Peer-reviewed version available at Metabolites 2019, 9, 115; doi:10.3390/metabo9060115

Euglena central metabolic pathways and their subcellular locations

Sahutchai Inwongwan[1], Nicholas J. Kruger[1], R. George Ratcliffe[1] and Ellis C. O'Neill[1,*]

[1] Department of Plant Sciences, University of Oxford, South Parks Road, Oxford, Oxfordshire, OX1 3RB, UK
* Correspondence: ellis.oneill@plants.ox.ac.uk; Tel.: +44 (0)1865 275 024

Received: date; Accepted: date; Published: date

Abstract: Euglenids are a group of algae of great interest for biotechnology, with a large and complex metabolic capability. To study the metabolic network, it is necessary to know the subcellular locations of the component enzymes, but despite a long history of research into *Euglena*, the subcellular locations of many major pathways are only poorly defined. *Euglena* is phylogenetically distant from other commonly studied algae, they have secondary plastids bounded by three membranes, and they can survive after destruction of their plastids. These unusual features make it difficult to assume that the subcellular organization of the metabolic network will be equivalent to that of other photosynthetic organisms. Moreover, we show here that the presence of the secondary chloroplast means that it is not possible to make reliable predictions of the subcellular locations of enzymes in *Euglena* using existing informatics tools. In order to generate a model of the central metabolic pathway operating in *Euglena* we analysed biochemical and proteomic information from a variety of sources to assess the subcellular location of relevant enzymes. We use these assignments to propose the compartmentation of the core metabolic pathways in *Euglena*, a prerequisite for the further study of the metabolic network of *Euglena*. This model of the metabolic network shows that, other than photosynthesis, all major pathways present in the chloroplast are duplicated elsewhere in the cell, and that several biosynthetic pathways confined to plastids in higher plants are localized elsewhere in *Euglena*. Our model demonstrates how this organism can synthesise all the metabolites required for growth from simple carbon inputs, and can survive in the absence of chloroplasts.

Keywords: *Euglena*; central metabolic pathway; subcellular location

1. Introduction

Euglenids, a group of unicellular flagellate alga, have long been studied for their biochemistry, physiology, anatomy and industrial potential due to the remarkable metabolic plasticity that allows them to grow in a wide range of conditions [1]. *Euglena* can harness energy heterotrophically, mixotrophically and photo-autotrophically, and its cultivation is relatively easy, fast and well established. Euglenids can be found in a broad range of ecological niches including fresh water, brackish water, snow, high and low pH conditions, aerobic and anaerobic environments [2]. *Euglena gracilis* is the most studied species of *Euglena* and is regarded as a useful model organism for studying cell biology and biochemistry [3,4]. *Euglena* were once considered one of the most ambiguous groups in terms of evolution and metabolic operation, due to the combination of both "plant-" and "animal-" like features [5]. They are now classified into the kingdom Excavata, superphylum Discoba, subphylum Euglenozoa [6,7] *Euglena* is one of the very few plastid-containing organisms for which

complete loss of the chloroplast is not lethal. Even the human parasitic apicomplexans retain their plastids for the synthesis of isoprenoids, fatty acids and heme, while in non-photosynthetic, parasitic plants plastids are necessary for aromatic amino acid biosynthesis and involved in starch synthesis [8]. Whilst these plastid-localised pathways can be targeted to kill such organisms, *Euglena* can survive complete loss of the plastid and the biochemical explanation for this remains to be established.

The genome of *E. gracilis* is estimated to be around 500 Mb in size, with large amounts of highly repetitive sequences [9], which leads to difficulty in genome sequencing and analysis. The structural complexity of the genome has arisen from a series of horizontal gene transfers and endosymbiosis events throughout its evolutionary history, which has caused difficulty in classifying Euglenids using modern molecular techniques [10]. A study of the distribution of the homologues of 2770 expressed sequence tags (ESTs) from *E. gracilis* has shown that Euglenids are most closely related to the kinetoplastids [11]. *Euglena* first split from the ancestral Euglenozoa, a eukaryotic protozoa that had mitochondria derived from an alphaproteobacterium, around a billion years ago [12]. After the endosymbiotic transfer of genes from a since-lost red algal endosymbiont to the nuclear genome [13], a eukaryotic green alga endosymbiont was incorporated [14], bringing many genes involved in the function and maintenance of the chloroplast. The transcriptome of *Euglena* suggests that many other genes were acquired from diverse distantly related species and the genetic control mechanisms in *Euglena* involve genes which are as sophisticated as those in higher eukaryotes [15].

Euglena is considered to be a promising organism for industrial application due to its ability to produce various nutrients and bioactive compounds, such as, proteins, polyunsaturated fatty acids, vitamin A, vitamin C and β-1,3-glucan [16-20]. The application of *Euglena* in environmental engineering has been studied for wastewater treatment systems, energy sources and bioindicators for environmental pollutants. *Euglena* sp. isolated from sewage treatment plants had higher nutrient removal capability and growth rate compared to other algae [21]. These results indicated that *Euglena* could be considered as a viable source for biofuel production from wastewaters.

There is no doubt that *E. gracilis* is an interesting organism in terms of its evolution, metabolic capacity and application and has thus been the subject of intense study. Due to its extraordinary metabolic capacity, investigating and understanding the *Euglena* metabolic network could help expand the applications of this organism and help to shed light on several mysteries of evolution and secondary endosymbiosis. Investigation of the metabolism of *Euglena*, requires the definition of the metabolic network, whether at genome scale for flux balance analysis, or at the level of core metabolism for metabolic flux analysis. This would allow the metabolic phenotype of the organism to be investigated in much the same way as in highly compartmented plant cells [22]. In organisms with complex evolution like *Euglena*, even though the central metabolic pathways are conserved, the characteristics and subcellular localisation of the enzymes involved in the pathway can differ. This is particularly true for *Euglena*, where the secondary chloroplast has a relatively recent evolutionary origin and a unique third plastid membrane, giving rise to a novel subcellular compartment in this intermembrane space.

Here, we provide an overview of the central metabolic pathways in *Euglena gracilis*, highlighting unique features. We assess the reported subcellular location of enzyme activities and proteins in *Euglena* and propose a model of the organisation of the central metabolic network.

2. Pathway localisation from sequence information

Even though *Euglena* has long been studied for its biotechnological potential, its genetic and metabolic capacities are poorly established due to the size and complexity of its genome. In the absence of an annotated genome sequence for any species of *Euglena*, transcriptome sequencing has been used as the preliminary alternative to genome structure analysis, with the aim of providing data on gene expression and regulation under different conditions [23,24].

2.1 Metabolic pathways in Euglena

The earliest reported extensive transcriptomic analysis of *E. gracilis* studied cells grown in dark and light conditions and illustrated the versatile metabolic capacity of *Euglena* [23]. All the core pathways of carbohydrate metabolism and photosynthesis were identified, including glycolysis, gluconeogenesis, the tricarboxylic acid cycle (TCA), the pentose phosphate pathway (PPP) and the Calvin cycle. In addition, the pathways for production of other major classes of compounds including carotenoids, thylakoid glycolipids, fatty acids and isoprenoids were also identified in the transcriptome. Besides the evidence for lipid, amino acid, carbohydrate and vitamin metabolism, the transcriptome also revealed the capacity of *E. gracilis* to produce multifunctional polydomain proteins that relate to those from both fungi and bacteria and may have been obtained by horizontal gene transfer during its evolution [15]. Furthermore, the transcriptome showed the capacity for polyketide and non-ribosomal peptide biosynthesis [25], along with capacities for using the pathways for vitamin C, vitamin E and glutathione metabolism to respond to stresses. A subsequent comparative study of the transcriptome of *E. gracilis* under aerobic and anaerobic conditions investigated the regulatory system of wax ester metabolism [24]. The metabolic network of *Euglena mutabilis* has been reconstructed using assembled transcript sequences and topology gap filling [26]. The initial draft network was incomplete with many missing reactions and could not simulate the heterotrophic growth of *E. mutabilis* in the dark [26], despite the long documented capacity of this species to do so. In combination, these studies demonstrate that the genome of *Euglena* has features in common with genomes from both phototrophic and heterotrophic organisms, and these features provide *Euglena* with the metabolic capacity to adapt to a wide variety of conditions. These studies also demonstrate that transcript abundance does not vary greatly under different growth conditions and does not correlate with protein abundance. Thus exploration of the metabolic capacity of *Euglena* using an exclusively transcriptomic approach might not be enough to understand pathway control.

2.2 Metabolic pathways in the Euglena plastid

The chloroplast genome of *E. gracilis* has been sequenced [27] and is substantially similar to higher plants in its gene content, although the structure and evolution is different [28]. As with other organisms, the acquisition of the plastid came with many gene loses and gene transfers from the endosymbiont to the host genome [29]. The expression level of plastid genes was found to respond to environmental stimuli [30] and the rate of protein synthesis by the *E. gracilis* plastid in the dark is extremely low compared to that in the light [31].

As in the primary plastids of other organisms, most of the *Euglena* secondary plastid proteome is encoded in the nuclear genome [32-34]. However, since the plastid of *Euglena* was acquired through secondary endosymbiosis of a photosynthetic eukaryote, its chloroplasts are surrounded by three membranes [35,36]. Thus, hundreds of plastidic proteins synthesized in the cytosol have to be transported through either three or four membranes to reach their destination in the plastid stroma

or the thylakoid lumen [37] and we have no knowledge of the metabolic capabilities of the unique intermembrane space, found in no other group of organisms.

2.3 Limitations of predicting the subcellular location of Euglena proteins

Most of the previously published studies of the subcellular compartmentation of *Euglena* enzymes have relied on subcellular fractionation of organelles and measurement of enzyme activity distributions. Very few studies have exploited complementary molecular techniques to investigate localisation in *Euglena*. In principle, eukaryotic protein subcellular location prediction tools, such as TargetP 1.1 [38], could be useful. To test this the protein sequences of selected marker enzymes (Table S1) with defined compartmentation were analysed using the subcellular location prediction work flow below (Figure 1). These included proteins known to be located in the chloroplast, mitochondria, cytosol or directed through the secretory pathway. The predicted full-length amino acid sequences of these marker proteins were deduced from the *E. gracilis* transcriptome [23]. TargetP 1.1 [38] was used to predict the subcellular localisation of all the matching *E. gracilis* protein sequences. Due to the potential presence of plant and non-plant targeting signals on *Euglena* proteins (arising from the complex evolutionary origin of *Euglena* genes), these analyses were conducted using both plant-based and non-plant-based prediction modes. Moreover, since transport of proteins into *Euglena* chloroplasts requires transit via the secretory pathway [37,39,40], any sequence that was predicted to contain a secretion signal based on the plant-based algorithm was subjected to extended analysis in which the signal sequence was removed and the prediction process repeated to establish the ultimate predicted location of the mature protein (Figure 1).

