universität wien

DISSERTATION

Reaction of the human parasite *Entamoeba histolytica* to metronidazole: mRNA expression analysis based on a compilation of annotation data from the genome project

angestrebter akademischer Grad

Doktorin der Naturwissenschaften (Dr.in rer.nat.)

Verfasserin / Verfasser:	Mag. Margit Tazreiter
Matrikel-Nummer:	9504086
Dissertationsgebiet (lt. Studienblatt):	Genetik und Mikrobiologie
Betreuerin / Betreuer:	Ao. Univ.Prof. Dr. Michael Duchêne

Wien, am 24.02.2010

Für meine Eltern,
die mir immer alles ermöglicht haben,

und für meinen Mann Wolfgang,
dessen Liebe mich trägt.

This thesis project was carried out at the Department of Specific Prophylaxis and Tropical Medicine at the Center for Pathophysiology, Infectiology and Immunology of the Medical University of Vienna, Austria, under the supervision of Ao.Univ.Prof. Dr. Michael Duchêne.

Contact:

Ao.Univ.Prof. Dr. Michael Duchêne
Kinderspitalgasse 15
1090 Vienna, Austria
Tel. 0043 1 4277 64874
michael.duchene@meduniwien.ac.at

The project was made possible by the Grant P15960-B08 of the Austrian Science Fund.

Table of contents

1 Introduction

1.1 *Entamoeba histolytica*

1.1.1 The parasite

The microaerophilic protozoan parasite *Entamoeba histolytica* generates a severe public health problem in developing countries, being the causative agent of amoebic dysentery and amoebic liver abcess. According to WHO reports it causes several million cases of disease and up to 100,000 deaths per annum, as a parasite disease being second only to malaria (WHO report on amoebiasis, 1997). The first description of the species was given by the Russian Fedor Lösch in 1875 (Lösch, 1875), who observed trophozoites in feces from a farmer with dysentery and named the parasite *Amoeba coli*. Not until 1903 the still valid name of *Entamoeba histolytica* was established by Fritz Schaudinn (Schaudinn, 1903), indicating its ability of host tissue destruction.

Although the disease is mainly found in warm and humid climates of the world, it has been observed in nearly every latitude (Brandt and Tamayo, 1970). Only in recent time infections are seen in countries where amoebiasis is not endemic, like Japan or Australia, where the parasite is transmitted *via* oral-anal sex among homosexual men (Stark *et al*., 2008). Often people with *E. histolytica* infections show no signs of disease and can clear the infection without any problems. However, 4-10% of these asymptomatic carriers develop disease over a year (Gathiram and Jackson, 1987).

The life cycle of *E. histolytica* is simple compared to other protozoans, consisting only of the amoeboid trophozoite and the infectious cyst stage. Cysts are usually transmitted *via* the fecal-oral route by consumption of contaminated water or food. After ingestion, the tetranucleated cysts pass the stomach without taking harm by the acidic environment because of their chitin wall, and hatch in the terminal ileum or proximal colon thereby producing eight small daughter trophozoites. These colonise the colon.

They can ingest bacteria, reproduce by binary fission and invade the host tissue where they feed on red blood cells and consequently can cause amoebic colitis or even amoebic liver abscess (Martinez-Palomo, 1982). Alternatively, trophozoites can encyst and leave the colon with stool, closing the infectious cycle.

The two main manifestations of human amoebiasis are the intestinal form, the colitis, where typical flask-shaped ulcers are formed in the submucosal tissues due to lateral spread, leading to diarrhoea and sometimes also fever and weight loss, as well as the most frequent extraintestinal occurrence, the amoebic liver abcess (ALA), forming often large solitary lesions in the liver (Stanley, 2003) that cause fever, hepatic tenderness and abdominal pain. Much rarer forms affect the respiratory tract or the brain. It has long been observed that the main part of ALA formation occurs in men between the age of 18 and 50, although the reasons for this phenomenon are not yet understood in detail. A recent study using a mouse model of disease showed that female mice had higher IFN-γ levels than male ones at the beginning of ALA formation (Lotter *et al*., 2006). Additionally, the importance of natural killer T cells in the host defence against ALA development was demonstrated.

1.1.2 Phylogeny

The genus *Entamoeba* belongs to the phylum *Entamoebidae*. It is described as distantly related to the genera *Mastigamoeba* and *Dictyostelium*, which together form the monophylum Conosa (Clark 2000, Bapteste *et al*., 2002). There are free-living amoebae like *E. moshkovskii*, gut commensals like *E. coli* as well as *E. gingivalis*, which causes oral infections. The reptilian parasite *E. invadens* is interesting for researchers because it is rather easy to induce encystation in axenic cultures of this species (Eichinger, 1997), whereas *E. histolytica* encystment until today has only been induced in monoxenic cultures of the parasite (Clark and Diamond, 2002).

Due to its apparent lack of cell organelles like mitochondria and Golgi apparatus, *E. histolytica* was long grouped with two other amitochondriate parasites, *Giardia intestinalis* and *Trichomonas vaginalis,* as phylum Archezoa, all of them also being microaerophilic. But already in 1995 the chaperonin 60 gene (cpn60) of *E. histolytica* was found to be homologous to the respective mitochondrial protein, what provided a first direct evidence for the secondary loss of mitochondria in *Entamoeba* (Clark and Roger, 1995).

Many different species of amoebae have been isolated from different hosts and the environment. One simple method to group them was to look at the number of nuclei in the cyst. New genetic methods use the analysis of the relationships between small subunit rRNA gene sequences (Clark *et al*., 2006a). Data obtained with this method suggest that tetranucleated cyst-producing species like *E. histolytica* are descended from uninucleated cyst-producers. *E. gingivalis*, which produces no cysts, appears to have secondarily lost this ability and may descend from a uninucleated cyst-producing species as well.

At the beginning of *Entamoeba* research all amoebic organisms found in human patients were addressed as *E. histolytica*. But already in 1925 Émile Brumpt suggested that at least one other species must exist which morphologically looks virtually identical but causes no disease (Brumpt, 1925). Consequently, he proposed the name of *E. dispar* for this species with clearly distinct properties. But this work found no response until several decades later molecular biological data suggested that there is a separate nonpathogenic species. The isoenzyme patterns of phosphoglucomutase (Sargeaunt *et al*., 1978) as well as those of hexokinases I and II (Farri *et al*., 1980) formed one part of the evidence for two different species. Support on the genetic level came from restriction fragment length polymorphism found in cDNA clones of pathogenic and nonpathogenic isolates (Tannich and Burchard 1991). Finally, in a redescription, *E.*

histolytica was separated from *E. dispar* (Diamond and Clark, 1993). This separation was widely accepted and still more evidence was found, *e.g.* the molecular biological explanation of the isoenzyme patterns of hexokinases and phosphoglucomutase (Ortner *et al*., 1997a; 1997b).

1.1.3 Cell biology

Entamoeba histolytica cysts are round with 10-15µm in diameter, comprising four nuclei surrounded by a rigid chitin wall. In contrast, the mobile trophozoites can reach 10-50µm in diameter (Stanley, 2003).

The cell structure of *E. histolytica* for a long time was regarded as very simple for a eukaryote, lacking for instance mitochondria, a Golgi-apparatus (GA) and an endoplasmatic reticulum (ER). One explanation was that amoebae were living fossils which had never acquired prokaryotic endosymbionts (Cavalier-Smith, 1991; Poole and Penny, 2007). But already in 1973 the so-called “hydrogenosome” had been discovered in *T. vaginalis*, a double-membrane organelle producing molecular hydrogen (Lindmark and Müller, 1973). Hrdy *et al*. (2004) showed that the 51-kDa and 24-kDa subunits of the NADH dehydrogenase module in complex I, the first step in the mitochondrial respiratory chain, are present in *T. vaginalis*. According to phylogenetic analysis they share common ancestry with the mitochondrial homologs. These findings support the theory that mitochondria and hydrogenosomes are aerobic and anaerobic homologs of the same endosymbiontically derived organelle.

Finally, in *E. histolytica* (Tovar *et al*., 1999; Mai *et al*., 1999) and in *Giardia intestinalis* (Tovar *et al*., 2003) the so called “mitosomes” were discovered, also termed “cryptons” in *E. histolytica* (Mai *et al*., 1999). These organelles are surrounded by a double membrane and are targeted by mitochondrial-related proteins, *e.g.* molecular chaperones Hsp10, Hsp60 and Hsp70 (Bakatselou *et al*., 2000; Tovar *et al*., 2007) and

an unusual ADP/ATP carrier (Chan *et al*., 2005). The function of mitosomes is still controversial, one hypothesis says that they play a role in iron-sulfur (Fe-S) protein maturation (Tovar *et al*., 2003; Maralikova *et al*., 2010) as well as that they contain a sulfate activation pathway, a function that in eukaryotes is generally localised in the cytoplasm and plastids (Mi-Ichi *et al*., 2009). These findings seem to support the theory that all eukaryotes, those with fully functional mitochondria as well as those with mitosomes and hydrogenosomes, are derived from one common ancestor (van der Giezen and Tovar, 2005; Richards and van der Giezen, 2006).

For a long time, *E. histolytica* seemed also to be devoid of a Golgi-system and an endoplasmatic reticulum (ER). But then advanced methods in staining and microscopy suggested that some of the numerous vesicles in the trophozoites could correspond to components of the ER and Golgi-complex (Mazzuco *et al*., 1997). Later on, different studies showed that the parasite contains vesicles that can be addressed as ER and Golgi-apparatus with different antibodies or by exploring their capability of transporting typical substances into their lumen, *e.g.* UDP-glucose (Gosh *et al*., 1999, Bredeston *et al*., 2005). So seemingly the function of these cell compartments is present even if the appearence of the organelles is somewhat different.

Entamoeba trophozoites contain proteasomes in their cytoplasm that seem to play an important role in cell differentiation and replication, as is shown by treatment of cells with proteasome inhibitors (Makioka *et al*., 2002). This fact makes them an interesting therapeutic target in *E. histolytica*, as well as in other pathogenic protozoan parasites (Paugam *et al*., 2003).

Entamoeba histolytica trophozoites have a pronounced ability for amoeboid motion and phagocytosis. This ability is a key feature in the pathogenicity of this parasite and has therefore been studied extensively (Meza *et al*., 2006). Chemotaxis seems to play an important role in this matter (Zaki *et al*., 2006). *E. histolytica*

trophozoites exhibit bipolarity, with the pseudopod at the front of the cell and the so-called uroid at the rear side (Bailey *et al*., 1992). The amoeba cytoskeleton is actin-rich, and both of the mentioned structures are formed by redistribution of F-actin (Guillén, 1996). Actin interacts with several other actin-binding proteins, such as the gelation factor ABP-120. It binds to *E. histolytica* F-actin and during movements of trophozoites localizes in the pseudopod and uroid region, suggesting a role for this protein in motility and the so-termed capping of surface receptors (Vargas *et al*., 1996). Capping at the surface of cells can be seen as a two-step process. First, receptors which have been exposed to ligands such as host antibodies and complement, are transferred to particular sites on the cell surface and form patches there. Then, these patches move rearwards to the uroid. Where they are concentrated a cap is formed which finally is eliminated with the uropodial structure itself (Arhets *et al*., 1995).

1.1.4 The genome

E. histolytica chromosomes do not condense during mitosis, making the identification of the exact number of chromosomes and the ploidy very difficult. Pulsed field gel electrophoresis predicts 14 chromosomes ranging from 0.3 to 2.2 Mbp and a possible ploidy of 4 (Willhoeft and Tannich, 1999). However, a recent study showed that the coupling between nuclear division and cytokinesis is abrogated, resulting in a wide range of different ploidy (Muhkerjee *et al*., 2009). The *E. histolytica* genome project was carried out at the Institute for Genomic Research (TIGR) in Rockville, Maryland, and at the Wellcome Trust Sanger Institute in Cambridge. It was built on a whole *Entamoeba* community effort to establish a basis for future research, for instance we provided our annotation databases (see 3.1 our publication on databases and 4.1 Conclusions). The draft genome sequence was published in 2005 (Loftus *et al*., 2005). The pathogenic strain HM-1:IMSS was sequenced by the whole-genome shotgun method. The

assembly is based on a 12.5-fold coverage of the genome. 9,938 genes were predicted to be contained in 23,751,783 base pairs (bp) distributed on 888 scaffolds. Genes have an average size of 1.17 kilobases (kb) and contain 49% of the genome sequence. Up to 25% of the forecast genes were predicted to contain introns, with 6% perhaps even containing multiple introns. A striking feature of the *E. histolytica* genome is the huge amount of tRNA-encoding arrays (Clark *et al*., 2006b). They make up around 10% of the whole genome and are mostly organized in linear arrays, 25 types of these have been identified, each containing up to five tRNA types per repeat unit. Several properties of the arrays suggest that they are located in subtelomeric regions and also fulfil a structural role, as has been proposed for the rDNA of *Dictyostelium discoideum* (Eichinger *et al*., 2005). The unusual organisation of rRNA-encoding genes is another trait of the parasite´s genome. They are located on extrachromosomal plasmids which can be found in around 200 copies per cell (Bhattacharya *et al*., 1998) and make up about 20% of the total DNA amount per cell.

Interestingly, the genome of *E. histolytica* contains a considerable amount of genes of assumed bacterial origin. Several important metabolic enzymes have been replaced by prokaryotic types. Examples that have been known for some time are fructose-1,6-bisphosphate aldolase (Sánchez *et al*., 2002), malic enzyme, acetyl-CoA synthetase (Field *et al*., 2000), aldehyde-alcohol dehydrogenase (ADH2) (Bruchhaus and Tannich, 1994) and the genes thought to be involved in metronidazole activation (see below), pyruvate:ferredoxin oxidoreductase (PFOR) and ferredoxin (FD) (Rosenthal *et al*., 1997). There are three different possible explanations for the presence of bacterial genes in eukaryotes. The amitochondriate fossil hypothesis says that these genes could be derived from an ancient anaerobe eukaryote without mitochondria. The hydrogen hypothesis suggests that fermentation enzymes were gained with the mitochondrial endosymbiont, and the lateral gene transfer (LGT)

hypothesis means that there was direct transfer of genes from prokaryotes to eukaryotes. Many publications find support for LGT (*e.g*. Alsmark *et al*., 2009; Andersson *et al*., 2003; Nixon *et al*., 2002).

Transposable elements make up a significant part of the *E. histolytica* genome. There are three subtypes of LINEs (long interspersed elements) EhLINE1, 2 and 3 as well as three subtypes of SINEs (short interspersed elements) EhSINE 1, 2 and 3, ranging in copy number from only a few to far over 100 (Bakre *et al*., 2005).

The publication of the *E. histolytica* draft genome sequence immensely enhanced research on this organism, making possible new approaches of study, *e.g.* gene expression analysis with microarrays or the analysis of the whole proteome of the parasite.

1.1.5 Metabolism

Life as a microaerophilic parasite strongly shaped the metabolism of *Entamoeba histolytica*. This environment rules out oxidative phosphorylation. As no tricarboxylic acid (TCA) cycle is found either, amoebae rely on substrate level phosphorylation (Reeves, 1984). Glycolysis appears to take place in the cytosol like in higher eukaryotes but different from trypanosomes where the first activating steps of glycolysis occur in the so-called glycosomes, maybe to prevent a toxic accumulation of intermediates (Haanstra *et al*., 2008). For a long time, *Entamoeba* was thought to lack any compartmentalised energy production and therefore was grouped as type I amitochondrate protist, contrary to *T. vaginalis* which contains hydrogenosomes and therefore was classified as type II amitochondriate protist (Martin and Müller, 1998).

The main part of ATP is generated through glycolysis, but the genome project found several genes, sometimes with bacterial background, that could contribute to ATP production by amino acid catabolism, *e.g.* aminotransferases (Clark *et al*., 2007;

Anderson and Loftus, 2005). Typically, after a first transamination step a 2-ketoacid is generated, that could be pyruvate or some other 2-ketoacid. This can be used with the help of PFOR (see below) to generate CoA-thioesters from which one ATP per molecule can be produced.

Amoebae live in the human gut, hence they have easy access to many preformed bacterial and host-derived organic compounds such as lipids or cholesterol (Das *et al*., 2002). It was noted very early that amoebae and other protozoan parasites can take up these compounds from the host and, in addition, that especially cholesterol has an influence on the virulence of parasites, also in *E. histolytica* (*e.g.* Ormerod and Venkatesan, 1982; Sharma, 1959). So it is an interesting question if amoebae still have genes of the cholesterol synthesis pathway in their genome. Furthermore, there are hints that amoebae and other amitochondriate protozoa like *Giardia* and *Trichomonas* are able to alter host-derived lipids or are capable of elongating/desaturating fatty acids. While compiling the annotation database for our project we searched the genome for genes coding for the enzymes of these pathways (see 4.1 Conclusions).

The enzyme pyruvate:ferredoxin oxidoreductase (PFOR) is of immense significance to *Entamoeba*, as the parasite lacks a pyruvate dehydrogenase or pyruvate decarboxylase. This bacterial-type iron-sulfur enzyme oxidatively decarboxylates pyruvate to acetyl-CoA. Electrons are transferred to ferredoxin which then reduces further substrates. In addition, it can probably activate metronidazole, the main antiamoebic drug (Müller, 1986) (see below, point 1.2).

1.1.6 Pathophysiology

Amoebic colitis can result in mucosal thickening to discrete ulcer formation up to necrosis and intestinal wall perforation (Stanley, 2003). The process always starts with adherence of the *E. histolytica* trophozoites to the colonic mucus and epithelium

(Bansal *et al*., 2009) mediated through the galactoase/N-acetylgalactosamine specific (Gal/GalNAc) lectin, a major pathogenic factor of amoebae comprising disulphide-linked 170 kDa and 35/31 kDa subunits (Mann, 2002). In some hamster cell lines that lack N-terminal galactose or N-acetylgalactosamine residues on their surface, a direct contact of amoebae with the cells is impossible and no cell killing can occur (Li *et al*., 1988). After adhesion of the amoebae, cytolysis is mediated by the amoebapores, small proteins that are highly effective in lysing lipid bilayers (Leippe, 1997). The attacked host cells undergo apoptosis (Huston *et al*., 2000 and 2003), as shown by the activation of caspase 3 in engulfed Jurkat T cells, contact via the amoeba Gal/GalNAc lectin being necessary.

The third major group of pathogenicity factors in *E. histolytica* are the cysteine proteinases. These enzymes which are found in all organisms play an important role in tissue invasion, *e.g.* they were shown to disturb enteric microvilli and thereby mediate an intimate contact between the amoebae and enteric cells (Lauwaet *et al*., 2003). After tissue invasion the amoebae enter the host blood stream and there encounter the complement system and host antibodies. These can be cleaved by excreted cysteine proteinases (Reed *et al*., 1989; Kelsall and Ravdin, 1993). Interestingly, the strict commensal *Entamoeba dispar* does not express the cysteine proteinase CP5. The gene encoding CP5 is present, but highly degenerated and therefore not expressed (Willhoeft *et al*., 1999). Most of the other pathogenicity factors are present in *E. dispar* but are only weakly expressed, *e.g.* one study showed differences in the structure and activity of pore forming peptides in *E. dispar* (Nickel *et al*., 1999). These discrepancies could to some extent explain the differences in pathogenicity of the two closely related species, but more recent studies show that many more factors seem to be involved in *Entamoeba* pathogenicity, for instance protection against oxidative stress by peroxiredoxin (Lejeune *et al*., 2009; Biller *et al*., 2009). Moreover, pathogenicity also

depends on the host response during the development of the disease (Lotter *et al*., 2009).

When trophozoites finally reach the liver, probably via the portal circulation, they cause typical abscesses consisting of areas of dead hepatocytes, liquefied cells and cellular debris (Stanley, 2003). The often surprisingly small number of amoebae relative to the size of the abscess could hint to an abilitiy to elicit apoptosis in hepatocytes without direct contact.

1.1.7 Diagnosis

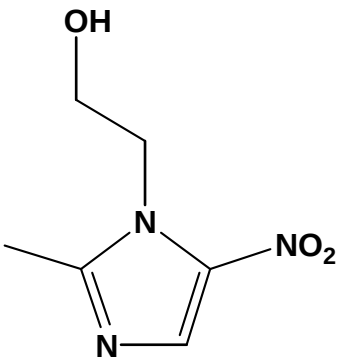
For many years the microscopic analysis of stool samples has been nearly the only method to diagnose an *Entamoeba* infection in patients with diarrhoea. However, only well-trained personnel can readily distinguish the *Entamoeba* trophozoites from leukocytes and from other intestinal protozoa (Stanley, 2003). Furthermore, even after the firm establishment of the two distinct species *E. histolytica* and *E. dispar*, it was not possible to differentiate between the two by microscopy due to their morphological similarity, leading to unnecessary treatment of people only harbouring *E. dispar*. Meanwhile, several more sophisticated methods for diagnosis and differentiation were developed, such as enzyme-linked immunosorbent assays (ELISAs), PCR and real-time PCR. Of these techniques, only ELISAs are widely used in endemic areas because they are not so costly and easy to perform, one example is the TechLab *E. histolytica* II test that was reported to be sensitive and specific (Buss *et al*., 2008). PCR and real-time PCR are very sensitive methods that can be also used to diagnose amoebic liver abscess and to detect several different pathogens at once in a sample (Singh *et al*., 2009; Roy *et al*., 2005). These methods will become more important in the future, *e.g.* for epidemiological studies. Meanwhile serology, looking for anti-*Entamoeba* antibodies, plays a large role in diagnosing invasive disease together with stool

examination. For detecting liver abscesses, serology and ultrasound scanning of the liver are the widely used methods.

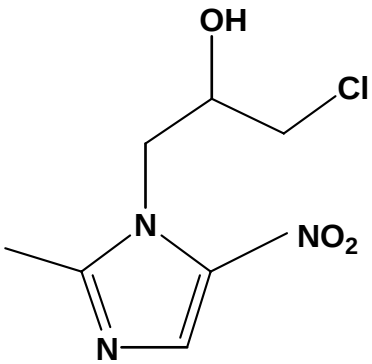
1.1.8 Prevention

Of course great efforts have been undertaken in the last decades to provide a vaccine against amoebic diseases (Lotter and Tannich, 2006). Several possible antigens have been recognized, *e.g.* the Gal/GalNac lectin (Petri *et al*., 2006) or passive protection by an antibody against lipophosphoglycans (LPG) (Marinets *et al*., 1997). Nevertheless, none of these candidate vaccines have yet reached clinical studies, so the actual availability of a vaccine is still many years away. What´s more, providing better sanitary conditions, clean water and clean food in developing countries would minimalise the risk for acquiring amoebiasis dramatically and also improve the life of the affected people in many more ways than finding a vaccine against one of many possible causes of disease. Unfortunately, this goal seems even further away than developing an amoebiasis vaccine.

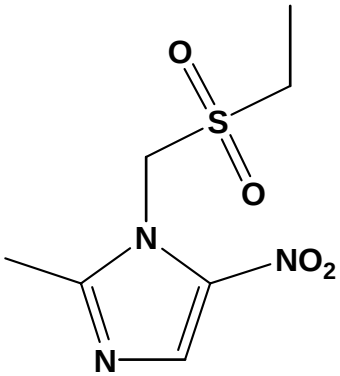
1.2 Treatment of parasitic diseases, the role of the nitroimidazoles



Metronidazole



Ornidazole



Tinidazole

For several decades, the 5-nitroimidazoles with their best known proponents metronidazole, ornidazole and tinidazole have been the cornerstone of antiamoebic treatment. These chemotherapeutical agents are generally administered against anaerobe parasites and bacteria, such as *Trichomonas vaginalis*, *Gardia lamblia* or *Bacteroides fragilis* (Ali and Nozaki, 2007; Fung and Doan, 2005). In the management of amoebic colitis, the three mentioned drugs are nearly equally effective, perhaps with a slight advantage for tinidazole and ornidazole, but in some countries, *e.g.* the USA, only metronidazole is available (Stanley, 2003). Normal treatment consists of the administration of 750 mg of metronidazole three times a day for 5-10 days. Subsequently, an additional treatment with the luminal agents paromomycin, iodoquinol or diloxanide furoate is strictly recommended for eradication of amoeba colonisation, as nitroimidazoles are not as effective against luminal amoebae.

The treatment of amoebic liver abscess in most cases is comparable to that of amoebic colitis, as even large abscesses can be cured by the administration of metronidazole alone. The additional use of surgical drainage in uncomplicated ALA is controversial, with authors advising against it (Akgun *et al*.,1999) and others wanting for more randomised trials to decide this matter as those available now do not possess the necessary quality (Chavez-Tapia *et al*., 2009). Comparable to amoebic colitis, metronidazole treatment must be followed by administration of a luminal agent to eradicate intestinal colonisation.

One drug that could become important in the future of amoebiasis treatment is nitazoxanide, a broad-spectrum antiparasitic drug with activity also against *Giardia* and *Trichomonas* as well as helminths (Rossignol *et al*., 2001). Another interesting group of substances are the alkylphosphocholines (alkyl-PC), originally developed as anti-cancer drugs, which were found to posses a strong antiamoebic activity (Seifert *et al*., 2001). Miltefosine, one of the alkyl-PCs, is already in use against visceral leishmaniasis

(caused by *Leishmania donovani*) in India, but alkyl-PCs are also effective against *Trypanosoma* spp. (Croft *et al*., 2003) and *Acanthamoeba* spp. (Walochnik *et al*., 2002).

Resistance to metronidazole is a problem of variable importance in the "amitochondriate" parasites *E. histolytica*, *G. intestinalis* and *T. vaginalis*. In *Entamoeba*, no clinical isolates with a high level of resistance are observed until now, although exposure to nonlethal drug levels could increase drug resistance (Upcroft and Upcroft, 2001). However, in laboratory strains modest resistance has been induced up to a concentration of 40 µM metronidazole (Wassmann *et al*., 1999), a lethal dose in unadapted amoebae. In this resistant line, the expression of superoxide dismutase and peroxiredoxin was increased, the expression of ferredoxin 1 was decreased, whereas the expression of PFOR decreased only slightly.

Metronidazole resistance in *Giardia* and *Trichomonas* is a severe clinical problem, as for instance in Giardia up to 20% of clinical cases show resistance (Ali and Nozaki, 2007). In these cases, PFOR is significantly down-regulated, what was thought to be the mechanism of metronidazole resistance in *Giardia*. Even multidrug-resistant strains can be grown. In *Trichomonas*, resistance *in vivo* can even develop with therapeutic levels of metronidazole and is also associated with the down-regulation or even loss of PFOR and ferredoxin, showing the differences of metronidazole resistance in the "amitochondirates".

All these data show the increasing problems with metronidazole resistance in anaoerobic parasites. As only a limited number of other chemotherapeuticals against these organisms are available and resistance against these has also been observed, it is clear that the search for new chemotherapeutic agents must go on or even be intensified.

Other protozoan parasites show similar developments concerning resistance. The probably most important example for these problems is the genus *Plasmodium*,

where the burden of disease and the number of deaths each year is far greater, but the regions with multiresistant strains grow (Nzila and Mwai, 2009), and even resistance against artemisinin has been observed (Noedl *et al*., 2008 and 2009). Therefore, the search for new chemotherapeuticals is an urgent duty.

Taken together, the crucial necessity for further research on chemotherapy of protozoan parasites is obvious. What is more, often the exact mode of action of widely used chemotherapeutic agents is not known. Regarding metronidazole, this gold standard drug is in use for more than 50 years now, and there has been a lot of research on its activation in the parasites, but the exact mode of action is still not fully understood and merits further experimental work.

Metronidazole in its genuine form is not active but needs to be reduced at the nitro group to become the active nitroradical anion that has cytotoxic properties (Müller, 1983). This nitroradical anion is further reduced to a nitrosoimidazole (Moreno and Docampo, 1985) that is able to react with sulfhydryl groups (West *et al*., 1982) and with DNA (Ludlum *et al*., 1988). Ongoing reduction leads to the formation of a hydroxylamine and finally to an amine. This activation pathway can only happen in the anaerobic environment *e.g.* in the parasites discussed here. In an aerobic environment like in humans the nitroradical anion is immediately reoxidized in a so-called futile cycle (Mason and Holtzmann, 1975), leading to the formation of superoxide radical anions and oxidative stress, but no severe cell damage. This reoxidation is the reason why nitroimidazoles are considered safe for use in humans, although the risk of cancer development because of the oxidative stress could not conclusively be ruled out by a number of studies (Bendesky *et al*., 2002).

The enzymes responsible for reduction of the nitro group have been examined extensively, and several candidates were found. Ferredoxin was proposed as the main reducing agent, this iron-sulfur protein is reduced itself by pyruvate:ferredoxin

oxidoreducase (PFOR) (Müller 1986). In *T. vaginalis*, both enzymes are located in the hydrogenosome, and activation of metronidazole was found to take place there (Chapman *et al*., 1985). Nevertheless, parasites in which ferredoxin was knocked out did not display full resistance to metronidazole, suggesting the existance of other pathways of metronidazole activation, a possible alternative being malic enzyme (Rasoloson *et al*., 2002). Recent findings from our group show that the flavin enzyme thioredoxin reductase can reduce and activate metronidazole and other nitroimidazoles at significant rates in *E. histolytica* (Leitsch *et al*., 2007) as well as in *T. vaginalis* (Leitsch *et al*., 2009; 2010). For further discussion see point 4 (Conclusions).

The mode of action *in vivo* of metronidazole is still not fully understood. It is supposed to react with DNA (Ludlum *et al*., 1988), leading to damage, but no binding to specific enzymes has been found until recently, as the mechanism of cytotoxicity was thought to be random.

The recent sequencing of the genomes of the anaerobic parasites *E. histolytica* (Loftus *et al*., 2005), *Giardia lamblia* (Morrison *et al*., 2007) and *T. vaginalis* (Carlton *et al*., 2007) really simplifies the research on mode of action of chemotherapeuticals, gene expression and many other aspects of the biology of these organisms. The availability of the complete sequences enables researchers to use various techniques for resolving scientific questions, *e.g.* microarray studies, qRT-PCR (quantitative real-time reverse transcription PCR) or two-dimensional gel electrophoresis (2DE).

1.3 Methods

1.3.1 The use of microarray studies in parasite research

DNA microarrays of the kind we know today were developed only some 15 years ago (Shalon *et al*., 1996). They mostly consist of a glass slide with modified surface properties onto which the target DNAs (genomic DNA, cDNA or oligonucleotides) are

spotted in a defined manner so to know which target DNA is found in which location. Then the probes of interest are prepared for hybridization onto the slide. For example, mRNAs are extracted from cells in two different conditions. The mRNAs are transcribed into cDNAs labeled with fluorescent dyes, mostly Cy3 and Cy5, and incubated with the microarray slides where hybridization of the complementary sequences can take place. After hybridization, excess probe is washed off, the slides dried and analysed with the help of a laser scanner microscope and camera. Data are stored and analysed with appropriate computer programs. A good overview of the technique is found in the NCBI factsheet on microarrays at http://www.ncbi.nlm.nih.gov/About/primer/microarrays.html. The greatest handicap of microarrays at the beginning were the high costs for these experiments so that only large institutes had the resources to really use this interesting method.

Perhaps the most intriguing feature of microarray studies is the fact that the whole genome can be studied in one experiment. The effect of, for instance, chemotherapeutical treatment, stress like heat shock or UV-radiation, or the change in nutritional conditions can be monitored on a whole genome scale. Therefore it is not surprising that with the availability of the sequence data through the *Entamoeba histolytica* genome project (Loftus *et al*., 2005) the number of published mircoarray studies grew rapidly. Different microarray platforms were used, *e.g.* arrays on the commercially available Affymetrix™ platform (www.affymetrix.com) versus selfspotted arrays, long or short oligonucleotide arrays versus cDNA arrays.

A variety of topics have already been investigated. The reaction of *E. histolytica* to stress by UV irradiation that caused DNA damage was studied by Weber *et al*. (2009). The overall change in gene expression 5 min after the UV-treatment was very limited, with only a few genes regulated stronger than 2-fold. Three hours after the UV-C irradiation the response was more pronounced but still limited to only a few genes

regulated stronger than 2-fold. The same research group has looked into amoebic response to heat shock (Weber *et al*., 2006). Surprisingly, a massive overall down-regulation of expression levels was found (471 genes), with only 40 genes up-regulated, including heat shock proteins and two cysteine proteinases 4 and 6, whereas a lot of pathogenicity factors including all other cysteine proteinases, peroxiredoxin and amoebapore C were down-regulated.

The impact of oxidative and nitrosative stress on *E. histolytica* HM-1:IMSS was surveyed by Vicente *et al*. (2009). There were a large number of genes modulated by both stressors, including heat shock proteins, ubiquitin-conjugation enzymes or protein kinases. Interestingly, when subjecting the non-pathogenic strain *E. histolytica* Rahman to the same stressors, there were marked differences in the response of the cells and they seemed to suffer greater damage from oxidative stress, perhaps partly explaining the decreased pathogenicity of this strain.

The same group examined developmentally regulated genes in *E. histolytica* upon stage conversion between trophozoite and cyst (Ehrenkaufer *et al*., 2007b). Using parasite strains from very recent human infections they found 672 genes enriched in the cyst stage including cysteine proteinases and transmembrane kinases. Of the 767 genes enriched in trophozoites many are thought to be important for tissue invasion like the Gal/GalNAc lectin light subunit. Some of the genes were also found to be developmentally regulated in *E. invadens*, hinting at a conserved mechanism in *Entamoeba* stage conversion.

Pathogenicity has always been an important topic in *Entamoeba* research, so several microarray studies focus on this issue. Tillack *et al*. (2007) looked at the peptidases of *E. histolytica* because especially the cysteine peptidases are regarded as major pathogenicity factors of this parasite. Using a small focussed array, they investigated seven different *E. histolytica* isolates for the expression levels of the 50

cysteine peptidases and more than 30 other peptidases they had found in the amoeba genome, under normal conditions and also under heat stress. Surprisingly, the overall expression level of these genes was rather low, with only a relatively small number significantly expressed, also when comparing different strains and stress conditions, so the role of peptidases in *Entamoeba* pathogenicity merits further investigation. Strain comparison has been the topic of a study by Davis *et al*. (2007) who found 353 genes differentially expressed between *E. histolytica* HM-1:IMSS and Rahman, with distinct patterns for cysteine proteinases, lectin light chains and AIG (avirulence induced gene) family members. Differences in the expression level of genes similar to the AIG1 plant antibacterial proteins were also found by Gilchrist *et al*. (2006), who investigated the impact of intestinal colonization and tissue invasion on the *E. histolytica* transcriptome. They compared trophozoites isolated from the colon of infected mice with in vitro cultures and found 523 transcripts significantly changed between the two cultures, including genes implicated in metabolism, oxygen defense, cell signalling or virulence.

Two recent papers look at the effects of histone modifications in the *E. histolytica* genome. Isakov *et al*. (2008) treated the parasites of strain HM-1:IMSS with 50nM trichostatin A (TSA), an inhibitor of histone deacetylase. They found enhanced cytopathic and hemolytic activity of the amoebae as well as increased peroxiredoxin expression, corresponding to increased resistance to oxidative stress. On the whole, 13 genes were up- and 31 genes down-regulated by TSA-treatment. Jacob, the most abundant glycoprotein of the amoebic cyst wall, was up-regulated threefold by TSA. Only one gene was commonly up- and two others down-regulated compared with the second TSA-study, conducted by Ehrenkaufer *et al*. (2007a). Here the *E. histolytica* strain 200:NIH was treated with 150 nM of TSA, resulting in 122 genes being up- and 41 genes being down-regulated. Interestingly, a strong overlap existed between genes up-regulated by TSA exposure and genes found induced in cysts. Amoeba treatment

with 5-azacytidine, an inhibitor of DNA-methyltransferase, did not regulate genes modulated by TSA, indicating that no substancial connection is found between these two methods of epigenetic modulation of gene expression. Additionally, the influence of short chain fatty acids on amoebae was investigated, as these molecules also alter histone acetylation patterns. But only a weak transcriptional change was seen (11 out of 9435 genes regulated).

Of course microarrays have also been used for different applications in other protozoan parasites. For instance, transcriptional and genomic analysis of *Plasmodium falciparum* field isolates compared to long-term in vitro cultivated parasites was performed by Mackinnon *et al*. (2009). Another recent study focusses on the effect of *Plasmodium* infection on host cell gene expression in the liver (Albuquerque *et al*., 2009).

In *Giardia lamblia*, gene expression of a clone resistant to metronidazole and nitazoxanide was in focus of a study by Müller *et al*. (2008). They found that resistance was correlated with changes in expression patterns of major surface antigens and also of heat shock proteins. Microarrays can also be used for the detection and genotyping of multiple protozoan parasites in environmental samples, as shown in a study by Wang *et al*. (2004). They designed a microarray able to distinguish between *E. histolytica* and *dispar*, 3 different *G. lamblia* assemblages and two types of *Cryptosporidium parvum*.

This short overview shows the wide variety of applications possible for microarray technique with regard to protozoan parasites. The focus of this doctoral thesis project lay on the construction of a focussed microarray to detect the effects of metronidazole treatment on *E. histolytica* gene expression. For the details of this study see point 2 (Aims of this study) and point 3 (Manuscripts).

1.3.2 Quantitative real-time RT-PCR (qRT-PCR) for investigating gene expression in parasites

Quantitative real-time reverse transcription - polymerase chain reaction, commonly abbreviated as qRT-PCR, is another fast evolving method that emerged more or less at the same time as microarrays (Heid *et al*., 1996). It is a further development of PCR technique that is able to reliably detect and measure the products of PCR during each cycle of the process. There are several methods in use. The TaqMan assay uses a forward and reverse primer as well as an oligonucleotide hybridisation probe that binds within the sequence of the gene of interest. This probe carries a fluorophore at the 5'-end and a quencher molecule at the 3'-end. As long as these two are in close proximity, the quencher inhibits fluorescence. During extension of the newly synthezised strand the 5' to 3' exonuclease activity of the TaqPolymerase degrades the hybridisation probe, setting free the reporter fluorophore from the quencher. The fluorescence that is detectable now is proportional to the amount of newly synthezised DNA (Heid *et al*., 1996).

