 55% and 37% sequence identity to the human and *P. falciparum* TATA-box binding protein C-terminal domains, respectively. This finding suggests that despite the differences relative to other eukaryotic counterparts, the *E. histolytica* TATA element is recognised by a typical eukaryotic TATA-box binding protein. However, this is still speculation, as the DNA-binding properties of the *E. histolytica* TATA-box binding protein have yet to be characterised.

The GAAC element appears to play a unique role in transcription initiation in *E. histolytica*. Using a nuclear run on assay, Singh and colleagues demonstrated that mutation of the GAAC element had a negative effect on transcription initiation in two *E. histolytica* promoters; hgl5 and ferredoxin (Singh and Rogers, 1998; Singh et al., 2002). Interestingly, the site of transcription initiation was altered in GAAC mutants in the hgl5 promoter (Singh and Rogers, 1998) but not the ferredoxin promoter (Singh et al., 2002). The new start-sites in the hgl5 promoter were in addition to the wild-type transcription initiation sites within the initiator. An unresolved issue is whether the GAAC element can function in the absence of a functional TATA and initiator to direct transcription initiation. Since mutation of the GAAC element appears to primarily affect the rate of transcription initiation, one possible function of the GAAC sequence would be to stabilise the binding of the basal transcription apparatus to the core promoter. In this case, it might function in an analogous manner to the higher eukaryotic downstream promoter element. In certain *Drosophila* and human TATA-less promoters, downstream promoter element has been shown to function cooperatively with the initiator to stabilise the binding of TFIID to the core promoter (reviewed in Burke et al., 1998). Recent analyses of the role of GAAC indicate that this region is necessary for higher-order nuclear protein complex assembly. Competitive gel shift assays showed that the same nuclear protein binds to the GAAC element in the *hgl5* and ferredoxin promoters (Singh et al., 2002). The function of GAAC in transcription activation is mediated by UREs. The URE3 motif fails to modulate expression of GAAC mutant minimal promoters while it does effect expression of promoter constructs in which GAAC is a wild type. On the other hand, the URE4 modulates the transcription initiation proportionally, in both GAAC mutant and wild type minimal promoters. Together, these studies support a role for GAAC in protein assembly and activation mediated by URE3. However, the different effects observed using two different promoters imply a context-dependence for GAAC-mediated gene regulation (Singh et al., 2002).

While the consensus sequences of the *E. histolytica* TATA and initiator differ significantly from their metazoan counterparts, the location and function of these elements strongly suggest that they are true homologues of the higher eukaryotic TATA-box and initiator (Purdy et al., 1996; Singh et al., 1997). The GAAC element, however, appears to be unique to *E. histolytica* as no other eukaryotic core promoter element has been identified with a similar sequence or location.

5.2.2. Promoters in G. intestinalis

Unlike other eukaryotes, including *Entamoeba* and *Trichomonas*, promoter regions in *G. intestinalis* appear to be remarkably compact. This is likely due to the high degree of gene density within the *Giardia* genome, resulting in very short intergenic regions, often less than 250 bp (Boothroyd et al., 1987; Smith et al., 1998). These short intergenic regions, therefore, likely contain all of the sequence elements required for transcriptional regulation. The first potential promoter elements in *G. intestinalis* were identified in a weight matrix analysis of the 5′ flanking sequences of seven cytoskeleton protein-encoding genes (Holberton and Marshall, 1995). This analysis revealed three potential promoter elements within the 70 bp upstream of the transcription start-site. The first was an AT-rich element surrounding the transcription start-site, which had the consensus $(Py)_{n=1-4}(A)_{n=2-5}$(Holberton and Marshall, 1995). The transcription start-sites of these genes were usually mapped to the first or second A following the last Py base. Given the position of this element, it has been proposed to function as an initiator-like element to direct the location of transcription start-sites (Holberton and Marshall, 1995). The other two elements identified were 6 and 18 bp elements located from −20 to −70 bp upstream of the transcription start-site (Holberton and Marshall, 1995). These elements were both AT-rich and contained several CAAT-like motifs. With the development of transfection methods for *Giardia* (Yee and Nash, 1995; Singer et al., 1998; Sun et al., 1998), several promoters have now been examined in detail and these analyses have shown that this weight matrix analysis was quite successful in identifying sequence elements important for promoter activity.

So far, the promoters for three *Giardia* genes have been analysed in detail, from the ras-related nuclear protein (*ran*), glutamate dehydrogenase (*GDH*), and α2-tubulin genes. Analysis of the GDH gene promoter revealed that 44 bp upstream of the translation start codon were required for full promoter activity and identified three sequence elements within this region (Yee et al., 2000). The first was a CAAT-like sequence, mutation of which had a minor effect on promoter activity. The second was an element between the CAAT sequence and the transcription start-site, which has a critical AG dinucleotide (Yee et al., 2000). The most important sequence, however, was the AT-rich sequence surrounding the transcription start-site. This sequence, ATTTTAAAAT, matched the consensus of the proposed *Giardia* initiator and mutation of this element essentially abolished promoter activity (Yee et al., 2000).

Examination of the *ran* gene promoter demonstrated that a 32 bp region was sufficient for full promoter activity (Sun and Tai, 1999). Two distinct elements were identified within this region, an AT-rich sequence at −51 from the translation start codon and a *Giardia* initiator sequence, ATTAAAA, at −30 (Sun and Tai, 1999). Interestingly, transcription start-sites were mapped to this initiator sequence in constructs containing either natural *ran* promoter or a synthetic construct in which only this 32 bp minimal *ran* promoter region was placed upstream of a luciferase reporter cassette (Sun and Tai, 1999). This suggests that the *Giardia* initiator is responsible for transcription start-site selection. Further analyses of distal regulatory elements in this promoter indicate a role for T-tract elements in modulating transcriptional activity (Ong et al., 2002). These studies have also revealed the presence of a bi-directional promoter that directs transcription of the *ran* gene from one strand and a PhD-finger protein gene from the other strand.

Analysis of the promoter of the α2-tubulin gene has shown that it has a similar structure to the GDH and *ran* promoters, with a 64 bp region required for full promoter activity (Elmendorf et al., 2001b). Within this region, two elements were identified, a 38 bp AT-rich sequence at −60 and an initiator element, AAATAAAA, containing the transcription start-site (Elmendorf et al., 2001b). As in the GDH promoter, mutation of the CAAT elements within the α2-tubulin promoter had a relatively minor effect on promoter activity, suggesting that these CAAT-like sequences may be components of a larger regulatory element (Elmendorf et al., 2001b). Interestingly, with the exception of several CAAT-like sequences within the GDH and α2-tubulin elements, and the overall AT-bias, comparison of the upstream promoter elements from these three genes fails to reveal highly conserved sequences.

Unlike the core promoter elements of higher eukaryotes, the *Giardia* initiator may be able to direct transcription in vivo by itself, in the absence of other promoter sequences. Elmendorf et al. analysed three constructs which contained solely a different *Giardia* initiator sequence upstream of a luciferase reporter gene (Elmendorf et al., 2001b). After stable transfection of *G. intestinalis* with these constructs, levels of luciferase activity were detected, that were significantly above background (Elmendorf et al., 2001b). One important caveat to this experiment is that the start-sites of the transcripts from these constructs were not mapped, therefore, it is unclear whether transcription initiated within the initiator sequences. Yet, if only an initiator element is sufficient for low levels of transcription in *Giardia*, it reveals a significant distinction in transcription between *Giardia* and other eukaryotes, as transcription directed solely by the core promoter elements of higher eukaryotes can only be detected using in vitro transcription assays and is not seen in vivo. In addition, given the degenerate nature of the *Giardia* initiator sequence, functional initiator elements are likely to be distributed at a high frequency throughout the *Giardia* genome resulting in the production of a large number of sterile transcripts. In fact this appears to be the case. Elmendorf and colleagues analysed random clones from two directional cDNA libraries and found that approximately 20% of the examined cDNAs represented polyadenylated sterile antisense transcripts (Elmendorf et al., 2001a). While a role for these transcripts in regulated gene expression cannot be ruled out, it appears more likely that they are the result of promiscuous transcription from cryptic promoters within the *Giardia* genome as the identities of the corresponding sense transcripts are very diverse (Elmendorf et al., 2001a). Whatever the function of these antisense transcripts within the cell, their presence does indicate that there may be substantial differences in the molecular mechanisms of transcription between *Giardia* and other eukaryotes. The presence of antisense RNA for up to 20% of cDNAs also leads to the prediction that RNA interference mechanisms will not be present or functional in *Giardia*.

One unresolved issue is how *Giardia* achieves the complex patterns of gene regulation necessary for its development given the apparent limited sequence of its promoter regions. To begin to address this issue, Knodler et al. examined the expression of glucosamine-6-phosphate isomerase (Gln6PI) which is the first enzyme required for the synthesis of the cyst wall polysaccharide, *N*-acetylgalactosamine (Knodler et al., 1999). In this study, two Gln6PI genes were isolated. The first Gln6PI-A, has a transcript with a short 5′ untranslated regions and is expressed at a low level during both vegetative growth and encystation. The second Gln6PI-B, encodes two transcripts, one, which is expressed constitutively and another whose expression is highly upregulated during encystation (Knodler et al., 1999). Interestingly, the non-regulated transcript has a 146 bp 5′ untranslated regions and appears to be 5′ capped, both of which are very unusual features for a *Giardia* transcript. The upregulated transcript, on the other hand, has a short 4–5 bp 5′ untranslated regions that is not capped (Knodler et al., 1999). It is not clear whether the upregulated transcript is a primary transcript or produced by post-

transcriptional processing of the longer, non-regulated transcript. The role of the 5′ flanking sequence of Gln6PI-B in its expression was examined by fusing either the sequence from −249 to −15, relative to the translation initiation codon, or −56 to +15 to a luciferase reporter gene. In both cases, there were large increases in luciferase activity in 24 h encysting cells as compared with vegetative trophozoites. However, the luciferase activity from the −249 to +15 construct, which includes the 5′-end of the non-regulated transcript, was 2.5-fold higher than the −56 to +15 construct (Knodler et al., 1999). As this experiment only measured luciferase enzyme activity, it cannot distinguish between increased transcription or RNA processing as the cause of the increased luciferase activity seen. However, one can speculate that the sequences between −249 and −56 may be involved in processing of the longer non-regulated transcript to the shorter, regulated form during encystation in order to provide a rapid increase in Gln6PI expression. Since the sequences between −56 and +15 have been shown to provide full promoter activity in other *Giardia* genes, they may then act to increase transcription. While further study will be required to determine the exact sequence elements within these regions that are critical for the regulated expression of Gln6PI, these results provide hints that *Giardia* may use unique mechanisms to control its gene expression.