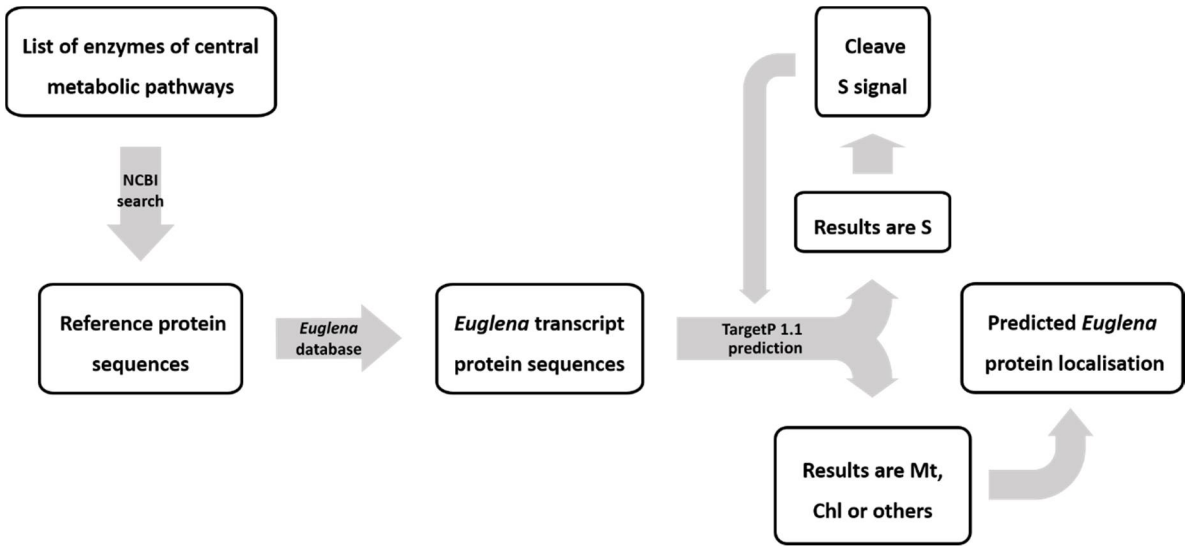

Figure 1. Subcellular location prediction work flow for *Euglena* proteins. Once transcripts are identified, their location is predicted using TargetP 1.1 in both plant and non-plant modes. If they are predicted to contain a secretion signal, this is removed to see if it is masking a second targeting signal, as is the case for targeting certain proteins to the chloroplast in *Euglena*. Abbreviations: Mt, mitochondria; Chl, chloroplast; others, cytosol; S, secretory pathway.

The TargetP analysis of the *E. gracilis* transcriptome showed that the algorithm accurately predicted most of the mitochondrial marker enzymes to be in mitochondria with high levels of reliability for both plant-based and non-plant-based criteria. In contrast, analysis of the chloroplast

marker enzymes, which was restricted to the plant-based prediction algorithm, resulted in almost none of the *Euglena* chloroplast marker proteins being predicted to be targeted to chloroplasts. The only exceptions were for one of the isoforms of fructose-bisphosphate aldolase, and one ribulose-bisphosphate carboxylase/oxygenase (small subunit) that was only predicted to be targeted to the chloroplast after removal of the secretory signal peptide. Many of the chloroplast marker enzymes were predicted to be in the mitochondria and the cytosol (Table S1). Proteins known to be in the cytosol were mostly not predicted to be targeted to any organelle or had only very weak mitochondrial or secretion targeting peptides. Proteins known to be in the Golgi, and which thus utilise the secretory pathway, were predominantly identified as being targeted for secretion with a high level of confidence, especially using the plant algorithm, although in some instances weak mitochondrial targeting was identified.

The limitations of the chloroplast targeting prediction of TargetP have been reported before [38]. The predictive power of TargetP 1.1 is based on the presence of N-terminal presequences, including chloroplast transit peptide (cTP), mitochondrial targeting peptide (mTP) or secretory pathway signal peptide (SP) [41]. However, the structure of cTP is not well characterized, especially in *Euglena*, and the prediction performance of chloroplast targeted proteins was reported to be less accurate than that for mitochondria, with occasional poor discrimination between mTP and cTP [42]. This lack of discrimination is partly due to some proteins using the same targeting sequence for both chloroplasts and mitochondria [38]. Thus, using TargetP to predict the location of proteins in *Euglena*, an evolutionarily complex organism with a secondary plastid, might not cover all the possible protein transport systems.

Apart from the evident limitations of TargetP as a protein localisation prediction tool in *Euglena*, protein targeting into chloroplasts of *Euglena* is likely to be inherently complex. In contrast to plants, the chloroplast of *Euglena* evolved from the secondary endosymbiosis, which led to the chloroplast being surrounded by three membranes [35,36]. A recent study of the *E. gracilis* chloroplast proteome identified three classes of chloroplast pre-protein based on targeted signal analysis. Class I and II proteins possess a bipartite topogenic signal (BTS), with class I proteins composed of a signal peptide (SP) followed by a stop-transfer signal (STS) and a transit peptide (TP), whilst class II proteins contain only an SP and TP [43]. The third class of chloroplast proteins was referred to as unclassified, with no signal sequence detected in the proteins. The transport mechanism used to import this unclassified category of proteins into the plastid remains unknown [40]. The transport of *Euglena* class I and II pre-proteins into the chloroplast involves the first step of co-translational transport into the endoplasmic reticulum (ER) lumen where the cleavage of the signal peptide occurs (Figure 2). The pre-proteins are subsequently transported to the outermost plastid membrane from the Golgi body via vesicles. However, the transport across the inner two membranes of the three-membrane-bound plastids in euglenophytes remains unclear [37,40,43]. The TOC/TIC-like pathway was believed to be involved in the inner membranes transport of the *Euglena* plastid due to the presence of plant-like targeting signal (TP) in the preproteins [44]. However, none of the TOC subunits have been detected in the transcriptome of *E. gracilis*, whereas homologues of several TIC subunits were identified [9]. A recent study analysing the structure of TP sequences in *E. gracilis* has suggested that it is possible for the TP to be recognised by the symbiont-derived ERAD-like machinery (SELMA) transport system, as is the case for diatoms [40,45].

Peer-reviewed version available at *Metabolites* **2019**, *9*, 115; doi:10.3390/metabo9060115

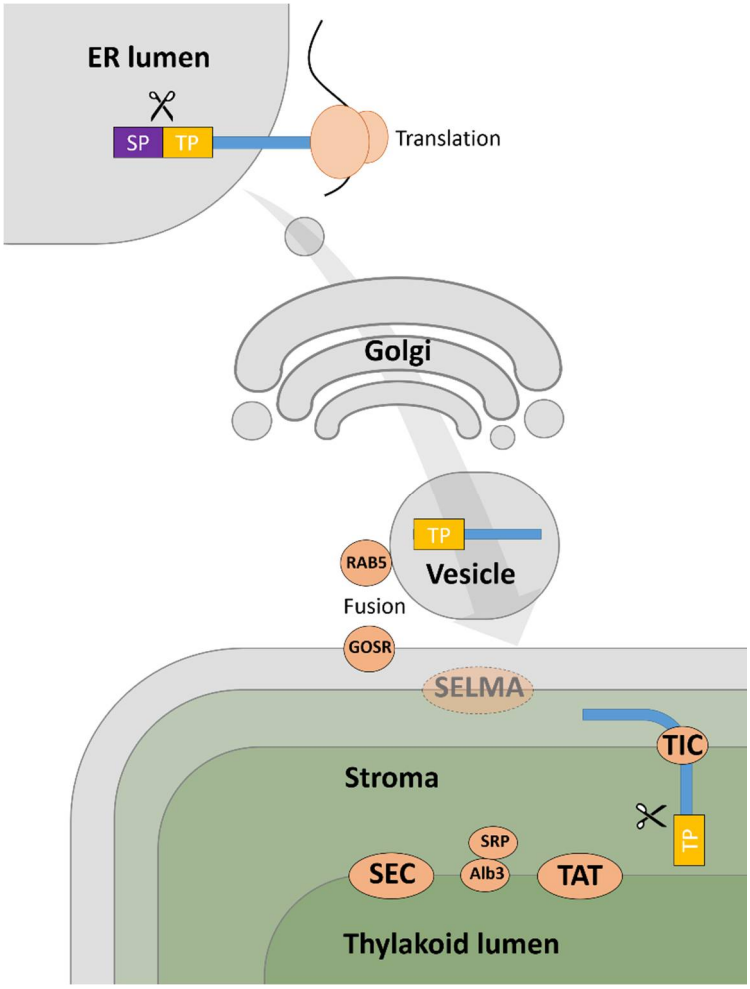

Figure 2. Protein transport into the secondary chloroplast of *Euglena*. Nuclear encoded chloroplast pre-proteins (blue strip) are synthesised into the lumen of the endoplasmic reticulum (ER) where the signal peptide (SP) is cleaved. Pre-proteins with transit peptides (TP) are subsequently transferred to the outermost chloroplast membrane through the Golgi body via vesicles. GOSR and RAB5 GTPase are proposed to mediate the fusion of the vesicle to the outermost membrane. After transport of proteins into the stroma, where the TP is removed, the mature protein can enter the thylakoid lumen via SEC, TAT or the Alb3/SRP pathway. This scheme only considers proteins possessing Class I and II targeting signals, as the transport of those with unclassified signals is not known [43].

It can be concluded that TargetP has limitations with predicting cTPs and does not specifically include protein targeting to the secondary plastid. Predicting chloroplast protein targeting in *Euglena* is likely to require more specific databases or algorithms, since the evolution of the *Euglena* chloroplast is different from that of plants. From this analysis, the prediction of mitochondria targeting with high reliability scores can be informative. However, due to the false predictions of chloroplast proteins to other locations, the prediction results cannot be fully relied upon and need to be carefully evaluated in conjunction with other evidence (Table S1, S2). To establish the subcellular localisation of the proteins in *Euglena*, enzymatic and biochemical analyses are unavoidable.

3. Pathway localisation from biochemical/proteomic information

3.1 Central metabolic pathways of Euglena

The central metabolic pathways are essential to all organisms, providing the precursors to other peripheral pathways in the cell, especially metabolites with carbon backbones that are derived from carbohydrate metabolism. In addition, under non-photosynthetic conditions, these pathways have a major role in producing the energy and reducing power for the cell. Carbohydrate metabolism pathways generally consist of glycolysis (Embden-Meyerhof-Parnas pathway), gluconeogenesis, the PPP, the Entner-Doudoroff (ED) pathway, and the TCA cycle. Notably, there is no evidence for the ED pathway in *Euglena.*

3.1.1 Glycolysis and gluconeogenesis

The intracellular distribution of the glycolytic enzymes in *Euglena* has been studied using fractionation in aqueous and non-aqueous media. This approach showed that most of the glycolytic enzymes are in the cytosol and that several of them are present in both the chloroplast and the cytosol [46,47]. By using sucrose density gradient centrifugation, it was found that phosphofructokinase, pyruvate kinase, triosephosphate isomerase and aldolase were present in the plastid cell fraction [48]. In addition, a recent proteomic study reported that several enzymes involved in glycolysis and gluconeogenesis were present in *Euglena* chloroplasts [40].

Hexose-phosphorylating enzymes

The activity of hexokinase (EC 2.7.1.1) was three times higher in *E. gracilis* grown on glucose than that on ethanol and acetate [49]. The activity of this enzyme in glucose media was also four times higher in heterotrophic cells than that in autotrophic cells [50]. *E. gracilis* was found to have glucokinase (EC 2.7.1.2) and fructokinase (EC 2.7.1.4) in different locations in both autotrophic and heterotrophic conditions. At 105,000 g separation, the glucokinase was present in the cell pellet while the fructokinase activity was only found in the supernatant [2,51]. Glucokinase is therefore concluded to be in organelles, whilst fructokinase is in the cytosol.

Phosphoglucoisomerase (EC 5.3.1.9)

The activity of this enzyme was detected in *E. longa* [2,52], although, the subcellular location has not been reported.

6-Phosphofructokinase (ATP-PFK, EC 2.7.1.11) and diphosphate--fructose-6-phosphate 1-phosphotransferase (PPi-PFK, EC 2.7.1.90)

In *E. gracilis*, 6-phosphofructokinase was reported to be located in both chloroplasts and the cytosol [48], while PPi-PFK was reported exclusively in the cytosol. During cell growth, the activity of PPi-PFK was 10-30 times higher than the activity of ATP-PFK [53].

