Mitochondrial-type iron–sulfur cluster biosynthesis genes (IscS and IscU) in the apicomplexan Cryptosporidium parvum

Michael J. LaGier,1† Jan Tachezy,3 Frantisek Stejskal,2 Katerina Kutisova3 and Janet S. Keithly1

1Wadsworth Center, New York State Department of Health, PO Box 22002, Albany, NY 12201-2002, USA
2First Faculty of Medicine, Charles University, Prague, Czech Republic
3Department of Parasitology, Faculty of Science, Charles University, Prague, Czech Republic

Several reports have indicated that the iron–sulfur cluster [Fe–S] assembly machinery in most eukaryotes is confined to the mitochondria and chloroplasts. The best-characterized and most highly conserved [Fe–S] assembly proteins are a pyridoxal-5′-phosphate-dependent cysteine desulfurase (IscS), and IscU, a protein functioning as a scaffold for the assembly of [Fe–S] prior to their incorporation into apoproteins. In this work, genes encoding IscS and IscU homologues have been isolated and characterized from the apicomplexan parasite Cryptosporidium parvum, an opportunistic pathogen in AIDS patients, for which no effective treatment is available. Primary sequence analysis (CpIscS and CpIscU) and phylogenetic studies (CpIscS) indicate that both genes are most closely related to mitochondrial homologues from other organisms. Moreover, the N-terminal signal sequences of CpIscS and CpIscU predicted in silico specifically target green fluorescent protein to the mitochondrial network of the yeast Saccharomyces cerevisiae. Overall, these findings suggest that the previously identified mitochondrial relict of C. parvum may have been retained by the parasite as an intracellular site for [Fe–S] assembly.

INTRODUCTION

The protist Cryptosporidium parvum (Phylum Apicomplexa) is a gut-dwelling parasite that infects both humans and animals. In AIDS patients, C. parvum is an opportunistic pathogen, causing significant morbidity and mortality by inducing chronic diarrhoeal disease (Chen et al., 2002). The lack of an effective therapy against human cryptosporidiosis has stimulated research on the basic biology of this parasite.

Iron–sulfur clusters [Fe–S] are ubiquitous, and [Fe–S]-containing proteins play critical metabolic, regulatory and signalling roles in essentially all forms of life, from archaea to man (Johnson et al., 1998; Lill & Kispal, 2000; Gerber & Lill, 2002). Furthermore, because [Fe–S] have a broad range of redox potentials, many [Fe–S]-containing proteins are key components in electron-transfer reactions. Although [Fe–S] have been proposed to be among the earliest catalytic centers in biochemical evolution (Camack, 1971, 1983), little is known about their biosynthesis in eukaryotes (Seeber, 2002). Nonetheless, the ancient character of [Fe–S] is supported by recent observations that proteins involved in [Fe–S] biosynthesis are highly conserved in prokaryotes and eukaryotes (Mühlenhoff & Lill, 2000; Tachezy et al., 2001; Tovar et al., 2003), including members of the Apicomplexa (Seeber, 2002).

Among the proteins involved in [Fe–S] assembly, the two best characterized and most highly conserved are a pyridoxal-5′-phosphate (PLP)-dependent cysteine desulfurase which catalyses the formation of L-alanine and elementary sulfur using L-cysteine as a substrate (IscS), and the protein IscU, which serves as a scaffolding for the assembly of [Fe–S] prior to their incorporation into apoproteins (Mühlenhoff & Lill, 2000).

In most eukaryotes [Fe–S] assembly and maturation is primarily confined to mitochondria and chloroplasts (reviewed by Lill & Kispal, 2000). Until very recently, the cellular localization of [Fe–S] assembly in ‘amitochondriate’ protists was completely unknown. Now, however, in addition to the hydrogenosomes of the parabasalid...
Trichomonas vaginalis (Tachez, 2001), mitochondrion-like organelles discovered in the diplomonad Giardia intestinalis have been shown to participate in [Fe–S] biosynthesis (Tovar et al., 2003). All of these organelles are thought to be descendents of a common endosymbiont which gave rise to both hydrogenosomes and mitochondria (Martin & Müller, 1998; Dyall & Johnson, 2000; Rotte et al., 2000; Martin et al., 2001). Putative mitochondrial relics have also been identified in the parasitic amoeba Entamoeba histolytica (mitosome, Tovar et al., 1999; crypton, Mai et al., 1999) and the microsporidian Trachipleistophora hominis (Williams et al., 2002). Although similar structures have not yet been observed in the microsporidian Encephalitozoon cuniculi, even this species contains genes of probable mitochondrial origin (Katinka et al., 2001).