Another method uses SYBR green fluorescence. This substance binds only to double-stranded DNA that is generated during the amplification process of the PCR cycles. The more dsDNA is present, the more SYBR green binds and the signal gets more intensive. A good overview of this method can be found at http://pathmicro.med.sc.edu/pcr/realtime-home.htm

The qRT-PCR method is rather easy to perform nowadays, with sophisticated equipment and many different chemicals available. But of course there are several important points to consider that can influence the quality and informative value dramatically, for instance assay design, the choice of internal controls or analysis tools. Therefore the establishment of best practice guidelines and accepted evaluation tools is

a vital issue for the future relevance of this technique (see for instance Bustin, 2009; Nolan *et al*., 2006).

The applications of real-time qRT-PCR are manifold, *e.g.* relative and absolute quantification of gene expression, variation analysis or the validation of microarray results. We also used real-time qRT-PCR to control our microarray results, for the details see point 3.4 (manuscript).

1.3.3 Two-dimensional gel electrophoresis (2DE) for the study of whole parasite proteomes

Gene expression profiling is one part of investigating how a cell reacts to stress, chemotherapeuticals etc. The other part concerns effects on the proteome, to which extent alterations in gene expression actually correlate with changes in protein abundance or *de novo* synthesis (de Sousa Abreu *et al*., 2009; Tian *et al*., 2004). So it is ideal to combine gene expression profiling with proteomics methods. The most commonly used method is 2DE (Görg *et al*., 2004), consisting of isoelectric focusing of solubilized proteins in an immobilized pH gradient gel according to their respective isoelectric point, and subsequently, separation of proteins according to their molecular weight on an SDS-polyacrylamide gel. In this way a high resolution of proteins can be realised, which afterwards can be visualised by staining with Coomassie Brilliant Blue, silver or fluorescent dyes. Relative quantitative analysis of the gels is possible with imaging programs, for qualitative analysis, the protein spots of interest can be cut from the gel and subjected to various mass spectrometric methods. With the help of genomic data the actual proteins can be elucidated.

This method can of course be used for parasite proteomes, what was greatly simplified by the dfferent genome projects published in the last years. For instance, a proteome reference map of *T. vaginalis* was published recently (Huang *et al*., 2009).

Pathogenicity is an important topic in this parasite, one study looked at the proteomes of one long-term grown and one freshly isolated strain of *T. vaginalis* that exhibited marked differences in pathogenicity (Cuervo *et al*., 2008). Nineteen proteins overexpressed in the highly virulent strain were characterised. In *Giardia* research, a recent review compares proteomic studies from the pre- and post-genomic era (Steuart, 2010). As stage conversion is of great interest in *Giardia*, one research group compared the proteomes of trophozoites versus cysts (Kim *et al*., 2009). Proteomic studies are also of great relevance in elucidating the processes in different stages of *Plasmodium* parasites, for instance in the liver (Tarun *et al*., 2008).

Concerning *Entamoeba histolytica* and 2DE gels, only a few studies have been published, as other research groups use other proteomics methods like 2D-DIGE. The first 2DE study was published in 2005 by Leitsch *et al*. and was a proof of principle for this method with this parasite, furthermore several proteins were identified. Tolstrup *et al*. (2006) published another protocol of 2DE in *E. histolytica*. Comparison of amoebae with different pathogenicity was the subject of two more research projects. In one case the *E. histolytica* HM-1:IMSS proteome was compared to avirulent *E. histolytica* Rahman (Davis *et al*., 2006), in the other study the former was compared to non-pathogenic *E. dispar* SAW 760 (Leitsch *et al*., 2006).

2DE complemented the transcriptomic analysis of the effect of metronidazole on *E. histolytica* in this study, for the results see point 3.4 (manuscript) and point 4, the Conclusions section.

2 Aims of this study

Three main goals of the thesis can be described as follows:

1) Compilation of databases containing genome information on *Entamoeba histolytica*. When we started this project, no genome annotation was available, so we had to prepare our own annotation databases as basic information necessary for the subsequent construction of our microarray. Furthermore in the course of information assemblage we hoped to get on overview about the genome of the parasite, *e.g.* how big is the proportion of genes with bacterial background, which metabolic pathways are present/absent in the genome, what substances can *E. histolytica* synthesise itself and which are taken up from the host.
2) Establishment of microarray technology in our laboratory for the accomplishment of this project and perhaps for future projects.
3) With the help of a focussed oligonucleotide microarray investigation of the gene expression of *E. histolytica* under metronidazole treatment. We wanted to find out how the parasite reacts to this chemotherapeutical stress, which genes or pathways are involved in the response to metronidazole. Additionally we were interested in the mode of action of metronidazole, what it does to the cell and why the cells cannot cope with this substance. Genes found to be involved in the response to metronidazole could point to processes in the establishment of resistance.

3 Manuscripts

3.1 *Entamoeba histolytica*: construction and applications of subgenomic databases.

Hofer M, Duchêne M.

Exp Parasitol. 2005; 110: 178-183.

Available online at www.sciencedirect.com
SCIENCE DIRECT®
Experimental Parasitology 110 (2005) 178–183
Experimental Parasitology
www.elsevier.com/locate/yexpr

Entamoeba histolytica: Construction and applications of subgenomic databases

Margit Hofer, Michael Duchêne *

Department of Specific Prophylaxis and Tropical Medicine, Center for Physiology and Pathophysiology, Medical University of Vienna, Kinderspitalgasse 15, A-1095 Vienna, Austria

Received 1 February 2005; received in revised form 15 March 2005; accepted 15 March 2005
Available online 22 April 2005

Abstract

Knowledge about the influence of environmental stress such as the action of chemotherapeutic agents on gene expression in *Entamoeba histolytica* is limited. We plan to use oligonucleotide microarray hybridization to approach these questions. As the basis for our array, sequence data from the genome project carried out by the Institute for Genomic Research (TIGR) and the Sanger Institute were used to annotate parts of the parasite genome. Three subgenomic databases containing enzymes, cytoskeleton genes, and stress genes were compiled with the help of the ExPASy proteomics website and the BLAST servers at the two genome project sites. The known sequences from reference species, mostly human and *Escherichia coli*, were searched against TIGR and Sanger *E. histolytica* sequence contigs and the homologs were copied into a Microsoft Access database. In a similar way, two additional databases of cytoskeletal genes and stress genes were generated. Metabolic pathways could be assembled from our enzyme database, but sometimes they were incomplete as is the case for the sterol biosynthesis pathway. The raw databases contained a significant number of duplicate entries which were merged to obtain curated non-redundant databases. This procedure revealed that some *E. histolytica* genes may have several putative functions. Representative examples such as the case of the δ-aminolevulinate synthase/serine palmitoyltransferase are discussed.

Index Descriptors and Abbreviations: EC, enzyme commission; LGT, lateral gene transfer; ORF, open reading frame

Keywords: *Entamoeba histolytica*; Genome project; Database; Enzyme; Cytoskeleton; Stress response

1. Introduction

Entamoeba histolytica causes up to 100,000 deaths per year and represents a major health problem especially in developing countries (WHO, 1997). The *E. histolytica* genome project (Loftus et al., 2005) carried out by the Wellcome Trust Sanger Institute in Hinxton, Cambridge (UK) and the Institute for Genomic Research (TIGR) in Rockville, Maryland (USA), creates many new opportunities for the scientific community. Hopes are high to gain new insights into the molecular mechanisms of amoebiasis and find new targets for chemotherapy.

In our laboratory we plan to monitor the effects of metronidazole and alkylphosphocholines on the gene expression of *E. histolytica* with the help of oligonucleotide microarrays. Microaerophilic *E. histolytica* reduces the nitro group of metronidazole to cytotoxic nitro radicals via its reduced ferredoxin (Wassmann et al., 1999). Little is known about the mode of action of alkylphosphocholines such as miltefosine, but it is believed to interact with the lipid metabolism of the cells (Croft et al., 2003). Hybridization of drug-treated and control cDNAs to a focussed microarray with all

* Corresponding author. Fax: +43 1 4277 64899.
E-mail address: michael.duchene@meduniwien.ac.at (M. Duchêne).

0014-4894/$ - see front matter © 2005 Elsevier Inc. All rights reserved.
doi:10.1016/j.exppara.2005.03.018

known enzymes of the redox system of the amoebae and their lipid metabolism enzymes could be an appropriate way to gain insight into the mechanism of action of the two chemotherapeutic agents. When we started our work, the genome sequences had not yet been annotated, therefore we had to compile our own database with genome information. A software program for such a database should be stable, create highly convertible data sets, be flexible, and easy to use and maintain. We decided to use the Microsoft Access database software. Here we describe the compilation of three different subgenomic databases containing enzymes, cytoskeleton genes, and stress genes found in the *E. histolytica* genome.

2. Materials and methods

The Microsoft Access 2000 database software was chosen as program for compiling our data. First we selected enzymes with known protein sequences from the web version of the Roche Applied Science (previously Boehringer–Mannheim) Biochemical Pathways (http://us.expasy.org/cgi-bin/show_thumbnails.pl). We followed the link to the ExPASy website (http://www. expasy.org/) with the Swiss-Prot and Trembl (http://www.expasy.org/sprot/) as well as the Enzyme (http://www.expasy.org/enzyme/) databases, where reference protein sequences for these enzymes were retrieved. These were used for a tblastn search with default settings against the TIGR (http://tigrblast.tigr.org/er-blast/index.cgi?project=eha1) and Sanger (http://www.sanger.ac.uk/cgi-bin/blast/submitblast/e_histolytica) DNA sequences of the *E. histolytica* genome project. Tblastn compares a protein query sequence against a nucleotide sequence database dynamically translated in all six reading frames (both strands) using the BLAST algorithm (Altschul et al., 1990). The tblastn output top hits were considered a significant match if their *P* values were in the range of 10^{-5}or better (lower than 9.9E−05), and a complete database entry was generated. We also recorded the cases where no significant match was found. The exact position of the homologous sequence in the contig was established with the help of the Expasy translate tool (http://www.expasy.org/tools/dna.html) as well as the UWGCG package (Program Manual for the Wisconsin Package Version 8.1-Unix, see References). The theoretical isoelectric point (p*I*) and molecular weight of the found homolog were calculated with the help of the ExPASy p*I*/Mw tool (http://www.expasy.org/tools/pi_tool.html). In total, a set of up to 30 parameters were recorded for each database entry. Biochemical pathways including all found *E. histolytica* enzyme homologs were established with the help of the Kyoto Encyclopedia of Genes and Genomes (KEGG) website (http://www.genome.jp/kegg/tool/search_pathway.html) and displayed in comparison to the Standard Data Set.

3. Results and discussion

3.1. Establishment and use of the annotation databases

The purpose of our database work was to get a collection of annotated *E. histolytica* genes which should be the basis of our focussed oligonucleotide microarray. Although we planned to specifically target enzymes of the lipid metabolism and those known to be involved in reactions to metronidazole chemotherapy, we also wanted to get an overall picture of the metabolism of *E. histolytica*. Therefore we searched for all enzymes in the Roche/Boehringer "Biochemical Pathways" chart. In addition, we tried to find homologs of a list of cytoskeleton genes and stress genes. For better organisation we compiled three independent databases of enzymes, cytoskeleton genes, and stress genes.

A number of data sets were extracted for each gene into our Microsoft Access database. Each homologous *E. histolytica* sequence found with a *P* value in the range of 10^{-5} or better was used to generate a complete database entry. The *P* value and the percentage of protein sequence identity were entered. For the reference sequence the accession number, the enzyme commission number, and the species were recorded, as well as how many hits had been found in total. Next, we entered the *E. histolytica* protein coding DNA sequence, the deduced open reading frame (ORF), the position of the ORF within the contig, if it was a full-length ORF, and the theoretical p*I*/Mw. References were given where *E. histolytica* genes had been published before, and sufficient space for comments was included, for example, we recorded if there were conflicts with the published sequence. Recently we also introduced a field for the date of last revision to keep track of changes.

For classification of the entries we introduced status numbers ranging from 1 to 6. Number 1 means a published *E. histolytica* gene sequence, 2 is a sequence with high degree of end-to end similarity, 3 a sequence with intermediate degree of similarity, 4 a sequence with weak similarity, 5 means that there was a sequence for searching but no homology found, and 6 indicates that the enzyme activity is known but there is no protein reference sequence in the databases for searching.

The intermediate result of our searches after going through the Biochemical Pathways chart was that out of 1536 enzyme genes searched for, 962 genes with at least some sequence similarity (status 1–4) were detected, searching for 295 cytoskeleton genes led to 281 hits (status 1–4), and 109 hits were obtained by searching for 120 putative stress-related genes.

Inspection of the databases revealed that the same *E. histolytica* genes sometimes had been detected in different searches. When performing a tblastn search with human alcohol dehydrogenase α-chain (EC 1.1.1.1; Accession No. P07327) for instance, we obtained three weakly homologous *E. histolytica* sequences which were entered

in our database. But subsequently, the same ORFs were found when searching with seven other dehydrogenases, for example, sorbitol dehydrogenase (EC 1.1.1.14; Q00796). So our databases were examined for redundant entries. The sequence with the closest similarity was kept as an entry whereas those with weaker *P* value and percentage of identity were merged with this entry. Information on all merged entries was entered in the comments field to be able to reconstruct the original findings. On the whole, this merging process shrinked the enzyme database from an initial number of 962 to 749 entries with status 1–4 which is a decrease of 213 or approximately 22% of the entries. At the moment 156 entries consist of two or more merged entries with the majority of 106 being made up from two former entries, 30 from three, 14 from four, and 6 from more than four previous entries.

In the same way, the cytoskeleton database was reduced by 28% from 281 to 202 entries, 41 entries were merged with 24 being made up from two, 8 from three, 4 from four and 5 from more than four previous entries. Myosin heavy chain (Raymond-Denise et al., 1993) was found to be the best hit for a number of reference gene sequences. The stress gene database was reduced from 120 to 117 entries.

During compilation of our enzyme database we found a large number of genes where only a tblastn search with a prokaryotic reference sequence gave significant hits. This remarkable occurrence of bacterial sequences in *E. histolytica* has been noted before, and several candidates for lateral gene transfer (LGT) have been described (e.g., Kumar et al., 1992; Rosenthal et al., 1997; Sanchez et al., 2002). The genome project paper identifies 96 genes for which LGT is most probable by stringent criteria (Loftus et al., 2005). We did not do the same analysis of our candidates for LGT, but our findings support the results of the genome paper.

The design of oligonucleotides for our microarray experiments was started after completion of the databases. We chose enzymes that might be influenced by metronidazole or alkylphosphocholines for our focused array, and added a strong group of housekeeping genes from glycolysis and cytoskeleton genes as internal controls. For the actual oligonucleotide design from the gene sequences, we used the software OligoWiz 1.0 which optimizes the length, melting temperature, and position of the oligonucleotide, and minimizes possible cross-hybridization with other sequences in the genome (Nielsen et al., 2003). Our initial microarray consists of 221 oligonucleotides, and preliminary hybridization experiments are currently performed (results not shown).

3.2. An example for an incomplete metabolic pathway: steroid biosynthesis

With the help of the Kyoto Encyclopedia of Genes and Genomes (KEGG) website, we could generate metabolic pathways from the enzyme commission (EC) numbers of all the enzymes we had been able to detect. In this way, a number of pathways, in particular those active in catabolism such as the glycolytic pathway, were reconstituted. Biosynthetic pathways were often absent or incomplete, as the parasitic microorganism can obtain a number of building blocks from its host.

For the case of the sterol biosynthesis pathway (Schroepfer, 1981), we obtained an interesting result. It has been reported that *E. histolytica* in culture needs cholesterol for growth (Reeves, 1984). In agreement with this finding there is no complete sterol biosynthesis pathway. No enzymes leading through mevalonate to isopentenyl- or dimethylallyl-diphosphate were found, also the mevalonate-independent methylerythritol 4-phosphate (MEP) pathway (Hunter et al., 2003; Rohmer et al., 1993) present in bacteria and plants was not detected. Unexpectedly, however, we found homologs of isopentenyldiphosphate isomerase (EC 5.3.3.2; 33.1% sequence identity to *Bacillus subtilis*, P50740) which catalyses the conversion between isopentenyldiphosphate and dimethylallyl diphosphate. This gene is one of the candidates of LGT from bacteria (Loftus et al., 2005). The two isomeric isoprenyl diphosphates undergo condensation to geranyl diphosphate, and to farnesyldiphosphate by addition of a further molecule of isopentenyl diphosphate. Both activities, dimethylallyltransferase (EC 2.5.1.1) and geranyl isopentenyldiphosphate transferase (EC 2.5.1.10) reside in a single polypeptide chain with 33.3% identity to *Methanococcus jannaschii* (Q58270). The gene coding for this bifunctional enzyme is also in the LGT candidate list. No presqualene synthase (EC 2.5.1.21) was detected, but a geranylgeranyldiphosphate synthase (EC 2.5.1.29). So the amoebae appear to be able to produce the C_{10}, C_{15}, and C_{20} isoprenoid compounds but not the C_{30} cholesterol precursor.

Another vitally important function of farnesyl- and geranylgeranyl-residues is to modify the hydrophobicity of proteins, e.g., for their interaction with membranes (Grünler et al., 1994). Indeed the genome contains sequences for the α- and β-chains of a protein farnesyltransferase (EC 2.5.1.58), which have been cloned and characterized on the recombinant protein level (Kumagai et al., 2004). Protein farnesylation is an important novel target of anti-parasitic chemotherapy (Maurer-Stroh et al., 2003) including anti-*E. histolytica* chemotherapy (Ghosh et al., 2004). Besides the protein farnesyltransferase, there are two genes encoding the identical sequence of a putative Type I protein geranylgeranyltransferase (EC 2.5.1.59) β chain and the α- and β-chains of a putative Type II (Rab-) geranylgeranyltransferase (EC 2.5.1.60). In addition, there are putative homologs of the CAAX prenyl protease (EC 3.4.24.84) and the prenyl cysteine carboxyl methyltransferase (EC 2.1.1.100) which further modify prenylated proteins.

So the described pathway from isopentenyldiphosphate may mainly serve to provide precursors for protein prenylation. Since cholesterol biosynthesis cannot be completed, the amoebae depend on cholesterol from the host. Currently the question is open, however, how the amoebae may acquire the isopentenyldiphosphate. Certainly more biochemical work is needed to understand these data.

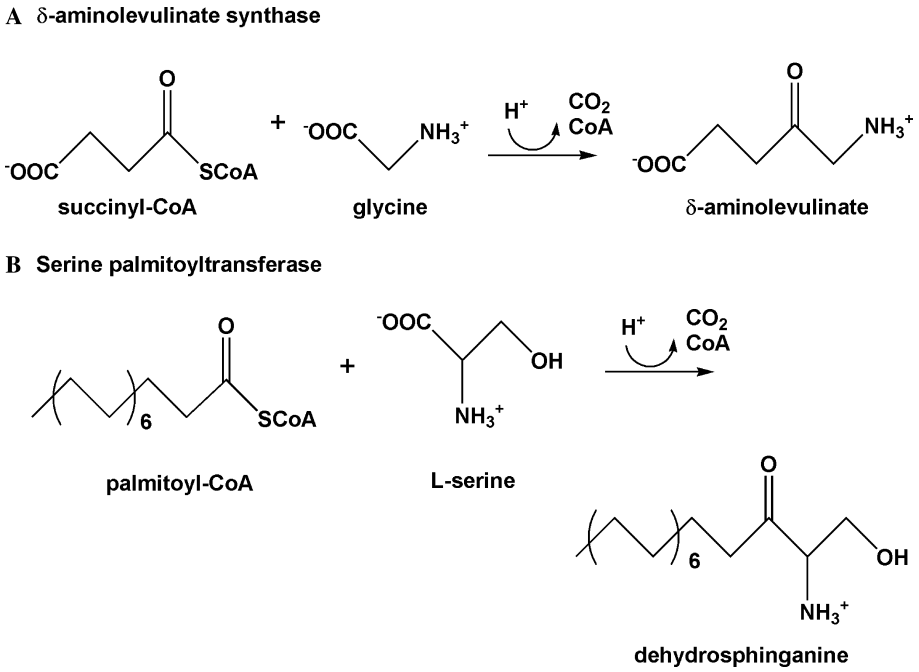


Fig. 1. (A) δ-Levulinate synthase (EC 2.3.1.37) is involved in heme biosynthesis and (B) serine palmitoyltransferase (EC 2.3.1.50) in sphingolipid biosynthesis. These two enzymes catalyze related reactions, their sequences, however, are only distantly related in humans. In contrast, there is only one homologous enzyme in *E. histolytica*, so far with unknown substrate specificity.

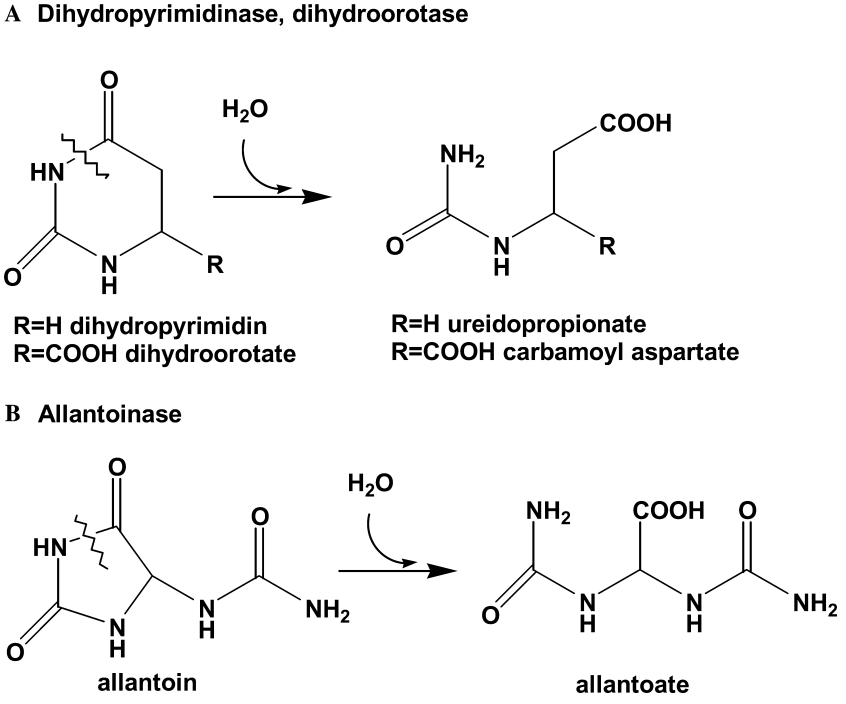


Fig. 2. The enzymes (A) dihydropyrimidinase (EC 3.5.2.2) and dihydroorotase (EC 3.5.2.3) and (B) allantoinase (EC 3.5.2.5) all catalyze the cleavage of amide bonds in different molecules, in *E. histolytica* there is only one homolog with unknown substrate specificity.

3.3. Unexpected examples of merged entries

In humans the enzyme δ-levulinate synthase is involved in the early steps of heme biosynthesis, whereas serine palmitoyltransferase is required for the synthesis of sphingolipids (Merrill, 2002). Although, at the first glance, the two enzymes appear to be unrelated, Fig. 1 shows that they both catalyze the head-to-head condensation of an amino acid with an acyl-CoA species liberating CoA and CO_2. In addition both use the cofactor pyridoxal phosphate. In humans δ-levulinate synthase (EC 2.3.1.37; P22557) and serine palmitoyltransferase (EC 2.3.1.50; O15269) are only 22.3% identical on the amino acid level, they share 27.7 and 25.4% sequence identities, respectively, with their closest *E. histolytica* homolog. Surprisingly, these two homologs are identical, there is only one gene in *E. histolytica*.

A further example shown in Fig. 2 is the case of the amide bond-cleaving enzymes dihydropyrimidinase (EC 3.5.2.2), dihydroorotase (EC 3.5.2.3), and allantoinase (EC 3.5.2.5). Searching for all three enzyme homologs is successful, but there is only one *E. histolytica* sequence. At this point the substrate specificity of this gene product is not known as is the case for the putative δ-levulinate synthase/serine palmitoyltransferase. In each of these cases more biochemical work with isolated or recombinant enzymes is required.

The presented work clearly demonstrates the evolving character of the genome project as well as the opportunities gained from it. Of course databases like ours always need to be modified because of new findings, e.g., closure of sequence gaps or characterization of the function of hypothetical genes. This permanent updating and curation will be essential to preserve the value of the databases. We hope that our funding will allow such work in the future.

Acknowledgments

This work was supported by Grant P15960 from the Austrian Science Fund, Vienna, Austria. We thank Martin Grabner (Vienna Biocenter EMBnet Austria) for help with bioinformatics. Preliminary sequence data was obtained from the Sanger Institute website at http://www.sanger.ac.uk and the Institute for Genomic Research (TIGR) website at http://www.tigr.org. Sequencing of *E. histolytica* was accomplished with support from the Wellcome Trust (Sanger Institute) and the National Institute of Allergy and Infectious Diseases (TIGR).

References

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J., 1990. Basic local alignment search tool. Journal of Molecular Biology 215, 403–410.

Croft, S.L., Seifert, K., Duchêne, M., 2003. Antiprotozoal activities of phospholipid analogues. Molecular and Biochemical Parasitology 126, 165–172.

Ghosh, S., Chan, J.M., Lea, C.R., Meints, G.A., Lewis, J.C., Tovian, Z.S., Flessner, R.M., Loftus, T.C., Bruchhaus, I., Kendrick, H., Croft, S.L., Kemp, R.G., Kobayashi, S., Nozaki, T., Oldfield, E., 2004. Effects of bisphosphonates on the growth of *Entamoeba histolytica* and *Plasmodium* species in vitro and in vivo. Journal of Medicinal Chemistry 47, 175–187.

Grünler, J., Ericsson, J., Dallner, G., 1994. Branch-point reactions in the biosynthesis of cholesterol, dolichol, ubiquinone and prenylated proteins. Biochimica et Biophysica Acta 1212, 259–277.

Hunter, W.N., Bond, C.S., Gabrielsen, M., Kemp, L.E., 2003. Structure and reactivity in the non-mevalonate pathway of isoprenoid biosynthesis. Biochemical Society Transactions 31, 537–542.

Kumagai, M., Makioka, A., Takeuchi, T., Nozaki, T., 2004. Molecular cloning and characterization of a protein farnesyltransferase from the enteric protozoan parasite *Entamoeba histolytica*. Journal of Biological Chemistry 279, 2316–2323.

Kumar, A., Shen, P.S., Descoteaux, S., Pohl, J., Bailey, G., Samuelson, J., 1992. Cloning and expression of an NADP(+)-dependent alcohol dehydrogenase gene of *Entamoeba histolytica*. Proceedings of the National Academy of Sciences USA 89, 10188–10192.

Loftus, B., Anderson, I., Davies, R., Alsmark, U.C.M., Samuelson, J., Amedeo, P., Roncaglia, P., Berriman, M., Hirt, R.P., Mann, B.J., Nozaki, T., Suh, B., Pop, M., Duchene, M., Ackers, J., Tannich, E., Leippe, M., Hofer, M., Bruchhaus, I., Willhoeft, U., Bhattacharya, A., Chillingworth, T., Churcher, C., Hance, Z., Harris, B., Harris, D., Jagels, K., Moule, S., Mungall, K., Ormond, D., Squares, R., Whitehead, S., Quail, M.A., Rabbinowitsch, E., Norbertczak, H., Price, C., Wang, Z., Guillén, N., Gilchrist, C., Stroup, S.E., Bhattacharya, S., Lohia, A., Foster, P.G., Sicheritz-Ponten, T., Weber, C., Singh, U., Mukherjee, C., El-Sayad, N.M., Petri Jr., W.A., Clark, C.G., Embley, T.M., Barrell, B., Fraser, C.M., Hall, N., 2005. The genome of the protist parasite *Entamoeba histolytica*. Nature 433, 865–868.

Maurer-Stroh, S., Washietl, S., Eisenhaber, F., 2003. Protein prenyltransferases: anchor size, pseudogenes and parasites. Biological Chemistry 384, 977–989.

Merrill Jr., A.H., 2002. De novo sphingolipid biosynthesis: a necessary, but dangerous, pathway. Journal of Biological Chemistry 277, 25843–25846.

Nielsen, H.B., Wernersson, R., Knudsen, S., 2003. Design of oligonucleotides for microarrays and perspectives for design of multi-transcriptome arrays. Nucleic Acids Research 31, 3491–3496.

Program Manual for the Wisconsin Package, Version 8, August 1994, Genetics Computer Group, 575 Science Drive, Madison, Wisconsin, USA 53711 and also for the Extension programs: Program Manual for the EGCG Package, Peter Rice, The Sanger Centre, Hinxton Hall, Cambridge, CB10 1RQ, England. Additional code by Peter Rice, The Sanger Centre, Hinxton, England and other members of the EGCG team.

Raymond-Denise, A., Sansonetti, P., Guillen, N., 1993. Identification and characterization of a myosin heavy chain gene (mhcA) from the human parasitic pathogen *Entamoeba histolytica*. Molecular and Biochemical Parasitology 59, 123–131.

Reeves, R.E., 1984. Metabolism of *Entamoeba histolytica* Schaudinn, 1903. Advances in Parasitology 23, 105–142.

Rohmer, M., Knani, M., Simonin, P., Sutter, B., Sahm, H., 1993. Isoprenoid biosynthesis in bacteria: a novel pathway for the early steps leading to isopentenyl diphosphate. Biochemical Journal 295, 517–524.

Rosenthal, B., Mai, Z., Caplivski, D., Ghosh, S., de la Vega, H., Graf, T., Samuelson, J., 1997. Evidence for the bacterial origin of genes encoding fermentation enzymes of the amitochondriate protozoan parasite *Entamoeba histolytica*. Journal of Bacteriology 179, 3736–3745.

Sanchez, L., Horner, D., Moore, D., Henze, K., Embley, T., Muller, M., 2002. Fructose-1, 6-bisphosphate aldolases in amitochondriate

protists constitute a single protein subfamily with eubacterial relationships. Gene 295, 51–59.
Schroepfer Jr., G.J., 1981. Sterol biosynthesis. Annual Review of Biochemistry 50, 585–621.
Wassmann, C., Hellberg, A., Tannich, E., Bruchhaus, I., 1999. Metronidazole resistance in the protozoan parasite *Entamoeba histolytica* is associated with increased expression of iron- containing superoxide dismutase and peroxiredoxin and decreased expression of ferredoxin 1 and flavin reductase. Journal of Biological Chemistry 274, 26051–26056.
World Health Organisation, 1997. Amebiasis. Weekly Epidemiological Records 72, 97–100.

3.2 The genome of the protist parasite *Entamoeba histolytica*.

Loftus B, Anderson I, Davies R, Alsmark UC, Samuelson J, Amedeo P, Roncaglia P, Berriman M, Hirt RP, Mann BJ, Nozaki T, Suh B, Pop M, Duchêne M, Ackers J, Tannich E, Leippe M, **Hofer M**, Bruchhaus I, Willhoeft U, Bhattacharya A, Chillingworth T, Churcher C, Hance Z, Harris B, Harris D, Jagels K, Moule S, Mungall K, Ormond D, Squares R, Whitehead S, Quail MA, Rabbinowitsch E, Norbertczak H, Price C, Wang Z, Guillén N, Gilchrist C, Stroup SE, Bhattacharya S, Lohia A, Foster PG, Sicheritz-Ponten T, Weber C, Singh U, Mukherjee C, El-Sayed NM, Petri WA Jr, Clark CG, Embley TM, Barrell B, Fraser CM, Hall N.

Nature. 2005; 433: 865-868.

The genome of the protist parasite *Entamoeba histolytica*

Brendan Loftus[1], Iain Anderson[1], Rob Davies[2], U. Cecilia M. Alsmark[3], John Samuelson[4], Paolo Amedeo[1], Paola Roncaglia[1], Matt Berriman[2], Robert P. Hirt[3], Barbara J. Mann[5], Tomo Nozaki[6], Bernard Suh[1], Mihai Pop[1], Michael Duchene[7], John Ackers[8], Egbert Tannich[9], Matthias Leippe[10], Margit Hofer[7], Iris Bruchhaus[9], Ute Willhoeft[9], Alok Bhattacharya[11], Tracey Chillingworth[2], Carol Churcher[2], Zahra Hance[2], Barbara Harris[2], David Harris[2], Kay Jagels[2], Sharon Moule[2], Karen Mungall[2], Doug Ormond[2], Rob Squares[2], Sally Whitehead[2], Michael A. Quail[2], Ester Rabbinowitsch[2], Halina Norbertczak[2], Claire Price[2], Zheng Wang[1], Nancy Guillén[12], Carol Gilchrist[5], Suzanne E. Stroup[5], Sudha Bhattacharya[11], Anuradha Lohia[13], Peter G. Foster[14], Thomas Sicheritz-Ponten[15], Christian Weber[12], Upinder Singh[16], Chandrama Mukherjee[13], Najib M. El-Sayed[1], William A. Petri Jr[5], C. Graham Clark[8], T. Martin Embley[3], Bart Barrell[2], Claire M. Fraser[1] & Neil Hall[2]*

[1]*TIGR, 9712 Medical Center Drive, Rockville, Maryland 20850, USA*
[2]*The Sanger Institute, The Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA, UK*
[3]*School of Biology, University of Newcastle, King George VI Building, Newcastle upon Tyne NE1 7RU, UK*
[4]*Department of Molecular and Cell Biology, Boston University Goldman School of Dental Medicine, 715 Albany Street, Boston, Massachusetts 02118, USA*
[5]*Departments of Internal Medicine & Microbiology, University of Virginia, Charlottesville, Virginia 22908, USA*
[6]*Department of Parasitology, National Institute of Infectious Diseases, 1-23-1 Toyama, Shinjuku-ku, Tokyo 162-8640, Japan*
[7]*Division of Specific Prophylaxis and Tropical Medicine, Center for Physiology and Pathophysiology, Medical University of Vienna, Kinderspitalgasse 15, A-1095 Vienna, Austria*
[8]*Department of Infectious and Tropical Diseases, London School of Hygiene and Tropical Medicine, Keppel Street, London WC1E 7HT, UK*
[9]*Department of Molecular Parasitology, Bernhard Nocht Institute for Tropical Medicine, Bernhard Nocht Str. 74, 20359 Hamburg, Germany*
[10]*Zoological Institute, University of Kiel, Olshausenstr. 40, 24098 Kiel, Germany*
[11]*School of Environmental Sciences, Jawaharlal Nehru University, New Delhi 110067, India*
[12]*Unite de Biologie Cellulaire du Parasitisme, INSERM U389, Institut Pasteur 28, rue du Dr Roux 75724, Paris Cedex 15, France*
[13]*Department of Biochemistry, Bose Institute, P1/12 CIT Scheme VIIM, Kolkata 700054, India*
[14]*Department of Zoology, The Natural History Museum, Cromwell Road, London SW7 5BD, UK*
[15]*Center for Biological Sequence Analysis, Technical University of Denmark, Building 208, DK-2800 Lyngby, Denmark*
[16]*Departments of Internal Medicine, Microbiology, and Immunology, Stanford University School of Medicine, Stanford, California 94305-5107, USA*

* Present address: TIGR, 9712 Medical Center Drive, Rockville, Maryland 20850, USA

Entamoeba histolytica **is an intestinal parasite and the causative agent of amoebiasis, which is a significant source of morbidity and mortality in developing countries[1]. Here we present the genome of *E. histolytica*, which reveals a variety of metabolic adaptations shared with two other amitochondrial protist pathogens: *Giardia lamblia* and *Trichomonas vaginalis*. These adaptations include reduction or elimination of most mitochondrial metabolic pathways and the use of oxidative stress enzymes generally associated with anaerobic prokaryotes. Phylogenomic analysis identifies evidence for lateral gene transfer of bacterial genes into the *E. histolytica* genome, the effects of which centre on expanding aspects of *E. histolytica*'s metabolic repertoire. The presence of these genes and the potential for novel metabolic pathways in *E. histolytica* may allow for the development of new chemotherapeutic agents. The genome encodes a large number of novel receptor kinases and contains expansions of a variety of gene families, including those associated with virulence. Additional genome features include an abundance of tandemly repeated transfer-RNA-containing arrays, which may have a structural function in the genome. Analysis of the genome provides new insights into the workings and genome evolution of a major human pathogen.**

Genome analysis was carried out on a 12.5-fold coverage genome assembly consisting of 23,751,783 base pairs (bp) distributed among 888 scaffolds. The 9,938 predicted genes average 1.17 kilobases (kb) in size and comprise 49% of the genome. One-quarter of *E. histolytica* genes are predicted to contain introns, with 6% of genes containing multiple introns. No homologues could be identified for a third of predicted proteins (31.8%) from the public databases (see Methods). *E. histolytica* chromosomes do not condense, and the uncertainty surrounding its ploidy and the extensive length variability observed between homologous chromosomes from different isolates makes the exact chromosome number difficult to determine. The chromosome size variation observed may be due to expansion and contraction of subtelomeric repeats, as in other protists[2,3], and it is tempting to speculate that in *E. histolytica* these regions consist of tRNA-containing arrays. Comprising almost 10% of the sequence reads, 25 types of long tandem array, each containing between one and five tRNA types per repeat unit, could be identified from the genome data. The full complement of tRNAs required for translation has been identified, and all but four of the tRNA genes are encoded exclusively in arrays. These unique tRNA gene arrays are thus predicted to be functional as well as potentially fulfilling a structural role in the genome. No association could be determined between codon usage and the relative copy numbers of their cognate tRNA species.

The metabolism of *E. histolytica* seems to have been shaped by secondary gene loss and lateral gene transfer (LGT), primarily from bacterial lineages (Fig. 1). *E. histolytica* is an obligate fermenter, using bacterial-like fermentation enzymes and lacking proteins of the tricarboxylic acid cycle and mitochondrial electron transport chain. An atrophic, mitochondrion-derived organelle has been identified in *E. histolytica*[4], and the genome data support the absence of a mitochondrial genome. Glucose is the main energy source; however, in place of the typical eukaryotic glucose transporters those of *E. histolytica* are related to the prokaryote glucose/ribose porter family, with the amino- and carboxy-terminal domains switched relative to their prokaryotic counterparts.