Fructose bisphosphate aldolase (EC 4.1.2.13)

There are two classes of aldolase found in *Euglena*: class I is located in the chloroplast and proplastid, and class II is located in the cytosol [54]. Class I enzyme activity was detected in the chloroplast proteome [40] and the class II cytosolic enzyme was shown to be more active when the *E. gracilis* culture was grown in the dark and is presumed to play the main role in heterotrophic glycolysis and gluconeogenesis [55].

Glyceraldehyde 3-phosphate dehydrogenase (G3P) dehydrogenase, EC 1.2.1.12)

E. gracilis contains both NAD-linked and NADP-linked G3P dehydrogenase, which are found in different subcellular locations [54,56]. The NAD-linked enzyme showed higher activity in

heterotrophic conditions and was located in the cytosol. On the other hand, the NADP-linked enzyme was shown to be located in chloroplasts and had higher activity in autotrophic cells [57]. Only the NADP-linked enzyme was detected in the proteome of *E. gracilis* chloroplasts [40].

Triosephosphate isomerase (EC 5.3.1.1)

As with fructose bisphosphate aldolase, two types of the isomerase were identified in *E. gracilis* using enzymatic activity profiling [56]. Type A triosephosphate isomerase was reported to function in the chloroplasts and proplastids of *E. gracilis*, while type B enzymes were located in the cytosol [58]. Sequences matching triosephosphate isomerase could also be detected in the *E. gracilis* chloroplast proteome [40].

Phosphoglycerate kinase (EC 2.7.2.3)

The activity of phosphoglycerate kinase was reported in isolated *E. gracilis* chloroplasts [59] and the enzyme was detected in *E. gracilis* chloroplast proteome [40], although, the presence in other subcellular locations has not been investigated.

Phosphoglycerate mutase (EC 5.4.2.11)

No specific studies of the activity of this enzyme have been reported in *Euglena*. However, the enzyme was recently reported to be present in the *E. gracilis* chloroplast proteome [40].

Enolase (EC 4.2.1.11)

The activity of enolase was previously detected in *E. gracilis* but the subcellular location was not described [47,60]. N-terminal targeting peptide analysis of cDNA clones of *E. gracilis* suggested that enolase could be present in both the cytosol and the chloroplast [61]. However, as shown in section 2.3, it is difficult to predict protein targeting into the chloroplasts of *Euglena* and, furthermore, enolase was not found in the chloroplast proteome of *E. gracilis* [40].

Pyruvate kinase (EC 2.7.1.40)

The activity of pyruvate kinase in *E. gracilis* was shown to be highly active in cultures grown on glucose [62]. This enzyme was reported to be located in both proplastids and the cytosol of *E. gracilis*, however, the activity of this enzyme was not detected in the mature chloroplast [48].

Fructose-1,6-bisphosphatase (EC 3.1.3.11)

Fructose-1,6-bisphosphatase is involved in gluconeogenesis and has been reported from *Euglena* [48,53]. The cytosolic fructose-1,6-bisphosphatase in *E. gracilis* was detected and characterized [63]. Recently, the enzyme was reported in the *E. gracilis* chloroplast proteome [40], where it is presumably involved in the Calvin cycle.

3.1.2 Pentose phosphate pathway

Oxidative phase

In contrast to higher plants and green algae, all the enzymes of the oxidative arm of the pentose phosphate pathway in *E. gracilis* were reported to be present in the cytosol, but not the chloroplast. Using non-aqueous fractionation, it was found that two dehydrogenases of the oxidative pentose phosphate pathway were absent from the *E. gracilis* plastid [46] and these enzymes were not detected in the proteome of the *E. gracilis* chloroplast [40]. In separate studies, the activity of 6-phosphogluconate dehydrogenase (EC 1.1.1.44) was confirmed to be in the cytosol [47], and glucose-6-phosphate dehydrogenase (EC 1.1.1.49) was reported to be located in the cytosol [2,47,64-66] and has been used as a cytosolic marker enzyme [67]. On the other hand, a single glucose-6-phosphate dehydrogenase was detected in the chloroplast proteome. However, the proteome was reported to be moderately contaminated with protein from other organelles [40] and thus, subcellular location of

the enzyme will need further investigations to confirm its location. This enzyme is specific for NADP in *Euglena* and induced by glucose, with low activity detected under heterotrophic growth in the absence of glucose [62]. There has been no specific study of *Euglena* 6-phosphogluconolactonase (EC 3.1.1.31).

Non-oxidative phase

All the enzymes involved in the non-oxidative section of the pentose phosphate pathway have been detected in *Euglena* and most of the enzymes were reported to localize to the chloroplast [2,40]. The activity of ribose 5-phosphate isomerase (EC 5.3.1.6) was reported in isolated *E. gracilis* chloroplasts [68]. The subcellular location of pentose-5-phosphate-3-epimerase (EC 5.1.3.1) has not been reported, although the activity of this enzyme was detected in heterotrophic, autotrophic and mixotrophic growth conditions, along with the activity of transketolase (EC 2.2.1.1) [69] and transaldolase (EC 2.2.1.2) [56]. Non-aqueous separation techniques showed the presence of transaldolase in *Euglena* chloroplasts and proplastids [48].

Notably, there are two isoforms of each enzyme of the non-oxidative PPP in the *E. gracilis* transcriptome, except transketolase which has three. For three of these enzymes, only one isoform was identified in the chloroplast proteome [40], whereas neither isozyme of transaldolase could be detected. This suggests that the other isoforms are present in another location within the cell and the lack of any detectable targeting signal indicates this is likely to be the cytosol. However, extensive study of this pathway has not been reported and further investigation would be needed to confirm the operation of the pathway in the cytosol.

3.1.3 Anaplerotic pathway: dicarboxylic acid bypass

Malate dehydrogenase (NADP-specific oxaloacetate-decarboxylating, EC 1.1.1.40) in *Euglena* is located in the cytosol but not in mitochondria, and is specific for NADP and L-malate [2]. The NAD-specific malic enzyme (EC 1.1.1.39) can only be detected in *E. gracilis* cultured with D-malate [70]. Recently, a proteomic study detected malate dehydrogenase (EC 1.1.1.40) in *E. gracilis* chloroplasts [40]. The activity of this enzyme varied widely with light and carbon sources. NADP-specific malic enzyme has 55 times greater activity in heterotrophic cells than in autotrophic cells. This result suggests a physiological role in *Euglena* for these enzymes in providing NADPH for cytosolic fatty-acid synthesis in the dark [71,72].

Phospho*enol*pyruvate carboxylase (PEP carboxylase, EC 4.1.1.31) was shown to have multiple isozymes which were active in different light conditions. It has been reported that PEP carboxylase functions for CO_2 fixation in *E. gracilis* grown in the dark and under CO_2 limited conditions [73,74]. The activity of PEP carboxykinase in *E. gracilis* is specific for GTP rather than ATP [75]. PEP carboxylase and phospho*enol*pyruvate carboxykinase (PEP carboxykinase, EC 4.1.1.32) are discrete, separate enzymes in *E. gracilis* [76]. PEP carboxykinase was reported to be located exclusively in the cytosol and the enzyme could not be detected in cells grown under autotrophic conditions [77]. In addition, the activity of PEP carboxykinase was detected in *E. gracilis* cultured with acetate or ethanol, but not with glucose [71]. Pyruvate carboxylase (EC 6.4.1.1) was also reported to be located in the cytosol [78]. The activity of this enzyme was found in cells grown under heterotrophic culture fed with glucose, but not with acetate or in autotrophic cells [2].

3.1.4 TCA cycle

The reactions of the TCA cycle occur in the mitochondria of *Euglena* in common with all other eukaryotic organisms [2]. Most of the enzymes involved in the TCA cycle are predicted to target to the mitochondria with high reliability (Table S2), in line with previous studies on the localisation of the TCA cycle.

Pyruvate dehydrogenase (NADP+) (EC 1.2.1.51)

In contrast to most organisms, *E. gracilis* lacks a conventional NAD^+ pyruvate dehydrogenase complex, and instead exploits an $NADP^+$-dependent pyruvate dehydrogenase to produce acetyl-CoA from pyruvate. This enzyme has been detected in the mitochondrial fractions of *E. gracilis* [79-81].

Citrate synthase (EC 4.1.3.7)

Citrate synthase activity was detected in both particulate and soluble fractions from bleached *E. gracilis* [47], indicating that the enzyme is located in cytosol and other cell compartments. Testing the activity of this enzyme from different organelle suspensions showed the presence of this enzyme in both mitochondria and microbodies (glyoxysome-like particles) [82,83].

Aconitase (EC 4.2.1.3)

The activity of aconitase was detected in *E. gracilis* [84,85]. However, the subcellular location of this enzyme has apparently never been investigated.

Isocitrate dehydrogenase (NAD-specific EC 1.1.1.41, NADP-specific EC 1.1.1.42)

NAD- and NADP-specific isozymes of isocitrate dehydrogenase have been characterised from *Euglena*. The activity of NAD-specific isocitrate dehydrogenase was detected in mitochondria and cytosol of *E. longa* [47,52]. The NAD-specific isozyme was detected solely in the mitochondria of the streptomycin-bleached *E. gracilis* [83,86,87]. The NADP-specific isozyme was reported in both mitochondria and the cytosol, with the activity of the mitochondrial enzyme being about 25% of that in the cytosolic type [83,87].

2-oxoglutarate decarboxylase (EC 4.1.1.71)

E. gracilis contains a 2-oxoglutarate decarboxylase that is dependent on thiamin pyrophosphate, in contrast to the more common CoA-dependent 2-oxoglutarate dehydrogenase complex which was not detected [88]. The thiamine pyrophosphate dependent activity which coverts 2-oxoglutarate to succinic semialdehyde is located solely in mitochondria [89].

Succinic semialdehyde dehydrogenase (EC 1.2.1.16)

NAD- and NADP-specific succinate semialdehyde dehydrogenase were detected in *E. gracilis* and reported to be in the mitochondria [81,90].

Succinate dehydrogenase (EC 1.3.5.1)

As with other eukaryotes, the succinate dehydrogenase in *E. gracilis* is tightly bound to the inner membrane of mitochondria and has been used as a marker enzyme for mitochondria in *Euglena* [91]. [67,82,83,86].

Fumarase (EC 4.2.1.2)

Using cell fractionation and enzyme activity assays, fumarase is routinely detected solely in *E. gracilis* mitochondria [48,82,83,86] and is commonly used as a soluble mitochondrial marker enzyme [91].

Malate dehydrogenase (EC 1.1.1.37)

In *E. gracilis*, malate dehydrogenase is found in both mitochondria and cytosol. The cytosolic enzyme had three times higher activity in heterotrophically grown cells than in photoautotrophic cells, whereas the activity of the mitochondrial isoform was largely uninfluenced by variation in growth conditions [71]. *E. gracilis* contains two forms of malate dehydrogenase, NAD-

linked and NADP-linked isozymes. Unlike in higher plants, where the NADP-linked malate dehydrogenase is present exclusively in chloroplasts, in *E. gracilis* the majority (81-91%) of both NAD-linked and NADP-linked activity were located in the cytosol with a smaller proportion (13-16%) found in mitochondria. The activity of the NAD-linked isozyme was reported to be about three times higher than that of the NADP dependent isozyme [92,93].