Recent data indicate that the ribosome-studded organelle of C. parvum sporozoites is a relict mitochondrion (Riordan et al., 2003). Significantly, the C. parvum mitochondrion is limited by a double membrane, contains internal membranous cristae-like compartments, and both a mitochondrial-type, nucleus-encoded chaperone protein (CpN60; GenBank AAC32614) and heat-shock protein 70 (Cp-mtHSP70; GenBank AY235430) are localized to it (Riordan et al., 2003; J. R. Slapeta & J. S. Keithly, unpublished observations). Although the physiological functions of the C. parvum mitochondrion, as well as other mitochondrial remnants, are unknown (Katinka et al., 2001; Williams et al., 2002), mitochondria-type [Fe–S] assembly machinery has been demonstrated for the amitochondriate protist G. intestinalis (Tovar et al., 2003), and suggested for T. vaginalis (Tachezy et al., 2001), E. histolytica (GenBank accession AY040613) and E. cuniculi (Katinka et al., 2001). Here we propose that IscS and IscU in the apicomplexan C. parvum may also contribute to [Fe–S] assembly within its relict mitochondrion.

As an initial step in addressing this hypothesis, nucleus-encoded IscS and IscU genes have been isolated from C. parvum, transcription in sporozoites documented, and the N-terminal signal sequences of these two [Fe–S] homologues reconstructed in a vector to deliver green fluorescent protein (GFP) to the yeast mitochondrial network.

METHODS

Parasite and nucleic acid isolation. Cryptosporidium parvum oocysts (IOWA strain) were collected from the faeces of infected calves (Fayer, 1997). Following in vitro excystation, released sporozoites were isolated and concentrated as described previously (Keithly et al., 1997). Genomic DNA (gDNA) or total RNA was isolated from sporozoites using a Puregene DNA isolation (Genta Systems) or Qiagen RNEasy kit (Qiagen), respectively (LaGier et al., 2001).

Molecular cloning. Partial sequences of C. parvum genes encoding homologues of proteins involved in [Fe–S] assembly were identified by scanning genomic DNA (gDNA) sequences deposited into the GenBank database as part of an ongoing C. parvum genomic sequence survey (GSS; Strong & Nelson, 2000). These initial C. parvum database searches were performed using BLAST (Altschul et al., 1997) 2.0 (Expected value set to 10) at the National Center for Biotechnology Information (NCBI) homepage (http://www3.ncbi.nlm.nih.gov/BLAST) using IscS (GenBank M98808) and IscU (GenBank NP014869) homologues from the yeast Saccharomyces cerevisiae (Mühlenhoff & Lil, 2000) as query sequences. gSS clones containing gDNA fragments of putative C. parvum IscS (GenBank AQ49762) and IscU (gSS contig #1309) homologues were identified from this database.

Next, PCR was used to amplify portions of both clones directly from isolated sporozoite gDNA. The resultant amplicons were extracted from agarose gels and utilized as probes to screen plasmid-based C. parvum gDNA libraries transformed into the bacterium Escherichia coli (Zhu et al., 2000). Two gDNA libraries of KSU-1 sporozoites constructed within pBluescript SK+ (Stratagene) at the EcoRI or HindIII sites were utilized for cloning. The final ORFs of both CpsIscS and CpsIscU were obtained by screening both gDNA libraries using gene-walking (LaGier et al., 2001; Zhu et al., 2000). All clones represent at least two identical overlapping clones obtained from each library. Each clone was sequenced twice on both strands by automated sequencing. Oligonucleotides for sequencing and PCR were synthesized, and automated sequencing was performed by the Molecular Genetics Core Facility at the Wadsworth Center (USA) or by Generi Biotech (Czech Republic). Sequence analyses, including clone assembly, protein translation and primary sequence analysis used the GCG Wisconsin Package UNIX version 9.1 (Genetics Computer Group), and Lasergene system (DNASTAR). The C. parvum IscS and IscU sequences have been submitted to GenBank under accession numbers AY029212 and AY078500, respectively.

Sequence alignments and phylogenetic analysis. Nucleotide and protein databases at NCBI were queried by BLAST 2.0 (Expected value set to 10) using the complete ORF of CpsIscS or CpsIscU. Sequences of interest were extracted from the various databases using the BlastAli program (http://www.joern-lewinn.de/). The IscS and IscU sequences of C. parvum were aligned to sequences from the eubacterium Azotobacter vinelandii, mitosme-bearing diplomonad G. intestinalis, hydrogenosome-bearing parabasalid Trichomonas vaginalis (IscS only), mitochondrion-bearing apicomplexan Plasmodium falciparum (IscS), yeast S. cerevisiae, and microsporidian Encephalitozoon cuniculi using the default parameters of CLUSTAL_X (Thompson et al., 2000). For phylogenetic analysis, CpsIscS was aligned to sequences from 24 selected taxa using CLUSTAL_X (Thompson et al., 2000) and was further edited visually using the ED program of MUST (Philippe, 1993). The alignment of the 24 IscS taxa resulted in 361 shared amino acid positions.

