Proc. Natl. Acad. Sci. USA
Vol. 87, pp. 6097–6101, August 1990
Biochemistry

Molecular analysis of the hydrogenosomal ferredoxin of the anaerobic protist *Trichomonas vaginalis*

(amino acid sequence/gene structure/organelle evolution)

PATRICIA J. JOHNSON*, CHRISTINE E. D'OLIVEIRA*, THOMAS E. GORRELL†‡, AND MIKLÓS MÜLLER†

*Department of Microbiology and Immunology, University of California School of Medicine, Los Angeles, CA 90024; and †Rockefeller University, New York, NY 10021

Communicated by Christian de Duve, May 24, 1990 (received for review March 8, 1990)

ABSTRACT **We have determined the primary structure of the [2Fe-2S]ferredoxin of the anaerobic protist *Trichomonas vaginalis*. This protein, situated in the hydrogenosome, is composed of 93 amino acids. A comparison of *T. vaginalis* ferredoxin with >80 other ferredoxins shows the closest similarity to [2Fe-2S]putidaredoxin of the aerobic bacterium *Pseudomonas putida* and a lesser one to mitochondrial [2Fe-2S]ferredoxins of vertebrates. This similarity is reflected in the overall primary structure and in the spacing of cysteine residues coordinating the iron-sulfur center. The primary structure, but not the environment of the iron-sulfur center, also shows similarity with [2Fe-2S]ferredoxins of photosynthetic organisms and halobacteria. We have cloned and analyzed the *T. vaginalis* ferredoxin gene. The gene is present in a single copy and devoid of introns. It gives rise to a transcript with unusually short 5′ and 3′ untranslated regions of 16 and 18 nucleotides, respectively. DNA sequence analysis of the gene predicts an additional 8 amino acids at the amino terminus which are absent from the purified protein. This amino-terminal region of the protein is characterized by properties typical of mitochondrial presequences.**

The presence of mitochondria is regarded as a universal hallmark of eukaryotic cells. Certain protists are an exception to this rule. Although some have no membrane-bounded organelles of energy metabolism, others contain hydrogenosomes, unusual organelles surrounded by two membranes (1, 2). These organelles are the site of oxidative decarboxylation of pyruvate with the formation of acetyl-CoA. This process is mediated by pyruvate:ferredoxin oxidoreductase and is coupled to anaerobic formation of H_2 catalyzed by hydrogenase. The electron carrier linking these reactions is an iron-sulfur protein, ferredoxin. In subsequent reactions the energy of the thioester bond of acetyl-CoA is conserved by substrate-level phosphorylation of ADP. This pathway and its enzymes render the hydrogenosomes markedly different from mitochondria (1, 2).

Hydrogenosomes were originally described from trichomonad flagellates, among them *Trichomonas vaginalis*, a parasite of humans (1, 2). Subsequently they have been detected in a number of free-living anaerobic ciliates (3, 4) and in ciliates and a fungus of the rumen (2). The biological nature and origin of hydrogenosomes remain enigmatic but an endosymbiotic origin is favored. The main difference in opinion centers on whether hydrogenosomes arose by conversion from mitochondria (4, 5) or through independent endosymbiosis involving anaerobic bacteria (1, 2, 6).

A comparison of hydrogenosomal constituents with those of other organelles and of prokaryotes should lead to a better understanding of the nature of hydrogenosomes. We have focused on ferredoxin, the main hydrogenosomal electron carrier. Ferredoxins of various prokaryotes, protists, and eukaryotic organelles have been extensively studied (e.g., refs. 7–10), providing rich material for comparative studies. Ferredoxins of trichomonad flagellates have been purified and characterized as [2Fe-2S]ferredoxins (11, 12). *T. vaginalis* ferredoxin has a midpoint potential of −320 mV (ref. 13; T.E.G., unpublished results) and an iron-sulfur center with axial symmetry (11, 13).