3.1.5 Glyoxylate cycle

The glyoxylate cycle is a modified form of the TCA cycle that is found in plants, bacteria, protists and fungi. The cycle has an important role in provision of precursors for gluconeogenesis and allows the cell to use other respiratory substrates when sugars are not available [94]. The subcellular location of the glyoxylate cycle in *Euglena* under different conditions is poorly defined, with studies suggesting that the cycle operates in either mitochondria or discrete microbodies (glyoxysome-like particles). Notably, the presence of microbodies in *E. gracilis* was reported to vary under different conditions [95]. Following cell fractionation on sucrose density gradients, the activities of isocitrate lyase (EC 4.1.3.1) and malate synthase (EC 2.3.3.9), enzymes unique to the glyoxylate cycle, were found in the microbody fraction of *E. gracilis* grown on acetate [83,86]. In contrast, using similar cell fractionation techniques and immunocytochemical analysis, both isocitrate lyase and malate synthase were localised to mitochondria in *E. gracilis* grown on ethanol in which microbodies could not be detected [96].

3.1.6 C2 metabolism

Ethanol, which can readily diffuse into the cell, is first oxidized to acetaldehyde by alcohol dehydrogenase (EC 1.1.1.1), and the acetaldehyde is then oxidised by acetaldehyde dehydrogenase (EC 1.2.1.10) to produce acetate. Both enzymes are found in *E. gracilis* mitochondria [97-99]. Acetate is taken up either by simple diffusion or active transport through monocarboxylate transporters and is then converted to acetyl-CoA by acetyl-CoA synthetase (EC 6.2.1.1), also located in *E. gracilis* mitochondria [100], and then metabolized through the TCA cycle or channelled into the glyoxylate cycle.

3.2. Subcellular locations of biomass production

The composition of *Euglena* biomass is similar to that of many organisms, with storage carbohydrates, proteins and lipids predominating. The amounts of the different components varies substantially depending on the growth stage, from almost 10% dry weight wax esters [101] under anaerobic growth to over 80 % paramylon under aerobic conditions [102].

3.2.1 Carbohydrate biosynthesis

Unlike most other photosynthetic organisms, such as plants and green algae, *Euglena* stores carbohydrate in the form of a crystalline β-1,3-glucan, called paramylon, instead of starch, and the soluble disaccharide trehalose, instead of sucrose. *Euglena* has a wide range of enzymes involved in carbohydrate metabolism but it is difficult to predict their substrates and products from sequence alone [103].

Paramylon

Paramylon is synthesized from UDP-glucose [104] using the membrane bound paramylon synthetase (beta-1,3-glucan beta-glucosyltransferase, EC 2.4.1.34) that was identified in the *E. gracilis*

mitochondrial fraction through differential centrifugation [105]. Based on transmission electron microscopy, paramylon was synthesised in vesiculated mitochondrial related membrane complexes (chondriomes). The matrix of these vesicles was dense with paramylon granules and extended into the cytosol. The vesicles developed, resulting in the membrane-bound paramylon grains found in the cytosol [50,106,107]. The endo-1,3-β-glucanases (EC 3.2.1.6 and EC 3.2.1.39), exo-1,3-β-glucanases (EC 3.2.1.58) and 1,3-β-glucan phosphorylases (EC 2.4.1.97) involved in glucan metabolism have been characterized [108-110], though the subcellular locations of these enzymes have not been defined.

Trehalose

In *Euglena gracilis*, trehalose synthesis was reported to have a role in the adaptation to osmotic stress [111,112]. Trehalose biosynthesis involves a two-step process through the sequential action of trehalose-phosphate synthase (TPS, EC 2.4.1.15) and trehalose-phosphate phosphatase (TPP, EC 3.1.3.12). It was found that the activities of TPS and TPP could not be separated and so a TPS/TPP enzyme complex of about 250kDa was suggested to be responsible for trehalose synthesis in *E. gracilis* [113]. In *Arabidopsis*, the bulk of the TPP was reported to be cytosolic [114,115]. However, the subcellular localisation of the TPS/TPP complex in *Euglena* has not been investigated. Analysis of the chloroplast proteome of *E. gracilis* [40] shows no evidence of the TPS and TPP suggesting it is more likely that the TPS/TPP complex is located in the cytosol (or conceivably mitochondria) rather than in chloroplasts.

3.2.2 Amino acid biosynthesis

The pathways of amino acid biosynthesis in *Euglena* have been poorly investigated, especially with regard to their subcellular localisation. The recent evidence from the proteomic analysis of *Euglena* chloroplasts suggested that their capacity for synthesis of amino acids is extremely limited, in contrast to plant and algal chloroplasts, which is the major subcellular site for synthesis of various amino acids [40]. Here we present a summary of the likely subcellular localisation of amino acid biosynthesis in *Euglena*.

Glycine and serine (glycolate pathway associated)

Glycine and serine are synthesised from glyoxylate, an intermediate of photorespiration and gluconeogenesis. Glycolate dehydrogenase (EC 1.1.99.14), the starting enzyme of the glycolate pathway, was reported to be located in both mitochondria and microbodies in *E. gracilis* [86]. Glutamate:glyoxylate aminotransferase (EC 2.6.1.4), which adds the amino group to form glycine [116], is found in mitochondria, the cytosol and microbodies [86,117]. A small proportion of the glyoxylate is converted to glycine by glutamate:glyxoylate aminotransferase in mitochondria, and the majority is split into CO_2 and formate. As in higher plants, the formate is then used to produce serine through condensation with glycine [118,119]. Folate coenzymes, which are involved in this C1 transfer, were reported to be located largely in the cytosol [87]. Glycine can also be produced through the cleavage of threonine by threonine aldolase (EC 4.1.2.5/48) [120] though the subcellular location of this activity has not been reported. The enzymes involved in serine biosynthesis from 3-phosphoglycerate have not been studied in detail in *Euglena*. However, recently, phosphoserine aminotransferase was identified in the *E. gracilis* chloroplast proteome, indicating the possibility of a plastidic serine biosynthesis pathway [40].

Methionine, cysteine and threonine

The activity of cobalamin-dependent methionine synthase (EC 2.1.1.13), producing methionine from N5-methyltetrahydrofolate and homocysteine, was reported to be distributed between the

cytosol (68.9%), chloroplast (18.4%) and mitochondria (9.5%) of phototrophic cells. The more stable, Mg-dependent, variant was reported to be found only in the cytosol [121]. Cysteine synthesis in *Euglena* has not been investigated in detail and the subcellular localisations of the enzymes associated with this synthesis pathway have not been elucidated. Two enzymes involved in the synthesis of cysteine (serine O-acetyltransferase and cysteine synthase) were reported in the *E. gracilis* transcriptome [120] and isoform A of cysteine synthase was detected in the *E. gracilis* chloroplast proteome [40]. Threonine is synthesized from aspartate via homoserine. Five enzymes involved in threonine biosynthesis in *E. gracilis* were reported to be expressed in different growth conditions [120]. However, the localisations of the enzymes involved in the synthesis pathway have not been elucidated.

Aromatic amino acids (phenylalanine, tyrosine and tryptophan)

Chorismate, the precursor to aromatic amino acids, is synthesised from D-erythrose 4-phosphate and phosphoenolpyruvate by the shikimate pathway in seven steps. Five reactions can be catalysed either by separate enzymes, as in plants [127], or by a pentafunctional enzyme, as in fungi [128]. There is evidence for both of these in the *E. gracilis* transcriptome [27].

In algal and plant cells, the aromatic amino acids are produced exclusively in the chloroplast but the protein analysis of isolated organelles of *E. gracilis* suggested that the shikimate pathway occurs in both the plastid and cytosol [122]. The preferred pathway depends on the growth conditions, with the cytosolic pathway used in the dark and the plastidic pathway in the light [122,123].

Chorismate is then converted into tyrosine and phenylalanine, via prephenate by dehydration, dehydrogenation and transamination. The enzymes catalysing these reactions are present in *E. gracilis* as unusual domain fusions, also found in thermophilic bacteria [23]. Tryptophan is synthesised from chorismate by a series of reactions via anthranilate. In *E. gracilis* all four reactions are carried out by a unique fusion protein rather than a series of separate enzymes, as in other organisms [15,120].

Together the data suggest that aromatic amino acid biosynthesis in *Euglena* is carried out by a combination of plant-, bacteria- and fungi-like enzymes, as well as unique proteins. The evidence suggests that these pathways are not exclusively located in the plastid, unlike in plants, supporting the dispensability of the plastid for their biosynthesis, as is known to be the case for chloroplast-bleached *Euglena* growing on simple media with a simple carbon source.

Branched-chain amino acids (valine, isoleucine and leucine)

Pyruvate and α-ketobutyrate are the precursors for valine, leucine and isoleucine biosynthesis in *Euglena*, as in other organisms [124]. In *E. gracilis*, α-ketobutyrate is synthesized by the action of two threonine dehydratases (EC 4.3.1.19 and EC 4.3.1.17) that are located in the cytosol [125]. The subsequent steps are catalysed by acetolactate synthase, dihydroxy-acid reductoisomerase, and branched-amino-acid aminotransferase, all of which are located in the mitochondria [124], suggesting the biosynthesis of branched-chain amino acids is located in mitochondria.

Arginine and proline

Arginine is synthesised by the sequential transfer of nitrogen on to glutamate semialdehyde. Arginine biosynthesis is likely to occur mostly in the cytosol in *Euglena*, as the majority of ornithine carbomyltransferase is located in cytosol and smaller portion in mitochondria [2]. Arginine metabolism follows the arginine dehydrolase pathway in which arginine is converted into citrulline and then ornithine, which occurs in the mitochondria [126]. Proline synthesis in *Euglena* has not been investigated. However, proline metabolism is tightly associated with arginine metabolism as

ornithine is the precursor of proline synthesis [127], suggesting that synthesis is likely to be located in the cytosol or mitochondria.

Lysine

Bacteria, plants, and algae synthesize lysine via the diaminopimelate (DAP) pathway, using aspartate and pyruvate as the precursors. On the other hand, fungi, synthesize lysine through the α -aminoadipate (AAA) pathway, which uses 2-oxoglutarate and acetyl-CoA [128,129]. Several enzymes involved in AAA pathway were detected in *Euglena*, including homocitrate synthase (EC 2.3.3.14), homoaconitate hydratase (EC 4.2.1.36) and homoisocitrate dehydrogenase (EC 1.1.1.87) [120]. However, the subcellular location of the AAA pathway has not been reported.

Histidine

Histidinol dehydrogenase, the enzyme catalysing the final step of histidine biosynthesis, has been detected in *E. gracilis* [120,130]. No other enzyme involved in this process was detected and the subcellular localisation of the enzymes involved in histidine biosynthesis have not been investigated.

Glutamate, glutamine, alanine, aspartate and asparagine

Aminotransferases and dehydrogenases play the main role in the synthesis of glutamate, alanine and aspartate from organic acids. For glutamate, the aspartate aminotransferase (glutamate: oxaloacetate aminotransferase) is present in mitochondria, chloroplasts, microbodies and cytosol, and was shown to be more active in dark growth conditions [82,86]. NADP-specific glutamate dehydrogenase was reported to be located solely in the cytosol of *E. gracilis*, instead of the mitochondria as in other organisms [131]. Similarly, glutamate synthase was reported to be localised to the cytosol in both wild-type and streptomycin-bleached *E. gracilis* strains [132]. Glutamine is synthesized from glutamate using glutamine synthetase, but the properties of this enzyme have not been studied in *Euglena* [133]. Asparagine synthetase, the enzyme that converts aspartate to asparagine, has not been reported from *Euglena.* The activities of alanine aminotransferase and alanine dehydrogenase were detected in *E. gracilis*, but the localisation of these enzymes has not been described [2,134,135].