As a phagocytic resident of the human gut, *E. histolytica* has access to many bacterial and host-derived preformed organic compounds. Most pathways for amino acid biosynthesis have been eliminated, except those for serine and cysteine, which are probably retained for the production of cysteine, the major intracellular thiol. The high levels of cysteine in *E. histolytica* may compensate for the lack of glutathione and its associated enzymes, a major component of oxidative stress resistance in many organisms[5]. *E. histolytica* lacks *de novo* purine, pyrimidine and thymidylate synthesis and must rely on salvage pathways, similar to *G. lamblia* and *T. vaginalis*[6]. In addition, *E. histolytica* appears to lack ribonucleotide reductase, a characteristic that it shares with *G. lamblia*[7]. *E. histolytica* is unable to synthesize fatty acids but retains the ability to synthesize a variety of phospholipids. The absence of identifiable pathways for the synthesis of isoprenoids and the sphingolipid head group aminoethylphosphonate suggest the existence of novel pathways. These pathways, once characterized, might represent attractive drug targets. Two unusual enzymes of fatty acid elongation are shared between *E. histolytica* and *G. lamblia*, including a predicted acetyl-CoA carboxylase with two carboxyltransferase domains[8]. We propose that this enzyme removes a carboxyl group from oxaloacetate and transfers it to acetyl-CoA to form malonyl-CoA and pyruvate. *E. histolytica* also has five members of a fatty acid elongase family, previously identified only in plants, green algae and *G. lamblia*[9,10]. Folate is

a cofactor essential for thymidylate synthesis and methionine recycling, and genome analysis reveals a complete lack of genes coding for known folate-dependent enzymes and folate transporters. Folate is also required for organelle protein synthesis in mitochondria and chloroplasts, and loss of the mitochondrial genome may have paved the way for the loss of these folate-dependent functions.

LGT is an important force in the evolution of prokaryotes but significantly less is known about its importance in eukaryotic evolution[11]. We conducted a phylogenetic screen of the *Entamoeba* genome for cases of relatively recent prokaryote to eukaryote LGT (see Methods), and for 96 genes we believe that this is the simplest explanation for the tree topologies obtained (see Supplementary Information). These genes are embedded among typically eukaryotic genes on *E. histolytica* scaffolds and do not seem to represent contaminating prokaryotic sequences. Most (58%) of the LGT genes encode a variety of metabolic enzymes, whereas most of the remaining genes (41%) encode proteins of unknown function (Supplementary Fig. 1). The major impact is in the area of carbohydrate and amino acid metabolism, where they have increased the range of substrates available for energy generation including tryptophanase and aspartase, which contribute to the use of amino acids. Several glycosidases and sugar kinases appear to have been acquired through LGT and would probably enable *E. histolytica* to use sugars other than glucose; for example, fructose and galactose. There is a strong bias in the data for a major donor being in the *Cytophaga–Flavobacterium–Bacteroides* (CFB) group of the phylum Bacteroidetes; however, this should be interpreted with caution, as current sampling of prokaryotic genomes is still relatively incomplete. It is clear that among the 96 genes, some result in significant enhancements to *E. histolytica* metabolism, thus contributing to its biology to a greater extent than indicated by the numbers alone.

E. histolytica feed on bacteria in the lumen of the colon and lyse host epithelial cells after invasion of the intestinal wall[12]. A number of amoebic virulence determinants have been characterized, including a multi-subunit GalGalNAc lectin involved in adhesion to host cells, cysteine proteases that degrade host extracellular matrix, and pore-forming peptides (amoebapores) capable of lysing target cells[12]. Analysis of the genome reveals redundancy in the genes encoding these virulence factors. Thirty homologues of the intermediate subunit and one homologue of the heavy subunit of the GalGalNAc lectin were identified. Ten new cysteine proteinases with predicted N-terminal transmembrane anchors, which might allow them to be localized on the amoeba cell surface, were identified. In addition to three new amoebapores a homologue of haemolysin III was identified, suggesting that, in addition to amoebapores, haemolysins may have a role in host cell lysis.

Vesicle trafficking has a role in *E. histolytica* pathogenesis through phagocytosis and the delivery of secreted hydrolytic enzymes and amoebapores to the cell surface[13]. *E. histolytica* lacks morphologically identifiable rough endoplasmic reticulum and the Golgi apparatus[14] but encodes the basic elements of the vesicle transport machinery common to other eukaryotic cells, with the coat complexes COPI, COPII, clathrin and retromer all being present. Rab and Arf protein family expansions reflect the increased complexity and number of vesicle fusion and recycling steps that have been associated with phagocytosis and pinocytosis in amoebae[15]. The cytoskeleton has a number of important roles in parasite motility, contact-dependant killing and phagocytosis of host intestinal epithelial cells[16]. This is reflected in expansions of Rho GTPases and their regulators RhoGAPs and RhoGEFs, which control a number of processes involving the actin cytoskeleton. Five proteins with a unique domain architecture containing both RhoGEF and Arf-GAP domains were identified, suggesting a mechanism for direct

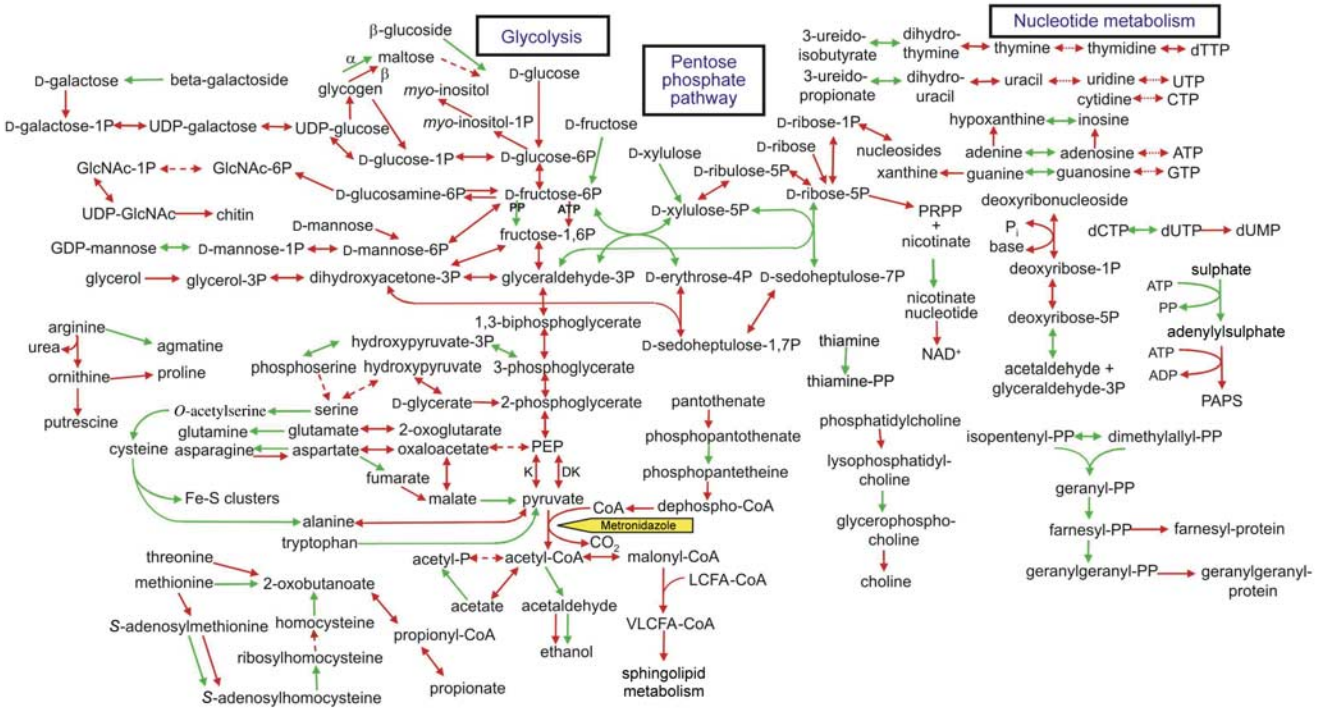


Figure 1 Predicted metabolism of *E. histolytica* based on analysis of the genome sequence data. Arrows indicate enzyme reactions. Glycolysis and fermentation are the major energy generation pathways. Green arrows represent enzymes encoded by genes that are among the 96 candidates for LGT into the *E. histolytica* genome. Broken arrows indicate enzymes for which no gene could be identified using searches of the genome data, although the activity is likely to be present. The yellow arrow points to the source of electrons for activation of metronidazole, the major drug for treatment of amoebic liver abscess. DK, pyruvate phosphate dikinase; GlcNAc, *N*-acetylglucosamine; GPI, glycosylphosphatidylinositol; K, pyruvate kinase; LCFA, long-chain fatty acid; PAPS, phosphoadenosine phosphosulphate; PEP, phosphoenolpyruvate; PP, pyrophosphate; PRPP, phosphoribosyl pyrophosphate; VLCFA, very-long-chain fatty acid.

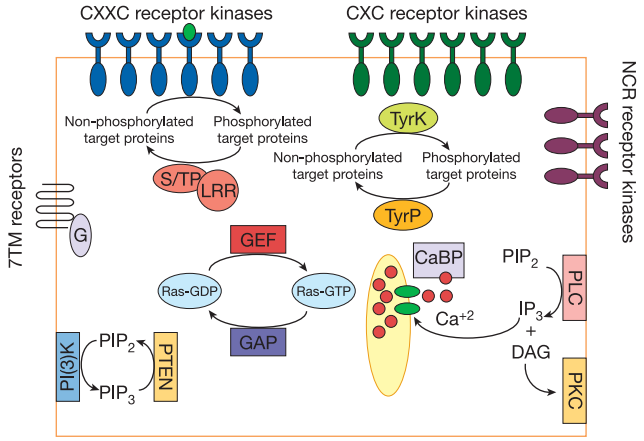


Figure 2 Predicted signal transduction mechanisms of *E. histolytica* based on analysis of the genome sequence data. *E. histolytica* possesses three types of receptor serine/threonine kinases: one group has CXXC repeats in the extracellular domain; a second has CXC repeats; and a third has non-cysteine rich (NCR) repeats. *E. histolytica* has cytosolic tyrosine kinases (TyrK), but not receptor tyrosine kinases. Some serine/threonine phosphatases (S/TP) have an attached LRR domain. CaBP, calcium-binding protein; DAG, diacylglycerol; G, G protein; GAP, GTPase-activating protein; GEF, guanine nucleotide exchange factor; IP_3, inositol-1,4,5-trisphosphate; PI(3)K, phosphatidylinositol-3-OH kinase; PIP_2, phosphatidylinositol-4,5-bisphosphate; PIP_3, phosphatidylinositol-3,4,5-trisphosphate; PKC, protein kinase C; PLC, phospholipase C; PTEN, phosphatase and tensin homologue; TyrP, tyrosine phosphatase; 7TM receptors, seven-transmembrane receptors.

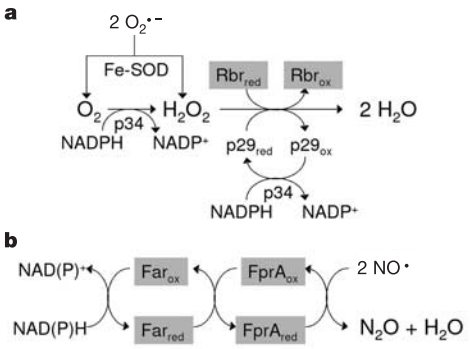


Figure 3 Predicted pathways for oxidative and nitrosative stress resistance in *E. histolytica*. Enzymes boxed and shaded have previously only been identified in anaerobic prokaryotes and amitochondrial protists. **a**, Superoxide is detoxified by an iron-containing superoxide dismutase (Fe-SOD). Molecular oxygen is reduced to hydrogen peroxide by the NADPH-flavin oxidoreductase (p34), which also transfers electrons to peroxiredoxin (p29). Rubrerythrin (Rbr) is predicted to convert hydrogen peroxide to water, although the source of electrons for rubrerythrin in *E. histolytica* is unknown. **b**, A-type flavoproteins (FprA) detoxify nitric oxide to nitrous oxide. FprA receives electrons from flavoprotein A reductase (Far).

communication between the regulators of vesicle budding and cytoskeletal rearrangement.

E. histolytica uses a complex mix of signal transduction systems in order to sense and interact with the different environments it encounters (Fig. 2). Almost 270 putative *E. histolytica* protein kinases representing members of all seven families of the eukaryotic protein kinase superfamily were identified[17]. These include tyrosine kinases with SH2 domains, tyrosine kinase-like protein kinases and 90 putative receptor Ser/Thr kinases. These Ser/Thr kinases are uncommon in protists, appear to be absent from *Dictyostelium* and have previously been described only in plants, animals and Choanoflagellates. The *E. histolytica* receptor Ser/Thr kinases all contain an N-terminal signal peptide, a predicted extracellular domain and a single transmembrane helix followed by a cytosolic tyrosine kinase-like domain. The receptor kinases fall into three groups on the basis of differences in their predicted extracellular domains. The first group of 50 receptor kinase proteins contains CXXC-rich repeats similar to those found in the intermediate subunit (Igl) of the Gal/GalNAc lectin and *G. lamblia* variant-specific surface proteins. A second group of 32 proteins encodes cysteine-rich domains containing CXC repeats. The third group of eight receptor kinase-like proteins lacks cysteine-rich extracellular domains. Although no immediate downstream effectors to the amoebic receptor kinases could be identified, *E. histolytica* contains greater than 100 protein phosphatases, which dephosphorylate proteins. An unusual feature of some of the phosphatases is the presence of varying numbers of leucine-rich repeat (LRR) domains that are involved primarily in protein–protein interactions and have not previously been associated with phosphatases. The *E. histolytica* genome encodes numerous putative seven-transmembrane receptors and trimeric G proteins, which are probably involved in mediating autocrine stimulation of encystation[18]. In contrast to autocrine stimulation of *Dictyostelium* sporulation, which uses secreted cyclic AMP, *E. histolytica* encystment is self-stimulated by secreted catecholamines[18]. Finally, *E. histolytica* has numerous cytosolic proteins involved in signal transduction, including Ras-family proteins, EF-hand calcium-binding proteins, phosphatidylinositol-3-OH kinase and MAP kinases. This represents the most varied set of signal-transduction-related proteins yet described in a single-celled eukaryote.

In contrast to life in the anoxic colon, *E. histolytica* encounters a relatively high-oxygen environment during invasive amoebiasis, and coping with this change is therefore an important virulence factor. The importance of this response is underscored by the redundancy of oxygen detoxification mechanisms. *E. histolytica* has four copies of flavoprotein A, which detoxifies nitric oxide and/or oxygen[19] (Fig. 3), and also contains rubrerythrin, which in anaerobic bacteria is protective against intracellular hydrogen peroxide[20] (Fig. 3). These oxidative and/or nitrosative stress resistance genes are shared with *G. lamblia* (with the exception of rubrerythrin) and *T. vaginalis*, but have generally been associated with anaerobic prokaryotes (Fig. 3).

E. histolytica is the first amoeba genome to be fully sampled, and comparisons with other genomes will assist in resolving fundamental issues relating to eukaryote and amoeba phylogeny, as well as how LGT affects eukaryotes. Despite a lack of representative genome sampling from amitochondrial protist lineages it is already clear that these unrelated anaerobic eukaryotes seem to use convergent metabolic strategies imposed by their environments. As a first insight into an amitochondrial protist genome, analysis of these data and particularly the bacterial-like proteins contained therein should illuminate future efforts aimed at the development of diagnostics and therapeutics of these luminal parasites. □

Methods

Genome sequencing and assembly

The *E. histolytica* genome sequence was generated by the whole-genome shotgun method. As the chromosomes of *E. histolytica* could not be resolved by pulsed field gel electrophoresis (PFGE) and the A + T content precluded making large or medium insert libraries in bacterial artificial chromosomes (BACs), we were required to use the whole-genome shotgun approach to sequence the genome. Genomic DNA was prepared from *E. histolytica* strain HM-1:IMSS (ATCC number 30459) grown axenically in TYI-S-33 medium[20]. At TIGR 390,000 reads were produced from a small (1.5–2.0 kb) and a medium insert library (8–10 kb) generated in the pHOS2 vector. At the Sanger Institute, 200,000 reads were generated from a pUC18 library with average insert size of 2.5 kb plus 6,500 reads from a BAC library with an average insert size of 10 kb (the high A + T content of the genomic DNA prevented cloning of larger fragments). To avoid assembly problems, reads containing episomal-derived rDNA or tRNA-containing sequences (170,000 reads (29%)) were excluded from the whole-genome assembly process. The average edited read length was 645 bp, giving an approximate 12.5-fold genome coverage. Genome assembly was carried out at the Sanger Centre using the program phusion[21]. All scaffolds smaller than

2 kb (327) were subsequently removed, leaving 1,425 scaffolds with a combined size of 25,393,225 bp. The remaining scaffolds were analysed to remove redundancy that may have resulted as a consequence of allelic differences or aneuploidy. We removed all scaffolds smaller than 5 kb that shared 98% or more nucleotide sequence identity over greater than 95% of their lengths. Removal of these scaffolds left 888 scaffolds remaining, with a total length of 23,751,783 bp. All scaffolds removed during the clean-up process as well as any singleton reads, although not used in the annotation process, were used in determining the presence or absence of genes in the *E. histolytica* genome. Unfortunately, there is no map to order the scaffolds generated by the assembly; however, the sequence generated by this project should assist in making maps for this genome in the future, and although the large-scale structure of the genome has been lost, the vast majority of the genes that have been predicted are full length with intact 3′ and 5′ untranslated regions.

Annotation

The Combiner algorithm was used for gene structure identification[22] using two genefinder programs, phat[23] and GlimmerHMM[24], trained using a set of published *E. histolytica* gene sequences, alignments of protein homologues to the genomic sequence and alignment of a set of *E. histolytica* complementary DNA sequences (provided by N. Guillén) to the genomic sequence. The Combiner gene predictions were then manually curated. Functional annotations for the predicted proteins were automatically generated using a combination of numerous sources of evidence including searches against a non-redundant protein database and identification of functional domains by searches against the Pfam database[25]. tRNAs were detected using the tRNAscan-SE[26] program with default parameters.

Identification of sequence homologues in other species

Sequence homologues from other species were identified by searching the predicted proteins from the *E. histolytica* genome against the publicly available nr database of GenBank using BlastP (http://www.ncbi.nlm.nih.gov/BLAST/) and filtering search results with an *e*-value of 10^{-5} or less, which was chosen because of the relatively large divergence between *E. histolytica* and other organisms for which the genomes have been sequenced and for which protein data are available.

Phylogenetic analysis

We modified a published suite of scripts and modules called PyPhy[27] to make an automated genome-wide primary screen for LGT. PyPhy was used to make bootstrap (100 replicates) consensus *p*-distance trees from edited alignments of 5,740 *E. histolytica* proteins; that is, those for which there were sufficient homologues (>4) in SwissProt and TrEMBL to make trees. The trees were analysed to identify cases where the nearest neighbour to the *E. histolytica* protein was a prokaryotic sequence. As an additional screen for LGT we identified all proteins for which a prokaryote was the top Blast hit. After manual inspection of the alignments, Blast outputs, tree support values and sequence identities, 279 cases of potential LGT were retained for more detailed phylogenetic analyses. Each candidate LGT was analysed by MrBayes[28] using the WAG matrix, a gamma correction for site rate variation and a proportion (pinvar) of invariant sites. The analyses were run for 600,000 generations and sampled every 100 generations, with the first 2,000 samples discarded as burn-in. A consensus tree was made from the remaining samples. Because posterior probabilities—the support values used by bayesian analysis to indicate confidence in groups—have been criticized[29], we also used bootstrapping to provide an additional indication of support for relationships. Each data set was bootstrapped (100 replicates) and used to make distance matrices under the same evolutionary model as in the bayesian analysis, using custom (P4) software (available on request). Trees were made from the distance matrices using FastME[30] and a bootstrap consensus tree made using P4. On the basis of these analyses we identified 96 genes in which the tree topology is consistent with prokaryote to eukaryote LGT. Blast summary statistics, trees and support values for these 96 candidate LGT are provided as Supplementary Information.

Received 26 October; accepted 2 December 2004; doi:10.1038/nature03291.

1. Stanley, S. L. Jr Amoebiasis. *Lancet* **361,** 1025–1034 (2003).
2. Pataraportikul, J. & Langsley, G. Chromosome size polymorphism in *Plasmodium falciparum* can involve deletions of the subtelomeric pPFrep20 sequence. *Nucleic Acids Res.* **16,** 4331–4340 (1988).
3. Melville, S. E., Gerrard, C. S. & Blackwell, J. M. Multiple causes of size variation in the diploid megabase chromosomes of African trypanosomes. *Chromosome Res.* **7,** 191–203 (1999).
4. Leon-Avila, G. & Tovar, J. Mitosomes of *Entamoeba histolytica* are abundant mitochondrion-related remnant organelles that lack a detectable organellar genome. *Microbiology* **150,** 1245–1250 (2004).
5. Fahey, R. C., Newton, G. L., Arrick, B., Overdank-Bogart, T. & Aley, S. B. *Entamoeba histolytica*: a eukaryote without glutathione metabolism. *Science* **224,** 70–72 (1984).
6. Abrahamsen, M. S. *et al.* Complete genome sequence of the apicomplexan, *Cryptosporidium parvum*. *Science* **304,** 441–445 (2004).
7. Baum, K. F., Berens, R. L., Marr, J. J., Harrington, J. A. & Spector, T. Purine deoxynucleoside salvage in *Giardia lamblia*. *J. Biol. Chem.* **264,** 21087–21090 (1989).
8. Jordan, I. K., Henze, K., Fedorova, N. D., Koonin, E. V. & Galperin, M. Y. Phylogenomic analysis of the *Giardia intestinalis* transcarboxylase reveals multiple instances of domain fusion and fission in the evolution of biotin-dependent enzymes. *J. Mol. Microbiol. Biotechnol.* **5,** 172–189 (2003).
9. James, D. W. Jr *et al.* Directed tagging of the *Arabidopsis* fatty acid elongation1 (FAE1) gene with the maize transposon activator. *Plant Cell* **7,** 309–319 (1995).
10. Azachi, M. *et al.* Salt induction of fatty acid elongase and membrane lipid modifications in the extreme halotolerant alga *Dunaliella salina*. *Plant Physiol.* **129,** 1320–1329 (2002).
11. Lawrence, J. G. & Hendrickson, H. Lateral gene transfer: when will adolescence end? *Mol. Microbiol.* **50,** 739–749 (2003).
12. Huston, C. D. Parasite and host contributions to the pathogenesis of amebic colitis. *Trends Parasitol.* **20,** 23–26 (2004).
13. Welter, B. H. & Temesvari, L. A. A unique Rab GTPase, EhRabA, of *Entamoeba histolytica*, localizes to the leading edge of motile cells. *Mol. Biochem. Parasitol.* **135,** 185–195 (2004).
14. Mazzuco, A., Benchimol, M. & De Souza, W. Endoplasmic reticulum and Golgi-like elements in *Entamoeba*. *Micron* **28,** 241–247 (1997).
15. Duhon, D. & Cardelli, J. The regulation of phagosome maturation in *Dictyostelium*. *J. Muscle Res. Cell Motil.* **23,** 803–808 (2002).
16. Voigt, H. & Guillen, N. New insights into the role of the cytoskeleton in phagocytosis of *Entamoeba histolytica*. *Cell. Microbiol.* **1,** 195–203 (1999).
17. Hunter, T. Protein kinase classification. *Methods Enzymol.* **200,** 3–37 (1991).
18. Coppi, A., Merali, S. & Eichinger, D. The enteric parasite *Entamoeba* uses an autocrine catecholamine system during differentiation into the infectious cyst stage. *J. Biol. Chem.* **277,** 8083–8090 (2002).
19. Gomes, C. M. *et al.* A novel type of nitric-oxide reductase. *Escherichia coli* flavorubredoxin. *J. Biol. Chem.* **277,** 25273–25276 (2002).
20. Sztukowska, M., Bugno, M., Potempa, J., Travis, J. & Kurtz, D. M. Jr Role of rubrerythrin in the oxidative stress response of *Porphyromonas gingivalis*. *Mol. Microbiol.* **44,** 479–488 (2002).
21. Mullikin, J. C. & Ning, Z. The phusion assembler. *Genome Res.* **13,** 81–90 (2003).
22. Allen, J. E., Pertea, M. & Salzberg, S. L. Computational gene prediction using multiple sources of evidence. *Genome Res.* **14,** 142–148 (2004).
23. Cawley, S. E., Wirth, A. I. & Speed, T. P. Phat–a gene finding program for *Plasmodium falciparum*. *Mol. Biochem. Parasitol.* **118,** 167–174 (2001).
24. Majoros, W. H., Pertea, M. & Salzberg, S. L. TigrScan and GlimmerHMM: two open-source ab initio eukaryotic gene-finders. *Bioinformatics* **20,** 2878–2879 (2004).
25. Bateman, A. *et al.* The Pfam protein families database. *Nucleic Acids Res.* **32,** D138–D141 (2004).
26. Lowe, T. M. & Eddy, S. R. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res.* **25,** 955–964 (1997).
27. Sicheritz-Ponten, T. & Andersson, S. G. A phylogenomic approach to microbial evolution. *Nucleic Acids Res.* **29,** 545–552 (2001).
28. Huelsenbeck, J. P. & Ronquist, F. MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics* **17,** 754–755 (2001).
29. Cummings, M. P. *et al.* Comparing bootstrap and posterior probability values in the four-taxon case. *Syst. Biol.* **52,** 477–487 (2003).
30. Desper, R. & Gascuel, O. Theoretical foundation of the balanced minimum evolution method of phylogenetic inference and its relationship to weighted least-squares tree fitting. *Mol. Biol. Evol.* **21,** 587–598 (2004).

Supplementary Information accompanies the paper on **www.nature.com/nature**.

Acknowledgements This work was supported by grants from the National Institute of Allergy and Infectious Disease and the Wellcome Trust.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to B.L. (bjloftus@tigr.org). Scaffold sequences have been deposited in GenBank under the project accession number AAFB00000000. Sequences and annotation are available at http://www.tigr.org/tdb/e2k1/eha1/.

Excitatory cortical neurons form fine-scale functional networks

Yumiko Yoshimura*, Jami L. M. Dantzker* & Edward M. Callaway

Systems Neurobiology Laboratories, The Salk Institute for Biological Studies, 10010 North Torrey Pines Road, La Jolla, California 92037, USA

* Present addresses: Department of Visual Neuroscience, Research Institute of Environmental Medicine, Nagoya University, Furo-cho, Chikusa-ku, Nagoya 464-8601, Japan (Y.Y.); Department of Neurology and Neurological Sciences, Stanford University, 300 Pasteur Drive, Room M016, Stanford, California 94305-5122, USA (J.L.M.D.)

The specificity of cortical neuron connections creates columns of functionally similar neurons spanning from the pia to the white matter[1–6]. Here we investigate whether there is an additional, finer level of specificity that creates subnetworks of excitatory neurons within functional columns. We tested for fine-scale specificity of connections to cortical layer 2/3 pyramidal neurons in rat visual cortex by using cross-correlation analyses of synaptic currents evoked by photostimulation. Recording simultaneously from adjacent layer 2/3 pyramidal cells, we find that when they are connected to each other (20% of all recorded pairs) they share common input from layer 4 and within layer 2/3. When adjacent layer 2/3 neurons are not connected to each other, they share very little (if any) common excitatory input from layers 4 and 2/3. In contrast, all layer 2/3 neurons share common excitatory input

3.3 Structure and content of the *Entamoeba histolytica* genome.

Clark CG, Alsmark UC, **Tazreiter M**, Saito-Nakano Y, Ali V, Marion S, Weber C, Mukherjee C, Bruchhaus I, Tannich E, Leippe M, Sicheritz-Ponten T, Foster PG, Samuelson J, Noël CJ, Hirt RP, Embley TM, Gilchrist CA, Mann BJ, Singh U, Ackers JP, Bhattacharya S, Bhattacharya A, Lohia A, Guillén N, Duchêne M, Nozaki T, Hall N.

Adv Parasitol. 2007; 65: 51-190.

Note:

Due to the great length of this review article it was impossible to include the whole manuscript in this dissertation. We decided to show the sections 4.1 to 4.6 on metabolism contributed by our group with the exception of section 4.4. We also included the Introduction.

CHAPTER 2

Structure and Content of the *Entamoeba histolytica* Genome

C. G. Clark,[*] U. C. M. Alsmark,[†] M. Tazreiter,[‡] Y. Saito-Nakano,[§] V. Ali,[¶] S. Marion,[‖,1] C. Weber,[‖] C. Mukherjee,[#] I. Bruchhaus,[**] E. Tannich,[**] M. Leippe,[††] T. Sicheritz-Ponten,[‡‡] P. G. Foster,[§§] J. Samuelson,[¶¶] C. J. Noël,[†] R. P. Hirt,[†] T. M. Embley,[†] C. A. Gilchrist,[||||] B. J. Mann,[||||] U. Singh,[##] J. P. Ackers,[*] S. Bhattacharya,[a] A. Bhattacharya,[b] A. Lohia,[#] N. Guillén,[‖] M. Duchêne,[‡] T. Nozaki,[¶] and N. Hall[c,2]

[*] Department of Infectious and Tropical Diseases, London School of Hygiene and Tropical Medicine, London WC1E 7HT, UK
[†] Division of Biology, Newcastle University, Newcastle NE1 7RU, UK
[‡] Department of Specific Prophylaxis and Tropical Medicine, Center for Physiology and Pathophysiology, Medical University of Vienna, A-1090 Vienna, Austria
[§] Department of Parasitology, National Institute of Infectious Diseases, Tokyo, Japan
[¶] Department of Parasitology, Gunma University Graduate School of Medicine, Maebashi, Japan
[‖] Institut Pasteur, Unité Biologie Cellulaire du Parasitisme and INSERM U786, F-75015 Paris, France
[#] Department of Biochemistry, Bose Institute, Kolkata 700054, India
[**] Bernhard Nocht Institute for Tropical Medicine, D-20359 Hamburg, Germany
[††] Zoologisches Institut der Universität Kiel, D-24098 Kiel, Germany
[‡‡] Center for Biological Sequence Analysis, BioCentrum-DTU, Technical University of Denmark, DK-2800 Lyngby, Denmark
[§§] Department of Zoology, Natural History Museum, London, SW7 5BD, UK
[¶¶] Department of Molecular and Cell Biology, Boston University Goldman School of Dental Medicine, Boston, Massachusetts 02118
[||||] Department of Medicine, Division of Infectious Diseases, University of Virginia Health Sciences Center, Charlottesville, Virginia 22908
[##] Departments of Internal Medicine, Microbiology, and Immunology, Stanford University School of Medicine, Stanford, California 94305
[a] School of Environmental Sciences, Jawaharlal Nehru University, New Delhi 110067, India
[b] School of Life Sciences and Information Technology, Jawaharlal Nehru University, New Delhi 110067, India
[c] The Institute for Genomic Research, Rockville, Maryland 20850
[1] Present address: Cell Biology and Biophysics Program, European Molecular Biology Laboratory, 69117 Heidelberg, Germany
[2] Present address: School of Biological Sciences, University of Liverpool, Liverpool L69 7ZB, United Kingdom

Advances in Parasitology, Volume 65
ISSN 0065-308X, DOI: 10.1016/S0065-308X(07)65002-7

Contents

Abstract The intestinal parasite *Entamoeba histolytica* is one of the first protists for which a draft genome sequence has been published. Although the genome is still incomplete, it is unlikely that many genes are missing from the list of those already identified. In this chapter we summarise the features of the genome as they are currently understood and provide previously unpublished analyses of many of the genes.

1. INTRODUCTION

Entamoeba histolytica is one of the most widespread and clinically important parasites, causing both serious intestinal (amoebic colitis) and extra-intestinal (amoebic liver abscess) diseases throughout the world. A recent World Health Organization estimate (WHO, 1998) places *E. histolytica* second after *Plasmodium falciparum* as causing the most deaths annually (70,000) among protistan parasites.

Recently a draft of the complete genome of *E. histolytica* was published (Loftus *et al.*, 2005) making it one of the first protist genomes to be sequenced. The *E. histolytica* genome project was initiated in 2000 with funding from the Wellcome Trust and the National Institute of Allergy and Infectious Diseases to the Wellcome Trust Sanger Institute and The Institute for Genomic Research (TIGR) in the UK and the USA, respectively. The publication describing the draft sequence concentrated on the expanded gene families, metabolism and the role of horizontal gene transfer in the evolution of *E. histolytica*. In this chapter we summarise

the structure and content of the *E. histolytica* genome in comparison to other sequenced parasitic eukaryotes, provide a description of the current assembly and annotation, place the inferred gene content in the context of what is known about the biology of the organism and discuss plans for completing the *E. histolytica* genome project and extending genome sequencing to other species of *Entamoeba*.

The fact that the genome sequence is still a draft has several important consequences. The first is that a few genes may be missing from the sequence data we have at present, although the number is likely to be small. For example, at least one gene (amoebapore B) is not present in the genome data despite it having been cloned, sequenced and the protein extensively characterised well before the start of the genome project. The second consequence is that the assembly contains a number of large duplicated regions that may be assembly artefacts, meaning that the number of gene copies is overestimated in several cases. These problems cannot as yet be resolved but should be eventually as more data becomes available. Nevertheless, it is important to remember these issues when reading the rest of this chapter.

As the number of genes in *E. histolytica* runs into several thousands, it is not possible to discuss all of them. However, we have generated a number of tables that identify many genes and link them to their entries in GenBank using the relevant protein identifier. Only a few tables are included in the text of this chapter, but the others are available online as supplementary material, http://pathema.tigr.org/pathema/entamoeba_resources.shtml. The *E. histolytica* genome project data are being 'curated' at the J. Craig Venter Institute (JCVI, formerly TIGR), and it is on that site that the most current version of the assembled genome will be found. The 'Pathema' database will hold the data and the annotation (http://pathema.tigr.org/). The gene tables are also linked to the appropriate entry in the Pathema database, and the links will be maintained as the genome structure is refined over time.

Reference is made throughout the text to other species of *Entamoeba* where data are available. *Entamoeba dispar* is the sister species to *E. histolytica* and infects humans without causing symptoms. *Entamoeba invadens* is a reptilian parasite that causes invasive disease, primarily in snakes and lizards, and is widely used as a model for *E. histolytica* in the study of encystation, although the two species are not very closely related (Clark *et al.*, 2006b). Genome projects for both these species are under way at TIGR, and it is anticipated that high-quality draft sequences will be produced for both in the near future. It is hoped that the *E. dispar* sequence will prove useful in identifying genomic differences linked to disease causation while that of *E. invadens* will be used to study patterns of gene expression during encystation. Small-scale genome surveys have been performed for two other species: *Entamoeba moshkovskii*, which is primarily a free-living species although it occasionally infects humans,

and *Entamoeba terrapinae*, a reptilian commensal species, http://www.sanger.ac.uk/Projects/Comp_Entamoeba/

2. GENOME STRUCTURE

2.1. The *E. histolytica* genome sequencing, assembly and annotation process

The first choice to be made in the genome project was perhaps the easiest—the identity of the strain to be used for sequencing. A significant majority of the existing sequence data prior to the genome project was derived from one strain: HM-1:IMSS. This culture was established in 1967 from a rectal biopsy of a Mexican man with amoebic dysentery and axenised shortly thereafter. It has been used widely for virulence, immunology, cell biology and biochemistry in addition to genetic studies. In an attempt to minimise the effects of long-term culture cryopreserved cells that had been frozen in the early 1970s were revived and this uncloned culture used to generate the DNA for sequencing.

Before undertaking a genome scale analysis, it is important to understand the quality and provenance of the underlying data. The *E. histolytica* genome was sequenced by whole genome shotgun approach with each centre generating roughly half of the reads. Several different DNA libraries containing inserts of different sizes were produced using DNA that had been randomly sheared and sequences were obtained from both ends of each cloned fragment. The Phusion assembler (Mullikin and Ning, 2003) was used to assemble the 450,000 short reads into larger contigs (contiguous sequences), resulting in 1819 genome fragments that were ~12× deep, which means that each base has been sequenced 12 times, on average. While the genome shotgun sequence provides high coverage of each base, it is inevitable that there will be misassemblies and sequencing errors in the final consensus particularly towards each end of the contigs. Another problem with draft sequence is that it contains gaps, and while most of these will be small and will mostly contain repetitive non-coding 'junk' sequence, some of the gaps will probably contain genes. This makes it impossible to be absolutely certain of the absence of particular genes in *E. histolytica* and, in some cases, the presence or absence of particular biological pathways. Due to the high repeat content and low GC content (24.1%) of the *E. histolytica* genome, closure of the remaining gaps is likely to be a lengthy process. Therefore, it was decided to undertake and publish an analysis of the genome draft following assembly of the shotgun reads.

Annotation of the protein coding regions of the genome was initially carried out using two genefinders [GlimmerHMM (Majoros *et al.*, 2004)

and Phat (Cawley *et al.*, 2001)] previously used successfully on another low G + C genome, that of *P. falciparum*. The software was re-trained specifically for analysis of the *E. histolytica* genome. The training process involved preparing a set of 600 manually edited genes to be used as models with the subsequent genefinding then being carried out on all of the assembled contigs to generate a 'complete' gene set. Predicted gene functions were generated automatically by homology searches using public protein and protein-domain databases, with subsequent refinement of identifications being carried out by manual inspection. For particular genes and gene families of special interest, members of the *Entamoeba* scientific community were involved throughout this process as expert curators with each individual assisting in the analysis and annotation of their genes of interest. Therefore although the manual curation of the genome has not been systematic, those areas of biology that are of primary interest to the *Entamoeba* community have been annotated most thoroughly. The publication of the genome by Loftus *et al.* therefore represents a 'first draft' of the complete genome sequence and the level of annotation is similar to the initial publications of other genomes such as *Drosophila* (Adams *et al.*, 2000; Myers *et al.*, 2000) and human (Lander *et al.*, 2001).