In this communication, we describe the primary structure of this protein and of its gene.§ We show that *T. vaginalis* ferredoxin belongs to a family of [2Fe-2S]ferredoxins that occur in certain aerobic bacteria and vertebrate mitochondria. Furthermore, our analysis of the trichomonad ferredoxin gene establishes a framework for studying the molecular biology of these primitive eukaryotes.

MATERIALS AND METHODS

Organism. *T. vaginalis* strain C1 (ATCC 30001) was cultured and harvested as described earlier (11).

Polyclonal Antiserum and Protein Sequencing. *T. vaginalis* ferredoxin was purified as described (11). Polyclonal antibodies reacting with this protein were developed in New Zealand White rabbits. The sera were tested with enzyme-linked immunoassay using either the purified protein or whole cell extracts on Western blots.

For amino acid sequencing, apoferredoxin was prepared by precipitation of the native protein with trichloroacetic acid (11). The apoprotein was fragmented with CNBr and, after propylation (14), by treatment with trypsin. The peptides were separated by HPLC. The intact protein and the peptides were carboxymethylated before sequencing. Gas-phase sequencing was performed on the whole protein, yielding results through residue 58, on a CNBr fragment, yielding results for residues 33–85, and a tryptic fragment, giving results for residues 76–93.

Recombinant DNA Methods. A *T. vaginalis* cDNA library was prepared in λgt11 by means of standard techniques (15). The library was screened for the expression of fusion proteins that were immunologically reactive with the antiferredoxin serum. The cDNA was excised, ligated into pBluescript (Stratagene), and sequenced by the method of Sanger *et al.* (16) with the help of a Sequenase kit (United States Biochemical).

A genomic DNA library was constructed in λEMBL3 from a partial *Mbo* I digest of *T. vaginalis* DNA. A ferredoxin genomic clone was isolated from the library by hybridization with ^{32}P-labeled ferredoxin cDNA, according to standard procedures. A 1.3-kilobase (kb) *Cla* I fragment, containing the ferredoxin gene and about 900 base pairs (bp) of flanking

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

‡Present address: EPCO Inc., Woodbury, CT 06798.
§The sequence reported in this paper has been deposited in the GenBank data base (accession no. M33717).

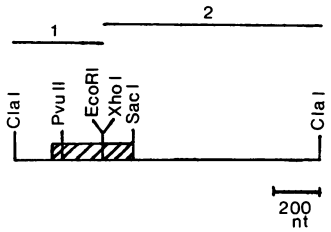

FIG. 1. Restriction map of *T. vaginalis* ferredoxin genomic subclone pGFD-2. The hatched box represents the coding region of the gene. Fragments 1 and 2, shown above the map, were used to probe DNA and RNA blots and to map the ends of the ferredoxin transcript. nt, Nucleotides.

sequences, was subcloned into pEMBL9. This subclone (pGFD-2; Fig. 1) was partially sequenced as described above.

DNA Extraction and Southern Hybridization. High molecular weight DNA was isolated from *T. vaginalis* as described (17). DNA (5 μg) was digested with restriction endonucleases, size-fractionated in agarose gels, and transferred onto nitrocellulose filters. The filters were hybridized at 65°C to random-primed ^{32}P-labeled DNA probes in 3× SSC (1× SSC = 150 mM NaCl/15 mM sodium citrate, pH 7), 1% sodium dodecyl sulfate, 0.1% sodium pyrophosphate, 5× Denhardt's solution (1× Denhardt's solution = 0.02% Ficoll/0.02% polyvinylpyrrolidone/0.02% bovine serum albumin), and 100 μg of denatured salmon sperm DNA per ml. Filters were washed at a final stringency of 0.3× SSC at 65°C.

Isolation and Analysis of RNA. Total RNA was isolated by a modified guanidinium thiocyanate procedure (17). Poly(A)$^+$ RNA, used to construct a λgt11 cDNA library (15), was purified on poly(dT) columns. Northern blots were prepared from total RNA, size-fractionated in 1% agarose/formaldehyde gels, and transferred to nitrocellulose filters. The filters were hybridized and subsequently washed as described above for DNA blot hybridizations.