Tetrapyrrole biosynthesis

Tetrapyrrole, the core of heme and chlorophyll, is synthesised from δ-aminolevulinic acid (ALA). Heterotrophs tend to synthesize ALA from glycine and succinyl-CoA via the Shemin pathway in the mitochondia [136], whilst photoautotrophs make ALA from glutamate in the C5 pathway, located in the chloroplast [137]. *E. gracilis* is known to utilise both routes [138], and the transcriptome shows a bacterial-derived Shemin pathway and a green algae-related C5 pathway, presumably obtained with the chloroplast [23]. These have been identified in the mitochondria and chloroplasts of *E. gracilis* respectively [139]. This again supports the multiple locations of core metabolic pathways that are plastid localised in other photosynthetic organisms.

3.2.3 Lipid biosynthesis

The subcellular locations of the enzymes involved in lipid metabolism in *Euglena* are poorly investigated. As in other organisms *Euglena* produces the lipid building block malonyl-CoA from CO_2 and acetyl-CoA using acetyl-CoA carboxylase, which forms a multienzyme complex with phosphoenolpyruvate carboxylase and malate dehydrogenase in the cytosol [140]. Malonyl-CoA is then used to synthesise fatty acid using fatty acid synthases (FAS), of which three types have been reported in *E. gracilis*. FAS I and FAS III were reported to function in heterotrophic growth conditions. The properties of FAS III has not been investigated in detail. The structure of FAS I was similar to

yeast and mammalian enzymes, and was located in cytosol [2]. On the other hand, FAS II resembled the plant and bacterial enzymes, and was located in the chloroplasts of *E. gracilis* [19]. In addition to these three types of FAS, a fatty acid biosynthesis system was found in the mitochondria of *Euglena* that is involved in wax-ester synthesis [19].

4. Discussion

By combining multiple strands of evidence, including biochemical, proteomic and bioinformatic data, we propose a model for the subcellular localisation of the reactions of the network of central carbon metabolism in *E. gracilis* (Figure 3). Many of these pathways are found in similar subcellular locations to those in other, well-characterised organisms. Glycolysis, which catalyses the initial breakdown of sugars produced by photosynthesis or absorbed from the medium, is present in the cytosol and plastids, as commonly found in green plants. The products of this pathway feed into the TCA cycle, which is mitochondrial, as in other eukaryotes. The enzymes commonly associated with microbodies in higher plants are additionally present in the mitochondria, and it is often difficult to separate these two groups of organelles in *Euglena*. The site of synthesis of many amino acids is unclear, though several appear to be synthesised in the mitochondria from TCA cycle intermediates. Lipids can be made in several cellular compartments, though for different purposes, such as the mitochondrial lipids which are directed towards wax ester biosynthesis and plastid lipids that are used to make photosynthetic glycolipids.

The location of many metabolic processes in *Euglena* differs substantially from those found in other photosynthetic organisms. For instance, in *Euglena* the complete PPP is present in the cytosol, with a duplicated non-oxidative phase present in the plastid. A plant-like pathway for aromatic amino acid biosynthesis is present in the plastids. However, unlike plants, in *Euglena* an additional pathway, similar to that found in fungi, is located in the cytosol. Tetrapyrroles, essential prosthetic groups of both the respiratory and photosynthetic electron transport chain proteins, are synthesised in both the chloroplast and mitochondria in *Euglena*.

Overall, these results indicate that, aside from the reactions of photosynthesis, all the metabolic pathways found in the *Euglena* plastid are also found elsewhere in the cell. This includes the biosynthesis of isoprenoids, for which two pathways are found in other plastid-containing organisms, the methylerythritol phosphate pathway found in the plastids and the mevalonate pathway in the cytosol. Although we have not found evidence for the location of these pathways in *Euglena*, the methylerythritol phosphate pathway only contributes to carotenoid biosynthesis in *E. gracilis*, and phytol is instead made by the mevalonate pathway [141], unlike in other studied organisms. The unusual and well-established ability of *E. gracilis* to survive on a simple carbon source when their chloroplasts have been destroyed can be rationalised from the subcellular localisation and duplication of these various critical pathways.

The complicated evolutionary history of *Euglena* means it is not trivial to predict the likely subcellular locations of the various metabolic pathways, or to decide whether the pathways will be similar to those in free-living heterotrophs, or plants, or be entirely different. Precise information is missing for some biosynthesis pathways and the lack of understanding of *Euglena* chloroplast protein targeting restricts the prediction of the subcellular location of some *Euglena* proteins. Despite these limitations, overall, the model is similar to plants and green algae, but has some important differences. The development of this model will lead to the ability to predict the metabolic phenotypes of *Euglena* under various growth conditions.

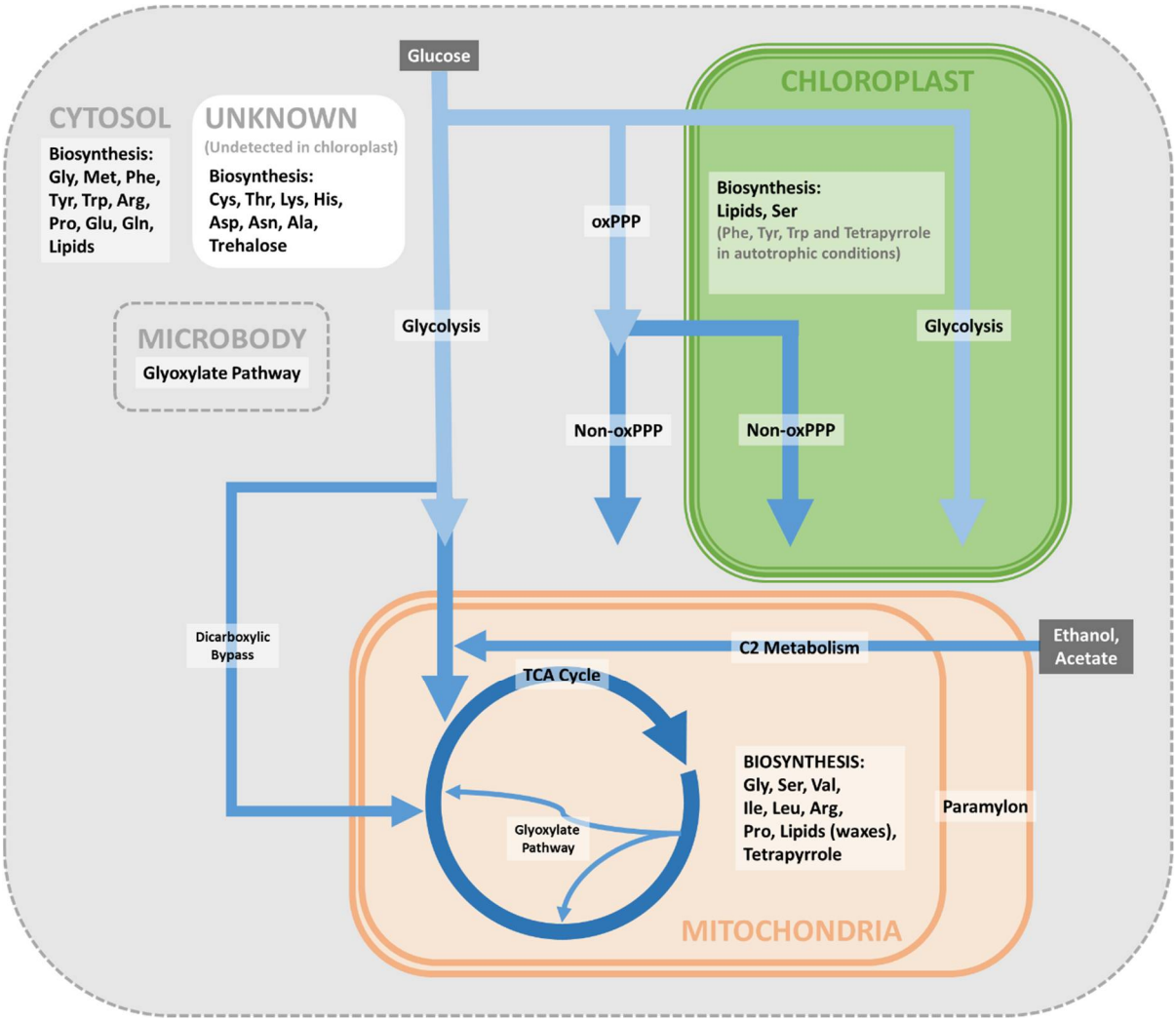

Figure 3. Proposed distribution of central metabolic pathways in *Euglena*. Abbreviations: oxPPP, oxidative pentose phosphate pathway; Non-oxPPP, non-oxidative pentose phosphate pathway

5. Conclusions

The subcellular compartmentation of many major metabolic pathways has been intensively studied in yeast and in plants. For many, more distantly related organisms, most information is typically inferred by extrapolation from these thoroughly examined species. Drawing on a range of *Euglena* biochemical and proteomic data, we propose a model for the organisation of central metabolism in *E. gracilis*. These analyses reveal unique features of this alga that diverge significantly from expectations derived from well-studied organisms. The most striking difference is the duplication within *Euglena* of various biosynthetic pathways solely present in the plastids of plants, contributing to the ability of *Euglena* to lose its plastid entirely and survive on simple carbon sources. We propose that this is due to the requirement of the heterotrophic ancestor to synthesise all necessary cellular components before the acquisition of the secondary plastid. In this context, it seems likely that the plastid pathways are duplicating pathways that were originally present in the Euglenid progenitor.

Supplementary Materials: The following are available online at www.mdpi.com/xxx/s1, Table S1: Subcellular location prediction of *E. gracilis* marker proteins using TargetP 1.1, Table S2: TargetP 1.1 subcellular location prediction of *E. gracilis* metabolic pathway components.

Author Contributions: Conceptualization, S.I., N.J.K., R.G.R. and E.C.O.; data gathering and analysis, S.I.; visualization, S.I.; writing—review and editing, S.I., N.J.K., R.G.R. and E.C.O.

Funding: The Development and Promotion of Science and Technology Talents Project (Royal Government of Thailand scholarship) funded this research. E.C.O. is supported by a Violette and Samuel Glasstone Fellowship.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Einicker-Lamas, M.; Mezian, G.A.; Fernandes, T.B.; Silva, F.L.S.; Guerra, F.; Miranda, K.; Attias, M.; Oliveira, M.M. *Euglena gracilis* as a model for the study of Cu^{2+} and Zn^{2+} toxicity and accumulation in eukaryotic cells. *Environ. Pollut.* **2002**, *120*, 779-786.
2. Kitaoka, S.; Nakano, Y.; Miyatake, K.; Yokota, A. Enzymes and their functional location In *The Biology of Euglena*, Buetow, D.E., Ed. Academic Press Limited: London, **1989**; Vol. 4, pp 2-119.
3. Ruiz, L.B.; Rocchetta, I.; Ferreira, V.d.S.; Conforti, V. Isolation, culture and characterization of a new strain of *Euglena gracilis*. *Phycol. Res.* **2004**, *52*, 168-173.
4. dos Santos Ferreira, V.; Rocchetta, I.; Conforti, V.; Bench, S.; Feldman, R.; Levin, M.J. Gene expression patterns in *Euglena gracilis*: Insights into the cellular response to environmental stress. *Gene* **2007**, *389*, 136-145.
5. Wolken, J.J. *Euglena; an Experimental Organism for Biochemical and Biophysical Studies*. Institute of Microbiology, Rutgers: New Brunswick, **1961**; p 173 p.
6. Adl, S.M.; Simpson, A.G.B.; Lane, C.E.; Lukeš, J.; Bass, D.; Bowser, S.S.; Brown, M.W.; Burki, F.; Dunthorn, M.; Hampl, V., *et al.* The revised classification of eukaryotes. *J. Eukaryotic Microbiol.* **2012**, *59*, 429-514.
7. Keeling, P.J. Diversity and evolutionary history of plastids and their hosts. *Am. J. Bot.* **2004**, *91*, 1481-1493.
8. Sato, S. The apicomplexan plastid and its evolution. *Cell. Mol. Life Sci.* **2011**, *68*, 1285-1296.
9. Ebenezer, T.E.; Zoltner, M.; Burrell, A.; Nenarokova, A.; Novák Vanclová, A.M.G.; Prasad, B.; Soukal, P.; Santana-Molina, C.; O'Neill, E.; Nankissoor, N.N., *et al.* Transcriptome, proteome and draft genome of *Euglena gracilis*. *BMC Biology* **2019**, *17*, 11.
10. Linton, E.W.; Karnkowska-Ishikawa, A.; Kim, J.I.; Shin, W.; Bennett, M.S.; Kwiatowski, J.; Zakryś, B.; Triemer, R.E. Reconstructing euglenoid evolutionary relationships using three genes: Nuclear SSU and LSU, and chloroplast SSU rDNA sequences and the description of *Euglenaria* gen. nov. (Euglenophyta). *Protist* **2010**, *161*, 603-619.
11. Ahmadinejad, N.; Dagan, T.; Martin, W. Genome history in the symbiotic hybrid *Euglena gracilis*. *Gene* **2007**, *402*, 35-39.
12. Parfrey, L.W.; Lahr, D.J.G.; Knoll, A.H.; Katz, L.A. Estimating the timing of early eukaryotic diversification with multigene molecular clocks. *Proc. Natl. Acad. Sci.* **2011**, *108*, 13624-13629.
13. Maruyama, S.; Suzaki, T.; Weber, A.P.; Archibald, J.M.; Nozaki, H. Eukaryote-to-eukaryote gene transfer gives rise to genome mosaicism in euglenids. *BMC Evol. Biol.* **2011**, *11*, 105.