5.2.3. Promoters in T. vaginalis

Promoters of protein-encoding genes in *T. vaginalis* have a similar structure as those of both *E. histolytica* and higher eukaryotes, with a conserved core promoter region and distal regulatory elements. As in *E. histolytica*, and unlike in higher eukaryotes, the core promoter region of *T. vaginalis* genes is highly conserved. In fact, it was this conservation that led to the discovery of the first promoter element in *T. vaginalis*. Quon et al. examined the upstream flanking sequences of seven *T. vaginalis* genes and found that while no typical TATA-box sequence was seen 25 to 30 bp upstream of the transcription start-site, there was a highly conserved sequence surrounding the transcription start-site with the consensus, T C A + 1 Py (T/A) (Quon et al., 1994). This element has been found in all *T. vaginalis* genes examined so far and its sequence matches that of the metazoan initiator core promoter element, Py Py A + 1 N (T/A) Py Py (Smale and Baltimore, 1989; Liston and Johnson, 1999). The mammalian initiator was first identified in the TATA-less promoter of the murine terminal deoxynucleotidyltransferase (TdT) gene (Smale and Baltimore, 1989). Functionally, the initiator is a core promoter element analogous to the TATA-box as it can direct accurate transcription initiation by RNA polymerase II in the absence of other control elements and mediate a high level of transcription when paired with an upstream activator (reviewed in Smale, 1997). An analysis of the sequence requirements for initiator function in mammalian cells has shown that the most important nucleotides are the Py at −1, A at +1 and (T/A) at +3 (Javahery et al., 1994). Mutational analysis of the *T. vaginalis* initiator element has shown that it is both a structural and functional homologue of its metazoan counterpart (Liston and Johnson, 1999). This study revealed that the same nucleotides that were critical for mammalian initiator function were also the most important for function of the *T. vaginalis* initiator. In addition, the murine TdT initiator could replace the endogenous *T. vaginalis* initiator with no loss of promoter activity. Primer extension analysis of the transcription start-sites from construct containing several mutations in the *T. vaginalis* initiator also demonstrated that the initiator was responsible for accurate start-site selection in *T. vaginalis* cells (Liston and Johnson, 1999). Thus, the initiator represents the first cognate core promoter element found in both deep-branching and metazoan eukaryotes, suggesting that it evolved early during eukaryotic evolution.

One of the key issues in understanding initiator function in higher eukaryotes has been the identity of the protein(s) that recognises the initiator during transcription initiation. Although several proteins, such as YY1, USF, TFII-I, and RNA polymerase II itself, have been shown to bind sites that coincide with the initiators of specific metazoan genes, most of the available evidence supports a role for TFIID in initiator recognition during transcription initiation in mammalian cells (reviewed in Smale, 1997; Burke et al., 1998; Chalkley and Verrijzer, 1999). In addition, the TFIID subunits $TAF_{II}150$ and $TAF_{II}250$ are required for initiator function and recently a $TAF_{II}150$-$TAF_{II}250$ heterodimer has been shown to specifically recognise the initiator in a binding site selection assay (Verrijzer et al., 1995; Kaufmann et al., 1996, 1998; Chalkley and Verrijzer, 1999). The exact mechanism of initiator binding by these transcription activating factors, as well as the identity of the initiator-binding domain(s), has yet to be defined. Since *T. vaginalis* appears to rely exclusively on the initiator to direct transcription initiation (Liston and Johnson, 1999), it is likely that the trichomonad transcriptional machinery is highly optimised for initiator function and is thus an excellent system in which to examine the mechanisms of initiator-mediated transcription.

Recently, a 39 kDa protein that specifically recognises the *T. vaginalis* initiator has been isolated by DNA-affinity chromatography (Liston et al., 2001). This protein, named IBP39 for initiator binding protein 39 kDa, was shown to have initiator-specific binding activity in both mobility shift and DNase I footprinting assays and was responsible for the initiator-binding activity seen in crude *T. vaginalis* nuclear extracts. In addition, nucleotides that are critical for initiator function in vivo were also shown to be necessary for IBP39 binding (Liston et al., 2001). Taken together, these data are consistent with a role for IBP39 in initiator recognition by the *T. vaginalis* transcription machinery. Interestingly, analyses of the amino acid sequence of IBP39 have failed to identify homologues in other organisms or known DNA-binding domains (Liston et al., 2001). Given the similarities

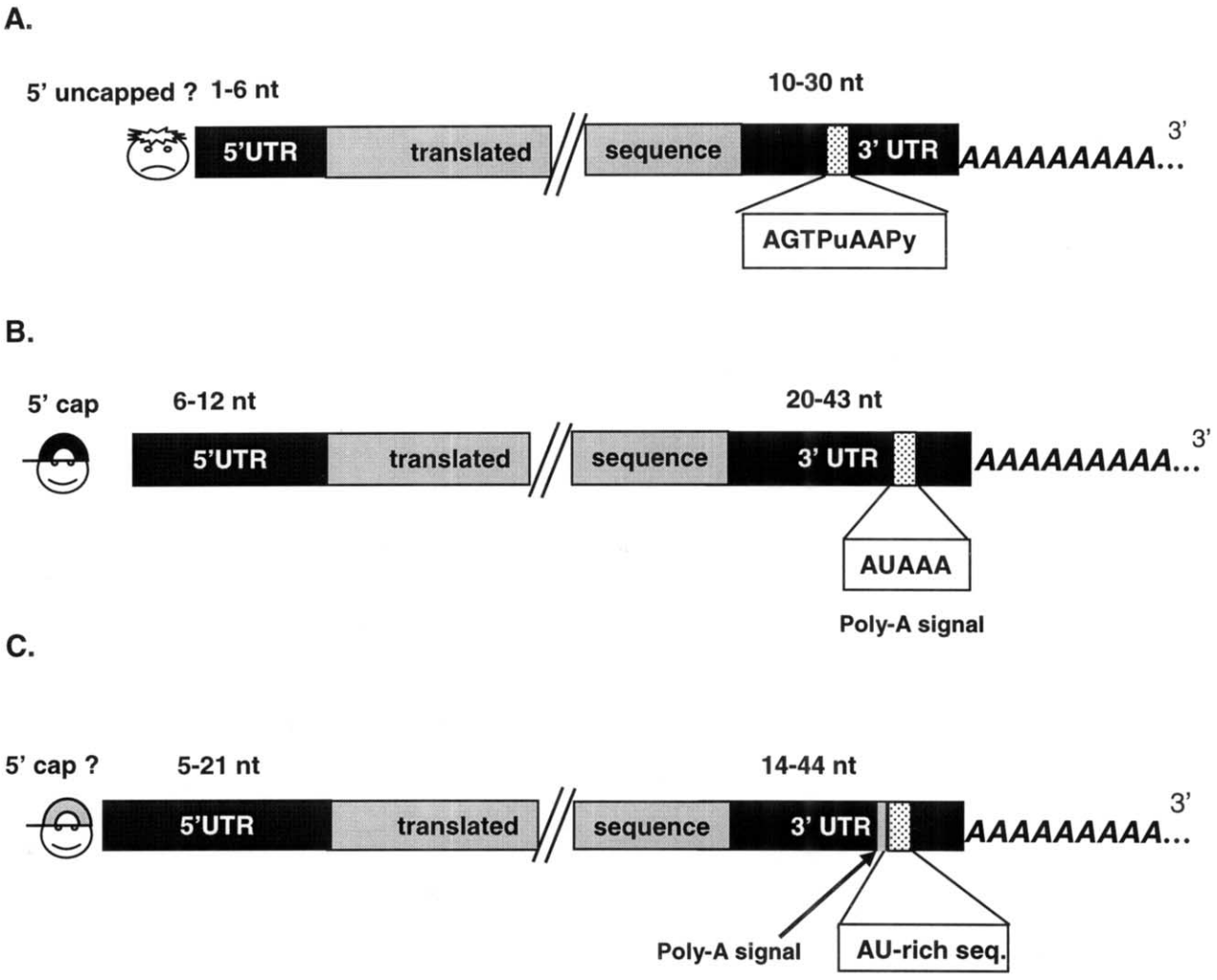

Fig. 2. Schematic comparison of mRNA structure of *G. intestinalis* (A), *T. vaginalis* (B), and *E. histolytica* (C).

between their initiators, it seems likely that trichomonads and metazoans would share a common mechanism for initiator recognition. However, there are several possible explanations for the lack of similarity between IBP39 and either $TAF_{II}150$ or $TAF_{II}250$ at the sequence level. IBP39 may have limited homology to the as yet unidentified initiator-recognition domains of $TAF_{II}150$ and/or $TAF_{II}250$ that a computer search might not recognise. Alternatively, since two studies have shown that TFIID is not sufficient to direct initiator-mediated transcription in a highly purified in vitro transcription system (Kaufmann et al., 1996; Martinez et al., 1998), IBP39 may be a homologue of a metazoan transcription component that has yet to be characterised. Lastly, as the *T. vaginalis* transcription machinery presumably does not have to distinguish between multiple core promoter elements, as does the metazoan transcription apparatus, it is possible that trichomonads have evolved a unique mechanism for initiator recognition. Therefore, IBP39 may be found to function like a transcription activating factor without sharing sequence homology with other eukaryotic transcription activating factors. Interestingly, as described above, conserved sequence elements surround the transcription start-sites of *E. histolytica*, *G. intestinalis*, and *T. vaginalis*. While these elements do not share sequence similarity with each other, it appears that they may play a similar role during transcription initiation. Therefore, it is possible that these three protists share a common mechanism for recognition and selection of transcription start-sites. Further characterisation of the transcription machinery of these three parasites should clarify this issue and lead to a greater understanding of the evolution of gene regulation in eukaryotes.

In addition to the initiator, several upstream promoter elements have been identified through the detailed examination of two *T. vaginalis* promoters. Analysis of the promoter from the α-succinyl CoA synthetase (αSCS) gene revealed four regions within 1051 bp of the transcription start-site which contained important positive and negative regulatory elements (Liston et al., 1999). The most critical region was from −251 to −51 as deletion of this region abolished promoter activity. Linker-scanning mutagenesis of this region revealed two critical sequence elements, from

−98 to −84 and from −80 to −69 (Liston et al., 1999). Neither of these two elements match the sequence of known eukaryotic promoter elements, suggesting that these elements are unique to *T. vaginalis*. Interestingly, the −98 to −84 element was also found upstream of two other hydrogenosomal enzymes, malic enzyme A (MaeA) and pyruvate-ferredoxin oxidoreductase B, showing that this element may be common to other *T. vaginalis* hydrogenosomal genes (Liston et al., 1999). Analysis of the MaeA and pyruvate-ferredoxin oxidoreductase B promoters will be required, however, to determine whether this sequence is also capable of modulating expression of these genes.