Phylogenetic relationships were analysed by the neighbour-joining (NJ) and maximum-parsimony (MP) methods using PHYLIP, version 3.6 (Felsenstein, 1989), and by the maximum-likelihood (ML) method using the PROML program in MOLPHY, version 2.3 (Adachi & Hasegawa, 1996). The ML tree was constructed by local rearrangement of an NJ tree using the Jones–Taylor–Thornton model of amino acid substitutions with the F-option to account for amino acid frequencies in the dataset. User-defined trees were analysed to compare alternative topologies (Kishino & Hasegawa, 1989). The local bootstrap proportion value was calculated for each internal branch of the ML tree using a local rearrangement option of the PROML program. Bootstrap support (bootstrap proportion, BP) for distance and parsimony analyses was based on 100 re-sampled datasets using SEQBOOT, PHYLIP, version 3.6.

Prediction of mitochondrial targeting sequences. Two in silico methods were used for predicting N-terminal mitochondrial targeting peptides, TargetP (Emanuelsson et al., 2000) and Mitoprot (Claros & Vincens, 1996; Emanuelsson & von Heijne, 2001). Specifically, potential N-terminal peptides were identified for both
CpIscS and CpIscU using TargetP (http://www.cbs.dtu.dk/services/TargetP/) and MitoProt (http://www.mips.biochem.mpg.de/cgi-bin/proj/medgen/mito.filter). In the analyses of CpIscS and CpIscU, TargetP was used in the 'winner-takes-all' mode without setting a specificity cut-off for targeting, and MitoProt was implemented using default parameters. The entire ORFs of CpIscS or CpIscU were used as query templates during analysis with both MitoProt and Target P. MitoProt was also used for the prediction of putative signal peptide cleavage sites (default settings). Prediction of z-helical structures was performed using the secondary structure program Predator at the Predict Protein server (http://www.embl-heidelberg.de/predictprotein/predictprotein.html). The reliability classes of TargetP decrease from a score of 1 (97% specificity) to 5 (53% specificity). The reliability feature of TargetP is an indication of the level of certainty in a prediction for a given sequence (Emanuelsson et al., 2000).

**Southern blotting.** For Southern blot analysis, 2-5 μg of sporozoite gDNA per lane was digested with restriction endonucleases, separated by electrophoresis in 0.75% agarose gels and transferred to Zeta-probe Nylon membranes in a 0.4 M NaOH solution (Bio-Rad). Fragments of the CpIscS and CpIscU ORFs, 758 and 447 bp respectively, were PCR-amplified from isolated sporozoite gDNA and were labelled with [32P]dATP (Boehringer Mannheim), and used as probes for Southern blot analyses under conditions of high stringency (Maniatis et al., 1989). Next, the isolated gDNA fragments of CpIscS and CpIscU were random-primer labelled with [α-32P]dATP (Boehringer Mannheim), and used as probes for Southern blot analyses under conditions of high stringency (Maniatis et al., 1989).

**CpIscS and CpIscU mRNA detection.** RT-PCR was used to detect CpIscS and CpIscU transcripts in C. parvum. Initially, total isolated sporozoite RNA was synthesized into cDNA using the standard protocol included with the SuperScript First-Strand Synthesis System for RT-PCR (Invitrogen). Briefly, each RT reaction contained 2 μg total sporozoite RNA, 100 ng random hexamers, 10 mM DTT, 0.5 mM dNTPs, 5 mM MgCl2, 10× RT reaction buffer, 40 U RNaseOUT ribonuclease inhibitor, and 50 U SuperScript II RT. The first-strand cDNA synthesis was carried out at 42°C for 50 min, followed by incubation at 70°C for 15 min. Prior to cDNA synthesis, DNase-treated RNA was tested for the presence of gDNA contamination using standard PCR (Maniatis et al., 1989) and oligonucleotide primers specific for the C. parvum 18S rRNA gene (Abrahamsen & Schroeder, 1999). More specifically, following extraction from sporozoites, 250 ng total RNA was PCR amplified for 35 cycles, and agarose gel electrophoresis was used to confirm the absence of gDNA in RNA preparations as indicated by a lack of detectable 18s rRNA amplicons. The amplification of cDNA via standard PCR (Maniatis et al., 1989) for 35 cycles (94°C for 1 min, 50°C for 1 min, 72°C for 1 min) was performed using primers specific for CpIscS or CpIscU (Table 1). Each reaction contained 0.2 mM gene-specific primers, 2 μl gDNA-free cDNA, 0.8 mM dNTPs, 2 mM MgCl2, 10× PCR reaction buffer (Promega), and 5 U Taq polymerase (Promega). Following agarose gel electrophoresis, amplicons of the expected sizes were sequenced to confirm identity.