14. Martin, W.; Somerville, C.C.; Loiseaux-de Goer, S. Molecular phylogenies of plastid origins and algal evolution. *Journal of Molecular Evolution* **1992**, *35*, 385-404.
15. O'Neill, E.C.; Trick, M.; Henrissat, B.; Field, R.A. *Euglena* in time: Evolution, control of central metabolic processes and multi-domain proteins in carbohydrate and natural product biochemistry. *Perspectives in Science* **2015**, *6*, 84-93.
16. Kusmic, C.; Barsacchi, R.; Barsanti, L.; Gualtieri, P.; Passarelli, V. *Euglena gracilis* as source of the antioxidant vitamin E effects of culture conditions in the wild strain and in the natural mutant WZSL. *J. Appl. Phycol.* **1998**, *10*, 555-559.
17. Ogbonna, J.C.; Tomiyama, S.; Tanaka, H. Heterotrophic cultivation of *Euglena gracilis* Z for efficient production of alpha-tocopherol. *J. Appl. Phycol.* **1998**, *10*, 67-74.
18. Takeyama, H.; Kanamaru, A.; Yoshino, Y.; Kakuta, H.; Kawamura, Y.; Matsunaga, T. Production of antioxidant vitamins, beta-carotene, vitamin C, and vitamin E, by two-step culture of *Euglena gracilis* Z. *Biotechnol. Bioeng.* **1997**, *53*, 185-190.
19. Inui, H.; Miyatake, K.; Nakano, Y.; Kitaoka, S. Wax ester fermentation in *Euglena gracilis*. *Febs Letters* **1982**, *150*, 89-93.
20. Krajčovič, J.; Matej, V.; Schwartzbach, S.D. Euglenoid flagellates: A multifaceted biotechnology platform. *Journal of Biotechnology* **2015**, *202*, 135-145.
21. Mahapatra, D.M.; Chanakya, H.N.; Ramachandra, T.V. *Euglena* sp. as a suitable source of lipids for potential use as biofuel and sustainable wastewater treatment. *J. Appl. Phycol.* **2013**, *25*, 855-865.
22. Kruger, N.J.; Ratcliffe, R.G. Fluxes through plant metabolic networks: Measurements, predictions, insights and challenges. *Biochem. J.* **2015**, *465*, 27-38.
23. O'Neill, E.C.; Trick, M.; Hill, L.; Rejzek, M.; Dusi, R.G.; Hamilton, C.J.; Zimba, P.V.; Henrissat, B.; Field, R.A. The transcriptome of *Euglena gracilis* reveals unexpected metabolic capabilities for carbohydrate and natural product biochemistry. *Molecular Biosystems* **2015**, *11*, 2808-2820.
24. Yoshida, Y.; Tomiyama, T.; Maruta, T.; Tomita, M.; Ishikawa, T.; Arakawa, K. De novo assembly and comparative transcriptome analysis of *Euglena gracilis* in response to anaerobic conditions. *BMC genomics* **2016**, *17*, 182.
25. O'Neill, E.C.; Saalbach, G.; Field, R.A. Chapter five - gene discovery for synthetic biology: Exploring the novel natural product biosynthetic capacity of eukaryotic microalgae. In *Methods in Enzymology*, O'Connor, S.E., Ed. Academic Press: **2016**; Vol. 576, pp 99-120.
26. Prigent, S.; Frioux, C.; Dittami, S.M.; Thiele, S.; Larhlimi, A.; Collet, G.; Gutknecht, F.; Got, J.; Eveillard, D.; Bourdon, J., *et al.* Meneco, a topology-based gap-filling tool applicable to degraded genome-wide metabolic networks. *PLoS Comput. Biol.* **2017**, *13*, e1005276.
27. Hallick, R.B.; Hong, L.; Drager, R.G.; Favreau, M.R.; Monfort, A.; Orsat, B.; Spielmann, A.; Stutz, E. Complete sequence of *Euglena gracilis* chloroplast DNA. *Nucleic Acids Research* **1993**, *21*, 3537-3544.
28. Turmel, M.; Gagnon, M.-C.; O'Kelly, C.J.; Otis, C.; Lemieux, C. The chloroplast genomes of the green algae *Pyramimonas*, *Monomastix*, and *Pycnococcus* shed new light on the evolutionary history of prasinophytes and the origin of the secondary chloroplasts of euglenids. *Mol. Biol. Evol.* **2009**, *26*, 631-648.
29. Martin, W.; Herrmann, R.G. Gene transfer from organelles to the nucleus: How much, what happens, and why? *Plant Physiol.* **1998**, *118*, 9-17.
30. Geimer, S.; Belicová, A.; Legen, J.; Sláviková, S.; Herrmann, R.G.; Krajčovič, J. Transcriptome analysis of the *Euglena gracilis* plastid chromosome. *Curr. Genet.* **2009**, *55*, 425-438.

31. Miller, M.E.; Jurgenson, J.E.; Reardon, E.M.; Price, C.A. Plastid translation in organello and in vitro during light-induced development in *Euglena*. *J. Biol. Chem.* **1983**, *258*, 14478-14484.
32. Armbrust, E.V.; Berges, J.A.; Bowler, C.; Green, B.R.; Martinez, D.; Putnam, N.H.; Zhou, S.; Allen, A.E.; Apt, K.E.; Bechner, M., *et al.* The genome of the diatom Thalassiosira pseudonana: ecology, evolution, and metabolism. *Science* **2004**, *306*, 79-86.
33. Douglas, S.; Zauner, S.; Fraunholz, M.; Beaton, M.; Penny, S.; Deng, L.-T.; Wu, X.; Reith, M.; Cavalier-Smith, T.; Maier, U.-G. The highly reduced genome of an enslaved algal nucleus. *Nature* **2001**, *410*, 1091-1096.
34. Gleason, F.K.; Hogenkamp, H.P.C. Ribonucleotide reductase from *Euglena gracilis*, a deoxyadenosylcobalamin-dependent enzyme. *J. Biol. Chem.* **1970**, *245*, 4894-4899.
35. Bachvaroff, T.R.; Sanchez Puerta, M.V.; Delwiche, C.F. Chlorophyll c–containing plastid relationships based on analyses of a multigene data set with all four chromalveolate lineages. *Mol. Biol. Evol.* **2005**, *22*, 1772-1782.
36. Yoon, H.S.; Hackett, J.D.; Bhattacharya, D. A single origin of the peridinin- and fucoxanthin-containing plastids in dinoflagellates through tertiary endosymbiosis. *Proc. Natl. Acad. Sci.* **2002**, *99*, 11724.
37. Bolte, K.; Bullmann, L.; Hempel, F.; Bozarth, A.; Zauner, S.; Maier, U.-G. Protein targeting into secondary plastids1. *J. Eukaryotic Microbiol.* **2009**, *56*, 9-15.
38. Emanuelsson, O.; Brunak, S.; von Heijne, G.; Nielsen, H. Locating proteins in the cell using TargetP, SignalP and related tools. *Nat. Protoc.* **2007**, *2*, 953-971.
39. Durnford, D.G.; Schwartzbach, S.D. Protein targeting to the plastid of *Euglena*. In *Euglena: Biochemistry, Cell and Molecular Biology*, Schwartzbach, S.D.; Shigeoka, S., Eds. Springer International Publishing: Cham, **2017**; pp 183-205.
40. Novák Vanclová, A.M.G.; Zoltner, M.; Kelly, S.; Soukal, P.; Záhonová, K.; Füssy, Z.; Ebenezer, T.E.; Lacová Dobáková, E.; Eliáš, M.; Lukeš, J., *et al.* Proteome of the secondary plastid of *Euglena gracilis* reveals metabolic quirks and colourful history. *bioRxiv* **2019**, 573709.
41. Darabi, M.; Seddigh, S. Computational study of biochemical properties of ribulose-1,5-bisphosphate carboxylase/oxygenase (RuBisCO) enzyme in C3 plants. *J. Plant Biol.* **2017**, *60*, 35-47.
42. Emanuelsson, O.; Nielsen, H.; Brunak, S.; von Heijne, G. Predicting subcellular localization of proteins based on their n-terminal amino acid sequence. *J. Mol. Biol.* **2000**, *300*, 1005-1016.
43. Záhonová, K.; Füssy, Z.; Birčák, E.; Novák Vanclová, A.M.G.; Klimeš, V.; Vesteg, M.; Krajčovič, J.; Oborník, M.; Eliáš, M. Peculiar features of the plastids of the colourless alga *Euglena longa* and photosynthetic euglenophytes unveiled by transcriptome analyses. *Sci. Rep.* **2018**, *8*, 17012.
44. Sláviková, S.; Vacula, R.; Fang, Z.; Ehara, T.; Osafune, T.; Schwartzbach, S.D. Homologous and heterologous reconstitution of Golgi to chloroplast transport and protein import into the complex chloroplasts of *Euglena*. *J. Cell Sci.* **2005**, *118*, 1651-1661.
45. Maier, U.G.; Zauner, S.; Hempel, F. Protein import into complex plastids: Cellular organization of higher complexity. *Eur. J. Cell Biol.* **2015**, *94*, 340-348.
46. Smillie, R.M. Formation and function of soluble proteins in ghloroplasts. *Can. J. Bot.* **1963**, *41*, 123-154.
47. Bégin-Heick, N. The localization of enzymes of intermediary metabolism in *Astasia* and *Euglena*. *Biochem. J.* **1973**, *134*, 607-616.
48. Dockerty, A.; Merrett, M.J. Isolation and enzymic characterization of *Euglena* proplastids. *Plant Physiol.* **1979**, *63*, 468-473.