Using linker-scanning mutagenesis, three regulatory elements, in addition to the initiator, have been identified within the 235 bp upstream of the transcription start-site of the ap65-1 gene (Tsai et al., 2001). Two of these sequences, at −121 to −102 and −52 to −39, were shown to be important for basal expression of ap65-1, while the third, at −109 to −56, was shown to be necessary for upregulation of promoter activity in response to iron. Tsai et al. also identified a DNA binding activity in nuclear extracts from *T. vaginalis* cells treated with iron which recognises the sequence, AGATAACGA, within the −109 to −56 iron response element (Tsai et al., 2001). This binding activity appears to be induced by iron, as it is not seen in extracts from untreated cells. This correlates with the finding that the expression of the key hydrogenosomal proteins involved in pyruvate metabolism is regulated by iron. It has been shown in *T. foetus* that transcription of genes encoding pyruvate-ferredoxin oxidoreductase and malic enzyme decreased under iron-restricted conditions (Vanacova et al., 2001b). Similar effect was observed in *T. vaginalis* (Tachezy, unpublished data). Interestingly, iron-dependent regulation has been also observed for putative adhesion factor ap 65-1 (Alderete et al., 1998). The ap 65-1 gene is almost identical to gene sequences encoding mallic enzyme suggesting a dual localisation and function of corresponding protein (Hrdy and Müller, 1995).

Taken together, these analyses demonstrate that promoters of protein-encoding genes in *T. vaginalis* have a similar bipartite structure as those of *E. histolytica* and higher eukaryotes, with a conserved core promoter and upstream gene-specific regulatory elements.

6. Structure of mRNA

Eukaryotic mRNAs are composed of the coding region surrounded by the start codon (AUG) and stop codon (UAA, UAG or UGA), and 5′ and 3′ untranslated regions. The ends of both untranslated regions are co-transcriptionally modified; the 5′ untranslated regions is typically modified by a methyl-guanosine cap and the 3′ untranslated regions is polyadenylated. A six nucleotide polyadenylation motif (AAUAAA) signals for addition of roughly 200 adenines by a polyA polymerase. This enzyme, as well as a capping enzyme, is known to associate with the RNA polymerase II complex.

When compared with other eukaryotes, *Entamoeba*, *Giardia*, and *Trichomonas* mRNAs possess unusually short 5′ and 3′ untranslated regions (see Fig. 2). The shortest 5′ untranslated regions, only one to six nucleotides, are found in *Giardia* (Adam, 1991; Yu et al., 1998; Svärd et al., 1999; Yee et al., 2000), whereas *Trichomonas* and *Entamoeba* 5′ untranslated regions are typically five to 20 nucleotides in length (Johnson et al., 1990; Bruchhaus et al., 1993; Lahti et al., 1994; Quon et al., 1994; Granger et al., 1997; McKie et al., 1998). Only a few exceptions have been noted for *Entamoeba* mRNAs (5′ untranslated regions of 420 nucleotides, 126 nucleotides, and 265 nucleotides) (Gangopadhyay et al., 1997a; Luna-Arias et al., 1999), no exceptions have been reported for trichomonad mRNAs so far and only one exception is known for *Giardia* mRNAs (Knodler et al., 1999). As mentioned previously, the gene encoding *Giardia* glucose-6-phosphoisomerase-B mRNA generates two types of transcripts; one that is expressed constitutively with a typical short 5′ untranslated regions and a second that appears during encystation. The 5′ untranslated regions of the encystation-specific glucose-6-phosphoisomerase-B mRNA is 146 nucleotides long, and unlike the majority of *Giardia* mRNAs (discussed subsequently) this message appears to contain a cap (Knodler et al., 1999).

The atypically short 5′ untranslated regions of amitochondriate mRNAs raises questions concerning the mechanism of translation initiation in these organisms. A major difference between eubacterial and eukaryotic mechanism of translation initiation is the requirement for base pairing between 16S rRNA and eubacterial RNAs, a process that proceeds via protein–protein and protein–RNA interactions for eukaryotic mRNAs. These protein–RNA interactions are mediated by the presence of a cap at the 5′end of eukaryotic mRNAs (Sachs et al., 1997). Several attempts have failed to detect caps at the 5′-ends of *G. intestinalis* mRNAs (Yu and Wang, 1996; Yu et al., 1998), suggesting that translation initiation mechanisms in *Giardia* may be similar to those found in eubacteria. Initiation in eubacteria requires the presence of the Shine–Dalgarno sequence in the mRNA; a six nucleotide purine-rich sequence in front of the coding region that is complementary to the 3′-end of 16S rRNA (Gold, 1988). Interestingly, there is an anti-Shine–Dalgarno-like sequence found at the 3′-end of 16S-like rRNA in *G. intestinalis* (Sogin et al., 1989); however, there is no functional Shine–Dalgarno present in the 5′ untranslated regions of giardial mRNAs (Adam, 1991; Yu et al., 1998). Another translation initiation signal in the mRNA of bacteria is a 13 nucleotide downstream box located 3′ of the initiation codon, which is also complementary to a region in the 16S rRNA (Sprengart et al., 1990). It promotes efficient protein synthesis even in the absence of Shine–Dalgarno (Sprengart et al., 1996). The downstream box has been identified in the

transcripts of the *Giardia* viral dsRNA (Yu et al., 1998). It complements a 15 nucleotide sequence near the 3′-end of the *Giardia* 16S-like rRNA (anti-downstream box). In transfection assays, this downstream box element significantly increased translation efficiency of viral mRNAs (Yu et al., 1998); however, whether this element is required for translation is not clear. Sequence comparison of all *G. intestinalis* mRNAs available in the Genbank in 1998 revealed that they contain a putative eight to 13 nucleotide downstream box homologue within the coding region (Yu et al., 1998). However, these homologs have not been shown to perform the downstream box function. The presence of conserved elements, such as caps and a downstream box, have not been rigorously addressed in *Entamoeba* and *Trichomonas*, leaving the mechanisms underlying translation in these organisms virtually unknown.

The 3′ untranslated regions of eukaryotic mRNAs provide a site for polyadenylation, a feature that is thought to play a role in mRNA stability and export from the nucleus. As mentioned above, the 3′ untranslated regions of amitochondriate mRNAs are polyadenylated and resemble the 5′ untranslated regions in length as they are unusually short, ranging from five to 43 nucleotides (Kirk-Mason et al., 1989; Bruchhaus et al., 1993; Liston and Johnson, 1998). Exceptions have been noted, such as the elongation factor-1 alpha mRNA (180 nucleotide 3′ untranslated regions) (De Meester et al., 1991) and a putative serine/threonine protein kinase (69 nucleotide 3′ untranslated regions) (Urban et al., 1996) in *E. histolytica* and two transcripts in *Giardia* encoding the encystation proteins ENC1 and ENC6, that possess >300 nucleotide 3′ untranslated regions and at least two distinct sites of polyadenylation (Que et al., 1996). All three organisms appear to use conserved signals to mark termination and polyadenylation of their mRNAs. Interestingly, only trichomonads use the characteristic eukaryotic polyadenylation signal, 'AAUAAA' (Lahti et al., 1992; Bagnara et al., 1996; Granger et al., 1997; Bricheux et al., 1998; Minotto et al., 2000). On the other hand, *Entamoeba* 3′ untranslated regions usually contain the conserved pentanucleotide transcription termination motif UA(A/U)UU and an A/U rich region close to the polyadenosine track (Bruchhaus et al., 1993; Purdy et al., 1996; Ostoa-Saloma et al., 2000). *Giardia* mRNAs have a consensus polyadenylation signal AGUPuAAPy (Peattie et al., 1989; Adam, 1991; Que et al., 1996). It should be noted that the function of these elements awaits verification by mutational analysis and are currently inferred by sequence conservation.

7. Translational machinery

Eukaryotic ribosomes typically have a sedimentation coefficient of 80S, comprised of a LSU with a coefficient of 60S and a SSU of 40S. Eubacterial ribosomes are smaller with a sedimentation coefficient of 70S (50S and 30S of LSU and SSU, respectively). In most eukaryotes, the LSU contains 28S, 5.8S, and 5S ribosomal rRNAs and about 50 proteins and the SSU is composed of the 18S rRNA and about 30 proteins (Wool et al., 1995; Mager et al., 1997). The genes that give rise to rRNAs are typically organised in tandem arrays that encode 18S, 5.8S and 28S rRNAs, with 100–5,000 copies per genome (Enright et al., 1996). Interestingly, the ribosomes of *Giardia* and *Trichomonas* exhibit certain features that are reminiscent of eubacteral ribosomes. Their ribosomes are similar in size to their eubacterial counterparts, with a sedimentation coefficient of 70S (Champney et al., 1992; Cavalier-Smith, 1993). Likewise, the large, small, and 5.8S rRNAs are smaller than in other eukaryotes (Sogin et al., 1989; Healey et al., 1990; Chakrabarti et al., 1992), and more closely aligned in size with eubacterial rRNAs. Despite these similiarities, a basic local alignment search tool (BLAST) analysis of more than 50 ribosomal protein sequences found by the *G. intestinalis* genome project revealed that genes encoding those proteins have a high homology to their partners from higher eukaryotes. *Giardia intestinalis* ribosomes contain at least 74 proteins detected by two-dimensional polyacrylamide gel electrophoresis (Shirakura et al., 2001). In agreement with the genome analysis, partial sequence analysis of 15 of these proteins and whole sequence analysis of the L29/L35 ribosomal protein (Wu and Hashimoto, 1999) demonstrated that the majority of them are encoded by genes that show a high sequence homology to their eukaryotic ribosomal counterparts (Eichler and Craig, 1995; Narcisi et al., 1998; Ghosh et al., 2001). Less is known about the composition of the 70S ribosomes described for trichomonads. Limited data indicate that these ribosomes contain fewer proteins (40–56) (Champney et al., 1992); however, more detailed analyses are needed.

Most of the data on the translational machinery of *Entamoeba* originate from microscopic observations. Ribosomes in *Entamoeba* are organised in helical arrays, which aggregate in large crystalline inclusions that constitute the chromatoid bodies seen under the light microscope in cysts. Unlike those described for *Giardia* and *Trichomonas*, *E. histolytica* ribosome sedimentation coefficients are similar to those of other eukaryotes, 80S for the whole ribosome, 60S for the LSU, and 40S for the SSU (Barker and Swales, 1972). As mentioned previously, the organisation of *Entamoeba* rDNA units is intriguing as these genes are located on extrachromosomal plasmids with a copy number of approximately 200 per cell (Bhattacharya et al., 1989). No rDNA genes have been identified in *Entamoeba* chromosomes. In contrast, *Giardia* and *Trichomonas* rRNAs genes are chromosomal, with a unit repeat of 5.5 kb in *Giardia* (Healey et al., 1990), and 5.8 kb in trichomonads (Chakrabarti et al., 1992). Copy number analysis of the rDNA unit of *T. foetus* has indicated only 12 copies (Chakrabarti et al., 1992), while in *Giardia* a presence of approximately 300 copies have been estimated (Edlind and Chakraborty, 1987).