2.2. Karyotype and chromosome structure

The current *E. histolytica* genome assembly is ~23.7 million basepairs (Mbp) in size (Table 2.1). This figure is not likely to be a very accurate measure. In part this is due to misassembly of repetitive regions, which will cause the genome to appear smaller and in part because of the possibility of aneuploidy in some regions of the genome, which would cause them to appear more than once in the assembly. Overall, however, this size is not inconsistent with data from pulse-field gels (Willhoeft and Tannich, 1999) and kinetic experiments (Gelderman *et al.*, 1971a,b) making the *E. histolytica* genome comparable in size (24 Mbp) to that of *P. falciparum* (23 Mbp) (Gardner *et al.*, 2002), *Trypanosoma brucei* (26 Mbp) (Berriman *et al.*, 2005) and the free-living amoeba *Dictyostelium discoideum* (34 Mbp) (Eichinger *et al.*, 2005).

The current assembly does not represent complete chromosomes. Analysis of pulse-field gels predicts 14 chromosomes ranging in sizes from 0.3 to 2.2 Mbp and possibly a ploidy of 4 (Willhoeft and Tannich, 1999). There is no current information regarding the size and nature of the centromeres, and there are no contigs that appear to contain likely centromeric regions based on comparisons with other organisms. A search for signature telomeric repeats within the data indicates that these are either not present in the genome, not present in our contigs, or are diverged enough to be unidentifiable. However, there is circumstantial evidence

only 33% sequence identity. At present, the function of the various flavodoxin-like molecules remains to be determined and deserves to be investigated fully, particularly as to whether they do indeed have antioxidant capacity.

4. METABOLISM

Biochemical analysis of *E. histolytica* metabolism has a long history (Reeves, 1984), dating back to shortly after the development of culture media that allowed the generation of substantial numbers of axenic cells. The genome sequence has confirmed most of the predicted metabolic pathways shown biochemically to be present or absent in *E. histolytica* in the past. As with most parasites, secondary loss of biosynthetic pathways is a recurring theme. However, a few surprises have also been uncovered. Every single enzyme involved in metabolism cannot realistically be discussed in this chapter. In this section, only the major energy generating and biosynthetic aspects of metabolism will be covered. Enzyme names, EC numbers and accession numbers are given in the the supplementary table for this section.

4.1. Energy metabolism

4.1.1. Glycolysis

E. histolytica lacks a functional tricarboxylic acid (TCA) cycle and oxidative phosphorylation. It is not able to convert organic substrates such as glucose into H_2O and CO_2, but has to rely on the energy generated by various types of substrate level phosphorylation (Reeves, 1984). Glycolysis is the major pathway of ATP generation, but in addition the genome project has identified a number of genes that could result in more ATP generation through the catabolism of amino acids. These enzymes will be described further below. As *E. histolytica* lacks compartmentalised energy generation, it has been classified as a type I amitochondriate protist (Martin and Müller, 1998) in contrast to the type II amitochondriate protists containing hydrogenosomes such as *T. vaginalis*. Nevertheless, it does contain a mitochondrial remnant, the mitosome (see Section 8).

In *E. histolytica*, glycolysis appears to be localised in the cytosol. This is in contrast to trypanosomes in which a major part is carried out in the glycosomes (Parsons, 2004) and the pathway is regarded as a potential target for chemotherapy (Opperdoes and Michels, 2001). The kinetic properties of recombinant *E. histolytica* glycolysis enzymes have recently been studied by Saavedra *et al.* (2005). Their analysis suggested that fructose-1,6-bisphosphate aldolase, phosphoglycerate mutase,

glyceraldehyde-3-phosphate dehydrogenase and pyruvate phosphate dikinase might be regulating the glycolytic flux.

4.1.1.1. Hexokinases Glucose taken up by *E. histolytica* is phosphorylated by two hexokinase (EC 2.7.1.1) isoenzymes (Hxk1 and Hxk2). The two *E. dispar* isoenzymes are shifted towards a slightly more basic pI, which is the basis of the classical biochemical method for distinguishing *E. histolytica* from *E. dispar* by starch gel electrophoresis (Farri *et al.*, 1980). The pI differences among the two *E. histolytica* isoforms (Ortner *et al.*, 1995) and between the two species (Ortner *et al.*, 1997b) are the result of genetic differences that lead to different amino acid sequences and charge differences. Hxk1 phosphorylates glucose and mannose, while Hxk2 phosphorylates mainly glucose and is much less active with mannose as a substrate (Kroschewski *et al.*, 2000).

4.1.1.2. Glucose-6-phosphate isomerase Glucose 6-phosphate is converted to fructose 6-phosphate by glucose-6-phosphate isomerase (EC 5.3.1.9). The genome has two genes for this enzyme, which code for proteins that differ only by a single insertion or deletion of seven amino acid residues. Glucose-6-phosphate isomerase is another of the enzymes for the classical differentiation of *Entamoeba* zymodemes by starch gel electrophoresis (Sargeaunt, 1987).

4.1.1.3. Phosphofructokinases The main phosphofructokinase activity in *E. histolytica* is pyrophosphate (PPi)-dependent (EC 2.7.1.90; Reeves *et al.*, 1976). There is a single gene (Deng *et al.*, 1998) encoding this 60 kDa enzyme. The gene is a candidate for lateral transfer from bacteria (Loftus *et al.*, 2005) (see Section 10). The enzyme is expressed at a 10-fold higher level and displays about 10-fold higher activity than a second phosphofructokinase of 48 kDa (XP_653373) (Chi *et al.*, 2001). The substrate specificity of the smaller enzyme is disputed. Whereas Bruchhaus *et al.* (1996) reported that this minor enzyme also used PPi as phosphate donor, Chi *et al.* (2001) found only an ATP-dependent activity. The 48 and 60 kDa enzymes are highly divergent with <20% sequence identity. Interestingly, the specificity of the 60 kDa phosphofructokinase can be changed from PPi to ATP by mutation of a single amino acid residue (Chi and Kemp, 2000). The authors concluded that ATP rather than PPi was the primordial high energy compound. In the genome, there are 2 additional genes encoding isoforms of the 48 kDa enzyme, which have not been studied at the protein level.

4.1.1.4. Fructose-1,6-bisphosphate aldolase Fructose 1,6-bisphosphate is cleaved to glyceraldehyde 3-phosphate and dihydroxyacetone 3-phosphate by fructose-1,6-bisphosphate aldolase (EC 4.1.2.13). The enzyme, a class II

aldolase (Marsh and Lebherz, 1992), has been cloned (XP_650373) and exhibits strong sequence similarity to eubacterial aldolases (Sanchez *et al.*, 2002). A second gene (XP_655966) encodes a protein differing from the first by a single deletion of 28 amino acids flanked by short divergent stretches. These bacterial-type aldolases are also found in *T. vaginalis*, *G. intestinalis* and other protists (Sanchez *et al.*, 2002). *E. histolytica* has no gene coding for a class I aldolase like those found in animals, which might make aldolase an interesting target for chemotherapy.

4.1.1.5. Triose-phosphate isomerase Triose-phosphate isomerase (EC 5.3.1.1) converts dihydroxyacetone 3-phosphate into glyceraldehyde 3-phosphate. The gene was previously cloned (Landa *et al.*, 1997), and is highly similar to the annotated gene product. This dimer-forming enzyme represents the first *E. histolytica* protein for which the structure has been solved by X-ray crystallography (Rodriguez-Romero *et al.*, 2002).

4.1.1.6. Glyceraldehyde-3-phosphate dehydrogenase Glyceraldehyde-3-phosphate dehydrogenase (EC 1.2.1.12) oxidises and phosphorylates glyceraldehyde 3-phosphate to 1,3-bisphosphoglycerate in two coupled reactions using NAD^+ as cofactor (Reeves, 1984). The genome project revealed 5 putative genes, 3 of which encode the identical protein sequence of 36.0 kDa and a predicted pI of 7.04. The fourth gene product, XP_648981, differs from these 3 only by a 13 amino acid deletion, while XP_650370 is a clearly distinct 34.8 kDa isoform with a lower predicted pI of 5.80. Interestingly, the isoforms XP_650356 and XP_650370 of different pI are encoded within the same contig.

4.1.1.7. Phosphoglycerate kinase Phosphoglycerate kinase has an unusual substrate (Reeves and South, 1974), transferring the high energy phosphate group from 3-phosphoglyceroyl phosphate to GDP leading to the formation of GTP (EC 2.7.2.10). There is one candidate gene encoding a 45 kDa protein.

4.1.1.8. Phosphoglycerate mutase Phosphoglycerate mutase (Reeves, 1984) isomerises 3-phosphoglycerate to 2-phosphoglycerate (EC 5.4.2.1). Five divergent putative genes for this enzyme are found in the genome. Two gene products of 62 kDa were classified as 2,3-bisphosphoglycerate-independent phosphoglycerate mutases (XP_649031 and XP_654182); they differ only at their C-termini and display significant similarity to bacterial phosphoglycerate mutases. The three other genes are very divergent. XP_651808 was identified as a candidate for lateral gene transfer (LGT) (Loftus *et al.*, 2005) (see Section 10). The remaining two gene

products XP_649053 and XP_657284 are related to genes found in both prokaryotes and eukaryotes.

4.1.1.9. Enolase (2-phosphoglycerate dehydratase) Enolase (EC 4.2.1.11) converts 2-phosphoglycerate to phosphoenolpyruvate. The gene has been cloned (Beanan and Bailey, 1995) and the protein characterised (Hidalgo *et al.*, 1997) previously. The 47 kDa gene product is a typical eukaryotic enolase (XP_649161). A carboxy-terminally truncated incomplete ORF is also found.

4.1.1.10. Pyruvate, orthophosphate dikinase and pyruvate kinase In *E. histolytica*, both activities forming ATP and pyruvate from phosphoenolpyruvate have been found. The exergonic pyruvate kinase reaction uses ADP (Saavedra *et al.*, 2004), and the pyruvate, orthophosphate dikinase uses AMP and PPi in a slightly endergonic reaction (Varela-Gomez *et al.*, 2004). The dikinase activity is found in C4 plants where it is involved in phosphoenolpyruvate generation for gluconeogenesis. In *E. histolytica* it was discovered long before the pyruvate kinase (Reeves, 1968).

The cloning of pyruvate, orthophosphate dikinase (EC 2.7.9.1) was reported by two groups. The published sequences (Bruchhaus and Tannich, 1993; Saavedra Lira *et al.*, 1992) are highly similar or identical to XP_657332 and XP_654666. In addition there are two shorter related ORFs.

In the genome three putative pyruvate kinase genes (EC 2.7.1.40) have been identified. The three are identical except for an amino-terminal deletion in XP_648240 and an internal deletion in XP_653635.

4.1.1.11. Pyruvate:ferredoxin oxidoreductase (PFOR) and ferredoxin PFOR (EC 1.2.7.1) is an enzyme of major importance to *E. histolytica*, as the parasite lacks NAD^+-dependent pyruvate dehydrogenase and pyruvate decarboxylase (Reeves, 1984). No evidence for the latter two genes was found in the genome, confirming the biochemical results. PFOR oxidatively decarboxylates pyruvate to acetyl-CoA. The electrons are transferred to ferredoxin which, in its reduced form, can activate and reduce metronidazole, the major anti-amoebic drug (Müller, 1986). The activated form of metronidazole can potentially react with a number of biomolecules and is able to cleave the parasite DNA. In human cells, metronidazole is not activated and is much less toxic. In *T. vaginalis*, down-regulation of PFOR is one mechanism of producing metronidazole resistance (Kulda, 1999); however, PFOR expression appears unaltered in partially resistant *E. histolytica* (Samarawickrema *et al.*, 1997; Wassmann *et al.*, 1999). All eukaryotic PFOR genes, including that of *E. histolytica*, appear to have been acquired during an ancient LGT event from bacteria

(Horner *et al.*, 1999; Rotte *et al.*, 2001). There are two putative PFORs in the *E. histolytica* genome, displaying minor sequence differences.

The genome contains seven ferredoxin genes in total with five quite divergent sequences. All are related to eubacterial and archaeal ferredoxins (Nixon *et al.*, 2002). The gene pairs XP_655183/XP_655182 and XP_654311/XP_652694 are identical. The other three gene products represent more divergent ORFs. The deduced proteins have similar molecular masses, between 6.1 and 8.8 kDa, and different predicted isoelectric points between 4.2 and 8.6 kDa.

4.1.1.12. Acetyl-CoA synthetase (acetate thiokinase) The normal fate of acetyl-CoA in mitochondriate organisms is entry into the tricarboxylic acid cycle. However, this pathway is absent from *E. histolytica*. Instead, the cleavage energy of the thioester bond of acetyl-CoA can be used to generate one ATP molecule. One of the known acetyl-CoA synthetases generates ATP from ADP and Pi (EC 6.2.1.13). Such an enzyme has been characterised by Reeves *et al.* (1977) and cloned (Field *et al.*, 2000), and reported to be a 77 kDa protein. The common acetyl-CoA synthetase activity that produces ATP from AMP and PPi (EC 6.2.1.1) appears to be absent in *E. histolytica*.

4.1.1.13. Aldehyde and alcohol dehydrogenases The *E. histolytica* genome encodes a complex system of alcohol or aldehyde dehydrogenases. In total, there are 25 predicted genes, 3 of which are on the list of LGT candidates.

Alcohol dehydrogenase ADH1 was the first alcohol dehydrogenase to be characterised in *E. histolytica* (Reeves *et al.*, 1971) and is a NADPH-dependent enzyme (EC 1.1.1.2). The gene was previously cloned (Kumar *et al.*, 1992); in the genome 3 genes are almost identical to that sequence, while 1 (XP_652772) has 67% identity.

Fermentation in *E. histolytica* uses the bifunctional NADH-dependent enzyme ADH2, which belongs to the ADHE family and has both alcohol dehydrogenase and aldehyde dehydrogenase activities (Lo and Reeves, 1978). Under anaerobic conditions, reduction of the acetyl-CoA generated by PFOR to ethanol is one way to regenerate the NAD^+ used by glyceraldehyde-3-phosphate dehydrogenase. ADH2 first reduces acetyl-CoA to an enzyme-bound hemiacetal which is then hydrolysed to acetaldehyde (EC 1.2.1.10) and further reduced to ethanol (EC 1.1.1.1). If the enzyme is also able to work in the reverse direction, *E. histolytica* would be able to generate acetyl-CoA and energy from ethanol in the presence of oxygen. This would explain older reports of ethanol stimulated oxygen uptake in *E. histolytica* (Weinbach and Diamond, 1974). The enzyme is closely related to AdhE from *E. coli* and other bacteria (Reid and Fewson, 1994), and there is strong support for its aquisition by LGT (Andersson *et al.*, 2006;

Field *et al.*, 2000; Loftus *et al.*, 2005) (see Section 10). Like its bacterial homologue, ADH2 appears to form helical rods that sediment with membrane fractions (Avila *et al.*, 2002). Two groups have previously cloned ADH2 (Bruchhaus and Tannich, 1994; Yang *et al.*, 1994), and in total the genome contains five full-length ADH2 genes and one that is truncated. All share between 98 and 100% sequence identity.

In total, there are 11 alcohol dehydrogenase ADH3 genes in the genome, 2 of which have been reported previously (Kimura *et al.*, 1996; Rodriguez *et al.*, 1996). The recombinant enzyme characterised by Rodriguez *et al.* (1996) was NADPH-specific, like ADH1. There are five genes similar to these previously reported sequences. The rest of the ADH3 sequences fall into two groups of three similar sequences. All 11 ADH3 sequences are between 44 and 100% identical on the amino acid level. XP_649823 was originally on the list of LGT candidates (Loftus *et al.*, 2005), and a similarity to ADH3 sequences of gram-negative bacteria had been noted before (Nixon *et al.*, 2002). However, a related sequence is now known to exist in *T. vaginalis* also (see Section 10).

The genome encodes three additional distinct alcohol dehydrogenases. XP_656535 is a putative Zn-containing enzyme, and is on the list of LGT candidates. XP_652753 has been annotated as a Fe-containing alcohol dehydrogenase and XP_652262 simply as putative alcohol dehydrogenase.

One NADPH-dependent aldehyde dehydrogenase encoding gene (ALDH1) is present and was reported previously (Zhang *et al.*, 1994).

4.1.2. Energy storage: The glycogen metabolism

E. histolytica uses glycogen as its major energy store. Glycogen is a polymer of α-1,4-linked glucose chains with α-1,6 branch points, which in *E. histolytica* has a compact structure as suggested by branch points every 5–6 glucose residues (Bakker-Grunwald *et al.*, 1995). The cytoplasm of trophozoites contains numerous glycogen granules that were first observed by electron microscopy (Rosenbaum and Wittner, 1970) and later characterised biochemically (Takeuchi *et al.*, 1977). A glycogen phosphorylase activity (EC 2.4.1.1), associated with the glycogen granules, generates glucose 1-phosphate from orthophosphate and the linear portion of various glucopolysaccharides (Werries and Thurn, 1989). The genome contains at least six putative full-length and truncated genes encoding glycogen phosphorylases, two of which were cloned by Wu and Müller (2003). These authors noted a marked sequence divergence in those regions of the enzymes involved in regulation by phosphorylation and concluded that classical regulation by phosphorylation may not occur.

Glycogen phosphorylase degrades the linear chains only down to the α-1,6 branch points. The remaining core molecule is called limit dextrin.

Degradation can proceed further with the help of a debranching enzyme that has been purified (Werries *et al.*, 1990). It exhibits activities of both amylo-1,6-glucosidase (EC 3.2.1.33) and 4-α-glucanotransferase (EC 2.4.1.25). The genome contains two genes putatively encoding a full-length (XP_653608) and a truncated glycogen debranching enzyme. The deduced molecular mass of the large protein is 166 kDa, which corresponds to the biochemical data (Werries *et al.*, 1990).

Glucose 1-phosphate is isomerised to glucose 6-phosphate by phosphoglucomutase (EC 5.4.2.2) before entering the glycolytic pathway. The isoelectric points of the phosphoglucomutases from *E. histolytica* and *E. dispar* differ, and this was exploited for differentiation of the two species by starch gel electrophoresis (Sargeaunt *et al.*, 1978). The migration properties are reproduced by recombinant enzymes and are the result of primary sequence differences (Ortner *et al.*, 1997a). *E. histolytica* has one gene coding for this important enzyme, and in addition there are two distantly related members of the phosphoglucomutase/phosphomannomutase family.

Genes encoding the enzymes involved in glycogen biosynthesis in *E. histolytica* have been identified: a glycogen synthase (EC 2.4.1.11) of 155 kDa and 2 putative branching enzymes (EC 2.4.1.18). The glycogen precursor UDP-glucose is generated from UTP and glucose 1-phosphate by UTP:glucose-1-phosphate uridylyltransferase (EC 2.7.7.9). Two UTP-hexose-1-phosphate uridyltransferases have been characterised biochemically, a larger glucose 1-phosphate-specific enzyme of 45 kDa and a less specific enzyme of 40 kDa reported to use both galactose 1-phosphate and glucose 1-phosphate (Lobelle-Rich and Reeves, 1983). The genome contains one larger ORF encoding a putative UTP:glucose-1-phosphate uridylyltransferase of 54.7 kDa and 2 smaller ones encoding enzymes of 46.3 kDa with high similarity identified as UTP:*N*-acetyl-glucosamine-1-phosphate uridyltransferases. These enzymes are interesting in that they could possibly be involved in the activation of *N*-acetyl-glucosamine 1-phosphate as a precursor of the chitin cyst wall.

4.1.3. Catabolism of sugars other than glucose

4.1.3.1. Activation of fructose and galactose for glycolysis Neither Hxk1 nor Hxk2 can use fructose or galactose as a substrate, but there are 2 genes encoding bacterial-type enzymes that may do so, a 33 kDa fructokinase, which is one of the candidates for LGT to the *E. histolytica* lineage (see Section 10), and a 43 kDa galactokinase. The fructokinase groups with bacterial fructose 6-kinases (EC 2.7.1.4), and the galactokinase groups with galactose 1-kinases (EC 2.7.1.6). This substrate specificity has been noted before (Reeves, 1984). Fructose 6-phosphate enters as an intermediate of the glycolytic pathway (see Section 4.1.1.3). As described earlier (see Section

4.1.2), galactose 1-phosphate can be activated to UDP-galactose (Lobelle-Rich and Reeves, 1983) and then epimerised to UDP-glucose by UDP-glucose 4-epimerase (EC 5.1.3.2) (Reeves, 1984). In the genome, a single candidate 38 kDa ORF for the latter enzyme has been identified. The UDP-bound glucose can then be used either for the synthesis of glycogen or fed into the glycolysis pathway via glucose 1-phosphate and glucose 6-phosphate. This efficient pathway allows *E. histolytica* to grow on galactose instead of glucose (Reeves, 1984).

4.1.3.2. Anomerisation of aldoses The 1-position in the pyranose form of aldoses has a hydroxyl group that can be in either the α- or β-configuration. These forms can be interconverted by means of an aldose 1-epimerase (EC 5.1.3.3), an enzyme that has recently been characterised (Villalobo *et al.*, 2005). There is a single gene encoding this product.

4.1.3.3. Activation of pentoses Two gene candidates encoding pentose-activating enzymes have been identified in the *E. histolytica* genome: a 35 kDa ribokinase (EC 2.7.1.15) and a 56 kDa xylulokinase (EC 2.7.1.17). The latter is another bacterial-type sequence putatively acquired by LGT.

4.1.3.4. Interconversion of hexoses and pentoses The pathway of interconversion between hexoses and pentoses in *E. histolytica* was described many years ago (Reeves, 1984; Susskind *et al.*, 1982). A transketolase (EC 2.2.1.1) converts fructose 6-phosphate and glyceraldehyde 3-phosphate into xylulose 5-phosphate and erythrose 4-phosphate. Erythrose 4-phosphate and dihydroxyacetone phosphate are condensed by the glycolytic enzyme fructose-1,6-bisphosphate aldolase to sedoheptulose 1,7-bisphosphate, an extended substrate specificity of the aldolase. Phosphofructokinase then is able to remove a phosphate group forming diphosphate and sedoheptulose 7-phosphate. This molecule and glyceraldehyde 3-phosphate are then converted by transketolase to the pentoses ribose 5-phosphate and xylulose 5-phosphate. A transaldolase activity is absent (Reeves, 1984) consistent with there being no such gene in the genome. In contrast, 7 gene products were identified as likely transketolases: 3 highly similar proteins of 73 kDa and 4 truncated versions.

4.2. Amino acid catabolism

4.2.1. General features

As discussed earlier, glycolysis under anaerobic conditions can use only part of the energy contained in glucose for ATP generation. *E. histolytica* is capable not only of taking up amino acids (Reeves, 1984), but also using them for the generation of energy, as suggested by Zuo and Coombs (1995). The genome has revealed a number of unusual genes, often with bacterial

affinities, coding for enzymes of amino acid catabolism (Anderson and Loftus, 2005).

In many cases, the degradation of amino acids starts with a transamination reaction (EC 2.6.1.-) generating a 2-ketoacid. The *E. histolytica* genome has five ORFs identified as aminotransferases. These ORFs are distinct from each other with the exception of XP_655090 and XP_655099, which differ only by one insertion and are LGT candidates. So far there is no enzymological data on this group of enzymes, so their substrate specificities in *E. histolytica* are unknown.

Both amino acid degradation and glycolysis have 2-ketoacids as intermediates. Pyruvate is one common intermediate, as amino acid degradation can produce either pyruvate or other 2-ketoacids. PFOR (see Section 4.1.1.11) is known to have a relaxed specificity, and in addition to pyruvate it can oxidatively decarboxylate 2-ketobutanoate, oxaloacetate and 2-ketoglutarate (Samarawickrema *et al.*, 1997). The reaction generates CoA-thioesters with the potential of producing one ATP per molecule.

The amino acids asparagine, aspartate, serine, alanine, tryptophan, cysteine, threonine, methionine, glutamine and glutamate can all be transformed into one of these 2-ketoacids in one or very few steps. This underlines the major importance of the PFOR in the energy metabolism of *E. histolytica*. The enzyme is indispensable, and as it always generates reduced ferredoxin it will always activate metronidazole. Consequently, it would be very difficult for *E. histolytica* to become resistant to metronidazole.

4.2.1. Aspartate and asparagine

E. histolytica takes up asparagine and aspartate in the presence or absence of glucose (Zuo and Coombs, 1995). Four putative asparaginases (EC 3.5.1.1) are found in the genome. Three are identical and share only 48% amino acid identity with the fourth (XP_656586). Asparaginase mediates the formation of aspartate from aspargine by releasing ammonia. The predicted sequences appear to possess a signal sequence, as suggested by the TargetP programme (http://www.cbs.dtu.dk/services/TargetP/), which is reminiscent of a periplasmic isotype (EcA, type II) (Swain *et al.*, 1993) that is up-regulated under anaerobic and carbon-restricted conditions (Cedar and Schwartz, 1967).

Aspartate can be converted to fumarate and ammonia by aspartate ammonia-lyase (aspartase, EC 4.3.1.1). Addition of a water molecule by fumarase (EC 4.2.1.2) produces malate. The genome encodes a putative fumarase that is related to bacterial class I fumarases. The aspartase is a member of the bacterial class II fumarase/aspartase protein family (Woods *et al.*, 1988), and also on the list of LGT candidates.

Aspartate is also decomposed into oxaloacetate and ammonia by aspartate aminotransferase, with the concomitant production of

glutamate from 2-oxoglutarate. Oxaloacetate is then converted into malate via malate dehydrogenase (EC 1.1.1.37) and, since *E. histolytica* lacks both a functional TCA cycle and a phosphoenolpyruvate carboxykinase, the malate generated can be decarboxylated oxidatively to pyruvate by malic enzyme (EC 1.1.1.39). Both of these enzymes are present in *E. histolytica*. Two very similar genes have been identified as encoding malic enzyme and are LGT candidates.

4.2.2. Serine, threonine

Serine and threonine are also taken up by *E. histolytica* in the presence and absence of glucose (Zuo and Coombs, 1995). Serine can be deaminated by the pyridoxal phosphate-dependent serine dehydratase (L-serine ammonia-lyase, EC 4.3.1.17) to pyruvate and ammonia. The enzyme was characterised by Takeuchi *et al.* (1979) who showed that addition of serine to the culture medium stimulated oxygen consumption. In an analogous reaction, threonine dehydratase (threonine ammonia-lyase, EC 4.3.1.19) breaks down threonine to 2-oxobutanoate. Both ketoacids can then be oxidised by PFOR to acetyl-CoA or propionyl-CoA. Both catabolic reactions can be carried out by the same enzyme, as has been shown in yeast for example (Ramos and Wiame, 1982). In the *E. histolytica* genome annotation, four gene products have been annotated as threonine dehydratases, but none as serine dehydratase. XP_650405 and XP_652480 are identical while XP_655614 and XP_657171 share 95 and 37% identity with the others, respectively. The exact substrate specificities of these four putative serine/threonine dehydratases have not been reported.

Degradation of serine via the non-phosphorylated serine pathway, by the sequential reactions of L-serine: pyruvate aminotransferase (EC 2.6.1.51), D-glycerate dehydrogenase (EC 1.1.1.29) and D-glycerate kinase (EC 2.7.1.31) (Snell, 1986) results in the glycolytic intermediate 3-phosphoglycerate. The genome encodes several putative aminotransferases (see Section 4.2.1), but it is not yet known if serine is among their substrates. An unusual bacterial-type NADPH-dependent D-glycerate dehydrogenase was characterised by Ali *et al.* (2003), and there are two genes encoding D-glycerate dehydrogenases, one of which (XP_648124) is among the weaker LGT candidates (see Section 10). The genome also contains two genes encoding identical glycerate kinases. The enzyme has recently been characterised (V. Ali and T. Nozaki, unpublished data).

4.2.3. Methionine, homocysteine and cysteine

Methionine γ-lyase (EC 4.4.1.11) decomposes methionine to methanethiol (mercaptomethane), ammonia and 2-oxobutanoate. In *E. histolytica*, two methionine γ-lyases, EhMGL1 and EhMGL2, of similar molecular weights

have been characterised (Tokoro *et al.*, 2003). These two isoenzymes show marked differences in substrate specificity, isoelectric point, enzymological and biochemical parameters (Tokoro *et al.*, 2003). Both enzymes can also act on other amino acids. In addition to degrading methionine, both EhMGL1 (pI 6.01) and EhMGL2 (pI 6.63) can convert homocysteine to hydrogen sulphide, ammonia and 2-oxobutanoate. EhMGL2 also decomposes cysteine to hydrogen sulphide, ammonia and pyruvate, whereas EhMGL1 is only weakly active against cysteine. Decomposition of homocysteine by methionine γ-lyase is essential since this parasite lacks the other known enzymes capable of destroying this toxic amino acid. In the genome, three ORFs correspond to EhMGL1 and one to EhMGL2. So far, the only eukaryotes known to possess methionine γ-lyases are *E. histolytica* and *T. vaginalis* (Lockwood and Coombs, 1991). As the enzymes are absent from the human host and important for the generation of metabolic energy, they could be targets for chemotherapy (Coombs and Mottram, 2001; Tokoro *et al.*, 2003).

In addition to serving as a source of metabolic energy, another important role of methionine is as a donor of methyl groups via *S*-adenosylmethionine synthetase (synonymous with methionine adenosyltransferase, EC 2.5.1.6). Seven gene candidates were identified, four full-length and three truncated. The *S*-adenosylhomoserine left after the transfer of the activated methyl group can be hydrolysed by *S*-adenosylhomocysteine hydrolase (EC 3.3.1.1), giving adenosine and homocysteine. Two candidate genes with identical sequences and one truncated form are present.

However, *E. histolytica* lacks the remaining enzymes for the reverse transsulphuration pathway (forming cysteine from methionine) (Nozaki *et al.*, 2005), that is cystathionine β-synthase and cystathionine γ-lyase. In addition, *E. histolytica* lacks all enzymes involved in the forward transsulphuration (forming methionine from cysteine) including cobalamin-dependent methionine synthase (EC 2.1.1.13) or cobalamin-independent methionine synthase (EC 2.1.1.14), which suggests that *E. histolytica* is capable of neither converting homocysteine to cystathionine nor recycling homocysteine to methionine.

E. histolytica lacks the methylthioadenosine cycle enzymes except for two, 5′-methylthioadenosine/*S*-adenosyl homocysteine nucleosidase (EC 3.2.2.9) and aspartate aminotransferase (AT, EC 2.6.1.1). The significance of these two enzymes in *E. histolytica* is unknown.

4.2.4. Arginine

In *G. intestinalis* and *T. vaginalis* the arginine deiminase (EC 3.5.3.6) pathway is important for energy generation (Knodler *et al.*, 1994; Linstead and Cranshaw, 1983; Schofield and Edwards, 1994), generating one ATP molecule from the breakdown of arginine to ornithine. In contrast, no

arginine deiminase gene or dihydrolase pathway was detected in the *E. histolytica* genome.

In *E. histolytica*, arginine can either be degraded by arginase (EC 3.5.3.1) via ornithine or by arginine decarboxylase (EC 4.1.1.19) via agmatine. The arginine decarboxylase reaction uses up protons and may be involved in the acid resistance needed for the passage of cysts through the human stomach (Anderson and Loftus, 2005). Another function suggested for arginine degradation was that it depletes arginine as a substrate for human macrophages, preventing NO synthesis and amoebicidal activity (Elnekave *et al.*, 2003). Both enzymes could also be important for the generation of the polyamine putrescine (see Section 4.3). The genome contains a single gene encoding a 96 kDa polypeptide annotated as ornithine/arginine/lysine decarboxylase, the substrate specificity of which has not yet been examined on the recombinant protein level. There is a single gene encoding a putative 33 kDa arginase.

4.2.5. Glutamate, glutamine

In aerobic organisms, the 2-oxoglutarate generated from glutamate in a transaminase reaction enters the citric acid cycle for further catabolism. In *E. histolytica*, which also contains transaminases, 2-oxoglutarate can be oxidised by PFOR to give succinyl-CoA from which one molecule of ATP can be generated.

Several other gene products of *E. histolytica* could act on glutamine and glutamate. The genome lacks a glutaminase (EC 3.5.1.2) to carry out the simple hydrolysis of glutamine. Instead there is a putative glucosamine-fructose-6-phosphate aminotransferase (EC 2.6.1.16), which uses the energy in the amide group of glutamine to generate glucosamine 6-phosphate from fructose 6-phosphate. This product may be used for cyst wall biosynthesis.

4.2.6. Tryptophan

Tryptophan can be degraded to indole, pyruvate and ammonia by the PLP-dependent enzyme tryptophanase (EC 4.1.99.1), for which one candidate gene exists. To date, tryptophanase has only been found in bacteria and *T. vaginalis* and it is also on the list of LGT candidates.

4.2.7. Alanine: A possible special case

Alanine could potentially be transformed into pyruvate by alanine aminotransferase (synonymous with alanine:pyruvate transaminase, EC 2.6.1.2). However, *E. histolytica* is reported to excrete alanine (Zuo and Coombs, 1995), suggesting that this enzyme is not used under the culture conditions tested. Conceivably, the purpose of the excretion process may be to carry excess nitrogen out of the cell in the absence of a functional urea cycle.

4.2.8. Catabolism of other amino acids

Most of the enzymes for branched-chain amino acid metabolism are missing in *E. histolytica*, but leucine, isoleucine and valine could be transformed into 2-oxoisocaproate, 2-oxo-3-methylvalerate and 2-oxovalerate, respectively, by a putative branched-chain amino acid aminotransferase (EC 2.6.1.42), one of the aminotransferases mentioned earlier (see Section 4.2). This could produce ammonia or transfer the amino group to 2-oxoglutarate to form glutamate. Subsequent oxidative decarboxylation to give the respective CoA-derivatives could be envisaged, but so far no gene candidates for the necessary dehydrogenases have been identified.

One gene encodes a putative histidine ammonia-lyase (EC 4.3.1.3), which is responsible for the decomposition of histidine into urocanate and ammonia. Other than the formation of ammonia, the significance of this enzyme is not clear since the downstream enzymes involved in histidine catabolism from urocanate to glutamate were not found.

Currently, there is little information regarding the fate of the amino acids glycine, proline, phenylalanine, tyrosine and lysine in *E. histolytica*. No genes for the catabolic enzymes necessary were detected except for an LGT candidate bacterial-type 96 kDa broad-specificity ornithine/arginine/lysine decarboxylase that may be acting on lysine.

4.3. Polyamine metabolism

The absence of *S*-adenosyl-L-methionine decarboxylase (EC 4.1.1.50), which converts *S*-adenosyl methionine into decarboxylated *S*-adenosyl methionine, spermidine synthase (EC 2.5.1.16) and spermine synthase (EC 2.5.1.22), suggests a complete lack of polyamine metabolism in this parasite (Anderson and Loftus, 2005). However, as mentioned earlier, *E. histolytica* possesses genes encoding arginase and arginine decarboxylase. Both could be involved in the production of putrescine via agmatine and agmatinase (EC 3.5.3.11) or via ornithine and ornithine decarboxylase (EC 4.1.1.17). The high putrescine concentration in trophozoites demonstrated by NMR spectroscopy (9.5 mM) (Bakker-Grunwald *et al.*, 1995) reinforces the physiological significance of putrescine. However, the fate of putrescine is unknown as neither spermine nor spermidine has been demonstrated in *E. histolytica*.

There is controversy regarding the presence or absence of trypanothione, a spermidine-containing thiol, in *E. histolytica*. Trypanothione is a major thiol in trypanosomes and leishmania (Fairlamb and Cerami, 1992) and contains two molecules of glutathione joined by a spermidine linker. The first reports detected the presence of trypanothione in *E. histolytica* (Ondarza *et al.*, 1997) but were contradicted soon after (Ariyanayagam and Fairlamb, 1999). More recently another study reaffirmed its presence

(Ondarza *et al.*, 2005). However, the gene encoding trypanothione reductase reported from *E. histolytica* strain HK-9 (AF503571) has no homologue in the genome of HM-1:IMSS. Although this matter has not been resolved, there is general agreement that the major thiol in *E. histolytica* is cysteine (Fahey *et al.*, 1984).

The *E. histolytica* genome encodes a 46 kDa ornithine decarboxylase with similarity to both plant and vertebrate enzymes, and there is also the 96 kDa ornithine/arginine/lysine decarboxylase (see Section 4.2.4). Only the former enzyme has been characterised at the biochemical level (Arteaga-Nieto *et al.*, 2002) and has been shown to be insensitive to difluoromethylornithine (DFMO), as is *E. histolytica* (Gillin *et al.*, 1984).

The conversion of arginine into putrescine via agmatine, in a reaction initiated by arginine decarboxylase, is generally present in bacteria and plants. Although arginine decarboxylase is present in *E. histolytica*, agmatinase (EC 3.5.3.11), which further catalyses conversion of agmatine into putrescine and urea, appears absent. However, one gene identified as a 33 kDa arginase also shares 21% sequence identity with human mitochondrial agmatinase and therefore its substrates need to be examined on the biochemical level to see whether the enzyme can act on arginine, agmatine, or both. At present, the role of arginine decarboxylase in *E. histolytica* is not clear, although as mentioned earlier this enzyme may also be involved in acid resistance in *E. histolytica*.

4.4. Biosynthesis of amino acids

4.4.1. Cysteine and serine

One of the areas in which reduction of metabolism is most evident is in amino acid biosynthesis. Biosynthetic pathways for most amino acids other than serine and cysteine (Ali *et al.*, 2003, 2004a; Nozaki *et al.*, 1998a, 1999) have been lost in *E. histolytica*. Similarly, *P. falciparum*, which predominantly acquires amino acids from host haemoglobins, lacks biosynthesis of most amino acids (Gardner *et al.*, 2002). Intracellular concentrations of some amino acids (glutamate, leucine, valine and proline in descending order of abundance) are very high in *E. histolytica* ranging from 6 to 21 mM (Bakker-Grunwald *et al.*, 1995). In particular, the glutamate and proline concentrations are much higher in the cells than in the growth medium (21 and 7.3 mM vs. 5.9 and 1.8 mM, respectively). Glutamate accounts for over one-third of the total amino acid pool (Bakker-Grunwald *et al.*, 1995), and is likely to play a central role in homeostasis not only of amino acids but also of energy metabolism in general. Thus, it is likely that these amino acids are actively taken up by as-yet unidentified amino acid transporters.