49. Graves, J.R.; Lynn, B. Effects of different substrates on glucose uptake and hexokinase activity in *Euglena gracilis*. *J. Protozool.* **1971**, *18*, 543-546.
50. Bäumer, D.; Preisfeld, A.; Ruppel, H.G. Isolation and characterization of paramylon synthase from *Euglena gracilis* (euglenophyceae). *J. Phycol.* **2001**, *37*, 38-46.
51. Lucchini, G. Control of glucose phosphorylation in *Euglena gracilis* I. Partial characterization of a glucokinase. *Biochimica et Biophysica Acta - Enzymology* **1971**, *242*, 365-370.
52. Bégin-Heick, N. Oxygen toxicity and carbon deprivation in *Astasia longa*. *Can. J. Biochem.* **1970**, *48*, 251-258.
53. Miyatake, K.; Enomoto, T.; Kitaoka, S. Detection and subcellular distribution of pyrophosphate: D-fructose 6- phosphate phosphotransferase (PFP) in *Euglena gracilis*. *Agric. Biol. Chem.* **1984**, *48*, 2857-2859.
54. Willard, J.M.; Gibbs, M. Purification and characterization of the fructose diphosphate aldolases from *Anacystis* is nidulans and *Saprospira thermalis*. *Biochimica et Biophysica Acta - Enzymology* **1968**, *151*, 438-448.
55. Mo, Y.; Harris, B.G.; Gracy, R.W. Triosephosphate isomerases and aldolases from light- and dark-grown *Euglena gracilis*. *Arch. Biochem. Biophys.* **1973**, *157*, 580-587.
56. Latzko, E.; Gibbs, M. Enzyme activities of the carbon reduction cycle in some photosynthetic organisms. *Plant Physiol.* **1969**, *44*, 295-300.
57. Grissom, F.E.; Kahn, J.S. Glyceraldehyde-3-phosphate dehydrogenases from *Euglena gracilis*: Purification and physical and chemical characterization. *Arch. Biochem. Biophys.* **1975**, *171*, 444-458.
58. Bukowiecki, A.C.; Anderson, L.E. Multiple forms of aldolase and triose phosphate isomerase in diverse plant species. *Plant Sci. Lett.* **1974**, *3*, 381-386.
59. Forsee, W.T.; Kahn, J.S. Carbon dioxide fixation by isolated chloroplasts of *Euglena gracilis*: I. Isolation of functionally intact chloroplasts and their characterization. *Arch. Biochem. Biophys.* **1972**, *150*, 296-301.
60. Ammon, R.; Friedrich, G. The behavior of enzymes in *Euglena gracilis*. *Acta Biol. Med. Ger.* **1967**, *19*, 659-672.
61. Hannaert, V.; Brinkmann, H.; Nowitzki, U.; Lee, J.A.; Albert, M.-A.; Sensen, C.W.; Gaasterland, T.; M, M.; Michels, P.; Martin, W. Enolase from *Trypanosoma brucei*, from the amitochondriate protist *Mastigamoeba balamuthi*, and from the chloroplast and cytosol of *Euglena gracilis*: Pieces in the evolutionary puzzle of the eukaryotic glycolytic pathway. *Mol. Biol. Evol.* **2000**, *17*, 989-1000.
62. Ohmann, E. Die regulation der pyruvat-kinase in *Euglena gracilis*. *Arch. Microbiol.* **1969**, *67*, 273-292.
63. Ogawa, T.; Kimura, A.; Sakuyama, H.; Tamoi, M.; Ishikawa, T.; Shigeoka, S. Identification and characterization of cytosolic fructose-1,6-bisphosphatase in *Euglena gracilis*. *Biosci., Biotechnol., Biochem.* **2015**, *79*, 1957-1964.
64. Hovenkamp-Obbema, R. Effect of chloramphenicol on the development of proplastids in *Euglena gracilis*: II. The synthesis of carotenoids. *Z Pflanzenphysiol* **1974**, *73*, 439-447.
65. Kempner, E.S.; Miller, J.H. The molecular biology of *Euglena gracilis*: IV. Cellular stratification by centrifuging. *Exp. Cell Res.* **1968**, *51*, 141-149.
66. Ohmann, E.; Rindt, K.P.; Borriss, R. Glucose-6-phosphat-dehydrogenase in autotrophen mikroorganismen i. Die regulation der synthese der glucose-6-phosphat-dehydrogenase in *Euglena gracilis* und *Rhodopseudomonas spheroides* in abhängigkeit von den kulturbedingungen. *Z. Allg. Mikrobiol.* **1969**, *9*, 557-564.

67. Hovenkamp-Obbema, R.; Moorman, A.; Stegwee, D. Aminolaevulinate dehydratase in greening cells of *Euglena gracilis*. *Z Pflanzenphysiol* **1974**, *72*, 277-286.
68. Forsee, W.T.; Kahn, J.S. Carbon dioxide fixation by isolated chloroplasts of *Euglena gracilis*: II. Inhibition of CO_2 fixation by AMP. *Arch. Biochem. Biophys.* **1972**, *150*, 302-309.
69. Ma, Y.A.N.; Jakowitsch, J.; Deusch, O.; Henze, K.; Martin, W.; Löffelhardt, W. Transketolase from *Cyanophora paradoxa*: In vitro import into cyanelles and pea chloroplasts and a complex history of a gene often, but not always, transferred in the context of secondary endosymbiosis. *J. Eukaryotic Microbiol.* **2009**, *56*, 568-576.
70. Stern, J.R.; Hegre, C.S. Inducible D-malic enzyme in *Escherichia coli*. *Nature* **1966**, *212*, 1611-1612.
71. Peak, M.J.; Peak, J.G.; Ting, I.P. Isoenzymes of malate dehydrogenase and their regulation in *Euglena gracilis* Z. *Biochim. Biophys. Acta, Enzymol.* **1972**, *284*, 1-15.
72. Peak, M.J.; Peak, J.G.; Ting, I.P. Light-induced reduction in specific activity of malate enzyme in *Euglena gracilis* Z. *Biochem. Biophys. Res. Commun.* **1972**, *48*, 1074-1078.
73. Codd, G.A.; Merrett, M.J. The regulation of glycolate metabolism in division synchronized cultures of *Euglena*. *Plant Physiol.* **1971**, *47*, 640-643.
74. Laval-Martin, D.; Farineau, J.; Pineau, B.; Calvayrac, R. Evolution of enzymes involved in carbon metabolism (phosphoenolpyruvate and ribulose-bisphosphate carboxylases, phosphoenolpyruvate carboxykinase) during the light-induced greening of *Euglena gracilis* strains Z and ZR. *Planta* **1981**, *151*, 157-167.
75. Briand, J.; Calvayrac, R.; Laval-Martin, D.; Farineau, J. Evolution of carboxylating enzymes involved in paramylon synthesis (phosphoenolpyruvate carboxylase and carboxykinase) in heterotrophically grown *Euglena gracilis*. *Planta* **1981**, *151*, 168-175.
76. Ohmann, E.; Plhák, F. Reinigung und eigenschaften von phosphoenolpyruvat - carboxylase aus *Euglena gracilis*. *Eur. J. Biochem.* **1969**, *10*, 43-55.
77. Miyatake, K.; Ito, T.; Kitaoka, S. Subcellular location and some properties of phosphoenolpyruvate carboxykinase (PEPCK) in *Euglena gracilis*. *Agric. Biol. Chem.* **1984**, *48*, 2139-2141.
78. Yokota, A.; Hosotani, K.; Kitaoka, S. Mechanism of metabolic regulation in photoassimilation of propionate in *Euglena gracilis* z. *Arch. Biochem. Biophys.* **1982**, *213*, 530-537.
79. Inui, H.; Ono, K.; Miyatake, K.; Nakano, Y.; Kitaoka, S. Purification and characterization of pyruvate:$NADP^+$ oxidoreductase in *Euglena gracilis*. *J. Biol. Chem.* **1987**, *262*, 9130-9135.
80. Rotte, C.; Stejskal, F.; Zhu, G.; Keithly, J.S.; Martin, W. Pyruvate:NADP oxidoreductase from the mitochondrion of *Euglena gracilis* and from the apicomplexan *Cryptosporidium parvum*: A biochemical relic linking pyruvate metabolism in mitochondriate and amitochondriate protists. *Mol. Biol. Evol.* **2001**, *18*, 710-720.
81. Hoffmeister, M.; van der Klei, A.; Rotte, C.; van Grinsven, K.W.A.; van Hellemond, J.J.; Henze, K.; Tielens, A.G.M.; Martin, W. *Euglena gracilis* rhodoquinone:Ubiquinone ratio and mitochondrial proteome differ under aerobic and anaerobic conditions. *J. Biol. Chem.* **2004**, *279*, 22422-22429.
82. Collins, N.; Merrett, M.J. Microbody-marker enzymes during transition from phototrophic to organotrophic growth in *Euglena*. *Plant Physiol.* **1975**, *55*, 1018-1022.
83. Graves Jr, L.B.; Trelease, R.N.; Grill, A.; Becker, W.M. Localization of glyoxylate cycle enzymes in glyoxysomes in *Euglena*. *J. Protozool.* **1972**, *19*, 527-532.
84. Cook, J.R.; Carver, M. Partial photo-repression of the glyoxylate by-pass in *Euglena*. *Plant Cell Physiol.* **1966**, *7*, 377-383.
85. Cook, J.R.; Heinrich, B. Unbalanced respiratory growth of *Euglena*. *Microbiology* **1968**, *53*, 237-251.

86. Collins, N.; Merrett, M.J. The localization of glycollate-pathway enzymes in *Euglena*. *Biochem. J.* **1975**, *148*, 321-328.
87. Oda, Y.; Miyatake, K.; Nakano, Y.; Kitaoka, S. Subcellular location and some properties of isocitrate dehydrogenase isozymes in *Euglena gracilis*. *Agric. Biol. Chem.* **1981**, *45*, 2619-2621.
88. Shigeoka, S.; Onishi, T.; Maeda, K.; Nakano, Y.; Kitaoka, S. Occurrence of thiamin pyrophosphate-dependent 2-oxoglutarate decarboxylase in mitochondria of *Euglena gracilis*. *FEBS Lett.* **1986**, *195*, 43-47.
89. Shigeoka, S.; Nakano, Y. Characterization and molecular properties of 2-oxoglutarate decarboxylase from *Euglena gracilis*. *Arch. Biochem. Biophys.* **1991**, *288*, 22-28.
90. Tokunaga, M.; Nakano, Y.; Kitaoka, S. Separation and properties of the NAD-linked and NADP-linked isozymes of succinic semialdehyde dehydrogenase in *Euglena gracilis*. *Biochim. Biophys. Acta, Enzymol.* **1976**, *429*, 55-62.
91. Barry, D.; Merrett, M.J. Malate dehydrogenase isoenzymes in division synchronized cultures of *Euglena*. *Plant Physiol.* **1973**, *51*, 1127-1132.
92. Isegawa, Y.; Nakano, Y.; Kitaoka, S. Submitochondrial location and some properties of NAD^+ - and $NADP^+$-linked malate dehydrogenase in *Euglena*. *Agric. Biol. Chem.* **1984**, *48*, 549-552.
93. Miyatake, K.; Washio, K.; Yokota, A.; Nakano, Y.; Kitaoka, S. Occurrence of two NAD^+-linked malate dehydrogenase isozymes in *Euglena gracilis*. *Agric. Biol. Chem.* **1985**, *49*, 859-860.
94. Nakazawa, M. C2 metabolism in *Euglena*. *Adv. Exp. Med. Biol.* **2017**, *979*, 39-45.
95. Graves, L.B.; Hanzely, L.; Trelease, R.N. The occurrence and fine structural characterization of microbodies in *Euglena gracilis*. *Protoplasma* **1971**, *72*, 141-152.
96. Ono, K.; Kondo, M.; Osafune, T.; Miyatake, K.; Inui, H.; Kitaoka, S.; Nishimura, M.; Nakano, Y. Presence of glyoxylate cycle enzymes in the mitochondria of *Euglena gracilis*. *J. Eukaryotic Microbiol.* **2003**, *50*, 92-96.
97. Rodríguez-Zavala, J.S.; Ortiz-Cruz, M.A.; Moreno-Sanchez, R. Characterization of an aldehyde dehydrogenase from *Euglena gracilis*. *J. Eukaryotic Microbiol.* **2006**, *53*, 36-42.
98. Yoval-Sánchez, B.; Jasso-Chávez, R.; Lira-Silva, E.; Moreno-Sánchez, R.; Rodríguez-Zavala, J.S. Novel mitochondrial alcohol metabolizing enzymes of *Euglena gracilis*. *J. Bioenerg. Biomembr.* **2011**, *43*, 519-530.
99. Ono, K.; Kawanaka, Y.; Izumi, Y.; Inui, H.; Miyatake, K.; Kitaoka, S.; Nakano, Y. Mitochondrial alcohol dehydrogenase from ethanol-grown *Euglena gracilis*. *J. Biochem.* **1995**, *117*, 1178-1182.
100. Graves Jr, L.B.; Becker, W.M. Beta-oxidation in glyoxysomes from *Euglena**. *J. Protozool.* **1974**, *21*, 771-774.
101. Tucci, S.; Vacula, R.; Krajcovic, J.; Proksch, P.; Martin, W. Variability of wax ester fermentation in natural and bleached *Euglena gracilis* strains in response to oxygen and the elongase inhibitor flufenacet. *Journal of Eukaryotic Microbiology* **2010**, *57*, 63-69.
102. Grimm, P.; Risse, J.M.; Cholewa, D.; Müller, J.M.; Beshay, U.; Friehs, K.; Flaschel, E. Applicability of *Euglena gracilis* for biorefineries demonstrated by the production of α-tocopherol and paramylon followed by anaerobic digestion. *J. Biotechnol.* **2015**, *215*, 72-79.
103. O'Neill, E.C.; Kuhaudomlarp, S.; Rejzek, M.; Fangel, J.U.; Alagesan, K.; Kolarich, D.; Willats, W.G.T.; Field, R.A. Exploring the glycans of *Euglena gracilis*. *Biology* **2017**, *6*, 45.
104. Marechal, L.R.; Goldemberg, S.H. Laminaribiose phosphorylase from *Euglena gracilis*. *Biochemical and Biophysical Research Communications* **1963**, *13*, 106-107.