**Targeting of predicted C. parvum signal peptides in S. cerevisiae.** The predicted mitochondrial signal peptides of CpIscS and CpIscU were amplified via PCR from Iowa gDNA using oligonucleotide primers with 5’ restriction cut-sites, EcoRI, and BamHI (Table 1). The mitochondrion-targeted plasmid pY223-mtGFP (Westermann & Neupert, 2000) was a gift from B. Westermann (Ludwig-Maxilians-Universität, München, Germany). This plasmid contains a known mitochondrial pre-sequence from protein 9 (proteolipid subunit) of the F1 part of the F1F0 ATPase of Neurospora crassa ligated to GFP, and has an inducible galactose promoter. The first 168 bp (56 aa) of the CpIscS gene and the first 111 bp (37 aa) of the CpIscU gene, respectively, were generated by PCR from gDNA by adding EcoRI and BamHI linker sequences to the 5’ and 3’ end of these genes, respectively. The N. crassa mitochondrial pre-sequence was digested from the plasmid using EcoRI and BamHI, and then was replaced with the putative C. parvum pre-sequences. The new recombinant plasmids pCpIscS-56-GFP and pCpIscU-37-GFP were amplified in UltraCompetent XL-10 gold cells (Stratagene), purified, and sequenced to confirm identities. These were then transfected using the standard lithium acetate method (Burke et al., 2000) into S. cerevisiae strain W303 (Gaxiola et al., 1998), a gift of D. Kornitzer (Israel Institute of Technology, Haifa, Israel), and plated onto 2% galactose-containing histidine-free synthetic complete agar to induce expression of the GFP. Only cells transformed with the recombinant plasmids grew on selective plates, and samples from these colonies were prepared for examination by growing for 3 days at 30°C in histidine-free synthetic complete liquid broth and immobilizing with 0.5% low-melt agarose on slides for fluorescence microscopy. Yeast cells were observed with an Olympus BX 51 microscope equipped for differential interference contrast and fluorescent microscopy. All photomicrographs were processed under the same conditions.
identically following capture with a colour digital camera (Alpha Innotech) and documented using Adobe Photoshop 6.1 (Adobe Systems). For excitation of GFP, a 450–490 nm band pass filter was used, and emitted light observed with a 520 nm pass filter.

RESULTS

Cloning of CplscS and CplscU

The complete ORFs of CplscS and CplscU were isolated from EcoRI and HindIII gDNA libraries using colony-hybridization-based gene-walking (Zhu et al., 2000). The overlapping gene fragments, each representing at least two independent plasmid clones, were isolated from gDNA libraries and assembled to form the complete ORFs of CplscS and CplscU (GenBankAY029212 andAY078900, respectively). Like all C. parvum genes, CplscS and CplscU are AT-rich, 61±6 and 65±3 mol% respectively, within coding regions (Spano & Crisanti, 2000). The translated ORF of CplscS encodes a peptide of 438 aa with an $M_r$ of 48 700, while CplscU encodes 150 aa with an $M_r$ of 16 600. Not unexpectedly, the ORFs of CplscS and CplscU consist of single exons, like all C. parvum genes except $\beta$-tubulin (Caccio et al., 1997) and an ATP-binding ABC transporter (Zapata et al., 2002), each of which contains a single intron.

While a number of potential initiation codons were noted at the 5′ end of CplscS and CplscU, the first in-frame ATG triplets ($^{10\text{th}}$ATG and $^{10\text{th}}$ATG, respectively) have been designated the putative translation start sites. This interpretation is supported by the presence of purines at positions −3 and +4, matching consensus nucleotides for start codon sequences (Kozak, 1989) in eukaryotic mRNAs (data not shown).

CplscS and CplscU are single-copy genes

Genomic DNA isolated from C. parvum sporozoites was digested with EcoRI or HindIII, and then used for Southern blot analysis. Using 758 bp of the CplscS ORF as a probe, hybridizing fragments of 1·2 and 8·0 kb were observed (Fig. 1a). Identically prepared Southern blots probed with a 447 bp CplscU probe hybridized with fragments of 2·0 and 2·6 kb (Fig. 1b). Overall, these data suggest that CplscS and CplscU are single-copy genes.

Primary sequence analysis of CplscS

When the deduced amino acid sequence of CplscS was aligned with IscS homologues from several phylogenetically diverse organisms, including protists, bacteria and fungi, highly conserved motifs and functionally significant residues were observed (Fig. 2). CplscS contains all the conserved regions believed to mediate cysteine desulphurase activity, including the following: His$^{136}$, which contributes to the initial deprotonation of the substrate (Kaiser et al., 2000); the PLP-binding site including the Schiff-base-forming Lys$^{237}$, and Asp$^{189}$ and Gln$^{215}$, which bind the pyridine nitrogen and the phenolate oxygen of PLP, respectively. CplscS also contains residues contributing to the formation of six additional hydrogen bonds anchoring the phosphate group: Thr$^{105}$, His$^{237}$, Ser$^{234}$ and Thr$^{275}$ (Zheng et al., 1993); the substrate-binding site Cys$^{362}$, which provides a reactive cysteinyl residue (Zheng