8. Codon usage

In most organisms, codon usage is biased and the bias is often species-specific. Codons that are most highly represented are thought to be those preferentially recognised by the most efficient and/or abundant tRNAs, resulting in preferences that correlate with the level of gene expression. Indeed, it has been shown that highly expressed genes have high frequencies of these optimal codons (Bulmer, 1990). Analysis of 50,651 codons used in *E. histolytica* genes has shown a strong bias towards the usage of A or T in the third position (Char and Farthing, 1992). Surprisingly, additional analyses indicates that these codons are predominant in poorly expressed genes (Ghosh et al., 2000b); whereas highly expressed genes more often contain a C in the third codon position (Ghosh et al., 2000b; Romero et al., 2000). Whether genes containing the latter codon bias are optimally translated in this organism remains to be tested.

The analysis of codon usage of 65 *G. intestinalis* genes has revealed heterogeneity in the usage of synonymous codons. Computational comparison indicates that highly-expressed proteins use almost exclusively a particular subset of 21 codons, while the proteins expressed at low levels use a random selection of codons (Lafay and Sharp, 1999). All 61 sense codons are frequently used, but the overall codon usage is biased toward C or G ending codons belonging to the subset of 21 codons 'optimal' for translation (Lafay and Sharp, 1999). All three alternative termination codons are used in similar frequencies with a slight bias for TGA (Lafay and Sharp, 1999).

The analysis of the codon usage pattern in 29 genes of *T. vaginalis* found a strong biased toward pyrimidine bases (63.1%) in the third position (Meade et al., 1997). In addition, there is a purine bias in codon position 1 (63.9%), and an A/T bias in codon position 2 (60.8%). Interestingly, individual amino acids have a pronounced bias in codon usage, often preferring only a single codon. *Trichomonas* genes also show a strong preference in choice of termination codons, with 27 out of 29 utilising TAA as the stop codon (Meade et al., 1997). Analysis of these sequences also showed that 5′ and 3′ untranslated regions have a strong A/T bias (over 70%) which is not present in protein coding regions (Meade et al., 1997, Espinosa et al., 2001).

9. Prospective

Much is yet to be learned about gene expression and the molecular biology of anaerobic protist parasites. Studies thus far indicate that *Giardia* and *Trichomonas* possess several features that are reminiscent of eubacteria, such as 70S ribosomes, smaller rRNAs, lack of caps on giardial mRNAs and apparent lack of introns in trichomonad genes. *Entamoeba* tends to share less features with eubacteria, which is not surprising as it branches higher in eukaryotic trees, more distal to eubacterial and archaeal cousins. The basic structure of promoters of protein-coding genes and the RNA polymerase II which appears to transcribe these genes in all three organims are essentially eukaryotic in nature; however, each has evolved its own specific structural features which merit further study. Many questions remain to be answered. What is the nature of the polymerase complex that assembles on initiator-like elements in the promoters of these genes? Are the proteins within this complex evolutionarily conserved or divergent between organisms? What is the relationship between this complex in amitochondriate protists and the well-characterised transcriptional complexes in model eukaryotes and archaea? What is the functional significance of the extremely short 5′ and 3′ untranslated regions that are almost universally conserved on amitochondriate mRNAs? Do these unusual features primarily arise as a result of faster rates of evolution (Germot and Philippe, 1999; Stiller and Hall, 1999; Philippe et al., 2000b) or early-divergence of these lineages, a question that is highly debated, at present (Embley and Hirt, 1998; Keeling and Palmer, 2000; Philippe et al., 2000a). To what extent do unusual features reflect adaptations to a parasitic lifestyle? The nearly completed genome project for *G. intestinalis* and the ones that are underway for *E. histolytica* and planned for *T. vaginalis* will help begin to answer these questions. These data will also undoubtedly bring forward more intriguing questions. A combination of proteomics and in vitro and in vivo functional assays designed to directly test hypotheses formulated from analyses of genome sequences promises to offer exciting insights into the biology of amitochondriates in the years to come.

Acknowledgements

The authors thank members of the lab for helpful discussions and critical comments on the manuscript and acknowledge the support of the National Institutes of Health, the Burroughs–Wellcome Fund and the Grant Agency of the Czech Republic.

References

Abonyi, A., 1995. Examination of nonflagellate and flagellate round forms of *Trichomonas vaginalis* by transmission electron microscopy. Appl. Parasitol. 36, 303–310.

Adam, R.D., Nash, T.E., Wellems, T.E., 1988. The *Giardia lamblia* trophozoite contains sets of closely related chromosomes. Nucleic Acids Res. 16, 4555–4567.

Adam, R.D., 1991. The biology of *Giardia* spp. Microbiol. Rev. 55, 706–732.

Adam, R.D., Nash, T.E., Wellems, T.E., 1991. Telomeric location of *Giardia* rDNA genes. Mol. Cell Biol. 11, 3326–3330.

Adam, R.D., 2000. The *Giardia lamblia* genome. Int. J. Parasitol. 30, 475–484.

Adam, R.D., 2001. Biology of *Giardia lamblia*. Clin. Microbiol. Rev. 14, 447–475.

Alderete, J.F., Engbring, J., Lauriano, C.M., O'Brien, J.L., 1998. Only two

of the *Trichomonas vaginalis*triplet AP51 adhesines are regulated by iron. Microb. Pathog. 24, 1–16.

Ankri, S., Stolarsky, T., Mirelman, D., 1998. Antisence inhibition of expression of cysteine proteinases does not affect *Entamoeba histolytica* cytopathic or haemolytic activity but inhibits phagocytosis. Mol. Microbiol. 28, 777–785.

Ankri, S., Stolarsky, T., Bracha, R., Padilla-Vaca, F., Mirelman, D., 1999a. Antisence inhibition of expression of cysteine proteinases affects *Entamoeba histolytica*-induced formation of liver abscess in hamsters. Infect. Immun. 67, 421–422.

Ankri, S., Padilla-Vaca, F., Stolarsky, T., Koole, L., Katz, U., Mirelman, D., 1999b. Antisence inhibition of expression of the light subunit (35 kDa) of the Gal/GalNac lectin inhibits *Entamoeba histolytica* virulence. Mol. Microbiol. 33, 327–337.

Bagnara, A.S., Tucker, V.E., Minotto, L., Howes, E.R., Ko, G.A., Edwards, M.R., Dawes, I.W., 1996. Molecular characterisation of adenosylhomocysteinase from *Trichomonas vaginalis*. Mol. Biochem. Parasitol. 81, 1–11.

Barker, D.C., Swales, L.S., 1972. Characteristics of ribosomes during differentiation from trophozoite to cyst in axenic *Entamoeba* sp. Cell Differ. 1, 297–306.

Benne, R., Van den Burg, J., Brakenhoff, J.P., Sloof, P., Van Boom, J.H., Tromp, M.C., 1986. Major transcript of the frameshifted coxII gene from trypanosome mitochondria contains four nucleotides that are not encoded in the DNA. Cell 46, 819–826.

Benne, R., 1994. RNA editing in trypanosomes. Eur. J. Biochem. 221, 9–23.

Bernander, R., Palm, J.E.D., Svärd, S.G., 2001. Genome ploidy in different stages of the *Giardia lamblia* life cycle. Cell Microbiol. 3, 55–62.

Bhattacharya, S., Bhattacharya, A., Diamond, L.S., Soldo, A.T., 1989. Circular DNA of *Entamoeba histolytica* encodes ribosomal RNA. J. Protozool. 36, 455–458.

Binder, M., Ortner, S., Plaimauer, B., Födinger, M., Wiedermann, G., Scheiner, O., Duchene, M., 1995. Sequence and organization of an unusual histone H4 gene in the human parasite *Entamoeba histolytica*. Mol. Biochem. Parasitol. 71, 243–247.

Boothroyd, J.C., Cross, G.A.M., 1982. Transcripts coding for variant surface glycoproteins of *Trypanosoma brucei* have a short, identical exon at their 5′ end. Gene 20, 281–289.

Boothroyd, J.C., Wang, A., Campbell, D.A., Wang, C.C., 1987. An unusually compact ribosomal DNA repeat in the protozoan *Giardia lamblia*. Nucleic Acids Res. 15, 4065–4084.

Bracha, R., Nuchamowitz, Y., Leippe, M., Mirelman, D., 1999. Antisence inhibition of amoebapore expression in *Entamoeba*. Mol. Microbiol. 34, 463–472.

Bracha, R., Nuchamowitz, Y., Mirelman, D., 2000. Inhibition of gene expression in *Entamoeba* by the transcription of antisence RNA: effect of 5′ and 3′ regulatory elements. Mol. Biochem. Parasitol. 107, 81–90.

Bricheux, G., Coffe, G., Pradel, N., Brugerolle, G., 1998. Evidence for an uncommon a-actinin protein in *Trichomonas vaginalis*. Mol. Biochem. Parasitol. 95, 241–249.

Bruchhaus, I., Leippe, M., Lioutas, C., Tannich, E., 1993. Unusual gene organization in the protozoan parasite *Entamoeba histolytica*. DNA Cell Biol. 12, 925–933.

Brugerolle, G., 1975. Etude de la cryptopleuromitose et de la morphogénése de division chez *Trichomonas vaginalis* et chez plusieurs genres de Trichomonadines primitives. Protistologica 11, 457–468.

Bulmer, M., 1990. The effect of context on synonymous codon usage in genes with low codon usage bias. Nucleic Acids Res. 18, 2869–2873.

Burch, D.J., Li, E., Reed, S., Jackson, T.F., Stanley, S.L.J., 1991. Isolation of a strain specific *Entamoeba histolytica* cDNA clone. J. Clin. Microbiol. 29, 696–701.

Burge, C.B., Padgett, R.A., Sharp, P.A., 1998. Evolutionary fates and origins of U12-type introns. Mol. Cell 2, 773–785.

Burke, T.W., Willy, P.J., Kutach, A.K., Butler, J.E., Kadonaga, J.T., 1998. The DPE, a conserved downstream core promoter element that is functionally analogous to the TATA box. Cold Spring Harb. Symp. Quant. Biol. 63, 75–82.

Buß, H., Lioutas, C., Dobinsky, S., Nickel, R., Tannich, E., 1995. Analysis of the 170-kDa lectin gene promoter of *Entamoeba histolytica*. Mol. Biochem. Parasitol. 72, 1–10.

Byers, T.J., 1986. Molecular biology of DNA in *Acanthamoeba*, *Amoeba*, *Entamoeba*, and *Naegleria*. Int. Rev. Cytol. 99, 311–341.

Cavalier-Smith, T., 1993. Kingdom protozoa and its 18 phyla. Microbiol. Rev. 57, 953–994.

Chakrabarti, D., Dame, J.B., Gutell, R.R., Yowell, C.A., 1992. Characterization of the rDNA unit and sequence analysis of the small subunit rRNA and 5.8S rRNA genes from *Tritrichomonas foetus*. Mol. Biochem. Parasitol. 52, 75–83.

Chalkley, G.E., Verrijzer, C.P., 1999. DNA binding site selection by RNA polymerase II TAFs: a TAF(II)250-TAF(II)150 complex recognizes the initiator. EMBO J. 18, 4835–4845.

