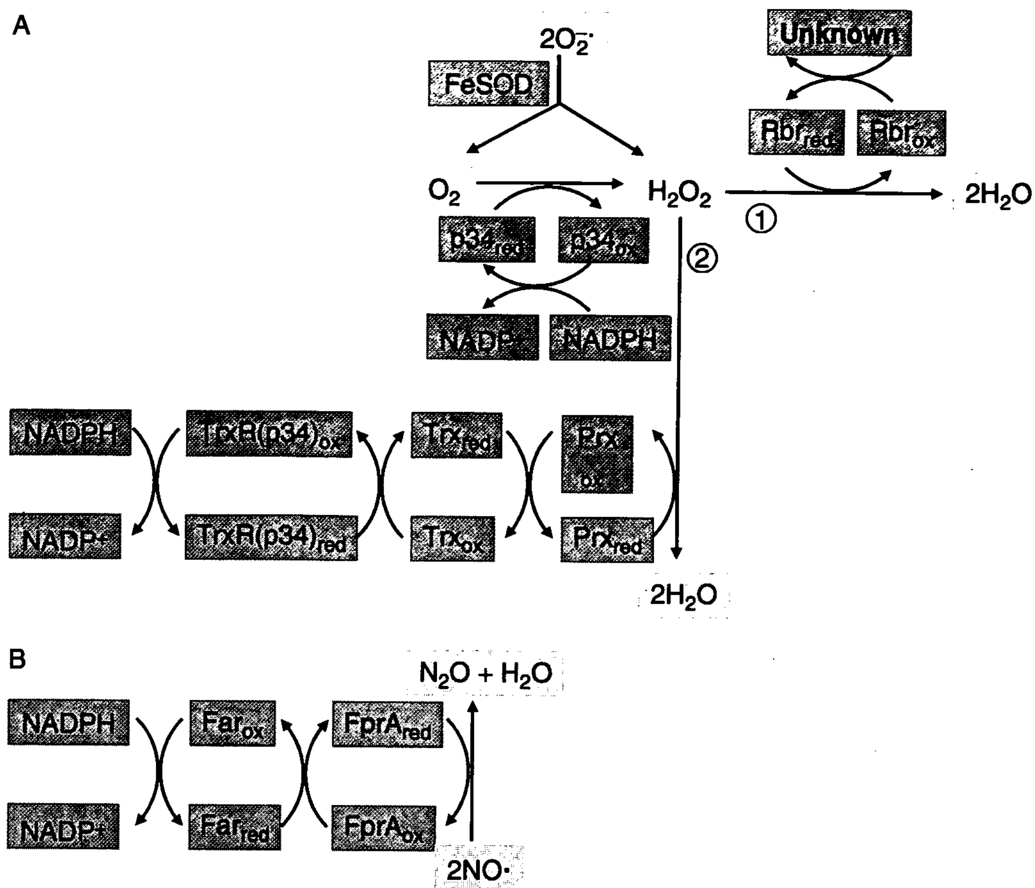

FIGURE 2.5 Predicted antioxidant system of *E. histolytica*. A. Superoxide radical anions are detoxified by an iron-containing superoxide dismutase (FeSOD). Molecular oxygen is reduced to hydrogen peroxide by a NADPH:flavin oxidoreductase (thioredoxin reductase, p34). Hydrogen peroxide is converted to water by rubrerythrin (Rbr). The nature of its redox partner is unknown. Hydrogen peroxide can also be converted to water via a classical thioredoxin redox system consisting of thioredoxin reductase (TrxR, p34), thioredoxin (Trx) and peroxiredoxin (Prx). B. Nitric oxide is reduced by an A-type flavoprotein (FprA) to nitrous oxide and water. For this reaction FprA receives electrons from NADH oxidase (Far).

aerobic cells (Sies, 1999). Instead, *E. histolytica* uses cysteine as its principal low molecular weight thiol (Ariyanayagam and Fairlamb, 1999; Fahey *et al.*, 1984; Nozaki *et al.*, 1999). As expected, coding sequences for enzymes that use glutathione as a cofactor, such as glutathione-S-transferase, glutathione-dependent peroxidase, glutathione reductase or glutaredoxin, are all absent from the *E. histolytica* genome. In addition, genes encoding catalases and peroxidases are also missing, as previously suggested (Sykes and Band, 1977; Weinbach and Diamond, 1974).

Other genes were identified that code for proteins involved in detoxification of H_2O_2, including one with homology to rubrerythrin. Rubrerythrin is a non-haeme iron protein thought to be able to reduce H_2O_2 as part of an oxidative stress protection system (Weinberg *et al.*, 2004). So far,