![Fig. 1. Southern blot of CplscS (a) and CplscU (b) using gDNA digested with EcoRI or HindIII, respectively. An arrowhead indicates the positions of the agarose wells, and the observed weights of bands are designated.](image)

![Fig. 2. Amino acid alignment of the putative Cryptosporidium parvum IscS with homologues from a diverse group of organisms including: the eubacterium Azotobacter vinelandii (Av, GenBank P05341); the microsporidian Encephalitozoon cuniculi (Ec, CAD26087); the diplomonad Giardia intestinalis (Gi, AAK39427); the mitochondria-bearing apicomplexan Plasmodium falciparum (Pf, Tachezy, 2001); the yeast Saccharomyces cerevisiae (Sc, M98808); and the parasabalan Trichomonas vaginalis (Tv1 andTv2, AF321005 andAF321006 respectively). Conserved lysine and other residues involved in PLP binding are indicated by filled and open ovals, respectively. The invariable cysteine is marked by an open square, while other substrate-binding residues are indicated by filled squares. An arrow indicates the conserved histidine involved in substrate deprotonation. Putative N-terminal presequences are underlined with dotted lines. The signature cysteines of eukaryotic IscS and C-terminal conserved residues typical for eukaryotic/eubacterial IscS are underlined with solid bold lines. An asterisk (*) indicates a completely conserved residue among all species, moderately conserved groups are indicated by a colon (:), and a period (.) indicates weakly conserved groups according to CLUSTAL_X.](image)
<table>
<thead>
<tr>
<th>Protein</th>
<th>Function</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Tv2</strong></td>
<td><strong>Fe–S</strong> cluster genes in Cryptosporidium parvum</td>
<td></td>
</tr>
</tbody>
</table>
Table 2. Potential mitochondrial targeting sequences identified in C. parvum

<table>
<thead>
<tr>
<th>C. parvum protein</th>
<th>Yeast homologue</th>
<th>Potential mitochondrial targeting sequence (C. parvum)</th>
<th>MitoProt score</th>
</tr>
</thead>
<tbody>
<tr>
<td>IscS</td>
<td>NfsIp</td>
<td>^MIVHRYCRWAPSVVRGISKLAFSSMSSIAAKRRP^37</td>
<td>0.99</td>
</tr>
<tr>
<td>IscU</td>
<td>IscU1p/IscU2p</td>
<td>^MLQLRQIDKRILKKCVPICQRLFYS^27</td>
<td>0.96</td>
</tr>
<tr>
<td>mt-HSP70</td>
<td>SsqIp</td>
<td>^MSMIINSSENGVNSSGIAARILKRSLPLVFSRYMS^36</td>
<td>0.99</td>
</tr>
<tr>
<td>Ferredoxin CpFd1</td>
<td>YahIp</td>
<td>^MVPNLIWRISRISSRVFSAIPYFSKRTLFLSFKRF^35</td>
<td>0.99</td>
</tr>
</tbody>
</table>

et al., 1994), as well as Arg389, Asp189 and Asn64, which anchor the cysteine with a salt bridge and hydrogen bond (Kaiser et al., 2000); and a C-terminal signature sequence that differentiates proteobacterial and eukaryotic IscS from homologues in other organisms (Tachezy et al., 2001). This signature consists of a 20–21 aa consensus sequence (Fig. 2, underlined residues) and is thought to be essential for IscS–IscU protein–protein interaction (Urbina et al., 2001). CpIscS also contains the conserved Cys113 in the substrate deprotonation region (Fig. 2, underlined) which distinguishes eukaryotic IscS from other organisms, including proteobacteria (Tachezy et al., 2001).

Of particular interest is the fact that detailed sequence analysis of CpIscS revealed the N-terminal residues of the ORF to constitute a potential mitochondrial signal sequence. Software (MitoProt) designed to identify mitochondrial signal sequences, including putative cleavage sites, estimated that the first 37 aa of CpIscS represent a bona fide (0.99 probability score, P) mitochondrial targeting peptide. A second algorithm, TargetP (see Methods), showed a high probability score (0.849, reliability class 2) for mitochondrial targeting of CpIscS. Furthermore, the first 37 aa of CpIscS include residues associated with typical physicochemical features of N-terminal mitochondrial targeting peptides, including: an Arg residue at the position relative to the predicted cleavage site (^3^RP/AY^37, Fig. 2, Table 2), an enrichment of Arg (3), Ser (5) and Ala (4) residues, a lack of any acidic residues within this region, more than two positively charged residues (Pfanner, 2000) and the potential of the first 37 aa of CpIscS to form amphiphilic α-helical structures (^VHRYCRWAPSVVR^16 and ^KLAFFS^25, Fig. 2) according to Predator software (see Methods).