Retention of the serine and cysteine biosynthetic pathways when the others have been lost is likely related to the physiological importance of

cysteine, which is the major intracellular thiol of this parasite. The cysteine biosynthetic pathway consists of two major steps, catalysed by serine acetyltransferase (EC 2.3.1.30), which produces *O*-acetylserine from serine and acetyl-coenzyme A, and cysteine synthase (EC 2.5.1.47), which subsequently transfers an alanyl moiety from *O*-acetylserine to sulphide to produce cysteine. *E. histolytica* possesses three genes each for cysteine synthase and serine acetyltransferase. Cysteine synthases 1 and 2 were considered to be allelic isotypes (Nozaki *et al.*, 1998b), while cysteine synthase 3 appears to be distinct, with only 83% identity to cysteine synthases 1 and 2. In contrast, all three serine acetyltransferase genes seem to be distinct, showing only 48–73% identity (V. Ali and T. Nozaki., unpublished data). It was previously shown that cysteine synthases 1 and 2 and serine acetyltransferase 1 are unique in that (a) they do not form a heterocomplex, in contrast to other organisms (Bogdanova and Hell, 1997; Droux *et al.*, 1998) and (b) serine acetyltransferase 1 is sensitive to allosteric inhibition by both L-cysteine and L-cystine (Nozaki *et al.*, 1999). Since all variants of these two enzymes lack organelle-targeting sequences, the significance of the multiple isotypes is unknown. It is important to determine subcellular distribution and specific functions of these isotypes to understand the significance of the redundancy. As this pathway is absent in humans, it is a rational target for development of new chemotherapeutic drugs against amoebiasis.

Serine is synthesised *de novo* utilising the glycolytic intermediate 3-phosphoglycerate, in a pathway that includes three sequential reactions catalysed by D-phosphoglycerate dehydrogenase (EC 1.1.1.95), phospho-L-serine aminotransferase (EC 2.6.1.52), and *O*-phospho L-serine phosphatase (EC 3.1.3.3). Although the final enzyme has not yet been enzymologically and functionally analysed, the first two enzymes have been characterised (Ali and Nozaki, 2006; Ali *et al.*, 2004a).

4.4.2. Interconversion of glutamate–glutamine and aspartate–asparagine

The single step interconversions of glutamate and glutamine, catalysed by glutamate synthase (EC 1.4.1.13) and glutamine synthetase (EC 6.3.1.2), and of aspartate and asparagine by asparagine synthase (EC 6.3.5.4) are found in *E. histolytica*. There are two isotypes of glutamine synthetase with 47% amino acid identity and five candidate genes. NADPH-dependent glutamate synthase (EC 1.4.1.13) catalyses the formation of two glutamates from glutamine and 2-oxo-glutarate in bacteria, yeasts and plants, and together with glutamine synthetase is involved in ammonia fixation under ammonia-restricted conditions. NADPH-dependent glutamate synthase is normally composed of two large and two small subunits (Petoukhov *et al.*, 2003). Although three genes encoding the small subunit are present, the large subunit appears to be absent in *E. histolytica*. These

putative NADPH-dependent glutamate synthase small subunits share 80% amino acid identity and show 44% amino acid identity to homologues from the Archaea. The similarity to archaeal-type glutamate synthase (Nesbo *et al.*, 2001) suggests that the *E. histolytica* small subunits may function as a glutamate synthase without the large subunit, as shown for gltA from the archaean *Pyrococcus* (Jongsareejit *et al.*, 1997).

The two enzymes that catalyse interconversion between aspartate and asparagine, aspartate ammonia ligase (EC 6.3.1.1) and asparaginase (EC 3.5.1.1; see Section 4.2.1), are present in *E. histolytica*. Two types of aspartate ammonia ligases, AsnA and AsnB, are known from other organisms: the former utilises only ammonia, while the latter uses both ammonia and glutamine as amide donors in a reverse reaction. Mammals possess only AsnA, whereas prokaryotes have both AsnA and AsnB (Boehlein *et al.*, 1996; Nakamura *et al.*, 1981). Interestingly, *E. histolytica* possesses only the AsnB homologue. Thus, the amoebic enzyme is likely involved in the formation of glutamate from glutamine, in addition to asparagine formation from aspartate.

4.4.3. Synthesis of glutamate and aspartate

Glutamate can be formed from 2-oxo-glutarate and ammonia in a reversible reaction catalysed by glutamate dehydrogenase (EC 1.4.1.2), which is present in *E. histolytica*. It is known that this enzyme plays a dominant role in ammonia fixation under ammonia-non-restricted conditions as this reaction consumes no ATP. In addition, glutamate dehydrogenase is also involved in gluconeogenesis from glutamate.

Aspartate ammonia-lyase (synonymous with aspartase, EC 4.3.1.1), which decomposes aspartate into fumarate and ammonia in a reversible reaction, is also present in *E. histolytica* (see Section 4.2.1).

4.5. Lipid metabolism

For *E. histolytica*, the lack of oxidative phosphorylation means that the high energy content of lipids such as fatty acids cannot be exploited. Therefore, lipids such as phospholipids and cholesterol are primarily membrane components in *E. histolytica* (Das *et al.*, 2002; Sawyer *et al.*, 1967). Although these components are mainly acquired from their food or from the human host, *E. histolytica* does have some capability for biosynthesis, as well as extending and remodelling lipids, and for attaching lipids to proteins.

4.5.1. Lipid biosynthetic capabilities

4.5.1.1. Polyisoprene biosynthesis and protein prenylation Cholesterol is an important membrane constituent generated from C_5 isoprene precursors. *E. histolytica* trophozoites in axenic culture need cholesterol in their

growth medium (Reeves, 1984), and it is likely that they acquire it from their human host. Reeves (1984) even cites several studies which show that hypercholesteremia in the host increases the damage inflicted by amoebic infection. *E. histolytica* lacks several enzymes for the classical sterol biosynthesis pathway (Schroepfer, 1981). The first stage of sterol biosynthesis is the formation of isopentenyl- or dimethylallyl diphosphate. In the *E. histolytica* genome no candidate genes for the generation of these intermediates were found, neither for the mevalonate pathway nor for the mevalonate-independent methylerythritol 4-phosphate (MEP) pathway that operates in bacteria and plants (Hunter *et al.*, 2003; Rohmer *et al.*, 1993). In a later step towards cholesterol synthesis, two molecules of C_{15} farnesyl diphosphate are dimerised to give C_{30} presqualene diphosphate by squalene synthetase (EC 2.5.1.21). This enzyme activity and those catalysing the subsequent steps also appear to be absent. The genome data thus support the long-standing conclusion that cholesterol biosynthesis is absent from *E. histolytica*.

Unexpectedly, the *E. histolytica* genome appears to encode enzymes involved in the intermediate stages of cholesterol biosynthesis from C_5 isopentenyl diphosphate to C_{15} farnesyl diphosphate. The latter compound, and the larger C_{20} compound geranylgeranyl diphosphate, may serve as precursors for the hydrophobic modification of GTP-binding proteins allowing them to bind to membranes (Grunler *et al.*, 1994). Protein prenylation is a ubiquitous process. It is important in human cell biology, health and disease (McTaggart, 2006), but it is also essential for parasites such that protein farnesylation has been proposed as a potential novel target for anti-parasitic chemotherapy (Maurer-Stroh *et al.*, 2003), including anti-*E. histolytica* chemotherapy (Ghosh *et al.*, 2004).

The first enzyme in this pathway is the isopentenyl-diphosphate δ-isomerase that catalyses the conversion of isopentenyl diphosphate to dimethylallyl diphosphate (EC 5.3.3.2). There is a single gene encoding this enzyme that is of presumed bacterial origin and is on the list of LGT candidates. The two isomeric C_5 isoprenyl diphosphates undergo condensation to C_{10} geranyl diphosphate, catalysed by geranyl-diphosphate synthase (EC 2.5.1.1). Farnesyl-diphosphate synthase (EC 2.5.1.10) then adds another C_5 unit to give C_{15} farnesyl diphosphate. Finally geranylgeranyl-diphosphate synthase (EC 2.5.1.29) adds another C_5 prenyl unit to give C_{20} geranylgeranyl diphosphate. The genome contains five putative prenyl transferase genes, which all have been annotated as geranylgeranyl-diphosphate synthases. Their sequences are highly similar, with the exception that the ORFs are disrupted in two of them (XP_650479 and XP_655958). These prenyl transferases appear to be of bacterial origin as well, and XP_650913 is on the list of LGT candidates. When searching for geranyl-diphosphate synthase or farnesyl-diphosphate synthase in the *E. histolytica* genome, the closest

matches are for the same genes, so that the substrate specificity of these enzymes is unclear and needs to be examined biochemically.

The *E. histolytica* genome contains one sequence each for the α and β chains of protein farnesyltransferase (EC 2.5.1.58), which were previously cloned and characterised as recombinant proteins (Kumagai *et al.*, 2004).

In addition to the protein farnesyltransferase, a protein geranylgeranyltransferase I (EC 2.5.1.59) β chain has recently been cloned and expressed together with the protein farnesyltransferase α chain (Makioka *et al.*, 2006). The heterodimeric molecule had protein geranylgeranyltransferase activity of unusually broad substrate specificity. The α and β chains of the protein (Rab-)geranylgeranyltransferase II (EC 2.5.1.60) have also been cloned, as cDNAs (M. Kumagai, A. Makioka, T. Takeuchi and T. Nozaki, unpublished data).

The *E. histolytica* genome encodes candidate enzymes for the modification of prenylated proteins. There are two highly divergent proteins both identified as CAAX prenyl proteases (EC 3.4.24.84). CAAX is the carboxy terminus of the substrate protein in which C is the prenylated cysteine residue, A is an aliphatic amino acid and X is the terminal residue. The proteases cleave after the modified cysteine. After the processing step, a prenylcysteine carboxyl methyltransferase (EC 2.1.1.100) methylates the carboxy-terminal residue; there are two divergent candidate genes for this enzyme.

Taken together, the *E. histolytica* genome contains all the necessary genes to encode the pathway from isopentenyl diphosphate to a processed farnesylated or geranylgeranylated protein. The source of the starting material, isopentenyl diphosphate, remains unknown at this time, but there may be a previously unknown pathway for its synthesis or *E. histolytica* may be able to aquire it from its environment.

4.5.1.2. Fatty acid biosynthesis *E. histolytica* encodes an unusual 138 kDa acetyl-CoA carboxylase with 2 bacterial-type carboxylase domains, an acetyl-CoA carboxylase and a pyruvate carboxylase. Since no biotin carboxylase domain is found in the *E. histolytica* genome, it was proposed that the enzyme removes a carboxyl group from oxaloacetate and transfers it to acetyl-CoA, forming malonyl-CoA and pyruvate (Jordan *et al.*, 2003; Loftus *et al.*, 2005). This fusion protein has not been identified in any organisms other than *Giardia* and *Entamoeba*.

In the classical pathway of fatty acid biosynthesis, starting from acetyl-CoA sequential two-carbon units are added from malonyl-CoA. In each round of extension, the β-keto group is reduced in three steps before a new two-carbon unit is added. The whole pathway is carried out in a large fatty acid synthase complex, where the growing chain is linked to an acyl carrier protein. *E. histolytica* lacks this classical pathway. There are,

however, plant homologues of fatty acid chain elongases such as *Arabidopsis thaliana* KCS1 (Todd *et al.*, 1999). There are eight putative fatty acid elongases in the *E. histolytica* genome, and all are very similar to each other. These enzymes could be involved in elongation of fatty acids taken up from the host or food sources, but their function and substrate specificity are unknown at this time.

4.5.2. Phospholipid metabolism

Phospholipids amount to 60–70% of the total lipids in *E. histolytica* (Sawyer *et al.*, 1967). So far little information is available at the biochemical level on how phospholipids are synthesised, acquired or remodelled. The genome project has revealed a number of genes, indicating that the phospholipid metabolism could be more complex than expected.

4.5.2.1. Phospholipid biosynthesis In order to produce phospholipids one has to generate the important intermediate phosphadidate (1,2-diacylglycerol 3-phosphate) by phosphorylation and acylation of glycerol. *E. histolytica* contains one gene for a glycerol kinase (EC 2.7.1.30). The second step would be the transfer of the acyl group to glycerol-3-phosphate by glycerol-3-phosphate *O*-acyltransferase (EC 2.3.1.15), but no candidate gene for this enzyme has been found in the genome. There are, however, two potential 1-acylglycerol-3-phosphate *O*-acyltransferases (EC 2.3.1.51) that could attach the second acyl group. After the attachment of the acyl groups, and in preparation for the attachment of the activated aminoalcohols, the phosphate is removed by phosphadidate phosphatase (EC 3.1.3.4), for which there is one gene, resulting in a diacylglycerol.

The activation of ethanolamine (EC 2.7.1.82) or choline (EC 2.7.1.32) for attachment to the phosphadidate starts with phosphorylation. There are two genes identified as choline/ethanolamine kinases that share 37% amino acid identity. Next, ethanolamine phosphate and choline phosphate are converted into CDP-ethanolamine (EC 2.7.7.14) and CDP-choline (EC 2.7.7.15), respectively. The genome encodes 2 enzymes sharing 57% sequence identity that are identified as ethanolamine-phosphate cytidylyltransferases. The substrate specificity of these enzymes needs to be examined on the biochemical level. Finally, the activated ethanolamine or choline is attached to diacylglycerol by the enzymes ethanolaminephosphotransferase (EC 2.7.8.1) or diacylglycerol cholinephosphotransferase (EC 2.7.8.2) producing phosphatidylethanolamine or phosphatidylcholine, respectively. For these activities a total of eight possible genes are found that share varying degrees of sequence similarity.

In *E. histolytica*, an alternative pathway of phospholipid biosynthesis could involve the biosynthesis of phosphatidylserine. In this pathway, the phosphatidate itself is activated by CTP in a reaction catalysed by phosphatidate cytidylyltransferase (EC 2.7.7.41) resulting in CDP-diacylglycerol. Three genes have been identified. Phosphatidylserine synthase then catalyses the reaction of CDP-diacylglycerol with serine to give phosphatidylserine (EC 2.7.8.8); one gene has been found.

Some organisms can form phosphatidylethanolamine from phosphatidylserine using a decarboxylase, but such an enzyme appears to be absent from the *E. histolytica* genome. There are, however, several candidate methyltransferases of yet unknown substrate specificity, which might be able to generate phosphatidylcholine from phosphatidylethanolamine.

Taken together, large portions of the pathways needed to generate the most important phospholipids can be assembled from genes tentatively identified to date in the *E. histolytica* genome. The first acylation of glycerol 3-phosphate to lysophosphatidate remains an important gap. As *E. histolytica* could potentially aquire all the necessary phospholipids from the host, the functional relevance of the described biosynthetic pathways may not be high.

Finally, two additional interesting enzymes present in *E. histolytica* should be mentioned. The first was previously characterised using cDNA sequences and recombinant proteins as L-*myo*-inositol 1-phosphate synthase (EC 5.5.1.4; Lohia *et al.*, 1999). This enzyme catalyses the complicated isomerisation of glucose 6-phosphate to L-*myo*-inositol 1-phosphate. Inositol is found in phosphatidylinositol (PI) and in GPI-anchors of some membrane proteins, as well as playing a major role in signal transduction via the secondary messenger 1,4,5-inositol trisphosphate. There are three *myo*-inositol 1-phosphate synthase genes, all highly similar to each other and to the previously sequenced cDNA.

The second is phospholipid-cholesterol acyltransferase (EC 2.3.1.43), which transfers an acyl group from phospholipids such as phosphatidylcholine to cholesterol giving a cholesterol ester. The genome contains seven genes for this enzyme. So far nothing is known about the importance of cholesterol esters for *E. histolytica*.

4.5.2.2. Phospholipid degradation Phospholipids are degraded by phospholipases. Whereas phospholipases A1 (EC 3.1.1.32) and A2 (EC 3.1.1.4) cleave the acyl residues in the 1 or 2 position of the glycerol core, phospholipases C (EC 3.1.4.3) and D (EC 3.1.4.4) cleave at the phosphate, phospholipase C on the glycerol side, and phospholipase D on the aminoalcohol side. In *E. histolytica* phospholipase A activity has been implicated in virulence (Ravdin *et al.*, 1985), as it liberates toxic fatty acids and

lysophospholipids (Said-Fernandez and Lopez-Revilla, 1988). Phospholipases A have been found in two forms, a membrane-bound Ca-dependent form active at alkaline pH and a soluble Ca-independent form active at acid pH (Long-Krug *et al.*, 1985; Vargas-Villarreal *et al.*, 1998). The genome encodes 11 potential phospholipases A with predicted pI values between 4.8 and 8.8 and various degrees of sequence similarity. In addition, the *E. histolytica* genome encodes three potential phospholipases D.

Finally, there are two highly similar genes for phospholipases C, but these are homologous to PI-specific phospholipases C (EC 3.1.4.11) and most likely do not cleave PI or phosphatidylcholine but GPI-anchors instead. So far there are no studies using individual recombinant phospholipases, and it is not yet known how much these enzymes may contribute to the virulence of *E. histolytica*.

4.6. Coenzyme A biosynthesis and pantothenate metabolism

Analysis of the genome revealed a complete lack of known folate-dependent enzymes and folate transporters, suggesting this cofactor is not utilised by *E. histolytica*. This is at odds with a study on the nutritional requirements of *E. histolytica* in which folate was found to be essential for growth (Diamond and Cunnick, 1991). More experimental research will be needed to resolve this discrepancy. Most organisms require folate as a cofactor for several reactions of amino acid metabolism and for synthesis of thymidylate, a component of DNA. The microsporidian *E. cuniculi*, which possesses the smallest-known eukaryotic genome, still contains a folate transporter and several folate-dependent enzymes (Katinka *et al.*, 2001). In eukaryotes possessing mitochondria or chloroplasts, folate is required for the formylation of methionine on the initiator tRNA used for organelle protein synthesis. Although *E. histolytica* possesses a mitochondrion-derived organelle, the mitosome, there is no organellar genome (Leon-Avila and Tovar, 2004) and so no need for organellar protein synthesis. The most important metabolic consequences of the loss of folate metabolism for *E. histolytica* are therefore the absence of thymidylate synthesis and methionine recycling, although it remains possible that *E. histolytica* possesses folate-independent enzymes carrying out these steps.

Phosphopantothenoyl-cysteine decarboxylase (EC 4.1.1.36) and phosphopantothenoyl-cysteine synthetase (EC 6.3.2.5, synonymous with phosphopantothenate-cysteine ligase) exist as a fusion protein in *E. histolytica*, as in Bacteria and Archaea. The amino- and carboxyl-terminal domains possess decarboxylase and synthetase activity, respectively (Kupke, 2002, 2004; Kupke *et al.*, 2000; Strauss *et al.*, 2001). The role of this enzyme in coenzyme A biosynthesis is not well understood in *E. histolytica* as the other necessary enzymes are absent.

REFERENCES - METABOLISM SECTION

Ali, V., Shigeta, Y., and Nozaki, T. (2003). Molecular and structural characterization of NADPH-dependent D-glycerate dehydrogenase from the enteric parasitic protist *Entamoeba histolytica*. *Biochem. J.* **375,** 729–736.

Ali, V., Hashimoto, T., Shigeta, Y., and Nozaki, T. (2004a). Molecular and biochemical characterization of D-phosphoglycerate dehydrogenase from *Entamoeba histolytica*. A unique enteric protozoan parasite that possesses both phosphorylated and nonphosphorylated serine metabolic pathways. *Eur. J. Biochem.* **271,** 2670–2681.

Ali, V., and Nozaki, T. (2006). Biochemical and functional characterization of phosphoserine aminotransferase from *Entamoeba histolytica*, which possesses both phosphorylated and non-phosphorylated serine metabolic pathways. *Mol. Biochem. Parasitol.* **145,** 71–83.

Anderson, I. J., and Loftus, B. J. (2005). *Entamoeba histolytica*: Observations on metabolism based on the genome sequence. *Exp. Parasitol.* **110,** 173–177.

Andersson, J. O., Hirt, R. P., Foster, P. G., and Roger, A. J. (2006). Evolution of four gene families with patchy phylogenetic distributions: Influx of genes into protist genomes. *BMC Evol. Biol.* **6,** 27.

Ariyanayagam, M. R., and Fairlamb, A. H. (1999). *Entamoeba histolytica* lacks trypanothione metabolism. *Mol. Biochem. Parasitol.* **103,** 61–69.

Arteaga-Nieto, P., López-Romero, E., Teran-Figueroa, Y., Cano-Canchola, C., Luna Arias, J. P., Flores-Carreón, A., and Calvo-Méndez, C. (2002). *Entamoeba histolytica*: Purification and characterization of ornithine decarboxylase. *Exp. Parasitol.* **101,** 215–222.

Avila, E.E., Martinez-Alcaraz, E. R.,Barbosa-Sabanero, G.,Rivera-Baron, E. I., Arias-Negrete, S., and Zazueta-Sandova, L. R. (2002). Subcellular localization of the NAD^+-dependent alcohol dehydrogenase in *Entamoeba histolytica* trophozoites. *J. Parasitol.* **88,** 217–222.

Bakker-Grunwald, T., Martin, J. B., and Klein, G. (1995). Characterization of glycogen and amino acid pool of *Entamoeba histolytica* by C13-NMR spectroscopy. *J. Eukaryot. Microbiol.* **42,** 346–349.

Beanan, M. J., and Bailey, G. B. (1995). The primary structure of an *Entamoeba histolytica* enolase. *Mol. Biochem. Parasitol.* **69,** 119–121.

Boehlein, S. K., Schuster, S. M., and Richards, N. G. (1996). Glutamic acid gammamonohydroxamate and hydroxylamine are alternate substrates for *Escherichia coli* asparagine synthetase B. *Biochemistry* **35,** 3031–3037.

Bogdanova, N., and Hell, R. (1997). Cysteine synthesis in plants: Protein–protein interactions of serine acetyltransferase from *Arabidopsis thaliana*. *Plant J.* **11,** 251–262.

Bruchhaus, I., and Tannich, E. (1993). Primary structure of the pyruvate phosphate dikinase in *Entamoeba histolytica*. *Mol. Biochem. Parasitol*. **62,** 153–156.

Bruchhaus, I., and Tannich, E. (1994). Purification and molecular characterization of the NAD^+-dependent acetaldehyde/alcohol dehydrogenase from *Entamoeba histolytica*. *Biochem. J.* **303,** 743–748.

Cedar, H., and Schwartz, J. H. (1967). Localization of the two-L-asparaginases in anaerobically grown *Escherichia coli*. *J. Biol. Chem.* **242,** 3753–3755.

Chi, A., and Kemp, R. G. (2000). The primordial high energy compound: ATP or inorganic pyrophosphate? *J. Biol. Chem.* **275,** 35677–35679.

Chi, A. S., Deng, Z., Albach, R. A., and Kemp, R. G. (2001). The two phosphofructokinase gene products of *Entamoeba histolytica*. *J. Biol. Chem*. **276,** 19974–19981.

Clark, C. G., Kaffashian, F., Tawari, B., Windsor, J. J., Twigg-Flesner, A., Davies-Morel, M. C. G., Blessmann, J., Ebert, F., Peschel, B., Le Van, A., Jackson, C. J., Macfarlane, L., et al. (2006b). New insights into the phylogeny of *Entamoeba* species provided by analysis of four new small-subunit rRNA genes. *Int. J. Syst. Evol. Microbiol.* **56,** 2235–2239.

Coombs, G. H., and Mottram, J. C. (2001). Trifluoromethionine, a prodrug designed against methionine gamma-lyase-containing pathogens, has efficacy *in vitro* and *in vivo* against *Trichomonas vaginalis*. *Antimicrob. Agents Chemother.* **45,** 1743–1745.

Das, S., Stevens, T., Castillo, C., Villasenor, A., Arredondo, H., and Reddy, K. (2002). Lipid metabolism in mucous-dwelling amitochondriate protozoa. *Int. J. Parasitol.* **32,** 655–675.

Deng, Z., Huang, M., Singh, K., Albach, R. A., Latshaw, S. P., Chang, K. P., and Kemp, R. G. (1998). Cloning and expression of the gene for the active PPi-dependent phosphofructokinase of *Entamoeba histolytica*. *Biochem. J.* **329,** 659–664.

Diamond, L. S., and Cunnick, C. C. (1991). A serum-free, partly defined medium, PDM-805, for the axenic cultivation of *Entamoeba histolytica* Schaudinn, 1903 and other *Entamoeba*. *J. Protozool.* **38,** 211–216.

Droux, M., Ruffet, M. L., Douce, R., and Job, D. (1998). Interactions between serine acetyltransferase and O-acetylserine (thiol) lyase in higher plants: Structural and kinetic properties of the free and bound enzymes. *Eur. J. Biochem.* **255,** 235–245.

Elnekave, K., Siman-Tov, R., and Ankri, S. (2003). Consumption of L-arginine mediated by *Entamoeba histolytica* L-arginase (EhArg) inhibits amoebicidal activity and nitric oxide production by activated macrophages. *Parasite Immunol.* **25,** 597–608.

Fahey, R. C., Newton, G. L., Arrick, B., Overdank-Bogart, T., and Aley, S. B. (1984). *Entamoeba histolytica*: A eukaryote without glutathione metabolism. *Science* **224,** 70–72.

Fairlamb, A. H., and Cerami, A. (1992). Metabolism and functions of trypanothione in the Kinetoplastida. *Annu. Rev. Microbiol.* **46,** 695–729.

Farri, T. A., Sargeaunt, P. G., Warhurst, D. C., Williams, J. E., and Bhojnani, R. (1980). Electrophoretic studies of the hexokinase of *Entamoeba histolytica* groups I to IV. *Trans. R. Soc. Trop. Med. Hyg*. **74,** 672–673.

Field, J., Rosenthal, B., and Samuelson, J. (2000). Early lateral transfer of genes encoding malic enzyme, acetyl-CoA synthetase and alcohol dehydrogenases from anaerobic prokaryotes to *Entamoeba histolytica*. *Mol. Microbiol*. **38,** 446–455.

Gardner, M. J., Hall, N., Fung, E., White, O., Berriman, M., Hyman, R. W., Carlton, J. M., Pain, A., Nelson, K. E., Bowman, S., Paulsen, I. T., James, K., et al. (2002). Genome sequence of the human malaria parasite *Plasmodium falciparum*. *Nature* **419,** 498–511.

Ghosh, S., Chan, J. M., Lea, C. R., Meints, G. A., Lewis, J. C., Tovian, Z. S., Flessner, R. M., Loftus, T. C., Bruchhaus, I., Kendrick, H., Croft, S. L., Kemp, R. G., et al. (2004). Effects of bisphosphonates on the growth of *Entamoeba histolytica* and *Plasmodium* species in vitro and in vivo. *J. Med. Chem*. **47,** 175–187.

Gillin, F. D., Reiner, D. S., and McCann, P. P. (1984). Inhibition of growth of *Giardia lamblia* by difluoromethylornithine, a specific inhibitor of polyamine biosynthesis. *J. Protozool.* **31,** 161–163.

Grunler, J., Ericsson, J., and Dallner, G. (1994). Branch-point reactions in the biosynthesis of cholesterol, dolichol, ubiquinone and prenylated proteins. *Biochim. Biophys. Acta* **1212,** 259–277.

Hidalgo, M. E., Sanchez, R., Perez, D. G., Rodriguez, M. A., Garcia, J., and Orozco, E. (1997). Molecular characterization of the *Entamoeba histolytica* enolase gene and modelling of the predicted protein. *FEMS Microbiol. Lett*. **148,** 123–129.

Horner, D. S., Hirt, R. P., and Embley, T. M. (1999). A single eubacterial origin of eukaryotic pyruvate: Ferredoxin oxidoreductase genes: Implications for the evolution of anaerobic eukaryotes. *Mol. Biol. Evol.* **16,** 1280–1291.

Hunter, W. N., Bond, C. S., Gabrielsen, M., and Kemp, L. E. (2003). Structure and reactivity in the non-mevalonate pathway of isoprenoid biosynthesis. *Biochem. Soc. Trans.* **31,** 537–542.

Jongsareejit, B., Rahman, R. N., Fujiwara, S., and Imanaka, T. (1997). Gene cloning, sequencing and enzymatic properties of glutamate synthase from the hyperthermophilic archaeon *Pyrococcus* sp. KOD1. *Mol. Gen. Genet.* **254,** 635–642.

Jordan, I. K., Henze, K., Fedorova, N. D., Koonin, E. V., and Galperin, M. Y. (2003). Phylogenomic analysis of the *Giardia intestinalis* transcarboxylase reveals multiple instances of domain fusion and fission in the evolution of biotin-dependent enzymes. *J. Mol. Microbiol. Biotechnol.* **5,** 172–189.

Katinka, M. D., Duprat, S., Cornillot, E., Metenier, G., Thomarat, F., Prensier, G., Barbe, V., Peyretaillade, E., Brottier, P., Wincker, P., Delbac, F., El Alaoui, H., et al. (2001). Genome sequence and gene compaction of the eukaryote parasite *Encephalitozoon cuniculi*. *Nature* **414,** 450–453.

Kimura, A., Hara, Y., Kimoto, T., Okuno, Y., Minekawa, Y., and Nakabayashi, T. (1996). Cloning and expression of a putative alcohol dehydrogenase gene of *Entamoeba histolytica* and its application to immunological examination. *Clin. Diagn. Lab. Immunol.* **3,** 270–274.

Knodler, L. A., Edwards,M. R., and Schofield, P. J. (1994). The intracellular amino acid pools of *Giardia intestinalis*, *Trichomonas vaginalis*, and *Crithidia luciliae*. *Exp. Parasitol.* **79,** 117–125.

Kroschewski, H., Ortner, S., Steipe, B., Scheiner, O., Wiedermann, G., and Duchêne, M. (2000). Differences in substrate specificity and kinetic properties of the recombinant hexokinases HXK1 and HXK2 from *Entamoeba histolytica*. *Mol. Biochem. Parasitol.* **105,** 71–80.

Kulda, J. (1999). Trichomonads, hydrogenosomes and drug resistance. *Int. J. Parasitol.* **29,** 199–212.

Kumar, A., Shen, P. S., Descoteaux, S., Pohl, J., Bailey, G., and Samuelson, J. (1992). Cloning and expression of an $NADP^+$-dependent alcohol dehydrogenase gene of *Entamoeba histolytica*. *Proc. Natl. Acad. Sci. USA* **89,** 10188–10192.

Kumagai, M., Makioka, A., Takeuchi, T., and Nozaki, T. (2004). Molecular cloning and characterization of a protein farnesyltransferase from the enteric protozoan parasite *Entamoeba histolytica*. *J. Biol. Chem*. **279,** 2316–2323.

Kupke, T., Uebele, M., Schmid, D., Jung, G., Blaesse, M., and Steinbacher, S. (2000). Molecular characterization of lantibiotic-synthesizing enzyme EpiD reveals a function for bacterial Dfp proteins in coenzyme A biosynthesis. *J. Biol. Chem.* **275,** 31838–31846.

Kupke, T. (2002). Molecular characterization of the 40-phosphopantothenoylcysteine synthetase domain of bacterial dfp flavoproteins. *J. Biol. Chem*. **277,** 36137–36145.

Kupke, T. (2004). Active-site residues and amino acid specificity of the bacterial 40-phosphopantothenoylcysteine synthetase CoaB. *Eur. J. Biochem*. **271,** 163–172.

Landa, A., Rojo-Dominguez, A., Jimenez, L., and Fernandez-Velasco, D. A. (1997). Sequencing, expression and properties of triosephosphate isomerase from *Entamoeba histolytica*. *Eur. J. Biochem*. **247,** 348–355.

Leon-Avila, G., and Tovar, J. (2004). Mitosomes of *Entamoeba histolytica* are abundant mitochondrion-related remnant organelles that lack a detectable organellar genome. *Microbiology* **150,** 1245–1250.

Linstead, D., and Cranshaw, M. A. (1983). The pathway of arginine catabolism parasitic flagellate *Trichomonas vaginalis*. Mol. Biochem. Parasitol. **8,** 241–252.

Lo, H. S., and Reeves, R. E. (1978). Pyruvate-to-ethanol pathway in *Entamoeba histolytica*. *Biochem. J*. **171,** 225–230.

Lobelle-Rich, P. A., and Reeves, R. E. (1983). Separation and characterization of two UTPutilizing hexose phosphate uridylyltransferases from *Entamoeba histolytica*. *Mol. Biochem. Parasitol*. **7,** 173–182.

Lockwood, B. C., and Coombs, G. H. (1991). Purification and characterization of methionine gamma-lyase from *Trichomonas vaginalis*. *Biochem. J*. **279,** 675–682.

Loftus, B., Anderson, I., Davies, R., Alsmark, U. C., Samuelson, J., Amedeo, P., Roncaglia, P., Berriman, M., Hirt, R. P., Mann, B. J., Nozaki, T., Suh, B., *et al*. (2005). The genome of the protist parasite *Entamoeba histolytica*. *Nature 433,* 865–868.

Long-Krug, S. A., Fischer, K. J., Hysmith, R. M., and Ravdin, J. I. (1985). Phospholipase A enzymes of *Entamoeba histolytica*: Description and subcellular localization. *J. Infect. Dis.* **152,** 536–541.

Majoros, W. H., Pertea, M., and Salzberg, S. L. (2004). TigrScan and GlimmerHMM: Two open source ab initio eukaryotic gene-finders. *Bioinformatics* **20,** 2878–2879.
Makioka, A., Kumagai, M., Takeuchi, T., and Nozaki, T. (2006). Characterization of protein geranylgeranyltransferase I from the enteric protist *Entamoeba histolytica*. *Mol. Biochem. Parasitol.* **145,** 216–225.
Marsh, J. J., and Lebherz, H. G. (1992). Fructose-bisphosphate aldolases: An evolutionary history. *Trends Biochem. Sci.* **17,** 110–113.
Martin, W., and Müller, M. (1998). The hydrogen hypothesis for the first eukaryote. *Nature* **392,** 37–41.
Maurer-Stroh, S., Washietl, S., and Eisenhaber, F. (2003). Protein prenyltransferases: Anchor size, pseudogenes and parasites. *Biol. Chem.* **384,** 977–989.
McTaggart, S. J. (2006). Isoprenylated proteins. *Cell. Mol. Life Sci.* **63,** 255–267.
Müller, M. (1986). Reductive activation of nitroimidazoles in anaerobic microorganisms. *Biochem. Pharmacol.* **35,** 37–41.
Mullikin, J. C., and Ning, Z. (2003). The phusion assembler. *Genome Res*. **13,** 81–90.
Nakamura, M., Yamada, M., Hirota, Y., Sugimoto, K., Oka, A., and Takanami, M. (1981). Nucleotide sequence of the asnA gene coding for asparagine synthetase of *E. coli* K-12. *Nucleic Acids Res*. **9,** 4669–4676.
Nesbo, C. L., L'Haridon, S., Stetter, K. O., and Doolittle, W. F. (2001). Phylogenetic analyses of two ''archaeal'' genes in thermotoga maritima reveal multiple transfers between archaea and bacteria. *Mol. Biol. Evol.* **18,** 362–375.
Nixon, J. E. J., Wang, A., Field, J., Morrison, H. G., McArthur, A. G., Sogin, M. L., Loftus, B. J., and Samuelson, J. (2002). Evidence for lateral transfer of genes encoding ferredoxins, nitroreductases, NADH oxidase, and alcohol dehydrogenase 3 from anaerobic prokaryotes to *Giardia lamblia* and *Entamoeba histolytica*. *Eukaryot. Cell* **1,** 181–190.
Lohia, A., Hait, N. C., and Lahiri Majumder, A. (1999). L-myo-inositol 1-phosphate synthase from *Entamoeba histolytica*. *Mol. Biochem. Parasitol*. **98,** 67–79.
Nozaki, T., Ali, V., and Tokoro, M. (2005). Sulfur-containing amino acid metabolism in parasitic protozoa. *Adv. Parasitol.* **60,** 1–99.
Nozaki, T., Arase, T., Shigeta, Y., Asai, T., Leustek, T., and Takeuchi, T. (1998a). Cloning and bacterial expression of adenosine-5´-triphosphate sulfurylase from the enteric protozoan parasite *Entamoeba histolytica*. *Biochim. Biophys. Acta* **1429,** 284–291.
Nozaki, T., Asai, T., Kobayashi, S., Ikegami, F., Noji, M., Saito, K., and Takeuchi, T. (1998b). Molecular cloning and characterization of the genes encoding two isoforms of cysteine synthase in the enteric protozoan parasite *Entamoeba histolytica*. *Mol. Biochem. Parasitol*. **97,** 33–44.
Nozaki, T., Asai, T., Sanchez, L. B., Kobayashi, S., Nakazawa, M., and Takeuchi, T. (1999). Characterization of the gene encoding serine acetyltransferase, a regulated enzyme of cysteine biosynthesis from the protist parasites *Entamoeba histolytica* and *Entamoeba dispar*. Regulation and possible function of the cysteine biosynthetic pathway in *Entamoeba*. *J. Biol. Chem*. **274,** 32445–32452.
Ondarza, R. N., Tamayo, E. M., Hurtado, G., Hernandez, E., and Iturbe, A. (1997). Isolation and purification of glutathionyl-spermidine and trypanothione from *Entamoeba histolytica*. *Arch. Med. Res*. **28,** 73–75.
Ondarza, R. N., Hurtado, G., Iturbe, A., Hernandez, E., Tamayo, E., and Woolery, M. (2005). Identification of trypanothione from the human pathogen *Entamoeba histolytica* by mass spectrometry and chemical analysis. *Biotechnol. Appl. Biochem*. **42,** 175–181.
Opperdoes, F. R., and Michels, P. A. (2001). Enzymes of carbohydrate metabolism as potential drug targets. *Int. J. Parasitol.* **31,** 482–490.
Ortner, S., Plaimauer, B., Binder, M., Scheiner, O., Wiedermann, G., and Duchêne, M. (1995). Molecular analysis of two hexokinase isoenzymes from *Entamoeba histolytica*. *Mol. Biochem. Parasitol*. **73,** 189–198.
Ortner, S., Binder, M., Scheiner, O., Wiedermann, G., and Duchêne, M. (1997a). Molecular and biochemical characterization of phosphoglucomutases from *Entamoeba histolytica* and *Entamoeba dispar*. *Mol. Biochem. Parasitol.* **90,** 121–129.