Champney, W.S., Chittum, H.S., Samuels, R., 1992. Ribosomes from trichomonad protozoa have prokaryotic characteristics. Int. J. Biochem. 24, 1125–1133.

Char, S., Farthing, M.J., 1992. Codon usage in *Entamoeba histolytica*. Int. J. Parasitol. 22, 381–383.

Clemens, D.L., Johnson, P.J., 2000. Failure to detect DNA in hydrogenosomes of *Trichomonas vaginalis* by nick translation and immunomicroscopy. Mol. Biochem. Parasitol. 106, 307–313.

Dahmus, M.E., 1995. Phosphorylation of the C-terminal domain of RNA polymerase II. Biochim. Biophys. Acta 1261, 171–182.

Dan, M., Wang, C.C., 2000. Role of alcohol dehydrogenase E (ADHE) in the energy metabolism of *Giardia lamblia*. Mol. Biochem. Parasitol. 109, 25–36.

Dan, M., Wang, A.L., Wang, C.C., 2000. Inhibition of pyruvate-ferredoxin oxidoreductase gene expression in *Giardia lamblia* by a virus-mediated hammerhead ribozyme. Mol. Microbiol. 36, 447–456.

De Meester, F., Bracha, R., Huber, M., Keren, Z., Rozenblatt, S., Mirelman, D., 1991. Cloning and characterization of an unusual elongation factor-1 alpha cDNA from *Entamoeba histolytica*. Mol. Biochem. Parasitol. 44, 23–32.

Delgadillo, M.G., Liston, D.R., Niazi, K., Johnson, P.J., 1997. Transient and selectable transformation of the parasitic protist *Trichomonas vaginalis*. Proc. Natl Acad. Sci. USA 94, 4716–4720.

Dhar, S.K., Choudhury, N.R., Bhattacharaya, A., Bhattacharya, S., 1995. A multitude of circular DNAs exist in the nucleus of *Entamoeba histolytica*. Mol. Biochem. Parasitol. 70, 203–206.

Dhar, S.K., Roychoudhury, N., Mittal, V., Bhattacharya, A., Bhattacharya, S., 1996. Replication initiates at multiple dispersed sites in the ribosomal DNA plasmid of the protozoan parasite *Entamoeba histolytica*. Mol. Cell Biol. 16, 2314–2324.

Dhar, S.K., Vines, R.R., Bhattacharya, S., Petri, W.A. Jr, 1998. Ribosomal DNA fragments enhance the stability of transfected DNA in *Entamoeba histolytica*. J. Eukaryot. Microbiol. 45, 656–660.

Drmota, T., Král, J., 1997. Karyotype of *Trichomonas vaginalis*. Eur. J. Protistol. 33, 131–135.

Dvorak, J.A., Kobayashi, S., Alling, D.W., Hallahan, C.W., 1995. Elucidation of the DNA synthetic cycle of *Entamoeba* spp using flow cytometry and mathematical modeling. J. Eukaryot. Microbiol. 42, 610–616.

Dyall, S.D., Johnson, P.J., 2000. Origins of hydrogenosomes and mitochondria: evolution and organelle biogenesis. Curr. Opin. Microbiol. 3, 404–411.

Edlind, T.D., Chakraborty, P.R., 1987. Unusual ribosomal RNA of the intestinal parasite *Giardia lamblia*. Nucleic Acids Res. 15, 7889–7901.

Eichler, D.C., Craig, N., 1995. Processing of eukaryotic ribosomal RNA. Prog. Nucleic Acid Res. Mol. Biol. 49, 179–239.

Elmendorf, H.G., Singer, S.M., Nash, T.E., 2001a. The abundance of sterile transcripts in *Giardia lamblia*. Nucleic Acids Res. 29, 4674–4683.

Elmendorf, H.G., Singer, S.M., Pierce, J., Cowan, J., Nash, T.E., 2001b. Initiator and upstream elements in the alpha2-tubulin promoter of *Giardia lamblia*. Mol. Biochem. Parasitol. 113, 157–169.

Embley, T.M., Hirt, R.P., 1998. Early branching eukaryotes? Curr Opin. Genet. Dev. 8, 624–629.

Enright, C.A., Maxwell, E.S., Eliceiri, G.L., Sollner-Webb, B., 1996. 5′ETS rRNA processing facilitated by four small RNAs: U14, E3, U17, and U3. RNA 2, 1318.

Erlandsen, S.L., Rasch, E.M., 1994. The DNA content of trophozoites and cysts of *Giardia lamblia* by microdensitometric quantitation of feulgen staining and examination by laser scanning confocal microscopy. J. Histochem. Cytochem. 42, 1413–1416.

Espinosa, N., Hernández, R., López-Griego, L., Arroyo, R., López-Villasenor, I., 2001. Differences between coding and non-coding regions in the *Trichomonas vaginalis* genome: an actin gene as a locus model. Acta Trop. 78, 147–154.

Fast, N.M., Doolittle, W.F., 1999. *Trichomonas vaginalis* possesses a gene encoding the essential spliceosomal component, PRP8. Mol. Biochem. Parasitol. 99, 275–278.

Flegr, J., Cerkasov, J., Kulda, J., Tachezy, J., Stokrova, J., 1987. The dsRNA of *Trichomonas vaginalis* is associated with virus-like particles and does not correlate with metronidazole resistance. Folia Microbiol. (Praha) 32, 345–348.

Flegr, J., Cerkasov, J., Stokrova, J., 1988. Multiple populations of double-stranded RNA in two virus- harbouring strains of *Trichomonas vaginalis*. Folia Microbiol. (Praha) 33, 462–465.

Födinger, M., Ortner, S., Plaimauer, B., Wiedermann, G., Scheiner, O., Duchene, M., 1993. Pathogenic *Entamoeba histolytica*: cDNA cloning of a histone H3 with a divergent primary structure. Mol. Biochem. Parasitol. 59, 315–322.

Gangopadhyay, S.S., Ray, S.S., Sinha, P., Lohia, A., 1997a. Unusual genome organisation in *Entamoeba histolytica* leads to two overlapping transcripts. Mol. Biochem. Parasitol. 89, 73–83.

Gangopadhyay, S.S., Ray, S.S., Kennady, K., Pande, G., Lohia, A., 1997b. Heterogeneity of DNA content in axenically growing *Entamoeba histolytica* HM1:IMSS clone A. Mol. Biochem. Parasitol. 90, 9–20.

Germot, A., Philippe, H., 1999. Critical analysis of eukaryotic phylogeny: a case study based on the HSP70 family. J. Eukaryot. Microbiol. 46, 116–124.

Ghosh, S., Field, J., Rogers, R., Hickman, M., Samuelson, J., 2000a. The *Entamoeba histolytica* mitochondrion-derived organelle (crypton) contains double-stranded DNA and appears to be bound by a double membrane. Infect. Immun. 68, 4319–4322.

Ghosh, T.C., Gupta, S.K., Majumdar, S., 2000b. Studies on codon usage in *Entamoeba histolytica*. Int. J. Parasitol. 30, 715–722.

Ghosh, S., Ghosh, R., Das, P., Chattopadhyay, D., 2001. Expression and purification of recombinant *Giardia* fibrillarin and its interaction with small nuclear RNAs. Protein Expr Purif 21, 40–48.

Gilchrist, C.A., Streets, H.L., Ackers, J.P., Hall, R., 1995. Transient expression of luciferase in *Entamoeba histolytica* driven by the ferredoxin gene 5′ and 3′ regions. Mol. Biochem. Parasitol. 74, 1–10.

Gilchrist, C.A., Mann, B.J., Petri, W.A. Jr, 1998. Control of ferredoxin and Gal/GalNAc lectin gene expression in *Entamoeba histolytica* by a *cis*-acting DNA sequence. Infect. Immun. 66, 2383–2386.

Gilchrist, C.A., Holm, C.F., Hughes, M.A., Schaenman, J.M., Mann, B.J., Petri, W.A. Jr, 2001. Identification and characterization of an *Entamoeba histolytica* upstream regulatory element 3 sequence-specific DNA-binding protein containing EF-hand motifs. J. Biol. Chem. 276, 11838–11843.

Gold, L., 1988. Posttranscriptional regulatory mechanisms in *Escherichia coli*. Annu. Rev. Biochem. 57, 199–233.

Gomez, C., Perez, D.G., Lopez-Bayghen, E., Orozco, E., 1998. Transcriptional analysis of the EhPgp1 promoter of *Entamoeba histolytica* multidrug-resistant mutant. J. Biol. Chem. 273, 7277–7284.

Gómez-Conde, E., Mena-López, R., Hernández-Jaúregui, P., González-Camacho, M., Arroyo, R., 2000. *Trichomonas vaginalis*: chromatin and mitotic spindle during mitosis. Exp. Parasitol. 96, 130–138.

Granger, B.L., Warwood, S.J., Hayai, N., Hayashi, H., Owhashi, M., 1997. Identification of a neutrophil chemotactic factor from *Tritrichomonas foetus* as superoxid dismutase. Mol. Biochem. Parasitol. 89, 85–95.

Hajduk, S.L., Harris, M.E., Pollard, V.W., 1992. RNA editing in kinetoplast mitochondria. FASEB J. 7, 54–63.

Hamann, L., Nickel, R., Tannich, E., 1995. Transfection and continuous expression of heterologous genes in the protozoan parasite *Entamoeba histolytica*. Proc. Natl Acad. Sci. USA 92, 8975–8979.

Hamann, L., Buß, H., Tannich, E., 1997. Tetracycline-controlled gene expression in *Entamoeba histolytica*. Mol. Biochem. Parasitol. 84, 83–91.

Hashimoto, T., Nakamura, Y., Nakamura, F., Shirakura, T., Adachi, J., Goto, N., Okamoto, K., Hasegawa, M., 1994. Protein phylogeny gives a robust estimation for early divergences of eukaryotes: phylogenetic place of a mitochondria-lacking protozoan, *Giardia lamblia*. Mol. Biol. Evol. 11, 65–71.

Hashimoto, T., Hasegawa, M., 1996. Origin and early evolution of eukaryotes inferred from the amino acid sequences of translation elongation factors 1a/Tu and 2/G. Adv. Biophys. 32, 73–120.

Healey, A., Mitchell, R., Upcroft, J.A., Boreham, P.F.L., Upcroft, P., 1990. Complete nucleotide sequence of the ribosomal RNA tandem repeat unit from *Giardia intestinalis*. Nucleic Acids Res. 18, 4006.

Holberton, D.V., Marshall, J., 1995. Analysis of consensus sequence patterns in *Giardia* cytoskeleton gene promoters. Nucleic Acids Res. 23, 2945–2953.

Hou, G., Le Blancq, S.M., E, Y., Zhu, H., Lee, M.G.-S., 1995. Structure of a frequently rearranged rRNA-encoding chromosome in *Giardia lamblia*. Nucleic Acids Res. 23, 3310–3317.

Hoyne, G.F., Boreham, P.F., Parsons, P.G., Ward, C., Biggs, B., 1989. The effect of drugs on the cell cycle of *Giardia intestinalis*. Parasitology 99, 333–339.