In summary, the alignment and identification of all conserved amino acids required for cysteine desulphurase activity in CpIscS suggests that this gene belongs to the group I proteobacterial and eukaryotic mitochondria-type IscS homologues (Tachezy et al., 2001), and that it probably participates in parasite [Fe–S] assembly. Moreover, the identification of an N-terminal mitochondrial signal peptide in CpIscS suggests that this enzyme can target the relict mitochondrion of C. parvum sporozoites.

CpIscS phylogeny

The phylogenetic relationship among major IscS clusters was examined using 361 aa within conserved domain contigs from a dataset of 25 sequences that included IscS from C. parvum and four other protists (two T. vaginalis isoforms), six from proteobacteria, and 14 from other eukaryotic taxa (Fig. 3). The topology of this ML tree essentially agrees with that for the C. parvum mitochondrial genes AK2 (adenylate kinase 2; Riordan et al., 1999) and Cpn60 (Riordan et al., 2003), as well as to a previous tree of eukaryotic-proteobacterial IscS constructed using the ML method (Tachezy et al., 2001). The NJ and ML methods confirmed the close relationship between eukaryotic IscS with high bootstrap support (Fig. 3, BP = 100 and 94 %, respectively). Six branches were defined on the tree: (1) Protista, (2) Fungi, (3) Arabidopsis and Dictyostelium, (4) Metazoa, (5) Rickettsia, and (6) Proteobacteria. Consistent with previous analyses, the Proteobacteria were recovered as a sister group to the mitochondrion-bearing eukaryote clades (Fig. 3) with robust bootstrap support (BP = 99/100/100 %), suggesting a common ancestral origin for IscS in [Fe–S] assembly (Dyall & Johnson, 2000; Mullenhoff & Lill, 2000; Martin et al., 2001; Tachezy et al., 2001). Although CpIscS clustered with significant bootstrap support (BP = 75 %) together with IscS homologues of the protists G. intestinalis, T. vaginalis and P. falciparum within the mitochondrion-bearing eukaryotes (Fig. 3), the placements between and among protist taxa are not clearly resolved. The two most likely reasons for poor resolution among protists in this study are: (i) lack of species diversity in the IscS database, and (ii) presence of lineages which are often fast-evolving and which can be misplaced due to long-branch attraction artefact (Philippe & Laurent, 1998; Tachezy et al., 2001). Long-branch attraction artefact is probably responsible for relationships here among trichomonads, diplomons and apicomplexans (Fig. 3).

Because a global IscS/NifS phylogenetic tree based upon the alignment of 67 taxa resulting in 231 shared amino acid sequences has already been published by one of us (Tachezy et al., 2001), a repeat of this global analysis was not included here. Briefly, this ML tree showed that proteobacterial and mitochondria-type IscS group I homologues (BP = 100 %) are a sister group to, and distinct from, group II
homologues, which include very diverse prokaryotic and plastid-derived eukaryotic IscS. Therefore, both the global tree reconstruction (Tachezy et al., 2001) and our analysis for CpIscS are congruent with a common proteobacterial and eukaryotic mitochondrial ancestry (group I homologues), rather than a plastid origin (group II).

Primary sequence analysis of CpIscU

The primary sequence of CpIscU was compared to IscU homologues from four diverse taxa, including the eubacteria A. vinelandii and E. coli, the yeast S. cerevisiae (two isoforms), and the diplomonad G. intestinalis. Most importantly, CpIscU contains the three cysteine residues (positions 60, 85 and 128) previously shown to be essential for [Fe–S] scaffolding of S. cerevisiae IscU (Garland et al., 1999), and interaction with a mitochondrial chaperone (Hsp70; Tokumoto et al., 2002). CpIscU also possesses the conserved Asp62 residue thought to play a role in the release of transient [Fe–S] for eventual delivery to apoproteins (Mühlhoff & Lill, 2000).

Similar to CpIscS, phylogenetic analysis supported CpIscU clustering with mitochondria-type homologues (data not shown). Like G. intestinalis IscU (Tovar et al., 2003), the N-terminal sequence of CpIscU is rich in hydroxylated and basic amino acids, and MitoProt predicted the mitochondrial localization of CpIscU with high confidence (P > 0.96). Furthermore, the CpIscU predicted N-terminal cleavage site at Ser27 is downstream just one residue from Tyr35, the cleavage site of the human mitochondria-type IscU sequence (Tong & Rouault, 2000). According to Predator software (section 2.5), these 27 aa have the potential to form the amphiphilic z-helical structures 3QLRQLIDKRIL13 and 21CQRLFYSDTVHDHF34 necessary for targeting mitochondria (Fig. 4, Table 2). Therefore, although neither in silico method predicted a ‘typical’ cleavage site, these findings suggest that the N-terminal pre-sequence might target CpIscU to the relict mitochondrion of C. parvum. In addition, not all targeting sequences follow the ‘rules’ for cleavage (Taylor et al., 2001), and most of the prediction programs, including MitoProt and TargetP, are based upon targeting sequences from mammals, yeast, and other model organisms. Therefore, it is not too surprising that these in silico methods vary in their ability to exactly predict cleavage sites in protists. For example, neither C. parvum Cpn60 nor mt-HSP70 perfectly fit the cleavage rules, yet both have been shown by immunogold labelling to properly target the relict mitochondrion of C. parvum (Riordan et al., 2003; J. R. Slapeta & J. S. Keithly, unpublished results). Furthermore, unlike CpIscS and GiIscU, both Giardia IscS and the microsporidian Trachipleistophora mtHsp70 lack
typical N-terminal signals, but both are efficiently targeted into remnant mitochondrion-derived organelles (Tovar et al., 2003; Williams et al., 2002).