Ortner, S., Clark, C. G., Binder, M., Scheiner, O., Wiedermann, G., and Duchêne, M. (1997b). Molecular biology of the hexokinase isoenzyme pattern that distinguishes pathogenic *Entamoeba histolytica* from nonpathogenic *Entamoeba dispar*. *Mol. Biochem. Parasitol.* **86,** 85–94.

Parsons, M. (2004). Glycosomes: Parasites and the divergence of peroxisomal purpose. *Mol. Microbiol*. **53,** 717–724.

Petoukhov, M. V., Svergun, D. I., Konarev, P. V., Ravasio, S., van den Heuvel, R. H., Curti, B., and Vanoni, M. A. (2003). Quaternary structure of *Azospirillum brasilense* NADPHdependent glutamate synthase in solution as revealed by synchrotron radiation x-ray scattering. *J. Biol. Chem*. **278,** 29933–29939.

Ramos, F., and Wiame, J. M. (1982). Occurrence of a catabolic L-serine (L-threonine) deaminase in *Saccharomyces cerevisiae*. *Eur. J. Biochem*. **123,** 571–576.

Ravdin, J. I., Murphy, C. F., Guerrant, R. L., and Long-Krug, S. A. (1985). Effect of antagonists of calcium and phospholipase A on the cytopathogenicity of *Entamoeba histolytica*. *J. Infect. Dis*. **152,** 542–549.

Reeves, R. E. (1968). A new enzyme with the glycolytic function of pyruvate kinase. *J. Biol. Chem.* **243,** 3202–3204.

Reeves, R. E., Serrano, R., and South, D. J. (1976). 6-phosphofructokinase (pyrophosphate). Properties of the enzyme from *Entamoeba histolytica* and its reaction mechanism. *J. Biol. Chem*. **251,** 2958–2962.

Reeves, R. E., Montalvo, F. E., and Lushbaugh, T. S. (1971). Nicotinamide-adenine dinucleotide phosphate-dependent alcohol dehydrogenase. Enzyme from *Entamoeba histolytica* and some enzyme inhibitors. *Int. J. Biochem.* **2,** 55–64.

Reeves, R. E., Warren, L. G., Susskind, B., and Lo, H. S. (1977). An energy-conserving pyruvate-to-acetate pathway in *Entamoeba histolytica*. Pyruvate synthase and a new acetate thiokinase. *J. Biol. Chem*. **252,** 726–731.

Reeves, R. E. (1984). Metabolism of *Entamoeba histolytica* Schaudinn, 1903. *Adv. Parasitol.* **23,** 105–142.

Reeves, R. E., and South, D. J. (1974). Phosphoglycerate kinase (GTP). An enzyme from *Entamoeba histolytica* selective for guanine nucleotides. *Biochem. Biophys. Res. Commun*. **58,** 1053–1057.

Reid, M. F., and Fewson, C. A. (1994). Molecular characterization of microbial alcohol dehydrogenases. *Crit. Rev. Microbiol.* **20,** 13–56.

Rodriguez, M. A., Baez-Camargo, M., Delgadillo, D. M., and Orozco, E. (1996). Cloning and expression of an *Entamoeba histolytica* $NADP^+$-dependent alcohol dehydrogenase gene. *Biochim. Biophys. Acta* **1306,** 23–26.

Rodriguez-Romero, A., Hernandez-Santoyo, A., del Pozo Yauner, L., Kornhauser, A., and Fernandez-Velasco, D. A. (2002). Structure and inactivation of triosephosphate isomerase from *Entamoeba histolytica*. *J. Mol. Biol*. **322,** 669–675.

Rohmer, M., Knani, M., Simonin, P., Sutter, B., and Sahm, H. (1993). Isoprenoid biosynthesis in bacteria: A novel pathway for the early steps leading to isopentenyl diphosphate. *Biochem. J.* **295,** 517–524.

Rosenbaum, R. M., and Wittner, M. (1970). Ultrastructure of bacterized and axenic trophozoites of *Entamoeba histolytica* with particular reference to helical bodies. *J. Cell Biol.* **45,** 367–382.

Rotte, C., Stejskal, F., Zhu, G., Keithly, J. S., and Martin, W. (2001). Pyruvate: $NADP^+$ oxidoreductase from the mitochondrion of *Euglena gracilis* and from the apicomplexan *Cryptosporidium parvum*: A biochemical relic linking pyruvate metabolism in mitochondriate and amitochondriate protists. *Mol. Biol. Evol*. **18,** 710–720.

Saavedra, E., Olivos, A., Encalada, R., and Moreno-Sanchez, R. (2004). *Entamoeba histolytica*: Kinetic and molecular evidence of a previously unidentified pyruvate kinase. *Exp. Parasitol.* **106,** 11–21.

Saavedra, E., Encalada, R., Pineda, E., Jasso-Chavez, R., and Moreno-Sanchez, R. (2005). Glycolysis in *Entamoeba histolytica*. Biochemical characterization of recombinant glycolytic enzymes and flux control analysis. *FEBS J*. **272,** 1767–1783.

Saavedra Lira, E., Robinson, O., and Pérez Montfort, R. (1992). Partial nucleotide sequence of the enzyme pyruvate, orthophosphate dikinase of *Entamoeba histolytica* HM-1:IMSS. *Arch. Med. Res*. **23,** 39–40.

Said-Fernandez, S., and Lopez-Revilla, R. (1988). Free fatty acids released from phospholipids are the major heat-stable hemolytic factor of *Entamoeba histolytica* trophozoites. *Infect. Immun*. **56,** 874–879.

Samarawickrema, N. A., Brown, D. M., Upcroft, J. A., Thammapalerd, N., and Upcroft, P. (1997). Involvement of superoxide dismutase and pyruvate:ferredoxin oxidoreductase in mechanisms of metronidazole resistance in *Entamoeba histolytica*. *J. Antimicrob. Chemother*. **40,** 833–840.

Sanchez, L., Horner, D., Moore, D., Henze, K., Embley, T., and Müller, M. (2002). Fructose-1,6-bisphosphate aldolases in amitochondriate protists constitute a single protein subfamily´with eubacterial relationships. *Gene* **295,** 51–59.

Sargeaunt, P. G. (1987). The reliability of *Entamoeba histolytica* zymodemes in clinical diagnosis. *Parasitol. Today* **3,** 40–43.

Sargeaunt, P. G., Williams, J. E., and Grene, J. D. (1978). The differentiation of invasive and non-invasive *Entamoeba histolytica* by isoenzyme electrophoresis. *Trans. R. Soc. Trop. Med. Hyg.* **72,** 519–521.

Sawyer, M. K., Bischoff, J. M., Guidry, M. A., and Reeves, R. E. (1967). Lipids from *Entamoeba histolytica*. *Exp. Parasitol.* **20,** 295–302.

Schofield, P. J., and Edwards, M. R. (1994). Biochemistry—is *Giardia* opportunistic in its use of substrates? In ‘‘*Giardia*: FromMolecules to Disease’’ (R. C.A. Thompson, J. A. Reynoldson, and A. J. Lymbery, eds.), pp. 171–183. CAB International, Wallingford, UK.

Schroepfer, G. J., Jr. (1981). Sterol biosynthesis. Annu. Rev. Biochem. **50,** 585–621.

Snell, K. (1986). The duality of pathways for serine biosynthesis is a fallacy. *Trends Biochem. Sci.* **11,** 241–243.

Strauss, E., Kinsland, C., Ge, Y., McLafferty, F.W., and Begley, T. P. (2001). Phosphopantothenoylcysteine synthetase from *Escherichia coli*. Identification and characterization of the last unidentified coenzyme A biosynthetic enzyme in bacteria. *J. Biol. Chem*. **276,** 13513–13516.

Susskind, B. M., Warren, L. G., and Reeves, R. E. (1982). A pathway for the interconversion of hexose and pentose in the parasitic amoeba *Entamoeba histolytica*. *Biochem. J.* **204,** 191–196.

Swain, A. L., Jaskolski, M., Housset, D., Rao, J. K., and Wlodawer, A. (1993). Crystal structure of *Escherichia coli* L-asparaginase, an enzyme used in cancer therapy. *Proc. Natl. Acad. Sci. USA* **90,** 1474–1478.

Takeuchi, T., Weinbach, E. C., and Diamond, L. S. (1977). *Entamoeba histolytica*: Localization and characterization of phosphorylase and particulate glycogen. *Exp. Parasitol*. **43,** 107–114.

Takeuchi, T., Weinbach, E. C., Gottlieb, M., and Diamond, L. S. (1979). Mechanism of L-serine oxidation in *Entamoeba histolytica*. *Comp. Biochem. Physiol. B Comp. Biochem.* **62,** 281–285.

Todd, J., Post-Beittenmiller, D., and Jaworski, J. G. (1999). KCS1 encodes a fatty acid elongase 3-ketoacyl-CoA synthase affecting wax biosynthesis in *Arabidopsis thaliana*. *Plant J*. **17,** 119–130.

Tokoro, M., Asai, T., Kobayashi, S., Takeuchi, T., and Nozaki, T. (2003). Identification and characterization of two isoenzymes of methionine γ-lyase from *Entamoeba histolytica*: A key enzyme of sulfur-amino acid degradation in an anaerobic parasitic protist that lacks forward and reverse trans-sulfuration pathways. *J. Biol. Chem*. **278,** 42717–42727.

Varela-Gomez, M., Moreno-Sanchez, R., Pardo, J. P., and Perez-Montfort, R. (2004). Kinetic mechanismand metabolic role of pyruvate phosphate dikinase from *Entamoeba histolytica*. *J. Biol. Chem*. **279,** 54124–54130.

Vargas-Villarreal, J., Olvera-Rodriguez, A., Mata-Cardenas, B. D., Martinez-Rogriguez, H. G., Said-Fernandez, S., and Alagon-Cano, A. (1998). Isolation of an *Entamoeba histolytica* intracellular alkaline phospholipase A2. *Parasitol. Res*. **84,** 310–314.

Villalobo, E., Wender, N., and Mirelman, D. (2005). *Entamoeba histolytica*: Molecular characterization of an aldose 1-epimerase (mutarotase). *Exp. Parasitol*. **110,** 298–302.
Wassmann, C., Hellberg, A., Tannich, E., and Bruchhaus, I. (1999). Metronidazole resistance in the protozoan parasite *Entamoeba histolytica* is associated with increased expression of iron-containing superoxide dismutase and peroxiredoxin and decreased expression of ferredoxin 1 and flavin reductase. J. Biol. Chem. 274, 26051–26056.
Weinbach, E. C., and Diamond, L. S. (1974). *Entamoeba histolytica*. I. Aerobic metabolism. *Exp. Parasitol*. **35,** 232–243.
Werries, E., and Thurn, I. (1989). Breakdown of glucopolysaccharides in *Entamoeba histolytica* by phosphorylase. *J. Protozool.* **36,** 607–612.
Werries, E., Franz, A., and Geisemeyer, S. (1990). Detection of glycogen-debranching system in trophozoites of *Entamoeba histolytica*. *J. Protozool*. **37,** 576–580.
WHO(1998). ''The world health report 1998. Life in the 21st century: A vision for all.'' World Health Organization, Geneva, Switzerland.
Woods, S. A., Schwartzbach, S. D., and Guest, J. R. (1988). Two biochemically distinct classes of fumarase in *Escherichia coli*. *Biochim. Biophys. Acta* **954,** 14–26.
Wu, G., and Müller, M. (2003). Glycogen phosphorylase sequences from the amitochondriate protists, *Trichomonas vaginalis*, *Mastigamoeba balamuthi*, *Entamoeba histolytica* and *Giardia* intestinalis. *J. Eukaryot. Microbiol*. **50,** 366–372.
Yang, W., Li, E., Kairong, T., and Stanley, S. L., Jr. (1994). *Entamoeba histolytica* has an alcohol dehydrogenase homologous to the multifunctional adhE gene product of *Escherichia coli*. *Mol. Biochem. Parasitol*. **64,** 253–260.
Zhang, W. W., Shen, P. S., Descoteaux, S., and Samuelson, J. (1994). Cloning and expression of the gene for an $NADP^+$-dependent aldehyde dehydrogenase of *Entamoeba histolytica*. *Mol. Biochem. Parasitol.* **63,** 157–161.
Zuo, X., and Coombs, G. H. (1995). Amino acid consumption by the parasitic, amoeboid protists *Entamoeba histolytica* and *E. invadens*. *FEMS Microbiol. Lett*. **130,** 253–258.

3.4 *Entamoeba histolytica*: response of the parasite to metronidazole challenge on the levels of mRNA and protein expression.

Tazreiter M, Leitsch D, Hatzenbichler E, Mair-Scorpio GE, Steinborn R, Schreiber M, Duchêne M.

Exp Parasitol. 2008; 120: 403-10.

Experimental Parasitology 120 (2008) 403–410

Contents lists available at ScienceDirect

Experimental Parasitology

journal homepage: www.elsevier.com/locate/yexpr

Entamoeba histolytica: Response of the parasite to metronidazole challenge on the levels of mRNA and protein expression

Margit Tazreiter [a], David Leitsch [a], Evelyn Hatzenbichler [b], Georg E. Mair-Scorpio [c], Ralf Steinborn [c], Martin Schreiber [b], Michael Duchêne [a,*]

[a] *Department of Specific Prophylaxis and Tropical Medicine, Center for Physiology, Pathophysiology and Immunology, Medical University of Vienna, Kinderspitalgasse 15, A-1090 Vienna, Austria*
[b] *Department of Obstetrics and Gynecology, Medical University of Vienna, Währinger Gürtel 18-20, A-1090 Vienna, Austria*
[c] *Vetomics Core Facility, University of Veterinary Medicine, Veterinärplatz 1, A-1210 Vienna, Austria*

ARTICLE INFO

Article history:
Received 2 May 2008
Received in revised form 23 July 2008
Accepted 18 September 2008
Available online 26 September 2008

Keywords:
Entamoeba histolytica
Metronidazole
Microarray
Proteomics
2DE, two-dimensional electrophoresis
EC number, enzyme commission number
DTT, dithiothreitol
IPG, immobilized pH gradient
PFOR, pyruvate ferredoxin oxidoreductase
qRT-PCR, quantitative real-time reverse transcription polymerase chain reaction
REST, Relative Expression Software Tool
RIN, RNA integrity number

ABSTRACT

Entamoeba histolytica remains a cause of significant morbidity and mortality in many countries. Although metronidazole has been used for the treatment of amoebiasis for several decades, little is known on how the amoebae react to this challenge on the levels of mRNA and protein expression. In this study, we examined their response using a focused microarray, quantitative RT-PCR, and two-dimensional gel electrophoresis (2DE). The amoebae modestly increased the levels of mRNA coding for superoxide dismutase, peroxiredoxin, ferredoxin, thioredoxin reductase, and the galactose/*N*-acetylgalactosamine specific lectin light and heavy chains. The mRNAs encoding actin and the 70 kDa and 101 kDa heat shock proteins were decreased. All the changes occured within 1 h of exposition, with very little further changes. In addition, the proteome revealed only very few changes. Taken together, *E. histolytica* appears to make only modest mRNA and protein expression changes when confronted with an unknown chemical stress.

1. Introduction

Entamoeba histolytica causes amoebic dysentery and liver abscess and is responsible for widespread morbidity and mortality (World Health Organization, 1997). The nitroimidazole drug metronidazole has been used successfully to treat invasive amoebiasis since 1966 (Powell et al., 1966) and has since then remained the gold standard for amoebiasis treatment. Metronidazole needs to be reduced in the parasite in order to become active (Müller, 1983). In trichomonads, reduced ferredoxin formed in the pyruvate:ferredoxin oxidoreductase (PFOR) reaction is believed to reduce metronidazole. In vitro data corroborate this hypothesis (Lindmark and Müller, 1976). Since PFOR is also an abundant enzyme in *E. histolytica*, the same mechanism is suggested to exist in *E. histolytica*. Recently, we have been able to show that the flavin enzyme thioredoxin reductase is also able to reduce and activate metronidazole and other nitro compounds (Leitsch et al., 2007). The activated reaction products such as the nitroradical anion and further reduced intermediates then attack cellular components of the parasite, in particular proteins and other thiol containing compounds. We were also able to define for the first time specific protein targets of these activated metabolites, the activating enzyme thioredoxin reductase, as well as thioredoxin, superoxide dismutase and purine nucleoside phosphorylase (Leitsch et al., 2007). This work showed that the target proteins were chemically modified, most likely via protein thiol groups, because cysteine was able to rescue *E. histolytica* from the toxic effects of metronidazole.

The selective poisoning of *E. histolytica* by metronidazole in vitro and in the human host has been an extremely successful example of anti-parasite chemotherapy for more than 40 years. Although considerable information is available on the activation of metronidazole and other nitroimidazoles, and we are learning more about protein targets, little is known if and how *E. histolytica* responds to the chemical stress elicited by metronidazole by changing the levels of mRNA and protein expression.

* Corresponding author. Fax: +43 1 4277 64899.
E-mail address: michael.duchene@meduniwien.ac.at (M. Duchêne).

0014-4894/$ - see front matter © 2008 Elsevier Inc. All rights reserved.
doi:10.1016/j.exppara.2008.09.011

The development of transfection methods for *E. histolytica* (Nickel and Tannich, 1994; Purdy et al., 1994) has paved the way for the identification of a core promoter structure consisting of a TATA-like element with the consensus sequence GTATTTAAAG/C, followed by a GAAC element, and an AAAAATTCAA element marking the transcriptional start site (reviewed by Vanacova et al., 2003). The availability of the genome data (Loftus et al., 2005) and microarrays to measure the baseline level of transcription was recently exploited to predict more promoter elements, which could also be verified by specific interaction of proteins with the elements (Hackney et al., 2007). Some of the elements in combination render the downstream genes heat-inducible.

Entamoeba histolytica trophozoites respond strongly to elevated temperatures. Shifting the temperature to 40 °C for 5 h led to a significant induction of mRNA coding for Hsp70, the 70 kDa heat shock protein (Ortner et al., 1992). Recently, the mRNA expression changes upon heat shock (42 °C for 4 h) were examined in detail by microarray hybridization (Weber et al., 2006). Hsp70 and other heat shock proteins and chaperones were up-regulated as well as the cysteine proteinases Cp4 and Cp6, which are possibly involved in the degradation of damaged proteins. Surprisingly, the study revealed a vast down-regulation of mRNAs coding for as many as 42% of all unique proteins examined. So studying the heat shock response leads to a number of new questions, even complicated by an earlier observation that epigenetic regulation may also play a role in this response (Bernes et al., 2005).

The response of *E. histolytica* trophozoites to the mouse host in an intestinal infection model led to extensive mRNA expression changes for genes implicated in metabolism, oxygen defense, cell signaling, virulence, antibacterial activity, and DNA binding (Gilchrist et al., 2006). Even the addition of human type I collagen alone led to significant expression changes of genes involved in cell signaling, cytoskeletal reorganisation, and vesicle formation as well as inducing cysteine proteinase and amoebapore mRNAs (Debnath et al., 2004).

In the present work, we examined the response of the amoebae to metronidazole using a 213 gene focused microarray, quantitative RT-PCR, and two-dimensional gel electrophoresis. Taken together, *E. histolytica* was unable to mount a strong, coordinated response to this lethal stress. Transcription of a limited number of mRNAs was up-regulated within one hour, but there was no significant further rise of these mRNA levels. When we examined the soluble proteome of the amoebae, we did not detect any up-regulated proteins, only the products of modification by activated metronidazole metabolites as described above.

2. Materials and methods

2.1. Parasite culture

Entamoeba histolytica trophozoites (HM-1:IMSS) were grown axenically at 36.5 °C in TYI-S-33 medium supplemented with 10% (v/v) complement-inactivated bovine serum (Diamond et al., 1978). The medium in the 25 ml tissue culture flasks (BD Biosciences, Heidelberg, Germany) was changed every three days.

2.2. Metronidazole treatment

Six culture flasks of log phase *E. histolytica* cells were pooled, counted and then again distributed to six new flasks. Metronidazole (Anaerobex 0.5%, Gerot Pharmaceuticals, Vienna, Austria) was added to three culture flasks each to give a final concentration of 50 μM, three flasks remained untreated as control. After 1 h, 3 h, and 5 h, one treated and one untreated flask of cells were used to prepare total RNA. The drug concentration of 50 μM was used as it is somewhat lower as the peak plasma concentration (50–80 μM) after one oral dose of 500 mg (Lamp et al., 1999). At this concentration, after 5 h, the amoebae had rounded off but were still alive, whereas after 24 h, only about 2% had survived as judged by Trypan blue exclusion. The same treatment scheme was used for two-dimensional gel electrophoresis (2DE), and samples for microarray and proteome analysis were derived from the same batch of amoebae.

2.3. Extraction of total RNA

Isolation of total RNA was performed using Tri Reagent (Sigma–Aldrich, Vienna, Austria) according to the instructions of the supplier. Briefly, the sedimented cells were lysed with 1 ml Tri Reagent per $5–10 \times 10^6$ cells, proteins were extracted into the phenolic phase, and RNA was isopropanol precipitated from the aqueous phase. The resulting pellet was washed with 75% ethanol, air-dried, and dissolved in diethylpyrocarbonate (DEPC)-treated water. Samples were stored at −80 °C.

2.4. Generation of fluorescently labeled cDNA probes

For generation of labeled probes, 8 μg of total RNA was spiked with test or reference spike mix, respectively (Lucidea Universal ScoreCard, Amersham, GE Healthcare, Vienna, Austria), oligo(dT) and water added and annealing was allowed while cooling from 70 °C to 4 °C in a PCR machine. For cDNA synthesis the Atlas PowerScript Fluorescent Labeling Kit (BD Biosciences, Rockville, MD, USA) was used, the mastermix containing 5× cDNA synthesis buffer, 10× dNTP mix containing aminoallyl-dUTP, DTT, Powerscript Reverse Transcriptase and RNaseOUT (Invitrogen, Lofer, Austria). Transcription was allowed to proceed for 60 min at 45 °C, then the samples were heated to 70 °C for 5 min, cooled to 37 °C, and 0.2 μl RNase H (10 U μl^{-1}) were added and incubated for 15 min at 37 °C. The sample was then cleaned with 2 μl of QuickClean Enzyme Removal Resin (Clontech, Vienna, Austria), and the cDNA was ethanol-precipitated, washed with 70% ethanol, air-dried, and resuspended in 10 μl of 2× Fluorescent Labeling Buffer. Labeling was performed with Cy3 (reference) and Cy5 (test) Mono-Reactive Dye Pack (Amersham, GE Healthcare), respectively, for 60 min at room temperature. Final purifying of the labeled cDNA was done with the QIAquick PCR purification kit (Qiagen, Hilden, Germany), or, alternatively, with the BD purification kit contained in the Atlas system. Samples were stored in light-proof tubes until hybridization.

2.5. Construction of a focused microarray

The design for the low density oligonucleotide array was done with OligoWiz 2.0, a free academic program (www.cbs.dtu.dk/services/OligoWiz2/) from the Center for Biological Sequence Analysis at the Technical University of Denmark in Lyngby (Wernersson and Nielsen, 2005; Nielsen et al., 2003), applying standard parameters except for setting the low complexity threshold to 1.5. The oligonucleotides had a length between 45 and 55 nucleotides with an average of 47.1. A full list of all 213 genes, accession numbers, and oligonucleotide sequences can be found in Supplementary Table 1. All oligonucleotides were synthesized with a 5′-amino modification (Microsynth, Balgach, Switzerland), and spotted onto epoxy-modified CodeLink Activated Slides (GE Healthcare).

Spotting was performed on a GeneMachines OmniGrid 100 (BST Scientific, Singapore) with the following configuration: pin configuration 4×4, array size 8×9, spacing 400×400, spotting buffer 150 mM sodium phosphate, 0.001% SDS, pH 8.5, 25 μM oligonucleotide probe. Humidity was 45%, temperature 22 °C, post coupling was performed according to the CodeLink protocol. All oligonucleotides were spotted in triplicate. The whole set was spotted twice per slide, so each slide contained six spots of a given oligonucleotide.

2.6. Microarray hybridization and data analysis

Labeled Cy3 and Cy5 probes were combined in one tube and lyophilized. Meanwhile the CodeLink slides were blocked with pre-hybridization buffer (6× SSC, 2.5% SDS, 1% BSA, total of 40 ml) for 45 min at 42 °C, then washed with MilliQ water and 95% EtOH and dried in a centrifuge. Coverslips were washed with MilliQ water and isopropanol and spun dry. The dry, labeled cDNA was dissolved in 40 μl of pre-warmed (50 °C) hybridization buffer (30% formamide, 5× Denhart´s solution, 5× SSC, 1 mM sodium pyrophosphate, 50 mM Tris pH 7.4, 0.1% SDS), denatured at 95 °C for 4 min and then placed onto the slide. Hybridization was performed overnight for 16 h at 30 °C in the HybChamber from GeneMachines (Genomic Solutions, Zinsser Analytic, Frankfurt, Germany), moisturized with 10 μl 1× SSC. After hybridization, slides were removed from the chamber and washed in solutions of decreasing ionic strength (4× SSC at room temperature to remove the cover slip; 2× SSC, 0.1% SDS at hybridization temperature twice for 5 min each; 0.2× SSC at room temperature for 1 min; 0.1× SSC at room temperature for 1 min; followed by briefly dipping in pure water at room temperature) and then spin-dried at 2500 rpm for 4 min.

The microarrays were scanned with a GenePix 4000A scanner (Axon Instruments, Union City, CA, USA) at 10 μm resolution, analysis of the resulting images was performed with the GenePix Pro 4.1 software from Axon Instruments. Automated spot detection, manual and automated flagging as well as calculation of Cy5/Cy3 ratios was done with this program, normalization was performed according to the GenePix normalization factor such that the median of ratios value is 1. Data were exported to a Microsoft Excel calculation sheet and the median of the median of ratios of all detected spots per oligonucleotide (maximum of six) was calculated. Geometric means of medians for each gene and treatment time were obtained and after log-normal transformation, the significance of up-regulation or down-regulation of the respective genes was assessed by means of the Student´s *t*-test for independent samples. Values with p-values < 0.05 were considered as significant.

2.7. Quantitative real-time reverse transcription-PCR (qRT-PCR)

The design of primer sequences for qRT-PCR (Supplementary Table 2) was performed with the Primer Express, version 2.0, software (Applied Biosystems, Foster City, USA). Primers were designed to target single exons which requires careful examination of minus RT controls (see below). Each amplicon was tested for potential secondary structures and for specificity using the Mfold web server (Zuker, 2003) and the Basic Local Alignment Search Tool BLASTn (Altschul et al., 1990), respectively. Amplicon sizes were between 60 and 137 basepairs (Supplementary Table 2).

Total RNA amount was measured spectrophotometrically on the Hellma TrayCell (Hellma, Müllheim/Baden, Germany) in combination with the BioPhotometer 6131 (Eppendorf, Hamburg, Germany). The integrity of each RNA sample was assessed by the 2100 Expert software (version B.02.03.SI307) of the 2100 Bioanalyzer (Agilent Technologies, Santa Clara, USA) using the RNA 6000 Nano Chip Kit (Agilent Technologies) and 200 ng of input RNA. To overcome failure to compute an RNA integrity number (RIN) value by the Bioanalyzer software, the anomaly threshold settings for the 5S region (advanced settings in the Set Point Explorer) were adjusted manually within a range of +0.1– +0.2 of the default value. A RIN value of five was used as threshold for qRT-PCR analysis (Fleige and Pfaffl, 2006). Sample RINs ranged from 5.0 to 6.5.

The High-Capacity Reverse Transcription Kit (Applied Biosystems) was used for reverse transcription for 120 min at 37 °C, primed with random RT primers. For each cDNA duplicate 1.5 μg of RNA was reverse transcribed in a reaction volume of 20 μl. The 20 μl-qPCR consisted of 80 mM Tris–HCl (pH 9.4), 20 mM $(NH_4)_2SO_4$, 0.02% w/v Tween-20, 3.5 mM $MgCl_2$, 0.2 mM dNTPs, 200 nM of each primer, 0.4× EvaGreen dsDNA-binding dye (Biotium, Hayward, USA), 1 U of a Taq DNA polymerase chemically modified for "hot start" (Hot FIREPol, Solis Biodyne, Tartu, Estonia) and 1.5 μl cDNA template. The qPCR was conducted on the ABI PRISM 7900HT Sequence Detection System (Applied Biosystems) using a thermal protocol which consisted of an initial hot start at 95 °C for 15 min followed by 40 amplification cycles (95 °C for 15 s, 59 °C for 40 s, 72 °C for 20 s). A series of five eight-fold dilutions of a randomly selected control cDNA amplified in duplicate was used to generate a standard curve. Reaction efficiencies (E) calculated from the slope (s) of the standard curve using the formula $E = 10^{-1}/s$ ranged from 70% to 90%.

RNA specificity of a qRT-PCR signal was concluded if the difference between the threshold cycle (C_T) values of a sample and its "minus-RT" control was above 9.5. This ΔC_T requirement calculated for assays with lower amplification efficiencies ($E = 0.7$) would correspond to a systematic error of about 1%.

Calculation of gene expression ratios and evaluation of their statistical significance were performed using the Relative Expression Software Tool (REST) 2005 software including the Pair Wise Fixed Reallocation Randomisation Test (Pfaffl et al., 2002). To achieve accurate normalization for differences in the quantity and quality of total nucleic acid added to each reaction and the efficiency of the RT step, a multiple reference gene normalization strategy was applied. The housekeeping genes thioredoxin reductase and hexokinase 2 displaying minimal variation across the treatment and non-treatment control groups (1.4- and 1.5-fold, respectively.) were selected to normalize the gene of interest to their geometric mean. Finally, the n-fold expression of the gene of interest was given relative to a reference (three sham-treated control samples at the respective time point).

2.8. Cell harvest and preparation of lysates

Preparation of lysates and 2DE were carried out as described before (Leitsch et al., 2005). Briefly, cells were harvested by centrifugation, washed, and disrupted in a Dounce homogenizer. Proteins were precipitated in 10% trichloroacetic acid in acetone at −20 °C for at least 90 min. After precipitation, proteins were pelleted in the ultracentrifuge, the pellet was washed twice with 90% (v/v) acetone in H_2O, and air-dried. Proteins were resuspended at room temperature in 2DE solubilization buffer (7 M urea, 2 M thiourea, 4% (w/v) CHAPS (3-[(3-cholamidopropyl)dimethylammonio]-1-propanesulfonate), 1% (w/v) DTT (dithiothreitol), and 1% (v/v) Pharmalyte pH 3–10 ampholytes (GE Healthcare)). Insoluble material was removed by ultracentrifugation. Protein concentration was determined by Bradford assay (Bio-Rad Laboratories).

2.9. Two-dimensional polyacrylamide gel electrophoresis (2DE)

Fifty microgram protein for silver-stained, and 500 μg protein for Coomassie-stained gels, were diluted in rehydration buffer (7 M urea, 2 M thiourea, 0.5% (w/v) CHAPS (3-[(3-cholamidopropyl)dimethylammonio]-1-propanesulfonate), 15 mM DTT, and 0.2% (v/v) ampholytes pH 3–10) up to a total volume of 350 μl. Isoelectric focussing was performed with 17 cm pH 3–10 (nonlinear) IPG strips in a Bio-Rad IEF cell (Bio-Rad Laboratories), and in the second dimension 12.5% PAGE was run. Gels were silver-stained to visualize as many spots as possible, whereas gels were stained with Coomassie Brilliant Blue R-250 (Bio-Rad) for better quantification. Gels were scanned with an Epson 1680 Pro scanner and analyzed with the Melanie 2 D gel analysis software (GeneBio, Geneva, Switzerland).

3. Results

3.1. A focused microarray containing 213 oligonucleotide probes

To measure the mRNA expression changes we constructed a focused oligonucleotide microarray for 213 selected target genes. When we started the work, large amounts of sequence data was available from the Institute for Genomic Research (http://www.tigr.org) and the Sanger Institute (http://www.sanger.ac.uk), but the gene annotations had not been made. So we constructed a database from own annotations made by TBLASTN (Gertz et al., 2006) searches starting with eukaryotic or prokaryotic query sequences (Hofer and Duchêne, 2005). We finally selected 213 genes from our database and, in a few cases, from previously published genes. The largest group were genes coding for enzymes ($n = 110$), a second group codes for cytoskeletal genes initially thought to represent typical housekeeping genes ($n = 76$), a third group of stress genes ($n = 15$), and finally a small number of known or putative surface antigens ($n = 12$). On the one hand, our array was limited to cover only a rather narrow part of the genome, but it contained only selected genes with known homologs, in particular, genes possibly involved in activation of or resistance to metronidazole (Wassmann et al., 1999). Supplementary Table 1 displays the whole list of genes and deduced oligonucleotide probes, also including the updated "XM" genome project accession numbers and accession numbers relating to previous publications. It may be noted that there is a high number of lipid-associated enzymes. These were on the array for examining the reaction to the anti-amoebic agent miltefosine (Seifert et al., 2001).

The design of the oligonucleotide probes is described above. A compromise had to be found when designing the probes between short probes serving the intention of distinguishing members of a gene family and long probes permitting stronger hybridization allowing stringent washes. Our probes had an average A/T content of 64.4% and had an average length of 47.1 nucleotides, this required mild hybridization and washing conditions at only 30 °C.

3.2. Optimization of treatment and array conditions

After performing some preliminary experiments, a metronidazole concentration of 50 μM was chosen for all experiments. After 24 h, this concentration killed 98% of the amoebae, but after 5 h the amoebae were still alive and intact although they had lost their elongated shape. So when we analyzed amoebae treated for 1 h, 3 h, and 5 h in the final experiments, we examined them from the moment of drug addition to a time with visible changes. During this period we expected the maximum response of the cells to the lethal challenge.

As the cleaning step of the labeled cDNA is an important step, we compared two procedures, the QIAquick PCR purification kit and the BD purification kit contained in the Atlas system. Two sets of metronidazole-treated samples and controls were generated with a short exposure time of 1 h, and cDNAs were purified with the QIAquick kit or the BD kit. As we observed no significant systematic difference between the two samples (data not shown), we included both experiments in the final statistics. In all the further experiments, the labeled cDNAs were purified with the QIAquick kit.

Table 1A
Up-regulated genes upon metronidazole treatment of *E. histolytica*

Up-regulated genes	Accession Numbers	1 h	3 h	5 h	1 h	3 h	5 h
		Microarray			qRT-PCR		
Fe superoxide dismutase	XM_643735, M63816, X70852	1.6*	2.2****	1.4			
Thioredoxin reductase	XM_650656, XM_643726, X79603	1.8	1.8*	1.6			
Thioredoxin	XM_651634	n.a.	n.a.	n.a.	1.4	1.6*	2.1
Peroxiredoxin	XM_647979, XM_646911, XM_644418, XM_644379, XM_643430, XM_642815, XM_642754, EHU67158, EHU67157	2.4	2.5*	1.6			
Ferredoxin	XM_650091, XM_650090, J03970	2.1**	2.9**	1.8***			
Ferredoxin	XM_646192	2.2***	2.6****	2.5***	2.7***	2.0*	3.8***
Ferredoxin	XM_648587	1.8*	1.9	1.6			
Lectin heavy subunit	XM_651089, X89415, M60498	2.1	1.8**	2.3*	2.0*	2.6**	1.7
Lectin light subunit	XM_646286	2.5*	2.1***	2.1***	1.3	1.8*	3.3*
Hexokinase 1	XM_648528, X82197	1.5	1.8*	2.0			
Hexokinase 2	XM_650873, X82198	1.6	1.7	2.0			
Long chain fatty acid CoA ligase	XM_651318	2.3*	2.0*	2.1***	1.9*	4.2***	4.9***
Diacylglycerol kinase	XM_646834	1.6	1.8*	1.7			
GTP binding protein beta chain	XM_651512	1.7	1.7**	1.9			
Ubiquitin UBI1	XM_645071, XM_645435, X98567	2.9	1.9	1.9			
Ubiquitin	XM_650087, XM_643062, XM_649937	2.2	2.2	1.9			

n.a.: Not analyzed.
* $p \leqslant 0.05$.
** $p \leqslant 0.01$.
*** $p \leqslant 0.005$.
**** $p \leqslant 0.001$.

3.3. Metronidazole treatment elicits modest changes in mRNA expression in E. histolytica trophozoites

For the main experiment, new cultures were grown, and the amoebae now were treated with 50 μM metronidazole for 1 h, 3 h, and 5 h as described above, together with the untreated parallel cultures. Two independent experiments were performed labeling the cDNAs with Cy3 (control) and Cy5 (metronidazole-treated) fluorescent dyes. Each of the six pairs of labeled samples were hybridized to the microarray as described above. The analysis of the microarray data, described in Materials and Methods (Section 2.6.), led to the final data summarized in Table 1A (up-regulated genes) and Table 1B (down-regulated genes).

Taken together, there were only modest changes in mRNA expression in the amoebae for the genes selected for our microarray. Table 1A contains 15 genes with a tendency for up-regulation. Out of these, seven genes showed at least one statistically significant ($p < 0.05$) value $\geqslant 2.0$-fold change. These were the superoxide dismutase, peroxiredoxin, two ferredoxin genes, the Gal/GalNAc specific lectin light and heavy chains, and a long chain fatty acid CoA ligase. Only two genes encoding one of the ferredoxins and the lectin light chain were consistently above a factor of 2.0. Table 1B shows eight genes with a tendency towards down-regulation, six of which showed significant ($p < 0.05$) values of $\leqslant 0.5$ for at least one of the treatment times. These were genes encoding an actin, the 101 kDa heat shock protein, a short chain dehydrogenase, a fatty acid elongase, diaphanous protein, and a WD repeat protein. Although the changes observed were modest, those small changes occurred consistently.

For seven selected genes we also performed qRT-PCR using sequence-specific primers and the cDNAs from treated and untreated amoebae as a template. The right columns of Tables 1A and 1B show the relative expression values of these genes. The expression changes observed with this gold standard assay were also modest, and where the array had shown up-regulation or down-regulation, the qRT-PCR gave the same results. The factors were very similar in most cases, only for the 3 h and 5 h values for the long chain fatty acid CoA ligase gene, the qRT-PCR showed a markedly higher up-regulation than was observed in the microarray. Taken together, although the changes observed in the array experiments were not dramatic, they were clearly confirmed by qRT-PCR, the gold standard assay.