Hrdy, I., Müller, M., 1995. Primary structure of the hydrogenosomal malic enzyme of *Trichomonas vaginalis* and its relationship to homologous enzymes. J. Eukaryot. Microbiol. 42, 593–603.

Jansson, A., Gillin, F., Kagardt, U., Hagblom, P., 1994. Coding of hemolysins within the ribosomal RNA repeat on a plasmid in *Entamoeba histolytica*. Science 263, 1440–1443.

Javahery, R., Khachi, A., Lo, K., Zenzie-Gregory, B., Smale, S.T., 1994. DNA sequence requirements for transcriptional initiator activity in mammalian cells. Mol Cell Biol. 14, 116–127.

Johnson, P.J., d'Oliveira, C.E., Gorrell, T.E., Müller, M., 1990. Molecular analysis of the hydrogenosomal ferredoxin of the anaerobic protist *Trichomonas vaginalis*. Proc. Natl Acad. Sci. USA 87, 6097–6101.

Johnson, P.J., 2002. Spliceosomal introns in a deep-branching eukaryote: the splice of life. Proc. Natl Acad. Sci. USA 99, 3359–3361.

Kabnick, K.S., Peattie, D.A., 1990. In situ analyses reveal that the two nuclei of *Giardia lamblia* are equivalent. J. Cell Sci. 95, 353–360.

Kaufmann, J., Verrijzer, C.P., Shao, J., Smale, S.T., 1996. CIF, an essential cofactor for TFIID-dependent initiator function. Genes Dev. 10, 873–886.

Kaufmann, J., Ahrens, K., Koop, R., Smale, S.T., Muller, R., 1998. CIF150, a human cofactor for transcription factor IID-dependent initiator function. Mol. Cell Biol. 18, 233–239.

Keeling, P.J., Palmer, J.D., 2000. Parabasalian flagellates are ancient eukaryotes. Nature 405, 635–637.

Khoshnan, A., Provenzano, D., Alderete, J.F., 1994. Unique double-stranded RNAs associated with the *Trichomonas vaginalis* virus are synthesized by viral RNA-dependent RNA polymerase. J. Virol. 68, 7108–7114.

Khoshnan, A., Alderete, J.F., 1995. Characterization of double-stranded RNA satellites associated with the *Trichomonas vaginalis* virus. J. Virol. 69, 6892–6897.

Kirk-Mason, K.E., Turner, M.J., Chakraborty, P.R., 1989. Evidence for unusually short tubulin mRNA leaders and characterization of tubulin genes in *Giardia lamblia*. Mol. Biochem. Parasitol. 36, 87–99.

Knodler, L.A., Svärd, S.G., Silberman, J.D., Davids, B.J., Gillin, F.D., 1999. Developmental gene regulation in *Giardia lamblia*: first evidence for an encystation-specific promoter and differential 5′ mRNA processing. Mol. Microbiol. 34, 327–340.

Kooter, J.M., Borst, P., 1984. Alpha-amanitin-insensitive transcription of

variant surface glycoprotein genes provides further evidence for discontinuous transcription in trypanosomes. Nucleic Acids Res. 12, 9457–9472.

Korman, S.H., Le Blancq, S.M., Deckelbaum, R.J., Van der Ploeg, L.H.T., 1992. Investigation of human giardiasis by karyotype analysis. J. Clin. Invest. 89, 1725–1733.

Lafay, B., Sharp, P.M., 1999. Synonymous codon usage variation among *Giardia lamblia* and isolates. Mol. Biol. Evol. 16, 1484–1495.

Lahti, C.J., d'Oliveira, C.E., Johnson, P.J., 1992. β-succinyl-coenzyme A synthetase from *Trichomonas vaginalis* is a soluble hydrogenosomal protein with an amino-terminal sequence that resembles mitochondrial presequences. J. Bacteriol. 174, 6822–6830.

Lahti, C.J., Bradley, P.J., Johnson, P.J., 1994. Molecular characterization of the alpha-subunit of *Trichomonas vaginalis* hydrogenosomal succinyl CoA synthetase. Mol. Biochem. Parasitol. 66, 309–318.

Lang, B.F., Gray, M.W., Burger, G., 1999. Mitochondrial genome evolution and the origin of eukaryotes. Annu. Rev. Genet. 33, 351–397.

Le Blancq, S.M., Korman, S.H., Van der Ploeg, L.H.T., 1991a. Frequent rearrangements of rRNA-encoding chromosomes in *Giardia lamblia*. Nucleic Acids Res. 19, 4405–4412.

Le Blancq, S.M., Kase, R.S., Van der Ploeg, L.H., 1991b. Analysis of a *Giardia lamblia* rRNA encoding telomere with [TAGGG]n as the telomere repeat. Nucleic Acids Res. 19, 5790.

Le Blancq, S.M., Adam, R.D., 1998. Structural basis of karyotype heterogeneity in *Giardia lamblia*. Mol. Biochem. Parasitol. 97, 199–208.

Lee, M.G., Van der Ploeg, L.H., 1997. Transcription of protein-coding genes in trypanosomes by RNA polymerase I. Annu. Rev. Microbiol. 51, 463–489.

Lehker, M.W., Alderete, J.F., 1999. Resolution of six chromosomes of *Trichomonas vaginalis* and conservation of size and number among isolates. J. Parasitol. 85, 976–979.

Lemon, B., Tjian, R., 2000. Orchestrated response: a symphony of transcription factors for gene control. Genes Dev. 14, 2551–2569.

Lewit-Bentley, A., Rety, S., 2000. EF-hand calcium-binding proteins. Curr. Opin. Struct. Biol. 10, 637–643.

Lindell, T.J., Weinberg, F., Morris, P.W., Roeder, R.G., Rutter, W.J., 1970. Specific inhibition of nuclear RNA polymerase II by alpha-amanitin. Science 170, 447–449.

Lioutas, C., Tannich, E., 1995. Transcription of protein-coding genes in *Entamoeba histolytica* is insensitive to high concentrations of alpha-amanitin. Mol. Biochem. Parasitol. 73, 259–261.

Lioutas, C., Schmetz, C., Tannich, E., 1995. Identification of various circular DNA molecules in *Entamoeba histolytica*. Exp. Parasitol. 80, 349–352.

Liston, D.R., Johnson, P.J., 1998. Gene transcription in *Trichomonas vaginalis*. Parasitol. Today 14, 261–265.

Liston, D.R., Johnson, P.J., 1999. Analysis of a ubiquitous promoter element in a primitive eukaryote: early evolution of the initiator element. Mol. Cell Biol. 19, 2380–2388.

Liston, D.R., Carrero, J.C., Johnson, P.J., 1999. Upstream regulatory sequences required for expression of the *Trichomonas vaginalis* alpha-succinyl CoA synthetase gene. Mol. Biochem. Parasitol. 104, 323–329.

Liston, D.R., Lau, A.O., Ortiz, D., Smale, S.T., Johnson, P.J., 2001. Initiator recognition in a primitive eukaryote: IBP39, an initiator- binding protein from *Trichomonas vaginalis*. Mol. Cell Biol. 21, 7872–7882.

Lohia, A., Samuelson, J., 1993. Cloning of the Eh cdc2 gene from *Entamoeba histolytica* encoding a protein kinase p34cdc2 homologue. Gene 127, 203–207.

Luna-Arias, J.P., Hernandez-Rivas, R., de Dios-Bravo, G., Garcia, J., Mendoza, L., Orozco, E., 1999. The TATA-box binding protein of *Entamoeba histolytica*: cloning of the gene and location of the protein by immunofluorescence and confocal microscopy. Microbiology 145, 33–40.

Mager, W.H., Planta, R.J., Ballesta, J.G., Lee, J.C., Mizuta, K., Suzuki, K., Warner, J.R., Woolford, J., 1997. A new nomenclature for the cytoplasmic ribosomal proteins of *Saccharomyces cerevisiae*. Nucleic Acids Res. 25, 4872–4875.

Mair, G., Shi, H., Li, H., Djikeng, A., Aviles, H.O., Bishop, J.R., Falcone, F.H., Gavrilescu, C., Montgomery, J.L., Santori, M.I., Stern, L.S., Wang, Z., Ullu, E., Tschudi, C., 2000. A new twist in trypanosome RNA metabolism: *cis*-splicing of pre-mRNA. RNA 6, 163–169.

Manning-Cela, R., Meza, I., 1997. Up-regulation of action mRNA and reorganization of the cytoskeleton in *Entamoeba histolytica* trophozoites. J. Eukaryot. Microbiol. 44, 18–24.

Marinets, A., Müller, M., Johnson, P.J., Kulda, J., Scheiner, O., Wiedermann, G., Duchene, M., 1996. The sequence and organization of the core histone H3 and H4 genes in the early branching amitochondriate protist *Trichomonas vaginalis*. J. Mol. Evol. 43, 563–571.

Martinez, E., Ge, H., Tao, Y., Yuan, C.X., Palhan, V., Roeder, R.G., 1998. Novel cofactors and TFIIA mediate functional core promoter selectivity by the human TAFII150-containing TFIID complex. Mol. Cell Biol. 18, 6571–6583.

McArthur, A.G., Morrison, H.G., Nixon, J.E., Passamaneck, N.Q., Kim, U., Hinkle, G., Crocker, M.K., Holder, M.E., Farr, R., Reich, C.I., Olsen, G.E., Aley, S.B., Adam, R.D., Gillin, F.D., Sogin, M.L., 2000. The *Giardia* genome project database. FEMS Microbiol. Lett. 189, 271–273.

McKie, A.E., Edlind, T., Walker, J., Mottram, J.C., Coombs, G.H., 1998. The primitive protozoon *Trichomonas vaginalis* contains two methionine gamma-lyase genes that encode members of the gamma-family of pyridoxal 5′-phosphate-dependent enzymes. J. Biol. Chem. 273, 5549–5556.

Meade, J.C., Shah, P.H., Lushbaugh, W.B., 1997. *Trichomonas vaginalis*: analysis of codon usage. Exp. Parasitol. 87, 73–74.

Michel, B., Lizardi, P.M., Alagon, A., Zurita, M., 1995. Identification and analysis of the start site of ribosomal RNA transcription of *Entamoeba histolytica*. Mol. Biochem. Parasitol. 73, 19–30.

Miller, S.I., Wirth, D.F., 1988. *Trans* splicing in *Leishmania enriettii* and identification of ribonucleoprotein complexes containing the spliced leader and U2 equivalent RNAs. Mol. Cell Biol. 8, 2597–2603.

Minotto, L., Edwards, M.R., Bagnara, A.S., 2000. *Trichomonas vaginalis*: characterization, expression, and phylogenetic analysis of a carbamate kinase gene sequence. Exp. Parasitol. 95, 54–62.