S1 nuclease and exonuclease VII mapping was carried out as described (18). Appropriate fragments were isolated from pGFD-2 (Fig. 1), ^{32}P-labeled, and ethanol precipitated together with 25 μg of total *T. vaginalis* RNA. DNA·RNA hybrids were allowed to form in hybridization buffer for 15 hr at 42°C. After nuclease digestion, protected DNA fragments were separated on a denaturing 5% polyacrylamide gel.

RESULTS

Determination of the Amino Acid Sequence and DNA Cloning. The amino acid sequence of purified *T. vaginalis* ferredoxin was established by gas-phase sequencing, with only residues 63, 72, and 74 left unidentified (Fig. 2). The protein was found to consist of 93 amino acids. Only glycine was detected as the first residue (6.4 nmol for 6 nmol of protein analyzed). Glycine was also the first residue detected in partial liquid-phase sequencing of apoprotein samples purified in two additional experiments. The molecular mass of the purified apoprotein calculated from its primary structure is 9930, lower than reported earlier from less direct evidence (11).

To complete and extend these findings, we constructed a *T. vaginalis* cDNA expression library in λgt11. Three independent cDNA clones were isolated by screening with anti-ferredoxin serum. One clone (pCFD-1) was completely sequenced and shown to encode the last 60 amino acids of ferredoxin (results not shown). We also isolated a genomic clone from a library constructed in λEMBL3, using the *T. vaginalis* cDNA as a probe. A map of the subclone (pGFD-2), derived from the genomic clone, is shown in Fig. 1. This clone contains the entire coding region and 150 and 800 bp 5′ and 3′ of the ferredoxin gene, respectively. We have sequenced all but the 3′ terminal 500 bp of pGFD-2.

The amino acids predicted by DNA sequencing are identical to those identified by protein sequencing (Fig. 2). In addition, three internal amino acids, that had not been identified by protein sequencing, were identified.

Interestingly, the protein predicted by the DNA sequence analysis is longer than that determined by protein sequencing (Fig. 2). An ATG codon located eight amino acids before the first amino acid of the mature protein is the only possible initiation codon since there is a stop codon three nucleotides upstream from it.

Given the hydrogenosomal location of ferredoxin (1, 2, 11), these results suggest that this eight-residue amino-terminal extension may play a role in the targeting of this protein to the organelle. It is noteworthy that this octapeptide resembles leader sequences of mitochondrial proteins (19, 20) in its amino acid composition and in having the potential to form an amphiphilic α-helix (Fig. 3).