the nature of its redox partner is unknown in *E. histolytica*, and it remains to be determined whether protection against oxidative stress is indeed its main function. Another group of H_2O_2-detoxifying proteins identified in *E. histolytica* are peroxiredoxins. Peroxiredoxins are known from a wide variety of organisms. They are able to reduce H_2O_2 as well as peroxynitrite with the use of electrons provided by thiols. In addition to involvement in the detoxification of reactive oxygen species, peroxiredoxins seem to play a role in other processes such as signalling and differentiation (Hofmann *et al.*, 2002; Rhee *et al.*, 2005; Wood *et al.*, 2003a,b). All peroxiredoxins contain a conserved cysteine residue that undergoes a cycle of peroxide-dependent oxidation and thiol-dependent reduction during the reaction. The whole protein family can be divided into three classes based on the number and position of active site Cys residues (2-Cys, atypical 2-Cys and 1-Cys peroxiredoxins; Wood *et al.*, 2003a,b). In *E. histolytica* five different genes coding for peroxiredoxins were identified (Prx1–5). They all belong to the 2-Cys peroxiredoxin family. Four of them (Prx1–4) share 98% sequence identity and have an unusual N-terminal Cys-rich repeat (KECCKKECQEKECQEKECCC) of unknown function. In contrast, the fifth peroxiredoxin (Prx5) lacks the cysteine-rich N-terminal extension and shares only 30% identity with Prx1–4. Biochemical studies have shown that *E. histolytica* peroxiredoxins are able to detoxify H_2O_2 and cumene hydroperoxide (Bruchhaus *et al.*, 1997; Poole *et al.*, 1997). Moreover, up-regulation of peroxiredoxin and FeSOD was associated with metronidazole resistance in cultured *E. histolytica* trophozoites (Samarawickrema *et al.*, 1997; Wassmann *et al.*, 1999).

Reactions catalysed by peroxiredoxins are dependent on the presence of physiological thiols like thioredoxin (Rhee *et al.*, 2005; Wood *et al.*, 2003b). Thioredoxins are small proteins involved in thiol-redox processes (Holmgren, 2000). They contain two redox-active site cysteine residues of the motif CXXC (Watson *et al.*, 2004). Five genes coding for classical cytoplasmic thioredoxins were identified in the *E. histolytica* genome (Trx1–5). These thioredoxins have a length of 103–114 amino acids and share 25–47% sequence identity. Trx1–3 have identical active site motifs of the sequence WCGPC, whereas the active sites of Trx4 and Trx5 have the sequences SCPSC and WCKDC, respectively. In addition, another five thioredoxin-related proteins were identified (Trx6–10). All have a signal sequence of 15–19 amino acid residues and the active site motif WCGHC, which is also known from the active site of protein disulphide isomerases. However, in contrast to the latter group of enzymes, the *E. histolytica* thioredoxin-related molecules contain only one rather than two active-site motifs and only two of the proteins have an endoplasmic reticulum (ER) membrane retention signal (Freedman *et al.*, 2002). Thus it remains to be determined whether the thioredoxin-related molecules of *E. histolytica*

do constitute protein disulphide isomerases or whether they undertake other functions within the cell.

Thioredoxins are kept in the reduced state by the enzyme thioredoxin reductase, which catalyses the reduction of oxidised thioredoxin by NADPH using FAD and its redox-active disulphide (Nakamura, 2005). Two different genes with homology to thioredoxin reductases have been previously described from *E. histolytica* [thioredoxin reductase (TrxR) and NADPH:flavin oxidoreductase (p34)]. They share about 87% sequence identity and both contain the 2 conserved sequence motifs forming the FAD and NAD(P)H binding domains. p34 was shown to catalyse the NADPH-dependent reduction of oxygen to H_2O_2 as well as of disulphides like DTNB and cystine (Bruchhaus *et al.*, 1998; Lo and Reeves, 1980). Therefore, in addition to disulphide reductase activity the enzyme has H_2O_2-forming NADPH oxidase activity. It was also shown that p34 can transfer reducing equivalents to peroxiredoxin, converting the protein from its non-active, oxidised form back into its active, reduced form (Bruchhaus *et al.*, 1997). However, it is unlikely that peroxiredoxin is directly reduced by p34 *in vivo*. It is more likely that *E. histolytica* contains a classical thioredoxin redox system consisting of thioredoxin reductase, thioredoxin and peroxiredoxin (Poole *et al.*, 1997).

In addition to genes coding for proteins with homology to thioredoxin reductase, four other gene families were identified that encode various flavoproteins. One of these families includes four members that have between 53 and 61% sequence identity to A-type flavoproteins (flavorubredoxin/flavodiiron). A-type flavoproteins belong to a large family of enzymes that are widespread among anaerobic and facultatively anaerobic prokaryotes. In addition to bacteria, homologous genes are also found in the genomes of the pathogenic amitochondriate protistan parasites *T. vaginalis* and *G. intestinalis* (Andersson *et al.*, 2003; Sarti *et al.*, 2004). The A-type flavoproteins are made up of two independent structural modules. The N-terminal region forms a metallo-β-lactamase-like domain, containing a non-haeme di-iron site, whereas the C-terminal region is a flavodoxin-like domain, containing one FMN moiety. These enzymes have significant nitric oxide reductase activity (Gomes *et al.*, 2002; Sarti *et al.*, 2004). For *Escherichia coli* it is known that the nitric oxide reductase (FIRd) receives electrons from a NADH:oxidoreductase (FIRd-red). Consistent with that situation, the *E. histolytica* genome contains a gene encoding an NADH oxidase with 25% sequence identity to several bacterial FIRd-reds.

The three other *E. histolytica* gene families with homology to iron–sulphur flavoproteins (families B–D) are characterised by the presence of a flavodoxin-like domain forming a typical FMN binding site. Family B and family C consist of 3 members each, which share sequence identity of 42 and 46%, respectively. Family D consists of 2 members, which share

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