Mittal, V., Ramachandran, S., Sehgal, D., Bhattacharya, A., Bhattacharya, S., 1991. Sequence analysis of a DNA fragment with yeast autonomously replicating sequence activity from the extrachromosomal ribosomal DNA circle of *Entamoeba histolytica*. Nucleic Acids Res. 19, 2777.

Mittal, V., Sehgal, D., Bhattacharya, A., Bhattacharya, S., 1992. A second short repeat detected downstream of rRNA genes in the *Entamoeba histolytica* rDNA episome. Mol. Biochem. Parasitol. 54, 97–100.

Moshitch-Moshkovitch, S., Stolarsky, T., Petter, R., Mirelman, D., 1997. Upstream regions of rp L21 genes play a role in regulation of expression at the post-transcriptional level in *E. histolytica* and *E. dispar*. Arch. Med. Res. 28 Spec No, 36–38.

Moshitch-Moshkovitch, S., Petter, R., Levitan, A., Stolarsky, T., Mirelman, D., 1998. Regulation of expression of ribosomal protein L-21 genes of *Entamoeba histolytica* and *E. dispar* is at the post-transcriptional level. Mol. Microbiol. 27, 677–685.

Nagai, K., Oubridge, C., Ito, N., Avis, J., Evans, P., 1995. The RNP domain: a sequence-specific RNA-binding domain involved in processing and transport of RNA. Trends Biochem. Sci. 20, 235–240.

Narcisi, E.M., Glover, C.V.C., Fechheimer, M., 1998. Fibrillarin, a conserved pre-ribosomal RNA processing protein of *Giardia*. J. Eukaryot. Microbiol. 45, 105–111.

Nickel, R., Tannich, E., 1994. Transfectionandtransientexpressionofchloramphenicolacetyltransferasegeneintheprotozoanparasite*Entamoebahistolytica*. Proc. Natl Acad. Sci. USA 91, 7095–7098.

Nixon, J.E., Wang, A., Morrison, H.G., McArthur, A.G., Sogin, M.L., Loftus, B.J., Samuelson, J., 2002. A spliceosomal intron in *Giardia lamblia*. Proc. Natl Acad. Sci. USA 99, 3701–3705.

Ong, S.J., Huang, L.C., Liu, H.W., Chang, S.C., Yang, Y.C., Bessarab, I., Tai, J.H., 2002. Characterization of a bi-directional promoter for divergent transcription of a PHD-zinc finger protein gene and a ran gene in the protozoan pathogen *Giardia lamblia*. Mol. Microbiol. 43, 665–676.

Ortiz, D., del Carmen Dominguez-Robles, M., Villegas-Sepulveda, N., Meza, I., 2000. Actin induction during PMA and cAMP-dependent signal pathway activation in *Entamoeba histolytica* trophozoites. Cell Microbiol. 2, 391–400.

Ostoa-Saloma, P., Carrero, J.C., Petrossian, P., Hérión, P., Landa, A., Laclette, J.P., 2000. Cloning, characterization and functional expression of a cyclophilin of *Entamoeba histolytica*. Mol. Biochem. Parasitol. 107, 219–225.

Peattie, D.A., Alonso, H.J., Hein, A., Caulfield, J.P., 1989. Ultrastructural localization of giardins to the edges of disk microribbons of *Giardia lamblia* and the nucleotide and deduced protein sequence of alpha giardin. J. Cell Biol. 109, 2323–2335.

Perez, D.G., Gomez, C., Lopez-Bayghen, E., Tannich, E., Orozco, E., 1998. Transcriptional analysis of the EhPgp5 promoter of *Entamoeba histolytica* multidrug-resistant mutant. J. Biol. Chem. 273, 7285–7292.

Petter, R., Rozenblatt, S., Nuchamowitz, Y., Mirelman, D., 1992. Linkage between actin and ribosomal protein L21 genes in *Entamoeba histolytica*. Mol. Biochem. Parasitol. 56, 329–333.

Petter, R., Rozenblatt, S., Schechtman, D., Wellems, T.E., Mirelman, D., 1993. Electrophoretic karyotype and chromosome assignments for a pathogenic and a nonpathogenic strain of *Entamoeba histolytica*. Infect. Immun. 61, 3574–3577.

Philippe, H., Germot, A., Moreira, D., 2000a. The new phylogeny of eukaryotes. Curr. Opin. Genet. Dev. 10, 596–601.

Philippe, H., Lopez, P., Brinkmann, H., Budin, K., Germot, A., Laurent, J., Moreira, D., Müller, M., Le Guyader, H., 2000b. Early-branching or fast-evolving eukaryotes? An answer based on slowly evolving positions. Proc. R. Soc. Lond. B. Biol. Sci. 267, 1213–1221.

Plaimauer, B., Ortner, S., Wiedermann, G., Scheiner, O., Duchene, M., 1994. An intron-containing gene coding for a novel 39-kilodalton antigen of *Entamoeba histolytica*. Mol. Biochem. Parasitol. 66, 181–185.

Purdy, J.E., Mann, B.J., Pho, L.T., 1994. Petri W.A.Jr.,Transienttransfectionoftheentericparasite*Entamoebahistolytica*andexpressionoffireflyluciferase.Proc.Natl.Acad.Sci.USA91,, 7099–7103.

Purdy, J.E., Pho, L.T., Mann, B.J., Petri, W.A. Jr, 1996. Upstream regulatory elements controlling expression of the *Entamoeba histolytica* lectin. Mol. Biochem. Parasitol. 78, 91–103.

Que, X., Svärd, S.G., Meng, T.C., Hetsko, M.L., Aley, S.B., Gillin, F.D., 1996. Developmentally regulated transcripts and evidence of differential mRNA processing in *Giardia lamblia*. Mol. Biochem. Parasitol. 81, 101–110.

Quon, D.V., Delgadillo, M.G., Khachi, A., Smale, S.T., Johnson, P.J., 1994. Similarity between a ubiquitous promoter element in an ancient eukaryote and mammalian initiator elements. Proc. Natl Acad. Sci. USA 91, 4579–4583.

Quon, D.V.K., Delgadillo, M.G., Johnson, P.J., 1996. Transcription in the early diverging eukaryote *Trichomonas vaginalis*: an unusual RNA polymerase II and alpha-amanitin-resistant transcription of protein-coding genes. J. Mol. Evol. 43, 253–262.

Ramachandran, S., Bhattacharya, A., Bhattacharya, S., 1993. Nucleotide sequence analysis of the rRNA transcription unit of a pathogenic *Entamoeba histolytica* strain HM-1:IMSS. Nucleic Acids Res. 21, 2011.

Ramakrishnan, G., Ragland, B.D., Purdy, J.E., Mann, B.J., 1996. Physical mapping and expression of gene families encoding the N-acetyl D-galactosamine adherence lectin of *Entamoeba histolytica*. Mol. Microbiol. 19, 91–100.

Ramakrishnan, G., Vines, R.R., Mann, B.J., Petri, W.A., 1997. A tetracyclin-inducible gene expression system in *Entamoeba histolytica*. Mol. Biochem. Parasitol. 84, 93–100.

Romero, H., Zavala, A., Musto, H., 2000. Compositional pressure and translational selection determine codon usage in the extremely GC-poor unicellular eukaryote *Entamoeba histolytica*. Gene 242, 307–311.

Rudenko, G., Bishop, D., Gottesdiener, K., Van der Ploeg, L.H., 1989. Alpha-amanitinresistanttranscriptionofproteincodinggenesininsectand-bloodstreamform*Trypanosomabrucei*. EMBO J 8, 4259–4263.

Sachs, A.B., Sarnow, P., Hentze, M.W., 1997. Starting at the beginning, middle, and end: translation initiation in eukaryotes. Cell 89, 831–838.

Sanchez, L.B., Enea, V., Eichinger, D., 1994. Increased levels of polyadenylated histone H2B mRNA accumulate during *Entamoeba invadens* cyst formation. Mol. Biochem. Parasitol. 67, 137–146.

Schaenman, J.M., Driscoll, P.C., Hockensmith, J.W., Mann, B.J., Petri, W.A. Jr, 1998. An upstream regulatory element containing two nine basepair repeats regulates expression of the *Entamoeba histolytica* hgl5 lectin gene. Mol. Biochem. Parasitol. 94, 309–313.

Schaenman, J.M., Gilchrist, C.A., Mann, B.J., Petri, W.A. Jr, 2001. Identification of two *Entamoeba histolytica* sequence-specific URE4 enhancer-binding proteins with homology to the RNA-binding motif RRM. J. Biol. Chem. 276, 1602–1619.

Scharfetter, J.K., Ortner, S., Binder, M., Wiedermann, G., Scheiner, O., Duchene, M., 1997. Molecular cloning of histone H1 in *Entamoeba histolytica* and histone H2B in *histolytica* and *Entamoeba dispar*. Arch. Med. Res. 28, 14–16.

Sehgal, D., Bhattacharya, A., Bhattacharya, S., 1993. Analysis of a polymorphic locus present upstream of the rRNA transcription units in the extrachromosomal circle of *Entamoeba histolytica*. Mol. Biochem. Parasitol. 62, 129–130.

Sehgal, D., Mittal, V., Ramachandran, S., Dhar, S.K., Bhattacharya, A., Bhattacharya, S., 1994. Nucleotide sequence organisation and analysis of the nuclear ribosomal DNA circle of the protozoan parasite *Entamoeba histolytica*. Mol. Biochem. Parasitol. 67, 205–214.

Shirakura, T., Hashimoto, T., Nakamura, Y., Kamaishi, T., Cao, Y., Adachi, J., Hasegawa, M., Yamamoto, A., Goto, N., 1994. Phylogenetic place of a mitochondria-lacking protozoan, *Entamoeba histolytica*, inferred from amino acid sequences of elongation factor 2′. Jpn. J. Genet. 69, 119–135.

Shirakura, T., Maki, Y., Yoshida, H., Arisue, N., Wada, A., Sánchez, L.B., Nakamura, F., Müller, M., Hashimoto, T., 2001. Characterization of the ribosomal proteins of the amitochondriate protist, *Giardia lamblia*. Mol. Biochem. Parasitol. 112, 153–156.

Singer, S.M., Yee, J., Nash, T.E., 1998. Episomal and integrated maintenance of foreign DNA in *Giardia lamblia*. Mol. Biochem. Parasitol. 92, 59–69.

Singh, U., Rogers, J.B., Mann, B.J., Petri, W.A. Jr, 1997. Transcription initiation is controlled by three core promoter elements in the *hgl5* gene of the protozoan parasite *Entamoeba histolytica*. Proc. Natl. Acad. Sci. USA 94, 8812–8817.

Singh, U., Rogers, J.B., 1998. The novel core promoter element GAAC in the *hgl5* gene of *Entamoeba histolytica* is able to direct a transcription start site independent of TATA or initiator regions. J. Biol. Chem. 273, 21663–21668.

Singh, U., Gilchrist, C.A., Schaenman, J.M., Rogers, J.B., Hockensmith, J.W., Mann, B.J., Petri, W.A., 2002. Context-dependent roles of the *Entamoeba histolytica* core promoter element GAAC in transcriptional activation and protein complex assembly. Mol. Biochem. Parasitol. 120, 107–116.