The primary structure of *T. vaginalis* ferredoxin was compared with those of >80 ferredoxins of prokaryotic, protist, chloroplast, and mitochondrial origin. The greatest overall similarity was found with proteins belonging to two different classes of ferredoxins (10). The first consists of [2Fe-2S]ferredoxins that participate in mixed function oxidase reactions. Of these the most similar was putidaredoxin of *Pseudomonas putida* (21), revealing 27% identity plus 43% conservative substitutions in a 93-amino acid overlap (Fig. 4). The protein was also similar, but less so, to mitochondrial ferredoxins (22, 25, 26). There is a 28% identity with porcine ferredoxin (22) plus 53% conservative substitutions in a 72-amino acid overlap (Fig. 4). *T. vaginalis* ferredoxin shares

```
CGA TTT ATT GAA TTT TTT CTT TCT TAA TAA TTA TAT TTA ATA TAA ACT CTA CAC TAT
TAA AAG TTA AAT GGC CGA AGA TAA CTT GAT TTG ATA AAT CAC ATT CAA TTG ATT GAG
                                                    MET LEU SER GLN VAL CYS
CTT TGT ATT CAA AAT ATT TAC TTC ACT TCT CTT TAG CGA ATG CTC TCT CAA GTT TGC
        *
ARG PHE GLY THR ILE THR ALA VAL LYS GLY GLY VAL LYS LYS GLN LEU LYS PHE GLU
CGC TTT GGA ACA ATC ACA GCC GTC AAG GGT GGT GTC AAG AAG CAA CTC AAG TTC GAA

ASP ASP GLN THR LEU PHE THR VAL LEU THR GLU ALA GLY LEU MET SER ALA ASP ASP
GAT GAC CAG ACA CTC TTC ACA GTT CTT ACA GAA GCC GGC CTC ATG TCA GCT GAT GAC

THR CYS GLN GLY ASN LYS ALA CYS GLY LYS CYS ILE CYS LYS HIS VAL SER GLY LYS
ACA TGC CAG GGC AAC AAG GCT TGC GGC AAG TGC ATC TGC AAG CAC GTT TCC GGC AAG

                    •                       •       •
VAL ALA ALA ALA GLU ASP ASP GLU LYS GLU PHE LEU GLU ASP GLN PRO ALA ASN ALA
GTC GCT GCT GCT GAG GAT GAT GAG AAG GAA TTC CTC GAG GAT CAG CCA GCT AAC GCT

ARG LEU ALA CYS ALA ILE THR LEU SER GLY GLU ASN ASP GLY ALA VAL PHE GLU LEU
CGC CTT GCT TGC GCT ATC ACA CTC AGT GGT GAA AAC GAT GGT GCT GTT TTC GAG CTC
                    ↓↓↓
TAA ATA ATT GAA AGT TTA TTA AAT TGT TTT TGA TTT TTT TCC AAT ACT TAA GTT ACA

TTC AAA ATG AAT CGC TTT ATT TTT TGT TTT TAT GTT GTC ACA GTC TAT ATT TCG GAT

CAA CTG TAA TAG AGG
```

FIG. 2. DNA and protein sequences of *T. vaginalis* ferredoxin. The asterisk marks the first amino acid determined by direct sequencing of the protein. An amino-terminal extension of eight amino acids, predicted from the DNA sequence, is boxed. Three internal amino acids that were identified only by DNA sequencing are designated by dots. Arrows mark the beginning and end of the transcript. Possible regulatory sequences are underlined.

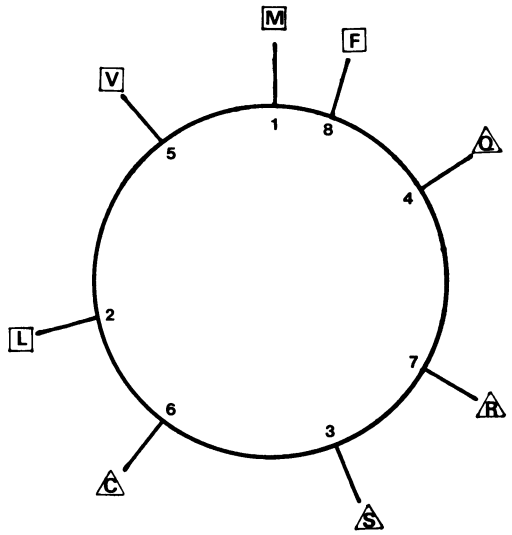

FIG. 3. Helical wheel projection showing an end-on view of a putative eight-amino acid α-helix formed by the amino-terminal part of *T. vaginalis* ferredoxin. Nonpolar residues are placed in squares; polar or charged ones are placed in triangles. Note that the helix has opposing polar and nonpolar surfaces.

with this group of proteins the CX_5CX_2C sequence containing the first three cysteine residues (10). A similarity was also noted with [2Fe-2S]ferredoxins of certain photosynthetic organisms and halophilic bacteria. The most pronounced is to ferredoxin II of cyanobacterium *A. sacrum* (23) (Fig. 4), with 28% identity and 40% conservative substitutions in a 93-amino acid overlap. *T. vaginalis* ferredoxin showed no similarity to the 2[4Fe-4S]ferredoxins of the anaerobic protist *Entamoeba histolytica* (27) and anaerobic bacteria (28).