3.4. Metronidazole treatment and protein levels in the soluble proteome of E. histolytica trophozoites

It was a main aim of the project to investigate the changes in metronidazole-treated amoebae both on the mRNA and protein levels, as in other organisms it had become clear that there is only a limited correlation with the protein and mRNA levels, for example in yeast (Gygi et al., 1999). So we knew from previous experiments that metronidazole treatment led only to few protein changes in the amoebae, and these changes were associated with the modification of five proteins in total by reduced metronidazole metabolites (Leitsch et al., 2007). These processes told us more about the mechanism of the drug´s action than about the response of the amoebae. So for the present work we repeated the experiments again using 50 μM of metronidazole for 1 h, 3 h, and 5 h, extracted the soluble proteome and performed 2DE. Both silver and Coomassie-stained gels were generated, but we did not find any appearing or disappearing, or strongly up- or down-regulated protein spots (gels not shown). The only changes we observed were again the appearance of metronidazole-modified proteins (Leitsch et al., 2007) in the gels from the treated amoebae. Fig. 1 shows the spots from Coomassie-stained gels that had been identified as thioredoxin, superoxide dismutase, thioredoxin reductase, actin, and the 70 kDa heat shock protein, all of which had shown some changes on the mRNA level. Only in the metronidazole-treated cells, thioredoxin and superoxide dismutase developed a modified sister spot shifted to the basic region (to the right in Fig. 1.), and thioredoxin reductase developed two extra spots. Although the three genes superoxide dismutase, thioredoxin reductase, and thioredoxin were found to be slightly up-regulated on the mRNA level, this was not reflected on the level of protein. Actin and Hsp70 had been found to be slightly down-regulated on the mRNA level. On the protein level, there is no visible effect in the 1 h and 3 h

Table 1B
Down-regulated genes upon metronidazole treatment of *E. histolytica*

Down-regulated genes	Accession numbers	1 h	3 h	5 h	1 h	3 h	5 h
		Microarray			qRT-PCR		
Actin	XM_649111	0.7	0.7	0.4***	1.0	0.6	0.6*
Actin	XM_647886	0.8	0.6	0.6***			
70 kDa heat shock protein	XM_646910	0.6	0.4	0.5			
101 kDa heat shock protein	XM_648360, XM_647538, XM_643852, XM_643174, XM_648976, XM_646042, XM_645824, XM_643226, XM_643122, XM_642790	0.8	0.3*	0.8			
Short chain dehydro-genase	XM_647586	0.5	0.5***	0.5**			
Fatty acid elongase	XM_649311	0.6	0.6****	0.4*	0.7	0.4*	0.5
Diaphanous protein	XM_650938	0.6	0.4*	0.6			
WD repeat protein	XM_645831	0.8	0.4***	0.7			

Note. RNAs from *E. histolytica* trophozoites treated for 1 h to 5 h with 50 μM metronidazole and controls were transcribed into either Cy3-labeled (control) or Cy5-labeled (treated) cDNAs and hybridized to a focused oligonucleotide microarray containing probes for 213 genes. Values for trophozoites treated for 1 h, 3 h, or 5 h were analyzed separately (3 left columns). Table 1A shows up-regulated genes, and Table 1B down-regulated genes. The results of the qRT-PCR for 7 selected genes are shown in the 3 right columns.

* $p \leqslant 0.05$.
** $p \leqslant 0.01$
*** $p \leqslant 0.005$.
**** $p \leqslant 0.001$.

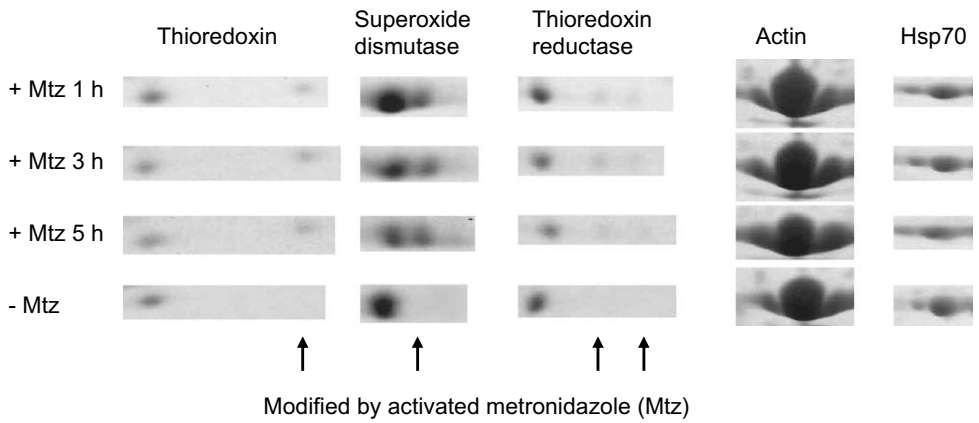


Fig. 1. Details from Coomassie-stained 2DE-gels displaying the soluble proteome of *E. histolytica* trophozoites exposed to 50 μM metronidazole. Upper three rows, cells treated with metronidazole (Mtz) for 1 h, 3 h, or 5 h; bottom row, control cells. The cutouts show thioredoxin, superoxide dismutase, thioredoxin reductase, actin and Hsp70 that had been identified before (Leitsch et al., 2005, 2007). In the gels from metronidazole-treated cells the cutouts annotated as thioredoxin, superoxide dismutase, and thioredoxin reductase (arrows) display weaker additional spots seen at the right (basic) side of the original spots. These are portions of the protein that have been chemically modified by activated metronidazole (Leitsch et al., 2007).

treated cells, but there appeared to be a slight down-regulation of Hsp70 after 5 h. Image analysis showed, however, that the spot volume was only decreased by 10% which is below the threshold of significance. Taken together, the amoebae are not able to transform the up-regulated mRNAs in a visible protein increase, underscoring the impression of a defensive, catabolic state of the cells.

4. Discussion

To encounter challenges like heat shock, the conditions favouring cyst formation, or contact with host molecules is a typical situation for *E. histolytica*, quite unlike the contact with metronidazole. So it may not be surprising that the amoebae do not mount a strong and coordinated response to the drug, and mRNA expression changes are only modest. Several of the up-regulated genes (superoxide dismutase, peroxiredoxin, thioredoxin reductase and thioredoxin) belong to the factors destined to protect the amoebae from oxidative stress. When metronidazole is activated by a one-electron reduction, a nitroradical anion is produced. This anion or its further reduction products can damage cellular components in a number of ways (Leitsch et al., 2007). In the presence of oxygen, the nitroradical anions can also generate oxidative stress by reacting with this oxygen to form superoxide radical anions. These can be detoxified in the amoebae by the combined action of superoxide dismutase and peroxiredoxin, therefore the up-regulation of these genes would be an appropriate reaction. When *E. histolytica* is slowly adapted in vitro to increased (10–40 μM) metronidazole concentrations, the amoebae also up-regulate both superoxide dismutase and peroxiredoxin (Samarawickrema et al., 1997; Wassmann et al., 1999).

In contrast to this potentially useful response, several ferredoxin genes are also up-regulated which could help in generating energy, but at the same time might lead to an increased activation of metronidazole by these newly generated ferredoxins. We also found thioredoxin reductase to be slightly up-regulated (Table 1A). On the one hand, thioredoxin reductase and thioredoxin together with peroxiredoxin can detoxify the hydrogen peroxide generated from superoxide anions by superoxide dismutase, on the other hand thioredoxin reductase is also able to reduce and activate metronidazole (Leitsch et al., 2007), to up-regulate this molecule would therefore be another counterproductive reaction of the amoebae. Indeed, in amoebae adapted to 40 μM metronidazole during very prolonged drug pressure, thioredoxin reductase and ferredoxin 1 were now found to be down-regulated (Wassmann et al., 1999).

An interesting finding is the up-regulation both heavy and light chains of the galactose/*N*-acetylgalactosamine specific lectin. We cannot provide an obvious explanation why the main adhesion factor of the amoebae should be up-regulated, however, it was demonstrated recently (Hughes et al., 2003) that the cytoplasmic domain of the lectin heavy chain interacts with the antioxidant peroxiredoxin. The authors concluded that the lectin may help to recruit peroxiredoxin to the cell surface, so a coordinated up-regulation of peroxiredoxin and lectin genes could be imagined.

The *E. histolytica* genome contains five highly similar full-length ubiquitin genes (Wostmann et al., 1996; Loftus et al., 2005) represented by two ubiquitin probes in the array. Both cDNAs were up-regulated in the metronidazole-treated amoebae by a factor of about two. A speculation about a possible up-regulation of degradation is certainly premature in view of the complexities of the ubiquitin-proteasome system (Demartino and Gillette, 2007).

The very mild up-regulation of both hexokinases may be seen as another weak attempt to increase energy generation. Metabolic modeling of *E. histolytica* glycolysis recently predicted that hexokinase together with 3-phosphoglycerate mutase would exert flux control of glycolysis at low substrate and product concentrations (Moreno-Sánchez et al., 2008), so hexokinase up-regulation might indeed increase the rate of glycolysis. Interestingly, *Trichomonas vaginalis* and *Tritrichomonas foetus*, in the process of developing anaerobic resistance to metronidazole, although shutting off PFOR and hydrogenase, nevertheless, increased their rate of glycolysis (Kulda et al., 1989).

The expression of both the 70 kDa and 101 kDa heat shock proteins was down-regulated. This was unexpected at first, as oxidative stress is able to up-regulate Hsp70 in *E. histolytica* (Akbar et al., 2004), nevertheless there are other examples, eg., the treatment of cancer cells with resveratrol which significantly down-regulates Hsp70 mRNA and protein (Chakraborty et al., 2008).

Finally two enzymes associated with lipid metabolism were found to be both up-regulated and down-regulated. The down-regulation of fatty acid elongase could be interpreted as an attempt to decrease anabolic processes to save energy. Also the mRNA levels of two actin isoforms decreased significantly as well as the expression of the diaphanous protein, a protein interacting with actin and profilin (Ganguly and Lohia, 2000).

Taken together, within the limited scope of our microarray, the response of the amoebae shows an attempt to up-regulate some factors needed for the response to oxidative stress which would be the most specific answer to metronidazole. Above that there

was a picture of a cell with a catabolic rather than anabolic state, a possible shift to intracellular protein degradation, to less chaperone activity, and a feeble attempt to up-regulate energy generation.

In the whole picture, an additional feature was noted. All the expression changes were manifested as early as after 1 h, in most cases, in the hours to follow, the up-regulation or down-regulation did not continue or even accelerate, but tapered off. The whole picture is of a cell which, though looking intact has already lost most of its metabolic strength.

This conclusion finally agrees well with the results from the comparison of the soluble proteomes of metronidazole-treated and untreated amoebae. It appears as if the cells are able to adapt their mRNA levels to a certain extent but are unable to translate these changes into the generation of the desired proteins. This may well be a consequence of the attack of activated metronidazole on the thiol-based redox system of the amoebae (Leitsch et al., 2007). Nevertheless, a lot needs to be learned about the exact targets and chain of events leading to the death of the amoebae, before we fully understand this classical example of successful chemotherapy.

Acknowledgments

This work was supported by Grant P15960 from the Austrian Science Fund. We thank Henrik Bjørn Nielsen for help with the design of the oligonucleotide probes using his program OligoWiz; and we thank Agnes Grünfelder for help with the analysis of the microarray data. The authors have no conflicting financial interests. Author contributions: MT (maiden name Margit Hofer) collected all the genes in silico, devised the probes, and performed the microarray hybridizations and analysis, also qRT-PCR, DL performed the proteome work, EH spotted the arrays and helped with the microarray work, GM performed the final qRT-PCR experiments, GM and RS planned the qRT-PCR experiments and performed the analysis, MS supervised the microarray work and provided optimized protocols, MD supervised the project and wrote the paper with major contributions of the other co-authors.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.exppara.2008.09.011.

References

Akbar, M.A., Chatterjee, N.S., Sen, P., Debnath, A., Pal, A., Bera, T., Das, P., 2004. Genes induced by a high-oxygen environment in *Entamoeba histolytica*. Molecular and Biochemical Parasitology 133, 187–196.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J., 1990. Basic local alignment search tool. Journal of Molecular Biology 215, 403–410.

Bernes, S., Siman-Tov, R., Ankri, S., 2005. Epigenetic and classical activation of *Entamoeba histolytica* heat shock protein 100 (EHsp100) expression. FEBS Letters 579, 6395–6402.

Chakraborty, P.K., Mustafi, S.B., Ganguly, S., Chatterjee, M., Raha, S., 2008. Resveratrol induces apoptosis in K562 (chronic myelogenous leukemia) cells by targeting a key survival protein, heat shock protein 70. Cancer Science 99, 1109–1116.

Debnath, A., Das, P., Sajid, M., McKerrow, J.H., 2004. Identification of genomic responses to collagen binding by trophozoites of *Entamoeba histolytica*. Journal of Infectious Diseases 190, 448–457.

Demartino, G.N., Gillette, T.G., 2007. Proteasomes: machines for all reasons. Cell 129, 659–662.

Diamond, L.S., Harlow, D.R., Cunnick, C.C., 1978. A new medium for the axenic cultivation of *Entamoeba histolytica* and other *Entamoeba*. Transactions of the Royal Society of Tropical Medicine and Hygiene 72, 431–432.

Fleige, S., Pfaffl, M.W., 2006. RNA integrity and the effect on the real-time qRT-PCR performance. Molecular Aspects of Medicine 27, 126–139.

Ganguly, A., Lohia, A., 2000. The diaphanous protein from *Entamoeba histolytica* controls cell motility and cytokinesis. Archives of Medical Research 31, S137–S139.

Gertz, E.M., Yu, Y.K., Agarwala, R., Schäffer, A.A., Altschul, S.F., 2006. Composition-based statistics and translated nucleotide searches: improving the TBLASTN module of BLAST. BioMed Central Biology 4, 41.

Gilchrist, C.A., Houpt, E., Trapaidze, N., Fei, Z., Crasta, O., Asgharpour, A., Evans, C., Martino-Catt, S., Baba, D.J., Stroup, S., Hamano, S., Ehrenkaufer, G., Okada, M., Singh, U., Nozaki, T., Mann, B.J., Petri Jr., W.A., 2006. Impact of intestinal colonization and invasion on the *Entamoeba histolytica* transcriptome. Molecular and Biochemical Parasitology 147, 163–176.

Gygi, S.P., Rochon, Y., Franza, B.R., Aebersold, R., 1999. Correlation between protein and mRNA abundance in yeast. Molecular and Cellular Biology 19, 1720–1730.

Hackney, J.A., Ehrenkaufer, G.M., Singh, U., 2007. Identification of putative transcriptional regulatory networks in *Entamoeba histolytica* using Bayesian inference. Nucleic Acids Research 35, 2141–2152.

Hofer, M., Duchêne, M., 2005. *Entamoeba histolytica*: construction and applications of subgenomic databases. Journal of Experimental Parasitology 110, 178–183.

Hughes, M.A., Lee, C.W., Holm, C.F., Ghosh, S., Mills, A., Lockhart, L.A., Reed, S.L., Mann, B.J., 2003. Identification of *Entamoeba histolytica* thiol-specific antioxidant as a GalNAc lectin-associated protein. Molecular and Biochemical Parasitology 127, 113–120.

Kulda, J., Kabickova, H., Tachezy, J., Cerkasovova, A., Cerkasov, J., 1989. Metronidazole resistant trichomonads: mechanisms of in vitro developed anaerobic resistance. In: Lloyd, D., Coombs, G.H., Paget, T.A. (Eds.), Biochemistry and Molecular Biology Of Anaerobic Protozoa. Harwood, Chur, pp. 137–160.

Lamp, K.C., Freeman, C.D., Klutman, N.E., Lacy, M.K., 1999. Pharmacokinetics and pharmacodynamics of the nitroimidazole antimicrobials. Clinical Pharmacokinetics 36, 353–373.

Leitsch, D., Kolarich, D., Wilson, I.B.H., Altmann, F., Duchêne, M., 2007. Nitroimidazole action in *Entamoeba histolytica*: a central role for thioredoxin reductase. PLoS Biology 5, e211.

Leitsch, D., Radauer, C., Paschinger, K., Wilson, I.B.H., Breiteneder, H., Scheiner, O., Duchêne, M., 2005. *Entamoeba histolytica*: analysis of the trophozoite proteome by two-dimensional polyacrylamide gel electrophoresis. Journal of Experimental Parasitology 110, 191–195.

Lindmark, D.G., Müller, M., 1976. Antitrichomonad action, mutagenicity, and reduction of metronidazole and other nitroimidazoles. Antimicrobial Agents and Chemotherapy 10, 476–482.

Loftus, B., Anderson, I., Davies, R., Alsmark, U.C., Samuelson, J., Amedeo, P., Roncaglia, P., Berriman, M., Hirt, R.P., Mann, B.J., Nozaki, T., Suh, B., Pop, M., Duchêne, M., Ackers, J., Tannich, E., Leippe, M., Hofer, M., Bruchhaus, I., Willhoeft, U., Bhattacharya, A., Chillingworth, T., Churcher, C., Hance, Z., Harris, B., Harris, D., Jagels, K., Moule, S., Mungall, K., Ormond, D., Squares, R., Whitehead, S., Quail, M.A., Rabbinowitsch, E., Norbertczak, H., Price, C., Wang, Z., Guillen, N., Gilchrist, C., Stroup, S.E., Bhattacharya, S., Lohia, A., Foster, P.G., Sicheritz-Ponten, T., Weber, C., Singh, U., Mukherjee, C., El-Sayed, N.M., Petri Jr., W.A., Clark, C.G., Embley, T.M., Barrell, B., Fraser, C.M., Hall, N., 2005. The genome of the protist parasite *Entamoeba histolytica*. Nature 433, 865–868.

Moreno-Sánchez, R., Encalada, R., Marín-Hernández, A., Saavedra, E., 2008. Experimental validation of metabolic pathway modeling. FEBS Journal 275, 3454–3469.

Müller, M., 1983. Mode of action of metronidazole on anaerobic bacteria and protozoa. Surgery 93, 165–171.

Nickel, R., Tannich, E., 1994. Transfection and transient expression of chloramphenicol acetyltransferase gene in the protozoan parasite *Entamoeba histolytica*. Proceedings of the National Academy of Sciences USA 91, 7095–7098.

Nielsen, H.B., Wernersson, R., Knudsen, S., 2003. Design of oligonucleotides for microarrays and perspectives for design of multitranscriptome arrays. Nucleic Acids Research 31, 3491–3496.

Ortner, S., Plaimauer, B., Binder, M., Wiedermann, G., Scheiner, O., Duchêne, M., 1992. Humoral immune response against a 70-kilodalton heat shock protein of *Entamoeba histolytica* in a group of patients with invasive amoebiasis. Molecular and Biochemical Parasitology 54, 175–183.

Pfaffl, M.W., Horgan, G.W., Dempfle, L., 2002. Relative expression software tool (REST) for group-wise comparison and statistical analysis of relative expression results in real- time PCR. Nucleic Acids Research 30, e36.

Powell, S.J., MacLeod, I., Wilmot, A.J., Elsdon-Dew, R., 1966. Metronidazole in amoebic dysentery and amoebic liver abscess. Lancet 2, 1329–1331.

Purdy, J.E., Mann, B.J., Pho, L.T., Petri Jr., W.A., 1994. Transient transfection of the enteric parasite *Entamoeba histolytica* and expression of firefly luciferase. Proceedings of the National Academy of Sciences USA 91, 7099–7103.

Samarawickrema, N.A., Brown, D.M., Upcroft, J.A., Thammapalerd, N., Upcroft, P., 1997. Involvement of superoxide dismutase and pyruvate:ferredoxin oxidoreductase in mechanisms of metronidazole resistance in *Entamoeba histolytica*. Journal of Antimicrobial Chemotherapy 40, 833–840.

Seifert, K., Duchêne, M., Wernsdorfer, W.H., Kollaritsch, H., Scheiner, O., Wiedermann, G., Hottkowitz, T., Eibl, H., 2001. Effects of miltefosine and other alkylphosphocholines on human intestinal parasite *Entamoeba histolytica*. Antimicrobial Agents and Chemotherapy 45, 1505–1510.

Vanacova, S., Liston, D.R., Tachezy, J., Johnson, P.J., 2003. Molecular biology of the mitochondriate parasites, *Giardia intestinalis*, *Entamoeba histolytica* and *Trichomonas vaginalis*. International Journal of Parasitology 33, 235–255.

Wassmann, C., Hellberg, A., Tannich, E., Bruchhaus, I., 1999. Metronidazole resistance in the protozoan parasite *Entamoeba histolytica* is associated with increased expression of iron-containing superoxide dismutase and peroxiredoxin and decreased expression of ferredoxin 1 and flavin reductase. Journal of Biological Chemistry 274, 26051–26056.

Weber, C., Guigon, G., Bouchier, C., Frangeul, L., Moreira, S., Sismeiro, O., Gouyette, C., Mirelman, D., Coppee, J.Y., Guillen, N., 2006. Stress by heat shock induces massive down-regulation of genes and allows differential allelic expression of the Gal/GalNAc lectin in *Entamoeba histolytica*. Eukaryotic Cell 5, 871–875.

Wernersson, R. and Nielsen, H. B., 2005. OligoWiz 2.0 – integrating sequence feature annotation into design of microarray probes. Nucleic Acids Research 33, Web Server issue W611-W615.

World Health Organization, 1997. WHO/PAHO/UNESCO report: a consultation with experts on amebiasis. Mexico City, Mexico 28-29 January, 1997. Epidemiological Bulletin PAHO 18, 13–14.

Wostmann, C., Liakopoulos, D., Ciechanover, A., Bakker-Grunwald, T., 1996. Characterization of ubiquitin genes and -transcripts and demonstration of a ubiquitin- conjugating system in *Entamoeba histolytica*. Molecular and Biochemical Parasitology 82, 81–90.

Zuker, M., 2003. Mfold web server for nucleic acid folding and hybridization prediction. Nucleic Acids Research 31, 3406–3415.

4 Conclusions

4.1 Genome annotation

Work on this thesis project was mainly devided in two parts. The first was a large computational part while setting up our annotation databases. These subgenomic databases of enzymes, cytoskeleton genes and stress genes found in the *E. histolytica* genome provided not only a good basis for the second part of the project, the actual microarray work, but also had the additional benefit of giving us insights into some interesting traits of the *E. histolytica* genome, like incomplete metabolic pathways (Hofer and Duchêne, 2005).

Living in the human gut gives *E. histolytica* easy access to host-derived organic compounds. Therefore many biosynthetic pathways for production of these compounds were secondarily lost, a typical trait of parasites (Clark *et al*., 2007). For instance, *E. histolytica* is not able to produce purines, pyrimidines or fatty acids *de novo*, but there are genes for putative fatty acid elongases which might extend existing fatty acids. During our database compilation we found eight very good homologous sequences to an *Arabidopsis thaliana* fatty acid elongase, a search with a Jojoba sequence also gave a good hit, all of them sharing some 50% protein sequence identiy. So amoebae can alter preformed compounds taken up from the host for their specific metabolic needs.

Cholesterol is also taken up from the host (Das *et al*., 2002). *E. histolytica* is not able to synthesize this important part of its membrane *de novo*. But we found several enzymes for isoprenlylation present in the genome, like protein farnesyl transferase or geranylgeranyl transferase. These enzymes were found to be active in the isoprenylation of small Ras GTPases, important molecular switches of signal transduction in cell proliferation and differentiation (Makioka *et al*., 2006; Kumagai *et al*., 2004). Most metabolic pathways for amino acid synthesis cannot be found, with the

exception of serine and cysteine (Loftus *et al*., 2005). Cysteine is the major intracellular thiol, and probably it is balancing the lack of glutathione as an important component of oxidative stress resistance (Fahey *et al*., 1984).

During our project we found a large number of amoeba genes that could be of bacterial origin. After compiling the first part of our database mainly containing metabolic enzymes we had 112 out of 772 entries that matched nearer with bacterial than with eukaryotic sequences. Our findings were confirmed by the stringent investigation of genes of possible bacterial origin in the genome paper (Loftus *et al*., 2005). In a strict phylogenetic screen 96 genes were found where lateral gene transfer (LGT) is the best explanation for their presence in *E. histolytica*. In the more recent and even stricter reassessment of these data (Clark *et al*., 2007) only 41 cases of LGT remained strongly supported. This is due to the fact that between the two publications the sampling of prokaryotic as well as eukaryotic genomes has increased, giving partly divergent results for phylogenetic analyses. But still the often sparse availabilty of genome data (pro- and eukaryotic) inflicts a strong bias in these data, so that in the future more changes could be seen. Nevertheless, the importance of LGT for *Entamoeba* metabolism remains.

These few samples show the usefulness of our databases for different applications, and we were happy to provide our annotations to the genome project consortium. For the whole genome project of course automated annotation with several adapted programs was done, but still a manual curation is necessary to clear ambiguous assignments and to improve the quality of the annotation. Our annotation databases were also the foundation for our part in the review of the *E. histolytica* genome structure and content (Clark *et al*., 2007), for which we provided a large part of the metabolism section.

4.2 *E. histolytica* reactions to metronidazole treatment

The second main part of this thesis project was the construction of a focussed microarray and its use for detecting the effects of metronidazole treatment on gene expression of amoeba. We compiled a final set of 213 genes for our array including glycolytic enzymes, various pathogenicity factors, all genes known to be involved in response to oxidative stress or heat shock as well as a number of cytoskeleton genes that were thought to be used as reference house keeping genes. As discussed in our microarray paper (Tazreiter *et al*., 2008) the observed changes in gene expression of *E. histolytica* under metronidazole treatment were unexpectedly few and with only small fold changes. One could argue that amoebae are only able to make a weak attempt of reacting to the drug because their standard enzymatic equipment is not capable of an effective defense. Metronidazole acts too quickly and negatively as to enable the cells to mount a coordinated defense. Other microaerophilic protist parasites with slightly divergent enzymatic configuration have better ways to react to the drug. For instance, *Trichomonas vaginalis* possesses a lactate dehydrogenase and is able to use it to get rid of excess reduction equivalents, therefore perhaps being less vulnerable to metronidazole. Consequently, resistance is a much graver problem in *T. vaginalis* treatment than in *E. histolytica*. Furthermore, trichomonads can downregulate thioredoxin reductase activity (Leitsch *et al*., 2009), thereby also minimizing the activation of metronidazole to its toxic nitroradical anions. As shown for both *E. histolytica* and *T. vaginalis*, thioredoxin and thioredoxin reductase are located in spatial proximity to metronidazole, enabling thioredoxin reductase to activate metronidazole, but on the other hand rendering themselves vulnerable to adduct formation with metronidazole (Leitsch *et al*., 2007). Analysis of the proteome of *E. histolytica* with 2DE gels showed that no significant up- or downregulation of proteins was found, thereby confirming the results of the limited scope of our microarray on a broader proteomic

basis. So it is possible to speculate that even if we had used a whole genome microarray, the extent of up- or downregulated genes found because of metronidazole treatment might not have been very high. Five enzymes in *E. histolytica* were found to be prone to adduct formation with metronidazole, namely the already mentioned thioredoxin and thioredoxin reductase, superoxide dismutase, purine nucleoside phosphorylase and a former hypothetical protein now designated metronidazole target protein 1 (Mtp1) (Leitsch *et al*., 2007).

The whole picture shows a cell that is not able to react properly to the chemotherapeutical stress. The small changes in transcriptional levels of a few genes cannot be translated to the upregulation of proteins. Fortunately for us, metronidazole is a very effective drug that *E. histolytica* finds only few possibilities to overcome.

5 Summary

The objective of this PhD thesis was to investigate the reaction of the microaerophilic protist parasite *Entamoeba histolytica* to the standard chemotherapeutical agent metronidazole on the gene expression level with the help of oligonucleotide microarrays. On the basis of early data from the *E. histolytica* genome project we first compiled three annotation databases currently containing genes encoding enzymes (1334 entries), cytoskeleton proteins (222 entries), and stress genes and redox factors (117 entries). We found a number of genes with an assumed bacterial origin and established an overview about metabolic pathways present in the *E. histolytica* genome.

In a second step we used our databases for the construction of oligonucleotide probes for our microarray. To investigate the reaction of amoebae to metronidazole we used cells of *E. histolytica* HM-1:IMSS treated with 50µM of metronidazole for 1 h, 3 h and 5 h, respectively, and untreated control cells. Fluorescently labeled cDNAs from these cells were hybridized to our focussed array representing 213 genes, and the data were analysed.

The transcriptional response of *E. histolytica* to metronidazole treatment was unexpectedly modest, with 15 genes at least partly upregulated. Only seven showed at least one value with a statistically significant fold change ≥2. These were superoxide dismutase (SOD), peroxiredoxin, two ferredoxin genes, the Gal/GalNAc specific lectin light and heavy chains and a long chain fatty acid CoA ligase. Modest down-regulation was found with eight genes, with six genes showing at least one time point with significant downregulation ≤ 0.5. These were genes encoding an actin isoform, the 101 kDa heat shock protein, a short chain dehydrogenase, a fatty acid elongase, diaphanous protein and a WD repeat protein. We used quantitative real time RT-PCR to

verify these transcriptional changes for seven selected genes, and these results confirmed our microarray data.

In a parallel PhD thesis the changes in protein levels were monitored with the help of 2DE gels. There again no significant up- or down-regulation of protein levels was found under metronidazole influence, but a modification of five distinct proteins by reduced metronidazole metabolites (Leitsch *et al*., 2007).

Taken together, the response of *E. histolytica* to metronidazole treatment on the mRNA level was rather modest, with only a few genes up- or down-regulated significantly. Several of these upregulated genes (SOD, peroxiredoxin, thioredoxin and thioredoxin reductase) protect the amoeba cells against oxidative stress that is generated by reduced metronidazole. But on the whole, *E. histolytica* appears to be unable to mount a strong specific response to inactivate metronidazole, or to prevent its activation.

6 Zusammenfassung

Das Ziel dieser Dissertation war die Untersuchung der Reaktion des mikroaerophilen einzelligen Parasiten *Entamoeba histolytica* auf die Behandlung mit dem klassischen Chemotherapeutikum Metronidazol auf mRNA-Ebene mithilfe von Oligonucleotid-Microarrays. Basierend auf dem *E. histolytica* – Genomprojekt erstellten wir zuerst drei Datenbanken, die die Annotation von aktuell 1334 Enzymen, 222 Zytoskelettgenen sowie 117 Stressgenen und Redox-Faktoren beinhalten. Wir fanden viele Gene, die möglicherweise bakteriellen Ursprungs sind, und etablierten einen Überblick über die in *E. histolytica* vorhandenen Stoffwechselwege.

In einem zweiten Schritt verwendeten wir unsere Datenbanken für das Design von Oligonucleotid-Proben für unseren Microarray. Um die Reaktion von Amöben auf Metronidazol zu untersuchen, verwendeten wir Zellen von *E. histolytica* HM-1:IMSS, die für 1 h, 3 h oder 5 h mit jeweils 50µM Metronidazol behandelt wurden, sowie Kontroll-Zellen. Die fluoreszenzmarkierte cDNA dieser Proben wurde auf unseren fokussierten, 213 Gene representierenden Microarray hybridisiert und die Daten wurden ausgewertet.

Erstaunlicherweise änderte sich die Menge der transkribierten mRNAs von *E. histolytica* durch die Metronidazol-Behandlung nur sehr geringfügig. Bei 15 Genen kam es zu einer zumindest teilweisen Erhöhung der Transkriptionsrate, nur sieben davon wiesen zumindest einen Wert mit einer signifikanten Erhöhung ≥2 auf. Dies waren Superoxiddismutase (SOD), Peroxiredoxin, zwei Ferredoxin-Gene, leichte- und schwere Ketten des Gal/GalNAc-spezifischen Lektins, sowie eine Fettsäure-CoA-Ligase mit Spezifität für lange Ketten. Bei acht Genen wurde ein leichtes Herunterfahren der Transkription gefunden, wobei der Wert bei sechs Genen signifikant auf ≤0,5 fiel. Dies waren ein Actin, das 101 kDa Hitzeschockprotein, eine kurzkettige Dehydrogenase, eine Fettsäure-Elongase, das sogenannte "diaphanous"-Protein sowie

ein "WD-repeat"-Protein. Zur Überprüfung unserer Ergebnisse führten wir für sieben Gene auch quantitative "real time RT-PCR" durch, welche die Microarray-Daten bestätigte.

In einer parallel laufenden Dissertation wurden die Veränderungen auf Protein-Ebene mithilfe von zweidimensionalen Gelen untersucht. Auch hier konnte bei keinem Protein eine signifikante Änderung der produzierten Menge aufgrund von Metronidazol-Einfluss gefunden werden. Alledings kam es zu Modifikationen bei fünf bestimmten Proteinen aufgrund der Anlagerung von reduzierten Metronidazol-Metaboliten (Leitsch *et al*., 2007).

Zusammenfassend gesehen reagierten *E. histolytica*-Zellen auf die Einwirkung von Metronidazol nur mit einer sehr geringen Anpassung der Transkription, nur wenige Gene wurden signifikant hinauf- oder hinunterreguliert. Einige dieser hochregulierten Gene (SOD, Peroxiredoxin, Thioredoxin und Thioredoxinreduktase) schützen die Amöbenzellen gegen oxidativen Stress, der von reduziertem Metronidazol verursacht wird. Das Gesamtbild zeigt Amöbenzellen, die zu keiner konzertierten sinnvollen Reaktion gegen Metronidazol imstande sind und auch die Aktivierung des Metronidazols nicht verhindern können.

7 References

Anderson, IJ and Loftus, BJ (2005). *Entamoeba histolytica*: Observations on metabolism based on the genome sequence. *Exp Parasitol* **110:** 173–177.

Andersson JO, Sjögren AM, Davis LA, Embley TM, Roger AJ (2003). Phylogenetic analyses of diplomonad genes reveal frequent lateral gene transfers affecting eukaryotes. *Curr Biol* **13:** 94–104.

Akgun Y, Tacyildiz IH, Cleik Y (1999). Amebic liver abscess: changing trends over 20 years. *World J Surg* **23:** 102–06.

Albuquerque SS, Carret C, Grosso AR, Tarun AS, Peng X, Kappe SH, Prudêncio M, Mota MM (2009). Host cell transcriptional profiling during malaria liver stage infection reveals a coordinated and sequential set of biological events. *BMC Genomics* **10:** 270.

Ali V, Nozaki T (2007). Current therapeutics, their problems, and sulfur-containing-amino-acid metabolism as a novel target against infections by “amitochondriate” protozoan parasites. *Clin Microbiol Rev* **20:** 164–87.

Alsmark UC, Sicheritz-Ponten T, Foster PG, Hirt RP, Embley TM (2009). Horizontal gene transfer in eukaryotic parasites: a case study of *Entamoeba histolytica* and *Trichomonas vaginalis*. *Methods Mol Biol* **532:** 489–500.

Arhets P, Gounon P, Sansonetti P, Guillén N (1995). Myosin II is involved in capping and uroid formation in the human pathogen *Entamoeba histolytica*. *Infect Immun* **63:** 4358–67.

Bailey GB, Day DB, McCoomer NE (1992). *Entamoeba* motility: dynamics of cytoplasmic streaming, locomotion and translocation of surface-bound particles, and organization of the actin cytoskeleton in *Entamoeba invadens*. *J Protozool* **39:** 267–72.

Bakatselou C, Kidgell C, Graham Clark (2000). A mitochondrial-type hsp70 gene of *Entamoeba histolytica*. *Mol Biochem Parasitol* **110:** 177–82.

Bakre AA, Rawal K, Ramaswamy R, Bhattacharya A, Bhattacharya S (2005). The LINEs and SINEs of *Entamoeba histolytica*: comparative analysis and genomic distribution. *Exp Parasitol* **110:** 207–13.

Bansal D, Ave P, Kerneis S, Frileux P, Boché O, Baglin AC, Dubost G, Leguern AS, Prevost MC, Bracha R, Mirelman D, Guillén N, Labruyère E (2009). An ex-vivo human intestinal model to study *Entamoeba histolytica* pathogenesis. *PLoS Negl Trop Dis* **3:** e551.

Bapteste E, Brinkmann H, Lee JA, Moore DV, Sensen CW, Gordon P, Durufle L, Gaasterland T, Lopez P, Muller M, Philippe H (2002). The analysis of 100 genes supports the grouping of three highly divergent amoebae: *Dictyostelium*, *Entamoeba*, and *Mastigamoeba*. *Proc Natl Acad Sci USA* **99:** 1414–9.

Bendesky A, Menéndez D, Ostrosky-Wegman P (2002). Is metronidazole carcinogenic? *Mutat Res* **511:** 133–44.

Bhattacharya, S., Som, I., and Bhattacharya, A. (1998). The ribosomal DNA plasmids of *Entamoeba*. *Parasitol Today* **14:** 181–85.

Biller L, Schmidt H, Krause E, Gelhaus C, Matthiesen J, Handal G, Lotter H, Janssen O, Tannich E, Bruchhaus I (2009). Comparison of two genetically related *Entamoeba histolytica* cell lines derived from the same isolate with different pathogenic properties. *Proteomics* **9:** 4107–20.

Brandt H, Tamayo RP (1970). Pathology of human amebiasis. *Human Pathology* **1:** 351–85.

Bredeston LM, Caffaro CE, Samuelson J, Hirschberg CB (2005). Golgi and endoplasmic reticulum functions take place in different subcellular compartments of *Entamoeba histolytica*. *J Biol Chem* **280:** 32168–76.

Bruchhaus I, Tannich E (1994). Purification and molecular characterization of the NAD(+)-dependent acetaldehyde/alcohol dehydrogenase from *Entamoeba histolytica. Biochem J* **303:** 743-8.

Brumpt, ME (1925). Ètude sommaire de l' "*Entamoeba dispar"* n. sp. Amibe a kystes quadrinuclées, parasite de l'homme. *Bull Acad Med* (Paris) **94:** 943–952.