Smale, S.T., Baltimore, D., 1989. The 'initiator' as a transcription control element. Cell 57, 103–113.

Smale, S.T., 1997. Transcription initiation from TATA-less promoters within eukaryotic protein-coding genes. Biochim. Biophys. Acta 1351, 73–88.

Smith, M.W., Aley, S.B., Sogin, M., Gillin, F.D., Evans, G.A., 1998. Sequence survey of the *Giardia lamblia* genome. Mol. Biochem. Parasitol. 95, 267–280.

Sogin, M., Gunderson, J., Elwood, J., Alonso, H.J., 1989. Phylogenetic meaning of the kingdom concept: an unusual ribosomal RNA from *Giardia lamblia*. Science 243, 75–77.

Sogin, M.L., Silberman, J.D., Hinkle, G., Morrison, H.G., 1996. Problems

with molecular diversity in the Eukarya. Symp. Soc. Gen. Microbiol. 54, 167–184.

Sprengart, M.L., Fatscher, H.P., Fuchs, E., 1990. The initiation of translation in *E. coli*: apparent base pairing between the 16S ribosomal RNA and downstream sequences of the mRNA. Nucleic Acids Res. 18, 1719–1723.

Sprengart, M.L., Fuchs, E., Porter, A.G., 1996. The downstream box: an efficient and independent translation initiation signal in *Escherichia coli*. EMBO J. 15, 665–674.

Stiller, J.W., Hall, B.D., 1999. Long-branch attraction and the rDNA model of early eukaryotic evolution. Mol. Biol. Evol. 16, 1270–1279.

Su, H.M., Tai, J.H., 1996. Genomic organization and sequence conservation in type I *Trichomonas vaginalis* viruses. Virology 222, 470–473.

Sun, C.H., Chou, C.F., Tai, J.H., 1998. Stable DNA transfection of the primitive protozoan pathogen *Giardia lamblia*. Mol. Biochem. Parasitol. 92, 123–132.

Sun, C.H., Tai, J.H., 1999. Identification and characterization of a ran gene promoter in the protozoan pathogen *Giardia lamblia*. J. Biol. Chem. 274, 19699–19706.

Sun, C.H., Tai, J.H., 2000. Development of a tetracycline controlled gene expression system in the parasitic protozoan *Giardia lamblia*. Mol. Biochem. Parasitol. 105, 51–60.

Sutton, R.E., Boothroyd, J.C., 1986. Evidence for *trans*-splicing in trypanosomes. Cell 47, 527–535.

Svärd, S.G., Rafferty, C., McCaffery, J.M., Smith, M.W., Reiner, D.S., Gillin, F.D., 1999. A signal recognition particle receptor gene from the early-diverging eukaryote, *Giardia lamblia*. Mol. Biochem. Parasitol. 98, 253–264.

Tannich, E., Horstmann, R.D., 1992. Codon usage in pathogenic *Entamoeba histolytica*. J. Mol. Evol. 34, 272–273.

Tessier, L.H., Keller, M., Chan, R.L., Fournier, R., Weil, J.H., Imbault, P., 1991. Short leader sequences may be transferred from small RNAs to pre-mature mRNAs by *trans*-splicing in *Euglena*. EMBO J. 10, 2621–2625.

Tovar, J., Fischer, A., Clark, C.G., 1999. The mitosome, a novel organelle related to mitochondria in the amitochondrial parasite *Entamoeba histolytica*. Mol. Microbiol. 32, 1013–1021.

Triana, O., Galanti, N., Olea, N., Hellman, U., Wernstedt, C., Lujan, H., Medina, C., Toro, G.C., 2001. Chromatin and histones from *Giardia lamblia*: a new puzzle in primitive eukaryotes. J. Cell Biochem. 82, 573–582.

Tsai, C.D., Liu, H.W., Tai, J.H., 2001. Characterization of an iron-responsive promoter in the protozoan pathogen *Trichomonas vaginalis*. J. Biol. Chem. 277, 5153–5162.

Turner, G., Müller, M., 1983. Failure to detect extranuclear DNA in *Trichomonas vaginalis* and *Tritrichomonas foetus*. J. Parasitol. 69, 234–236.

Upcroft, J.A., Boreham, P.F., Upcroft, P., 1989. Geographic variation in *Giardia* karyotypes. Int. J. Parasitol. 19, 519–527.

Upcroft, J.A., Healey, A., Murray, D.G., Boreham, P.F., Upcroft, P., 1992. A gene associated with cell division and drug resistance in *Giardia duodenalis*. Parasitology 104, 397–405.

Upcroft, J.A., Chen, N., Upcroft, P., 1996. Mapping variation in chromosome homologues of different *Giardia* strains. Mol. Biochem. Parasitol. 76, 135–143.

Upcroft, P., Upcroft, J.A., 1999. Organization and structure of the *Giardia* genome. Protist 150, 17–23.

Urban, B., Blasig, C., Forster, B., Hamelmann, C., Horstmann, R.D., 1996. Putative serine/threonine protein kinase expressed in complement-resistant forms of *Entamoeba histolytica*. Mol. Biochem. Parasitol. 80, 171–178.

Vanacova, S., Tachezy, J., Ullu, E., Tschudi, C., 2001a. Unusual diversity in alpha-amanitin sensitivity of RNA polymerases in trichomonads. Mol. Biochem. Parasitol. 115, 239–247.

Vanacova, S., Rasoloson, D., Rázga, J., Hrdy, I., Kulda, J., Tachezy, J., 2001b. Iron-induced chages in pyruvate metabolism of *Tritrichomonas foetus*and involvement of iron in expression of hydrogenosomal proteins. Microbiology 147, 53–62.

Verrijzer, C.P., Chen, J.L., Yokomori, K., Tjian, R., 1995. Binding of TAFs to core elements directs promoter selectivity by RNA polymerase II. Cell 81, 1115–1125.

Vines, R.R., Purdy, J.E., Ragland, B.D., Samuelson, J., Mann, B.J., Petri, W.A. Jr, 1995. Stable episomal transfection of *Entamoeba histolytica*. Mol. Biochem. Parasitol. 71, 265–267.

Wang, A.L., Wang, C.C., 1985a. A linear double-stranded RNA in *Trichomonas vaginalis*. J. Biol. Chem. 260, 3697–3702.

Wang, A.L., Wang, C.C., 1985b. Isolation and characterization of DNA from *Tritrichomonas foetus* and *Trichomonas vaginalis*. Mol. Biochem. Parasitol. 14, 323–335.

Wang, A.L., Wang, C.C., 1986. The double-stranded RNA in *Trichomonas vaginalis* may originate from virus-like particles. Proc. Natl Acad. Sci. USA. 83, 7956–7960.

Wang, A.L., Yang, H.M., Shen, K.A., Wang, C.C., 1993. Giardiavirus double-stranded RNA genome encodes a capsid polypeptide and a gag-pol-like fusion protein by a translation frameshift. Proc. Natl Acad. Sci. USA 90, 8595–8599.

White, T.C., Wang, C.C., 1990. RNA dependent RNA polymerase activity associated with the double-stranded RNA virus of *Giardia lamblia*. Nucleic Acids Res. 18, 553–559.

WHO, 1996. The World Health Report 1996, World Health Organization, Geneva.

WHO, 2001. Global Prevalence and Incidence of Selected Curable Sexually Transmitted Infections. Overview and Estimates. pp. 27–29.

WHO/PAHO/UNESCO, 1997. Report of a consultation of experts on amoebiasis. Weekly Epidemiol. Rep. WHO 72.pp. 97–99.

Willhoeft, U., Tannich, E., 1999. The electrophoretic karyotype of *Entamoeba histolytica*. Mol. Biochem. Parasitol. 99, 41–53.

Willhoeft, U., Tannich, E., 2000. Fluorescence microscopy and fluorescence in situ hybridization of *Entamoeba histolytica* nuclei to analyse mitosis and the localization of repetitive DNA. Mol. Biochem. Parasitol. 105, 291–296.

Willhoeft, U., Campos-Gongora, E., Touzni, S., Bruchhaus, I., Tannich, E., 2001. Introns of *Entamoeba histolytica* and *Entamoeba dispar*. Protist 152, 149–156.

Wool, I.G., Chan, Y.L., Gluck, A., 1995. Structure and evolution of mammalian ribosomal proteins. Biochem. Cell Biol. 73, 933–947.

Wu, G., Hashimoto, T., 1999. Sequence analysis of genes encoding ribosomal proteins of amitochondriate protists: L1 of *Trichomonas vaginalis* and L29 of *Giardia lamblia*. Parasitol. Int. 48, 135–144.

Wu, G., McArthur, A.G., Fiser, A., Sali, A., Sogin, M., Müller, M., 2000. Core histones of the amitochondriate protist, *Giardia lamblia*. Mol. Biol. Evol. 17, 1156–1163.

Yee, J., Nash, T.E., 1995. Transient transfection and expression of firefly luciferase in *Giardia lamblia*. Proc. Natl Acad. Sci. USA 92, 5615–5619.

Yee, J., Mowatt, M.R., Dennis, P.P., Nash, T.E., 2000. Transcriptional analysis of the glutamate dehydrogenase gene in the primitive eukaryote, *Giardia lamblia*. Identification of a primordial gene promoter. J. Biol. Chem. 275, 11432–11439.

Yu, D.C., Wang, A.L., Wang, C.C., 1995. Virus-mediated expression of firefly luciferase in the parasitic protozoan *Giardia lamblia*. Mol. Cell Biol. 15, 4867–4872.

Yu, D.C., Wang, C.C., 1996. Identification of *cis*-acting signals in the giardiavirus (GLV) genome required for expression of firefly luciferase in *Giardia lamblia*. RNA 2, 824–834.

Yu, D.C., Wang, A.L., Wang, C.C., 1996a. Stable coexpression of a drug-resistance gene and a heterologous gene in an ancient parasitic protozoan *Giardia lamblia*. Mol. Biochem. Parasitol. 83, 81–89.

Yu, D.C., Wang, A.L., Wang, C.C., 1996b. Amplification, expression, and packaging of a foreign gene by giardiavirus in *Giardia lamblia*. J. Virol. 70, 8752–8757.

Yu, D.C., Wang, A.L., Botka, C.W., Wang, C.C., 1998. Protein synthesis in *Giardia lamblia* may involve interaction between a downstream box

(DB) in mRNA and an anti-DB in the 16S-like ribosomal RNA. Mol. Biochem. Parasitol. 96, 151–165.

Yuh, Y.S., Liu, J.Y., Shaio, M.F., 1997. Chromosome number of *Trichomonas vaginalis*. J. Parasitol. 83, 551–553.

Yuh, Y.S., Liu, J.Y., Shaio, M.F., 1998. Chromosome number of *Trichomonas vaginalis*-errata. J.Parasitol. 84, 28.