Expression of \(\text{CpIscS}\) and \(\text{CpIscU}\)

\(\text{CpIscS}\) and \(\text{CpIscU}\) gene activity was demonstrated in sporozoites by the detection of transcripts by PCR-amplification of gDNA-free cDNA. Transcript detection used gene-specific primers spanning a region of the \(\text{CpIscS}\) or \(\text{CpIscU}\) ORF, respectively, including the predicted mitochondrial signal peptides (Table 1). Amplicons of the expected sizes were obtained from sporozoite cDNA, confirming parasite-mediated transcription of \(\text{CpIscS}\) and \(\text{CpIscU}\) (Fig. 5, lanes 2 and 3, respectively). These results are significant because they indicate that the N-terminal signal peptides of \(\text{CpIscS}\) and \(\text{CpIscU}\) are transcribed, further suggesting that they act as signals for mitochondrial relict targeting.

During our survey of the \(C.\ parvum\) genome for \(\text{IscS}\) and \(\text{IscU}\) homologues, two additional \(C.\ parvum\) genes encoding proteins associated with [Fe–S] assembly in other eukaryotes were identified. These genes encode a mitochondrial-type heat-shock protein 70 (\(\text{Cp-mtHSP70}\), GenBank AY235430), which appears to be important for [Fe–S] assembly as it acts as a molecular chaperone for \(\text{IscU}\) (Paschen & Neupert, 2001; Dutkiewicz et al., 2003) and an adrenodoxin-type [2Fe–2S] ferredoxin (\(\text{CpFd1}\), GenBank AY113180) which provides reducing equivalents for the assembly of [Fe–S] (Lange et al., 2000). Both \(\text{Cp-mtHSP70}\) and \(\text{CpFd1}\) encode
N-terminal mitochondrial targeting peptides (Table 2), which are transcribed in sporozoites of *C. parvum* (Fig. 5, lanes 4 and 5, respectively). Interestingly, recent data show that the N-terminal mitochondrial targeting sequence of Cp-mtHSP70 directs GFP to the mitochondrial network of *S. cerevisiae*, as well as to mitochondria of the apicomplexan *Toxoplasma gondii*. Furthermore, immunogold labelling shows that Cp-mtHSP70 localizes to the relict mitochondrion of *C. parvum* sporozoites (J. R. Slapeta & J. S. Keithly, unpublished data). Together with CpIscS and CpIscU, these additional data on Cp-mtHSP70 and CpFdI strengthen our hypothesis that the *C. parvum* relict mitochondrion may have been retained during reductive evolution for the assembly of [Fe–S].

**Targeting of GFP to yeast mitochondria using *C. parvum* signal peptides**

As mentioned previously, plasmid vectors were constructed to test for the ability of the N-terminal signal peptides of CpIscS and CpIscU to deliver proteins to the mitochondrial network of yeast. Both pCpIscS-56-GFP and pCpIscU-37-GFP (see Methods) can correctly target GFP to the mitochondrial network of transformed yeast, as indicated by a sinuous intracellular staining pattern (Fig. 6). This pattern is consistent with previous studies describing the mitochondria of *S. cerevisiae* as branched, tubular networks located below the cell cortex (Hoffman & Avers, 1973). As expected, yeast transfected with pYX223-mtGFP lacking the *N. crassa* mitochondrial targeting sequence showed a cytoplasmic localization of GFP, confirming the dependence of an N-terminal targeting sequence for trafficking of GFP to yeast mitochondria (data not shown). Intact pYX223-mtGFP was used to transfet yeast as a positive control. As shown in Fig. 6(f), yeast containing pYX223-mtGFP localize GFP to the mitochondrial network of *S. cerevisiae*.

**DISCUSSION**

Recent observations show that protists containing organelles derived from a common proteobacterial ancestor, e.g. *T. vaginalis* (hydrogenosomes, Tachezy et al., 2001), *E. histolytica* (mitosome, Tovar et al., 1999; crypton, Mai et al., 1999), *G. intestinalis* (mitosome, Tovar et al., 2003), and the microsporidian *Trachipleistophora hominis* (Williams et al., 2002) encode the highly conserved [Fe–S] assembly proteins IscS and IscU. Since the [Fe–S] assembly machinery is confined to the mitochondria of most eukaryotes, or to the chloroplasts of higher plants (Leon et al., 2002; Pilon-Smits et al., 2002), we proposed that the *C. parvum* mitochondrial relict may play a role in [Fe–S] biosynthesis.