105. Tomos, A.D.; Northcote, D.H. A protein-glucan intermediate during paramylon synthesis. *Biochem. J.* **1978**, *174*, 283-290.
106. Calvayrac, R.; Laval-Martin, D.; Briand, J.; Farineau, J. Paramylon synthesis by *Euglena gracilis* photoheterotrophically grown under low O_2 pressure. *Planta* **1981**, *153*, 6-13.
107. Briand, J.; Calvayrac, R. Paramylon synthesis in heterotrophic and photoheterotrophic *Euglena* (Euglenophyceae). *J. Phycol.* **1980**, *16*, 234-239.
108. Takeda, T.; Nakano, Y.; Takahashi, M.; Konno, N.; Sakamoto, Y.; Arashida, R.; Marukawa, Y.; Yoshida, E.; Ishikawa, T.; Suzuki, K. Identification and enzymatic characterization of an endo-1,3-β-glucanase from *Euglena gracilis*. *Phytochemistry* **2015**, *116*, 21-27.
109. Barras, D.R.; Stone, B.A. β-1,3-glucan hydrolases from *Euglena gracilis*: II. Purification and properties of the β-1,3-glucan exo-hydrolase. *Biochim. Biophys. Acta, Enzymol.* **1969**, *191*, 342-353.
110. Kuhaudomlarp, S.; Patron, N.J.; Henrissat, B.; Rejzek, M.; Saalbach, G.; Field, R.A. Identification of *Euglena gracilis* β-1,3-glucan phosphorylase and establishment of a new glycoside hydrolase (GH) family GH149. *J. Biol. Chem.* **2018**, *293*, 2865-2876.
111. Takenaka, S.; Kondo, T.; Nazeri, S.; Tamura, Y.; Tokunaga, M.; Tsuyama, S.; Miyatake, K.; Nakano, Y. Accumulation of trehalose as a compatible solute under osmotic stress in *Euglena gracilis* Z. *J. Eukaryotic Microbiol.* **1997**, *44*, 609-613.
112. Porchia, A.C.; Fiol, D.F.; Salerno, G.L. Differential synthesis of sucrose and trehalose in Euglena gracilis cells during growth and salt stress. *Plant Science* **1999**, *149*, 43-49.
113. Fiol, D.F.; Salerno, G.L. Trehalose synthesis in *Euglena gracilis* (euglenophyceae) occurs through an enzyme complex. *Journal of Phycology* **2005**, *41*, 812-818.
114. Schluepmann, H.; Paul, M. Trehalose metabolites in Arabidopsis-elusive, active and central. *Arabidopsis Book* **2009**, *7*, e0122-e0122.
115. Vandesteene, L.; Ramon, M.; Le Roy, K.; Van Dijck, P.; Rolland, F. A single active trehalose-6-P synthase (TPS) and a family of putative regulatory TPS-like proteins in *Arabidopsis*. *Mol Plant* **2010**, *3*, 406-419.
116. Murray, D.R.; Giovanelli†, J.; Smillie, R.M. Photoassimilation of glycolate, glycine and serine by *Euglena gracilis*. *J. Protozool.* **1970**, *17*, 99-104.
117. Yokota, A.; Komura, H.; Kitaoka, S. Different metabolic fate of two carbons of glycolate in *Euglena gracilis* Z. In *Advances in Photosynthesis Research: Proceedings of the VIth International Congress on Photosynthesis, Brussels, Belgium, August 1–6, 1983*, Sybesma, C., Ed. Springer Netherlands: Dordrecht, **1984**; pp 407-410.
118. Yokota, A.; Kawabata, A.; Kitaoka, S. Mechanism of glyoxylate decarboxylation in the glycolate pathway in *Euglena gracilis* Z. *Plant Physiol.* **1983**, *71*, 772-776.
119. Yokota, A.; Haga, S.; Kitaoka, S. Purification and some properties of glyoxylate reductase ($NADP^+$) and its functional location in mitochondria in *Euglena gracilis* z. *Biochem. J.* **1985**, *227*, 211-216.
120. Hasan, M.T.; Sun, A.; Mirzaei, M.; Te'o, J.; Hobba, G.; Sunna, A.; Nevalainen, H. A comprehensive assessment of the biosynthetic pathways of ascorbate, α-tocopherol and free amino acids in Euglena gracilis var. saccharophila. *Algal Research* **2017**, *27*, 140-151.
121. Isegawa, Y.; Watanabe, F.; Kitaoka, S.; Nakano, Y. Subcellular distribution of cobalamin-dependent methionine synthase in *Euglena gracilis* z. *Phytochemistry* **1993**, *35*, 59-61.
122. Reinbothe, C.; Ortel, B.; Parthier, B.; Reinbothe, S. Cytosolic and plastid forms of 5-enolpyruvylshikimate-3-phosphate synthase in *Euglena gracilis* are differentially expressed during light-induced chloroplast development. *Molecular and General Genetics MGG* **1994**, *245*, 616-622.

123. Richards, T.A.; Dacks, J.B.; Campbell, S.A.; Blanchard, J.L.; Foster, P.G.; McLeod, R.; Roberts, C.W. Evolutionary origins of the Eukaryotic shikimate pathway: gene fusions, horizontal gene transfer, and endosymbiotic replacements. *Eukaryotic Cell* **2006**, *5*, 1517-1531.

124. Oda, Y.; Nakano, Y.; Kitaoka, S. Utilization and toxicity of exogenous amino acids in *Euglena gracilis*. *Microbiology* **1982**, *128*, 853-858.

125. Oda, Y.; Nakano, Y.; Kitaoka, S. Occurrence and some properties of two threonine dehydratases in *Euglena gracilis*. *Microbiology* **1983**, *129*, 57-61.

126. Park, B.-S.; Hirotani, A.; Nakano, Y.; Kitaoka, S. Purification and someproperties of arginine deiminase in *Euglena gracilis* Z. *Agric. Biol. Chem.* **1984**, *48*, 483-489.

127. Koonin, E.V.; Galperin, M.Y. Evolution of central metabolic pathways: The playground of non-orthologous gene displacement. In *Sequence — Evolution — Function: Computational Approaches in Comparative Genomics*, Koonin, E.V.; Galperin, M.Y., Eds. Springer US: Boston, MA, **2003**; pp 295-355.

128. Torruella, G.; Suga, H.; Riutort, M.; Pereto, J.; Ruiz-Trillo, I. The evolutionary history of lysine biosynthesis pathways within Eukaryotes. *Journal of Molecular Evolution* **2009**, *69*, 240-248.

129. Xu, H.; Andi, B.; Qian, J.; West, A.H.; Cook, P.F. The α-aminoadipate pathway for lysine biosynthesis in fungi. *Cell Biochem. Biophys.* **2006**, *46*, 43-64.

130. Creaser, E.H.; Varela-Torres, R. Immunological comparisons of histidinol dehydrogenases. *Microbiology* **1971**, *67*, 85-90.

131. Tokunaga, M.; Nakano, Y.; Kitaoka, S. Subcellular localization of the GABA-shunt enzymes in *Euglena gracilis* strain Z. *J. Protozool.* **1979**, *26*, 471-473.

132. Miyatake, K.; Kitaoka, S. NADH-dependent glutamat synthase in *Euglena gracilis* z. *Agric. Biol. Chem.* **1981**, *45*, 1727-1729.

133. Fayyaz-Chaudhary, M.; Javed, Q.; Merrett, M.J. Effect of growth conditions on NADPH-specific glutamate dehydrogenase activity of *Euglena gracilis*. *New Phytol.* **1985**, *101*, 367-376.

134. Sulzman, F.M.; Edmunds, L.N. Persisting circadian oscillations in enzyme activity in non-dividing cultures of *Euglena*. *Biochem. Biophys. Res. Commun.* **1972**, *47*, 1338-1344.

135. Edmunds, L.N.; Walther, W.G.; Sulzman, F.M. At cellular level. *J. Interdiscip. Cycle Res.* **1972**, *3*, 107-108.

136. Shemin, D.; Rittenberg, D. The utilization of glycine for the synthesis of a porphyrin. *Journal of Biological Chemistry* **1945**, *159*, 567-568.

137. Beale, S.I.; Castelfranco, P.A. Biosynthesis of delta-aminolevulinic-acid in higher-plants.1. Accumulation of delta-aminolevulinic-acid in greening plant-tissues. *Plant Physiology* **1974**, *53*, 291-296.

138. Iida, K.; Mimura, I.; Kajiwara, M. Evaluation of two biosynthetic pathways to δ-aminolevulinic acid in *Euglena gracilis*. *Eur. J. Biochem.* **2002**, *269*, 291-297.

139. Weinstein, J.D.; Beale, S.I. Separate physiological roles and subcellular compartments for two tetrapyrrole biosynthetic pathways in *Euglena gracilis*. *J. Biol. Chem.* **1983**, *258*, 6799-6807.

140. Wolpert, J.S.; Ernst-Fonberg, M.L. Multienzyme complex for carbon dioxide fixation. *Biochemistry* **1975**, *14*, 1095-1102.

141. Kim, D.; Filtz, M.R.; Proteau, P.J. The methylerythritol phosphate pathway contributes to carotenoid but not phytol biosynthesis in *Euglena gracilis*. *Journal of Natural Products* **2004**, *67*, 1067-1069.