Buss S, Kabir M, Petri WA Jr., Haque R (2008). Comparison of two immunoassays for detection of *Entamoeba histolytica. J Clin Microbiol* **46:** 2778–9.

Bustin SA (2009). Why the need for qPCR publication guidelines?-The case for MIQE. *Methods* 2009 Dec 16. [Epub ahead of print].

Carlton JM, Hirt RP, Silva JC, Delcher AL, Schatz M, Zhao Q, Wortman JR, Bidwell SL, Alsmark UC, Besteiro S, Sicheritz-Ponten T, Noel CJe, Dacks JB, Foster PG, Simillion C, Van de Peer Y, Miranda-Saavedra D, Barton GJ, Westrop GD, Müller S, Dessi D, Fiori PL, Ren Q, Paulsen I, Zhang H, Bastida-Corcuera FD, Simoes-Barbosa A, Brown MT, Hayes RD, Mukherjee M, Okumura CY, Schneider R, Smith AJ, Vanacova S, Villalvazo M, Haas BJ, Pertea M, Feldblyum TV, Utterback TR, Shu CL, Osoegawa K, de Jong PJ, Hrdy I, Horvathova L, Zubacova Z, Dolezal P, Malik SB, Logsdon JM Jr, Henze K, Gupta A, Wang CC, Dunne RL, Upcroft JA, Upcroft P, White O, Salzberg SL, Tang P, Chiu CH, Lee YS, Embley TM, Coombs GH, Mottram JC, Tachezy J, Fraser-Liggett CM, Johnson PJ (2007). Draft genome sequence of the sexually transmitted pathogen *Trichomonas vaginalis. Science* **315:** 207–12.

Cavalier-Smith T (1991). Archamoebae: the ancestral eukaryotes? *Biosystems* **25:** 25–38.

Chan KW, Slotboom DJ, Cox S, Embley TM, Fabre O, van der Giezen M, Harding M, Horner DS, Kunji ER, Leon-Avila G, Tovar J (2005). A novel ADP/ATP transporter in the mitosome of the microaerophilic human parasite *Entamoeba histolytica*. *Curr Biol* **15:** 737–42.

Chapman A, Cammack R, Linstead D, Lloyd D (1985). The generation of metronidazole radicals in hydrogenosomes isolated from *Trichomonas vaginalis*. *J Gen Microbiol* **131:** 2141–4.

Chavez-Tapia NC, Hernandez-Calleros J, Tellez-Avila FI, Torre A, Uribe M (2009). Image-guided percutaneous procedure plus metronidazole versus metronidazole alone for uncomplicated amoebic liver abscess. *Cochrane Database Syst Rev* **21:** CD004886.

Clark CG (2000). The evolution of *Entamoeba*, a cautionary tale. *Res Microbiol* **151:** 599–60.

Clark CG, **Alsmark UC**, **Tazreiter M**, **Saito-Nakano Y**, **Ali V**, **Marion S**, **Weber C**, **Mukherjee C**, **Bruchhaus I**, **Tannich E**, **Leippe M**, **Sicheritz-Ponten T**, **Foster PG**, **Samuelson J**, **Noël CJ**, **Hirt RP**, **Embley TM**, **Gilchrist CA**, **Mann BJ**, **Singh U**, **Ackers JP**, **Bhattacharya S**, **Bhattacharya A**, **Lohia A**, **Guillén N**, **Duchêne M**, **Nozaki T**, **Hall N**. **(2007)**. Structure and content of the *Entamoeba histolytica* genome. *Adv Parasitol* **65:** 51–190.

Clark, CG & Diamond, LS (2002). Methods for cultivation of luminal parasitic protists of clinical importance. *Clin Microbiol Rev* **15:** 329–341.

Clark CG, Roger AJ (1995). Direct evidence for secondary loss of mitochondria in *Entamoeba histolytica*. *Proc Natl Acad Sci U S A* **92:** 6518–21.

Clark CG, Kaffashian F, Tawari B, Windsor JJ, Twigg-Flesner A, Davies-Morel MC, Blessmann J, Ebert F, Peschel B, Le Van A, Jackson CJ, Macfarlane L, Tannich E (2006a). New insights into the phylogeny of *Entamoeba* species provided by analysis of four new small-subunit rRNA genes. *Int J Syst Evol Microbiol*. **56:** 2235–9.

Clark, C. G., Ali, I. K., Zaki, M., Loftus, B. J., and Hall, N. (2006b). Unique organisation of tRNA genes in *Entamoeba histolytica*. *Mol. Biochem. Parasitol*. **146:** 24–29.

Croft SL, Seifert K, Duchêne M (2003). Antiprotozoal activities of phospholipid analogues. *Mol Biochem Parasitol* **126:** 165–72.

Cuervo P, Cupolillo E, Britto C, González LJ, E Silva-Filho FC, Lopes LC, Domont GB, De Jesus JB (2008). Differential soluble protein expression between *Trichomonas vaginalis* isolates exhibiting low and high virulence phenotypes. *J Proteomics* **71:** 109–22.

Das S, Stevens T, Castillo C, Villasenor A, Arredondo H, Reddy K (2002). Lipid metabolism in mucous-dwelling amitochondriate protozoa. *Int J Parasitol* **32:** 655–75.

Davis PH, Schulze J, Stanley SL Jr. (2007). Transcriptomic comparison of two *Entamoeba histolytica* strains with defined virulance phenotypes identifies new virulence factor candidates and key differences in the expression patterns of cysteine proteases, lectin light chains, and calmodulin. *Mol Biochem Parasitol* **151:** 118–28.

Davis PH, Zhang X, Guo J, Townsend RR, Stanley SL Jr (2006). Comparative proteomic analysis of two *Entamoeba histolytica* strains with different virulence phenotypes identifies peroxiredoxin as an important component of amoebic virulence. *Mol Microbiol* **61:** 1523–32

de Sousa Abreu R, Penalva LO, Marcotte EM, Vogel C (2009). Global signatures of protein and mRNA expression levels. *Mol Biosyst* **5:** 1512–26.

Diamond LS. and Clark CG (1993). A redescription of *Entamoeba histolytica* Schaudinn, 1903 (Emended Walker, 1911) separating it from *Entamoeba dispar* Brumpt, 1925. *J Eukaryot Microbiol* **40:** 340–4.

Ehrenkaufer GM, Eichinger DJ, Singh U (2007a). Trichostatin A effects on gene expression in the protozoan parasite *Entamoeba histolytica. BMC Genomics* **8:** 216.

Ehrenkaufer GM, Haque P, Hackney JA, Eichinger DJ, Singh U (2007b). Identification of developmentally regulated genes in *Entamoeba histolytica*: insights into mechanisms of stage conversion in a protozoan parasite. *Cell Mocrobiol* **9:** 1426–44.

Eichinger D (1997). Encystation of *Entamoeba* parasites. *Bioessays* **19:** 633–9.

Eichinger L, Pachebat JA, Glöckner G, Rajandream MA, Sucgang R, Berriman M, Song J, Olsen R, Szafranski K, Xu Q, Tunggal B, Kummerfeld S, Madera M, Konfortov BA, Rivero F, Bankier AT, Lehmann R, Hamlin N, Davies R, Gaudet P, Fey P, Pilcher K, Chen G, Saunders D, Sodergren E, Davis P, Kerhornou A, Nie X, Hall N, Anjard C, Hemphill L, Bason N, Farbrother P, Desany B, Just E, Morio T, Rost R, Churcher C, Cooper J, Haydock S, van Driessche N, Cronin A, Goodhead I, Muzny D, Mourier T, Pain A, Lu M, Harper D, Lindsay R, Hauser H, James K, Quiles M, Madan Babu M, Saito T, Buchrieser C, Wardroper A, Felder M, Thangavelu M, Johnson D, Knights A, Loulseged H, Mungall K, Oliver K, Price C, Quail MA, Urushihara H, Hernandez J, Rabbinowitsch E, Steffen D, Sanders M, Ma J, Kohara Y, Sharp S, Simmonds M, Spiegler S, Tivey A, Sugano S, White B, Walker D, Woodward J, Winckler T, Tanaka Y, Shaulsky G, Schleicher M, Weinstock G, Rosenthal A, Cox EC, Chisholm RL, Gibbs R, Loomis WF, Platzer M, Kay RR, Williams J, Dear PH, Noegel AA, Barrell B, Kuspa A. (2005). The genome of the social amoeba *Dictyostelium discoideum*. *Nature* **435**: 43–57.

Fahey RC, Newton GL, Arrick B, Overdank-Bogart T, Aley SB (1984). *Entamoeba histolytica*: a eukaryote without glutathione metabolism. *Science* **224:** 70–72.

Farri TA, **Sargeaunt PG**, **Warhurst DC**, **Williams JE**, **Bhojnani R (1980)**. Electrophoretic studies of the hexokinase of *Entamoeba histolytica* groups I to IV. *Trans R Soc Trop Med Hyg* **74:** 672–3.

Field J, Rosenthal B, Samuelson J. (2000). Early lateral transfer of genes encoding malic enzyme, acetyl.-CoA synthetase and alcohol dehydrogenases from anaerobic prokaryotes to *Entamoeba histolytica*. *Mol Microbiol* **38:** 446–55.

Fung HB, Doan DL (2005). Tinidazole: a nitroimidazole antiprotozoal agent. *Clin Ther* **27:** 1859–84.

Gathiram V, Jackson TF (1987). A longitudinal study of asymptomatic carriers of pathogenic zymodemes of *Entamoeba histolytica*. *S Afr Med J* **72:** 669–72.

Gilchrist CA, Houpt E, Trapaidze N, Fei Z, Crasta O, Asgharpour A, Evans C, Martino-Catt S, Baba DJ, Stroup S, Hamano S, Ehrenkaufer G, Okada M, Singh U, Nozaki T, Mann BJ, Petri WA Jr (2006). Impact of intestinal colonization and invasion on the *Entamoeba histolytica* transcriptome. *Mol Biochem Parasitol* **147:** 163–76.

Görg A, Weiss W, Dunn MJ (2004). Current two-dimensional electrophoresis technology for proteomics. *Proteomics* **4:** 3665–3685.

Gosh SK, Field J, Frisardi M, Rosenthal B, Mai Z, Rogers R, Samuelson J (1999). Chitinase secretion by encysting *Entamoeba invadens* and transfected *Entamoeba histolytica* trophozoites: localization of secretory vesicles, endoplasmic reticulum, and Golgi apparatus. *Infect Immun* **67:** 3073–81.

Guillén N (1996). Role of signalling and cytoskeletal rearrangements in the pathogenesis of *Entamoeba histolytica*. *Trends Microbiol* **4:** 191–7.

Haanstra JR, van Tuijl A, Kessler P, Reijnders W, Michels PA, Westerhoff HV, Parsons M, Bakker BM (2008). Compartmentation prevents a lethal turbo-explosion of glycolysis in trypanosomes. Proc Natl Acad Sci USA **105:** 17718–23.

Heid CA, Stevens J, Livak KJ, Williams PM (1996). Real time quantitative PCR. G*enome Res* **6:** 986–94.

Hofer M, Duchêne M (2005). *Entamoeba histolytica*: construction and applications of subgenomic databases. *Exp Parasitol* **110:** 178–83.

Hrdy I, Hirt RP, Dolezal P, Bardonova L, Foster PG, Tachezy J, Embley TM (2004). *Trichomonas* hydrogenosomes contain the NADH dehydrogenase module of mitochondrial complex I. *Nature* **432:** 618–22.

Huang KY, Chien KY, Lin YC, Hsu WM, Fong IK, Huang PJ, Yueh YM, Gan RR, Tang P (2009). A proteome reference map of *Trichomonas vaginalis*. *Parasitol Res* **104:** 927–33.

Huston CD, Boettner DR, Miller-Sims V, Petri WA Jr. (2003). Apoptotic killing and phagocytosis of host cells by the parasite *Entamoeba histolytica. Infect Immun* **71:** 964–72.

Huston CD, Houpt ER, Mann BJ, Hahn CS, Petri WA Jr. (2000). Caspase 3-dependent killing of host cells by the parasite *Entamoeba histolytica. Cell Microbiol* **2:** 617–25.

Isakov E, Siman-Tov R, Weber C, Guillen N, Ankri S (2008). Trichostatin A regulates peroxiredoxin expression and virulance of the parasite *Entamoeba histolytica*. *Mol Biochem Parasitol* **158:** 82–94.

Kelsall BL, Ravdin JI (1993). Degradation of human IgA by *Entamoeba histolytica*. *J Infect Dis* **168:** 1319–22.

Kim J, Bae SS, Sung MH, Lee KH, Park SJ (2009). Comparative proteomic analysis of trophozoites versus cysts of *Giardia lamblia. Parasitol Res* **104:** 475–9.

Kumagai M, Makioka A, Takeuchi T, Nozaki T (2004). Molecular cloning and characterization of a protein farnesyltransferase from the enteric protozoan parasite *Entamoeba histolytica*. *J Biol Chem* **279:** 2316–23.

Lauwaet T, Oliveira MJ, Callewaert B, De Bruyne G, Saelens X, Ankri S, Vandenabeele P, Mirelman D, Mareel M, Leroy A (2003). Proteolysis of enteric cell villin by *Entamoeba histolytica* cysteine proteinases. *J Biol Chem* **278**: 22650–6.

Leippe M (1997). Amoebapores. *Parasitology Today* **13:** 178–183.

Leitsch D, Kolarich D, Binder M, Stadlmann J, Altmann F, Duchêne M. (2009). *Trichomonas vaginalis*: metronidazole and other nitroimidazole drugs are reduced by the flavin enzyme thioredoxin reductase and disrupt the cellular redox system. Implications for nitroimidazole toxicity and resistance. *Mol Microbiol*. **72:** 518–536.

Leitsch D, Kolarich D, Wilson IB, Altmann F, Duchêne M. (2007). Nitroimidazole action in *Entamoeba histolytica*: a central role for thioredoxin reductase. *PLoS Biol*. **5:** e211.

Leitsch D, Kolarich D, Duchêne M (2010). The flavin inhibitor diphenyleneiodonium renders *Trichomonas vaginalis* resistant to metronidazole, inhibits thioredoxin reductase and flavin reductase, and shuts off hydrogenosomal enzymatic pathways. Mol Biochem Parasitol 2010 Jan 20 [Epub ahead of print].

Leitsch D, Radauer C, Paschinger K, Wilson IBH, Breiteneder H, Scheiner O, Duchêne M (2005). *Entamoeba histolytica*: analysis of the trophozoite proteome by two-dimensional polyacrylamide gel electrophoresis. *Exp Parasitol* **110:** 191–195.

Leitsch D, Wilson IBH, Paschinger K, Duchêne M (2006). Comparison of the proteome profiles of *Entamoeba histolytica* and its close but non-pathogenic relative *Entamoeba dispar*. *Wien Klin Wochenschr* **118** [Suppl3]**:** 37–41.

Li E, Becker A, Stanley SL Jr. (1988). Use of Chinese hamster ovary cells with altered glysosylation patterns to define the carbohydrate specificity of *Entamoeba histolytica* ahesion. *J Exp Med* **167:** 1725–30.

Lejeune M, Rybicka JM, Chadee K (2009). Recent discoveries in the pathogenesis and immune response toward *Entamoeba histolytica. Future Microbiol* **4:** 105–18.

Lindmark DG, Müller M (1973). Hydrogenosome, a cytoplasmic organelle of the anaerobic flagellate *Tritrichomonas foetus*, and its role in pyruvate metabolism. *J Biol Chem* **248:** 7724–8.

Loftus BJ, Anderson I, Davies R, Alsmark UCM, Samuelson J, Amadeo P, Roncaglia P, Berriman M, Hirt RP, Mann BJ, Nozaki T, Suh B, Pop M, Duchêne M, Ackers J, Tannich E, Leippe M, Hofer M, Bruchhaus I, Willhoeft U, Bhattacharya A, Chillingworth T, Churcher C, Hance Z, Harris B, Harris D, Jagels K, Moule S, Mungall K, Ormond D, Squares R, Whitehead S, Guillén N, Gilchrist C, Stroup SE, Bhattacharya S, Lohia A, Foster PG, Sicheritz-Ponten T, Weber C, Singh U, Mukherjee C, Petri WA Jr, Clark CG, Embley TM, Barrell B, Fraser C, Hall N, (2005). The genome of the protist parasite *Entamoeba histolytica. Nature* **433:** 865–8.

Lösch (Lesh) FD (1875). Massive development of amebas on the large intestine. Translated and reprinted in *AM J Trop Med Hyg* **24:** 383–92.

Lotter H, González-Roldán N, Lindner B, Winau F, Isibasi A, Moreno-Lafont M, Ulmer AJ, Holst O, Tannich E, Jacobs T (2009). Natural killer T cells activated by a lipopeptidophosphoglycan from *Entamoeba histolytica* are critically important to control amebic liver abscess. *PloS Pathog* **5(5):** e1000434.

Lotter H, Jacobs T, Gaworski I, Tannich E (2006). Sexual dimorphism in the control of amebic liver abscess in a mouse model of disease. *Infect Immun* **74:**118–24.

Lotter H, Tannich E (2006). The Current Status of an Amebiasis Vaccine. *Arch Med Res* **37:** 291–5.

Ludlum D, Colinas RJ, Kirk MC, Mehta JR (1988). Reaction of reduced metronidazole with guanosine to form an unstable adduct. *Carcinogenesis* **9:** 593–6.

Mackinnon MJ, Li J, Mok S, Kortok MM, Marsh K, Preiser PR, Bozdech Z (2009). Comparative transcriptional and genomic analysis of *Plasmodium falciparum* field isolates. *PloS Pathog* **5(10):** e1000644.

Mai Z, Ghosh S, Frisardi M, Rosenthal B, Rogers R, Samuelson J (1999). Hsp60 is targeted to a cryptic mitochondrion-derived organelle ("crypton") in the microaerophilic protozoan parasite *Entamoeba histolytica*. *Mol Cell Biol* **19:** 2198–205.

Makioka A, Kumagai M, Ohtomo H, Kobayashi S, Takeuchi T (2002). Effect of proteasome inhibitors on the growth, encystation, and excystation of *Entamoeba histolytica* and *Entamoeba invadens*. *Parasitol Res* **88:** 454–9.

Makioka A, Kumagai M, Takeuchi T, Nozaki T (2006). Characterization of protein geranylgeranyltransferase I from the enteric protist *Entamoeba histolytica. Mol Biochem Parasitol* **145:** 216–25.

Mann BJ (2002). Structure and function of the *Entamoeba histolytica* Gal/GalNAc lectin. *Int Rev Cytol* **216:** 59–80.

Maralikova B, Ali V, Nakada-Tsukui K, Nozaki T, van der Giezen M, Henze K, Tovar J (2010). Bacterial-type oxygen detoxification and iron-sulfur cluster assembly in amoebal relict mitochondria. *Cell Microbiol* **12:** 331-42.

Marinets A, Zhang T, Guillen N, Gounon P, Bohle B, Vollmann U, Scheiner O, Wiedermann G, Stanley SL, Duchêne M (1997). Protection against invasive amebiasis by a single monoclonal antibody directed against a lipophosphoglycan antigen localized on the surface of *Entamoeba histolytica*. *J Exp Med* **186:** 1557–65.

Martin, W, and Müller, M (1998). The hydrogen hypothesis for the first eukaryote. *Nature* **392:** 37–41.

Martínez-Palomo A (1982). The Biology of *Entamoeba histolytica. Research Studies Press* ISBN 0471104043 .

Mason RP, Holtzman JL (1975). The role of catalytic superoxide formation in the O2 inhibition of nitroreductase. *Biochem Biophys Res Commun* **67:**1267–74.

Mazzuco A, Benchimol M, De Souza W (1997). Endoplasmic reticulum and Golgi-like elements in *Entamoeba*. *Micron* **28:** 241–7.

Meza I, Talamás-Rohana P, Vargas MA (2006). The cytoskeleton of *Entamoeba histolytica*: structure, function, and regulation by signaling pathways. *Arch Med Res* **37:** 234–43.

Mi-Ichi F, Yousuf MA, Nakada-Tsukui K, Nozaki T (2009). Mitosomes in *Entamoeba histolytica* contain a sulfate activation pathway. *Proc Natl Acad Sci USA* **106:** 21731–6.

Moreno S, Docampo R (1985). Mechanism of toxicity of nitro compounds used in the chemotherapy of trichomoniasis. *Environ Health Perspect* **64:** 199–208.

Morrison HG, McArthur AG, Gillin FD, Aley SB, Adam RD, Olsen GJ, Best AA, Cande WZ, Chen F, Cipriano MJ, Davids BJ, Dawson SC, Elmendorf HG, Hehl AB, Holder ME, Huse SM, Kim UU, Lasek-Nesselquist E, Manning G, Nigam A, Nixon JE, Palm D, Passamaneck NE, Prabhu A, Reich CI, Reiner DS, Samuelson J, Svard SG, Sogin ML (2007). Genomic minimalism in the early diverging intestinal parasite *Giardia lamblia. Science* **317:** 1921–6.

Muhkerjee C, Majumder S, Lohia A (2009). Inter-cellular variation in DNA content of *Entamoeba histolytica* originates from temporal and spatial uncoupling of cytokinesis from the nuclear cycle. *PloS Negl Trop Dis* **3:** e409.

Müller J, Ley S, Felger I, Hemphill A, Müller N (2008). Identification of differentially expressed genes in a *Giardia lamblia* WB C6 clone resistant to nitazoxanide and metronidazole. *J Antimicrob Chemother* **62:** 72–82.

Müller M (1983). Mode of action of metronidazole on anaerobic bacteria and protozoa. *Surgery* **93:** 165–71.

Müller M (1986). Reductive activation of nitroimidazoles in anaerobic microorganisms. *Biochem Pharmacol* **35:** 37–41.

Nickel R, Ott C, Dandekar T, Leippe M (1999). Pore-forming peptides of *Entamoeba dispar*. Similarity and divergence to amoebapores in structure, expression and activity. *Eur J Biochem* **265:** 1002–7.

Nixon, JEJ, Wang A, Field J, Morrison HG, McArthur AG, Sogin ML, Loftus BJ, Samuelson J (2002). Evidence for lateral transfer of genes encoding ferredoxins, nitroreductases, NADH oxidase, and alcohol dehydrogenase 3 from anaerobic prokarytoes to *Giardia lamblia* and *Entamoeba histolytica*. *Eukaryot Cell* **1:** 181–90.

Noedl H, Se Y, Schaecher K, Smith BL, Socheat D, Fukuda MM, Artemisinin Resistance in Cambodia 1 (ARC1) Study Consortium (2008). Evidence of artemisinin-resistant malaria in western Camobodia. *N Engl J Med* **359(24):** 2619–20.

Noedl H, Socheat D, Satimai W (2009). Artemisinin-resistant malaria in Asia. *N Engl J Med* **361:** 541–1.

Nolan T, Hands RE, Bustin SA (2006). Quantification of mRNA using real-time RT-PCR. *Nat Protoc* **1:**1559–82.

Nzila A, Mwai L (2009). In vitro selection of *Plasmodium falciparum* drug-resistant parasite lines. *J Antimicrob Chemother* 2009 Dec 18, Epub ahead of print, doi:10.1093/jac/dkp449.

Ormerod WE, Venkatesan S (1982). Similarities of lipid metabolism in mammalian and protozoan cells: an evolutionary hypothesis for the prevalence of atheroma. *Microbiol Reviews* **46:** 296–307.

Ortner S, Clark CG, Binder M, Scheiner O, Wiedermann G, Duchêne M (1997a). Molecular biology of the hexokinase isoenzyme pattern that distinguishes pathogenic *Entamoeba histolytica* from nonpathogenic *Entamoeba dispar*. *Mol Biochem Parasitol* **86:** 85–94.

Ortner S, Binder M, Scheiner O, Wiedermann G, Duchêne M (1997b). Molecular and biochemical characterization of phosphoglucomutases from *Entamoeba histolytica* and *Entamoeba dispar*. *Mol Biochem Parasitol* **90:** 121–9.

Paugam A, Bulteau AL, Dupouy-Camet J, Creuzet C, Friguet B (2003). Characterization and role of protozoan parasite proteasomes. *Trends Parasitol*. **19:** 55–9.

Petri WA Jr, Chaudhry O, Haque R, Houpt E (2006). Adherence-blocking vaccine for amebiasis. *Ach Med Res* **37:** 288–91.

Poole A; Penny D (2007). Evaluating hypotheses for the origin of eukaryotes. *Bioessays* **29:** 74–84.

Rasoloson D, Vanácová S, Tomková E, Rázga J, Hrdy I, Tachezý J, Kulda J (2002). Mechanisms of in vitro development of resistance to metronidazole in *Trichomonas vaginalis*. *Microbiology* **148:** 2467–77.

Reed SL, Keene WE, McKerrow JH, Gigli I (1989). Cleavage of C3 by a neutral cysteine proteinase of *Entamoeba histolytica*. *J Immunol* **143:** 189–95.

Reeves RE (1984). Metabolism of *Entamoeba histolytica* Schaudinn, 1903. *Adv Parasitol* **23:**105–42.

Richards, TA and van der Giezen M (2006). Evolution of the Isd11-IscS complex reveals a single a-proteobacterial endosymbiosis for all eukaryotes. *Mol Biol Evol* **23:** 1341–44.

Rosenthal B, Mai Z, Caplivski D, Ghosh S, de la Vega H, Graf T, Samuelson J (1997). Evidence for the bacterial origin of genes encoding fermentation enzymes of the amitochondriate protozoan parasite *Entamoeba histolytica*. *J Bacteriol* **179:** 3736–45.

Rossignol JF, Ayoub A, Ayers MS (2001). Treatment of diarrhea caused by *Giardia intestinalis* and *Entamoeba histolytica* or *E. dispar*: a randomized, double-blind, placebo-controlled study of nitazoxanide. *J. Infect. Dis*. **184:** 381–4.

Roy S, Kabir M, Mondal D, Ali IK, Petri WA Jr, Haque R (2005). Real-time-PCR assay for diagnosis of *Entamoeba histolytica* infection. *J Clin Microbiol* **43:** 2168–72.

Sánchez L, Horner D, Moore D, Henze K, Embley T, Müller M (2002). Fructose-1,6-bisphosphate aldolases in amitochondriate protists constitute a single protein subfamily with eubacterial relationships. *Gene* **295:** 51–9.

Sargeaunt, P, Williams JE, Grene JD (1978). The differentiation of invasive and non-invasive *Entamoeba histolytica* by isoenzyme electrophoresis. *Trans R Soc Trop Med Hyg* **72:** 519–21.

Schaudinn F (1903). Untersuchungen über die Fortpflanzung einiger Rhizopoden. *Arbeiten aus dem Kaiserlichen Gesundheitsamte* **19:** 547–76.

Seifert K, Duchêne M, Wernsdorfer WH, Kollaritsch H, Scheiner O, Wiedermann G, Hottkowitz T, Eibl H (2001). Effects of miltefosine and other alkylphosphocholines on human intestinal parasite *Entamoeba histolytica*. *Antimicrob Agents Chemother* **45:** 1505–10.

Shalon D, Smith SJ, Brown PO (1996). A DNA microarray system for analyzing complex DNA samples using two-color fluorescent probe hybridization. *Genome Res* **6:** 639–45.

Sharma R (1959). Effect of cholesterol on the growth and virulence of *Entamoeba histolytica. Trans R Soc Trop Med Hyg* **53:** 278–81.

Singh A, Houpt E, Petri WA (2009). Rapid diagnosis of intestinal parasitic protozoa, with a focus on *Entamoeba histolytica. Interdiscip Perspect Infect Dis* **2009:** 547090.

Stanley, SL Jr.(2003). Amoebiasis. *Lancet* **361:**1025–34.

Stark D, van Hal SJ, Matthews G, Harkness J, Marriott D (2008). Invasive amebiasis in men who have sex with men, Australia. Emerg Infect Dis. 2008 July; **14:** 1141–3.

Steuart RF (2010). Proteomic analysis of *Giardia*: studies from the pre-and post-genomic era. *Exp Parasitol* **124:** 26–30.

Tannich, E; Burchard, GD (1991). Differentiation of pathogenic from non-pathogenic *Entamoeba histolytica* by restriction fragment analysis of a single gene amplified in vitro. *J.Clin.Microbiology* **29:** 250–5.

Tarun AD, Peng X, Dumpit RF, Ogata Y, Silva-Rivera H, Camargo N, Daly TM, Bergman LW, Kappe SH (2008).A combined transcriptome and proteome survey of malaria parasite liver stages. *Proc Natl Acad Sci USA* **105:** 305–10.

Tazreiter M, Leitsch D, Hatzenbichler E, Mair-Scorpio GE, Steinborn R, Schreiber M, Duchêne M (2008). *Entamoeba histolytica*: response of the parasite to metronidazole challenge on the levels of mRNA and protein expression. *Exp Parasitol* **120:** 403–10.

Tian Q, Stepaniants SB, Mao M, Weng L, Feetham MC, Doyle MJ, Yi EC, Dai H, Thorsson V, Eng J, Goodlett D, Berger JP, Gunter B, Linseley PS, Stoughton RB, Aebersold R, Collins SJ, Hanlon WA, Hood LE (2004). Integrated genomic and proteomic analyses of gene expression in mammalian cells. *Mol Cell Proteomics* **3:** 960–9.

Tillack M, Biller L, Irmer H, Freitas M, Gomes MA, Tannich E, Bruchhaus I (2007). The *Entamoeba histolytica* genome: primary structure and expression of proteolytic enzymes. *BMC Genomics* **8:** 170.

Tolstrup J, Krause E, Tannich E, Bruchhaus I (2006). Proteomic analysis of *Entamoeba histolytica*. *Parasitology* **134:** 289–98.

Tovar J, Cox SS, van der Giezen M (2007). A mitosome purification protocol based on percoll density gradients and its use in validating the mitosomal nature of *Entamoeba histolytica* mitochondrial Hsp70. *Methods Mol Biol*. **390:**167–77.

Tovar J, Fischer A, Clark CG (1999). The mitosome, a novel organelle related to mitochondria in the amitochondrial parasite *Entamoeba histolytica*. *Mol Microbiol* **32:** 1013–21.

Tovar J, Leon-Avila G, Sanchez LB, Sutak R, Tachezy J, van der Giezen M, Hernandez M, Muller M, Lucocq JM (2003). Mitochondrial remnant organelles of *Giardia* function in iron-sulphur protein maturation. *Nature* **426:**172–6.

Upcroft P, Upcroft JA (2001). Drug targets and mechanisms of resistance in the anaerobic protozoa. *Clin Microbiol Rev* **14:** 150–64.

van der Giezen M and Tovar J (2005). Degenerate mitochondria. *EMBO reports* **6:** 525–30.

Vargas M, Sansonetti P, Guillén N (1996). Identification and cellular localization of the actin-binding protein ABP-120 from *Entamoeba histolytica*. Mol Microbiol **22:** 849–57.

Vicente JB, Ehrenkaufer GM, Saraiva LM, Teixeira M, Singh U (2009). *Entamoeba histolytica* modulates a complex repertoire of novel genes in response to oxidative and nitrosative stresses: implications for amebic pathogenesis. *Cell Microbiol* **11:** 51–69.

Walochnik J, Duchêne M, Seifert K, Eibl HJ, Obwaller A, Aspöck H (2002). Cytotoxic activities of various alkylphosphocholines (APCs) against clinical isolates of *Acanthamoeba*. *Antimicrob Agents Chemother* **46:** 695–701.

Wang Z, Vora GJ, Stenger DA (2004). Detection and genotyping of *Entamoeba histolytica*, *Entamoeba dispar*, *Giardia lamblia*, and *Cryptosporidium parvum* by oligonucleotide microarray. *J Clin Microbiol* **42:** 3262–71.

Wassmann C, Hellberg A, Tannich E, Bruchhaus I (1999). Metronidazole resistance in the protozoan parasite *Entamoeba histolytica* is associated with increased expression of iron-containing superoxide dismutase and peroxiredoxin and decreased expression of ferredoxin 1 and flavin reductase. *J Biol Chem* **274:** 26051–6.

Weber C, Guigon G, Bouchier C, Frangeul L, Moreira S, Sismeiro O, Gouyette C, Mirelman D, Coppee JY, Guillen N (2006). Stress by heat shock induces massive down regulation of genes and allows differential allelic expression of the Gal/GalNAc lectin in *Entamoeba histolytica*. *Eukaryot Cell* **5:** 871–5.

Weber C, Marchat LA, Guillen N, López-Camarillo C (2009). Effects of DNA damage induced by UV irradiation on gene expression in the protozoan parasite *Entamoeba histolytica. Mol Biochem Parasitol* **164:** 165–9.

West SB, Wislocki PG, Fiorentini KM, Alvaro R, Wolf FJ, Lu AYH (1982). Drug residue formation from ronidazole, a 5-nitroimidazole . I. Characterization of *in vitro* protein alkylation*. Chem-Biol Interact* **41:** 265–279.

Willhoeft U, Hamann L, Tannich E (1999). A DNA sequence corresponding to the gene encoding cysteine proteinase 5 in *Entamoeba histolytica* is present and positionally conserved but highly degenerated in *Entamoeba dispar*. *Infect Immun* **167:** 5925–9.

Willhoeft U, Tannich E (1999). The electrophoretic karyotype of *Entamoeba histolytica*. *Mol Biochem Parasitol* **15:** 49–53.

World Health Organization (1997). WHO/PAHO/UNESCO report: a consultation with experts on amebiasis. MexicoCity, Mexico 28-29 January,1997. *Epidemiological Bulletin PAHO* **18:** 13–4.

Zaki M, Andrew N, Insall RH (2006). *Entamoeba histolytica* cell movement: a central role for self-generated chemokines and chemorepellents. *Proc Natl Acad Sci USA* **103:** 18751–6.

8 List of abbreviations

2D-DIGE	two-dimensional difference gel electrophoresis
2DE	two-dimensional gel electrophoresis
ALA	amoebic liver abscess
Alkyl-PC	alkylphosphocholine
Cy3/5	green (3) or red (5) fluorescent dye of the cyanine dye familiy
ELISA	enzyme-linked immunosorbent assay
ER	endoplasmatic reticulum
FD	ferredoxin
GA	Golgi apparatus
Gal/GalNAc lectin	galactose/N-acetylgalactosamine specific lectin
Hsp10	10 kDa heat shock protein
Hsp70	70 kDa heat shock protein
LGT	lateral gene transfer
LINE	long interspersed element
LPG	lipophosphoglycan
Mbp	mega base pairs
PCR	polymerase chain reaction
PFOR	pyruvate:ferredoxin oxidoreductase
qRT-PCR	quantitative real-time reverse transcription polymerase chain reaction
SINE	short interspersed element
SOD	superoxide dismutase
TCA	trichloroacetic acid
TSA	trichostatin A

Danksagung

Zuallererst möchte ich mich bei meinem Betreuer Michael Duchêne bedanken, der mir dieses interessante Thema anvertraut hat und mich während der Arbeit daran stets mit angeregten Diskussionen und Ratschlägen unterstützt hat. Besonders dankbar bin ich ihm für seine herzliche und freundschaftliche Betreuung sowie seinen ungebrochenen Glauben an die Fertigstellung dieser Arbeit auch während der 2 Jahre nach der Geburt meines Sohnes Jakob.
Ebenso bedanke ich mich bei Univ. Prof. Dr. Otto Scheiner sowie bei Univ. Prof. Dr. Ursula Wiedermann-Schmidt für die Möglichkeit, das Projekt am Institut durchzuführen.
Meinen Kollegen im Labor, David Leitsch, Marina Binder sowie Helen Melzer, Birgit Wagner und Nicola Mason danke ich sehr herzlich für die schöne Zeit, das gute Klima und viele hilfreiche Kleinigkeiten. Besonders David möchte ich für die gute Zusammenarbeit und Hilfe an den Schnittstellen unserer Arbeitsgebiete danken.
Für die unersetzliche Hilfe bei der Durchführung der Microarray-Experimente gilt mein Dank der Universitätsklinik für Frauenheilkunde mit Martin Schreiber, Agnes Grünfelder sowie ganz besonders Evelyn Hatzenbichler, die sehr viel Zeit in das Spotten der Arrays investiert hat. Die qRT-PCR Experimente wurden am Institut für Tierzucht und Genetik der Veterinärmedizinischen Universität Wien durchgeführt, wo ich mich sehr herzlich bei Ralf Steinborn und Georg Mair für ihre großzügige Hilfe und Unterstützung bedanken möchte.
Meine Familie und hier besonders meine Eltern haben mich mein ganzes Leben lang immer unterstützt und ermutigt, meinen Weg zu gehen. Für all ihre Liebe und Sorge gilt ihnen mein innigster Dank. Besonders danke ich auch meinen Schwiegereltern, die sich in der letzten Phase dieser Arbeit noch öfter als sonst liebevoll um unseren Sonnenschein Jakob gekümmert haben, ohne ihre Hilfe hätte ich nicht abschließen können.
Zuletzt danke ich meinem Mann Wolfgang. Du bist der Mittelpunkt meines Lebens geworden, ohne deine Liebe hätte nichts mehr Sinn.

Curriculum vitae

Name:	Margit Helga Tazreiter, geb. Hofer
Geburtsdatum:	19. Jänner 1977 in Eisenstadt
Eltern:	Helga und Emmerich Hofer, Landwirte in Ritzing, Mittelburgenland
Schulbildung:	Volksschule in Ritzing (1983-1987) Bundesrealgymnasium in Oberpullendorf (1987-1995)
Universitätsausbildung:	ab Oktober 1995 Studium der Biologie an der Universität Wien, Studienzweig Botanik März 2000 bis Juni 2002: Diplomarbeit bei Ao.Univ. Prof. Dr. Harald Greger in der Abteilung für Vergleichende und Ökologische Phytochemie am Institut für Botanik der Universität Wien 26. Juni 2002: Abschluss des Diplomstudiums Seit Jänner 2003: Doktoratsstudium unter der Betreuung von Ao.Univ. Prof. Dr. Michael Duchêne am Institut für Spezifische Prophylaxe und Tropenmedizin am Zentrum für Pathophysiologie, Infektiologie und Immunologie der Medizinischen Universität Wien
Familie:	8. Juli 2006: Hochzeit mit Wolfgang Tazreiter 9. September 2007: Geburt unseres Sohnes Jakob