Here we show by genetic analysis that the apicomplexan *C. parvum*, like other protists, encodes two key components of [Fe–S] assembly machinery – the PLP-dependent cysteine desulfurase (IscS), and IscU, a protein providing scaffolding for transient [Fe–S] formation (Lill & Kispal, 2000; Gerber & Lill, 2002). Both sequence and phylogenetic analyses suggest that CpIscS and CpIscU share a common ancestry from a proteobacterium. Our data are also consistent with the observation that most protist IscS
homologues belong to group I, which includes those of bacteria and eukaryotic mitochondria (Tachezy et al., 2001), rather than group II IscS homologues, which include a diverse group of bacteria and plastid-containing eukaryotes.

Both pCpIscS-56-GFP and pCpIscU-37-GFP (see Methods) delivered GFP into the S. cerevisiae mitochondrial network, indicating that CpIsCS and CpIscU contain N-terminal sequence elements that are interpreted by yeast as signals for import into the organelle. Interestingly, the CpIscU predicted N-terminal cleavage site at Ser27 is downstream just one residue from Tyr28, the cleavage site of the human IscU mitochondria-type isoform (Tong & Rouault, 2000), whereas the G. intestinalis IscU mitochondrial cleavage site at Glu27 is two residues upstream of the human site (Tovar et al., 2003). Therefore, it appears that although some divergence has occurred among protist [Fe–S] protein mitochondrial targeting signals, sufficient sequence information is retained to properly deliver nucleus-encoded IscU into the mitochondrial compartment.

[Fe–S] biogenesis also requires reduced iron (Scheffler, 1999), which is highly toxic for cells (Scheffler, 1999; Gerber & Lill, 2002), and therefore a reason for retaining mitochondrial compartments even if they have become altered (Roger et al., 1998; Katinka et al., 2001; Williams et al., 2002). Recent evidence has shown, for example, that Fe can be incorporated into the hydrogenosome of the cattle parasite Tritrichomonas foetus, and that this requires ferredoxin and other peptides within the [Fe–S] machinery (Suchan et al., 2003). Whether C. parvum might also incorporate iron into the relict mitochondrion is not yet known, but certainly the fact that this apicomplexan possesses IscS, IscU, ferredoxin and mtHsp70 homologues with mitochondrial targeting signals suggests that [Fe–S] biogenesis may be one of the essential biochemical pathways of the endosymbiont retained by parasitic protists after reductive evolution of the mitochondrial relict.

The localization of human IscS and IscU homologues is regulated through translation initiation at alternative in-frame AUG sites (Land & Rouault, 1998), and alternative splicing of a common pre-mRNA, respectively (Tong & Rouault, 2000). That is, human IscS and IscU can be targeted to mitochondria, the cytosol or nucleus (IscS only) as a consequence of these post-transcriptional events. Although alternative splicing is unlikely to regulate CplscU targeting since the gene consists of a single exon, only the intracellular localization of CplscS and CplscU in sporozoites can directly ascertain whether [Fe–S] assembly occurs within the relict mitochondrion of C. parvum.

In summary, sequence and phylogenetic analyses indicate that C. parvum IscS and IscU are mitochondrial isotypes, and that like the [Fe–S] protease sequences found in the genomes of the ‘amitochondriate’ protists Entamoeba histolytica (http://www.sanger.ac.uk/Projects/E_histolytica/; http://www.tigr.org/db/e2k1/eha1/), Giardia intestinalis (McArthur et al., 2000), Trichomonas vaginalis (Tachezy et al., 2001), and the microsporidian Encephalitozoon cuniculi (Katinka et al., 2001), arose by a symbiogenic event from a common proteobacterial ancestor. Although these data lend support to the hypothesis that one of the functions retained by the relict mitochondrion of C. parvum is the assembly and maturation of [Fe–S], this hypothesis awaits both definitive localization of CplscS and CplscU to the organelle and measurement of enzyme activities for both enzymes.

ACKNOWLEDGEMENTS

At Rockefeller University, we thank Dr L. B. Sanchez for the initial phylogenetic analyses and Dr M. Müller for critically reading the manuscript. At the Wadsworth Center we thank the Molecular Genetics Core Facility for oligonucleotide synthesis and sequencing, and Dr N. M. Cirino for his patience during the completion of the manuscript. This work was supported in part by funds from the National Institutes of Health (NIH) Fogarty Emerging Infectious Disease Training Grant 1D43 TWO 0091 (J. S. K.) and two NIH Fogarty International Center (FIC) awards SRO3-TWO 1507-02 (J. S. K./F. S.) and SRO3-TWO 5536-02 (M. Müller/J. T.).

REFERENCES


