reported sequences. The rest of the ADH3 sequences fall into two groups of 3 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. The cytoplasm of trophozoites contains numerous glycogen granules which were first observed by electron microscopy (Rosenbaum and Wittner, 1970) and later characterised biochemically (Takeuchi *et al.*, 1977). A phosphorylase activity (EC 2.4.1.1), associated with the glycogen granules, generates glucose 1-phosphate from orthophosphate and various glucopolysaccharides, such as glycogen (Werries and Thurn, 1989). The genome contains at least 6 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.

A debranching enzyme has been purified and is able to degrade glycogen-limit dextrin in the presence of phosphorylase (Werries *et al.*, 1990). It exhibits activities of both amylo-1,6-glucosidase (EC 3.2.1.33) and 4-alpha-

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 which 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 two 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 open reading frame encoding a putative UTP:glucose-1-phosphate uridylyltransferase of 54.7 kDa and two 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 (a) 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 lateral gene transfer 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 4.1.1 (c)). As described above (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 (b) Anomerisation of aldoses

The 1-position in the pyranose form of aldoses has a hydroxyl group that can be in either the □- or □-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 (c) 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 (EC2.7.1.17). The latter is another bacterial-type sequence putatively acquired by lateral gene transfer.

4.1.3 (d) 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: three highly similar proteins of 73 kDa and four truncated versions.

4.2 Amino acid catabolism

4.2.1 General features

As discussed above, 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 4.1.1 (k)) 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 TargetP program (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 phosphoenolpyruvate carboxykinase, the malate generated can be oxidatively decarboxylated 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 proprionyl-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 4 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 2 genes encoding D-glycerate dehydrogenases, one of which (XP_648124) is among the weaker LGT candidates (see section 10). The genome also contains 2 genes encoding identical glycerate kinases. The enzyme has recently been characterised by Ali and Nozaki (unpublished).

4.2.3 Methionine, homocysteine and cysteine

Methionine □-lyase (EC 4.4.1.11) decomposes methionine to methanethiol (mercaptomethane), ammonia, and 2-oxobutanoate. In *E. histolytica*, two methionine □□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 □-lyase is essential since this parasite lacks the other known enzymes capable of destroying this toxic amino acid. In the genome, three open reading frames correspond to EhMGL1 and one to EhMGL2. So far, the only eukaryotes known to possess methionine □□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), i.e. cystathionine □-synthase and cystathionine □-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* 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 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 4.3 below). 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, never in eukaryotes, 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 to 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 above (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 (EC4.3.1.3), 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 necessary catabolism enzymes were detected except for a 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 above, *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 physiological significance of putrescine. However, the fate of putrescine is unknown as neither spermine nor spermidine was demonstrated in *E. histolytica*.

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 on 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 to 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 above 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; Ali *et al.*, 2004a; Nozaki *et al.*, 1998a; Nozaki *et al.*, 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 6-21 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 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 synthase 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 synthase 1 and 2. In contrast, all three serine acetyltransferase genes seem to be distinct, showing only 48-73% identity (Ali and Nozaki, unpublished). It was previously shown that cysteine synthase 1/2 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 5 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, yeast, 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 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 4.2.1 above).

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 acquired mainly from their food or from the human host, *E. histolytica* does have some capability for biosynthesis, as well as extending and remodeling lipids, and for attaching lipids to proteins.

4.5.1 Lipid biosynthetic capabilities

4.5.1(a) 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 sythesis, two molecules of C_{15} farnesyl diphosphate are dimerised to give C_{30} presqualene diphosphate (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 delta-isomerase which 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 open reading frames 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 alpha and beta 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) beta chain has recently been cloned and expressed together with the protein farnesyltransferase alpha chain (Makioka *et al.*, 2006). The heterodimeric molecule had protein geranylgeranyltransferase activity of unusually broad substrate specificity. The alpha and beta chains of the protein (Rab-) geranylgeranyltransferase II (EC 2.5.1.60) have also been cloned, as cDNAs (Kumagai *et al.* unpublished results).

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 (b) Fatty acid biosynthesis

E. histolytica encodes an unusual 138 kDa acetyl-CoA carboxylase with two 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 beta-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 revealeda number of genes indicating that the phospholipid metabolism could be more complex than expected.

4.5.2 (a) 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 to CDP-ethanolamine (EC 2.7.7.14) and CDP-choline (EC 2.7.7.15), respectively. The genome encodes two 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 8 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 and in glycosylphosphatidylinositol- (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 7 genes for this