Molecular Analysis of the Ferredoxin Gene. Genomic DNA digested with one of seven restriction enzymes, blotted, and probed with the coding region of ferredoxin (fragment 1, Fig. 1) invariably shows a single hybridizing band indicating the presence of a single gene (representative results are shown in Fig. 5*A*). Similarly, analysis of *T. vaginalis* RNA shows a single, stable, polyadenylylated ferredoxin mRNA of ≈400 nucleotides (Fig. 5*B*).

We have used S1 nuclease protection (17) to map the 5′ end of the ferredoxin transcript. Using a 750-nucleotide probe extending 5′ of the coding region (fragment 1, Fig. 1), we observed a single band of ≈235 nucleotides (Fig. 6*A*). This band is absent from controls in which *T. vaginalis* RNA was replaced by yeast RNA. These results map the start of the

```
                        10        20        30
T.v.          GTITAVKGGVKKQLKFEDDQTLFTVLTEAGL-MSADDT
P.p.        SKVVYYSHDGTRRELDVADGVSLMQAAVSNGI-YDIVGD
S.s.  SSSEDKITVHFINRDGKTLTTQGKVGDSLLDVVIENNLDIDGFGA
A.s.       ATYKVTLINEEEGINAILEVADDQTILDAGEEAGL--DLPSS

      40        50        60        70
CQGNKACGKC---ICKHVSGKVAAAEDDEKEFLE----DQPANA--RLAC
CGGSASCATCHVYVNEAFTDKVPAANEREIGMLECVTAELKPNS--RLCC
CEGTLACSTCHLIFEDHIFEKLEAITDEENDMLDLAY-GLTDRS--RLGC
CRAG-SCSTC---AGKLVSG--AAPNQDDQAFLD----DQLAAGWVMTC

80        90
ATTLSGENDGAVFEL
QIIMTPELDGIVVDVPDRQW
QICLTKAMDNMTVRVPEAVADARES
VAYPTGDCTIMTHQESEVL
```

FIG. 4. Comparison of the amino acid sequence of purified *T. vaginalis* ferredoxin (T.v.), *P. putida* putidaredoxin (21) (P.p.), porcine mitochondrial ferredoxin (22) (S.s.), and *Aphanothece sacrum* ferredoxin II (23) (A.s.). The Protein Identification Resource of the National Biomedical Research Foundation (release 23) was searched for sequences showing similarities to *T. vaginalis* ferredoxin and amino acids were aligned for maximum similarity with the help of the FASTA program (24). Exact matches are boxed.

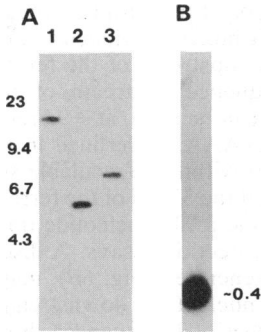

FIG. 5. Analysis of total genomic DNA (*A*) and total RNA (*B*) of *T. vaginalis*. (*A*) Southern blot analysis of DNA. DNA samples were digested with *Hin*dIII (lane 1), *Pst* I (lane 2), and *Xho* I (lane 3). DNA fragments were size-fractionated in agarose, transferred to nitrocellulose, and hybridized to a ^{32}P-labeled ferredoxin probe. (*B*) Northern blot analysis of RNA. RNA was size-fractionated in agarose/formaldehyde gels and was transferred and hybridized as in *A*. Sizes are in kilobases.

mRNA 15–20 nucleotides upstream of the initiation codon. Results of primer extension analysis excluded the possibility that the S1 data represent an intron in the 5′ untranslated region. These experiments mapped the 5′ end of the mRNA within a few nucleotides of that determined in S1 protection experiments, confirming the results shown in Fig. 6*A* (data not shown).

To map the 5′ end of the ferredoxin mRNA precisely, we used anchored polymerase chain reaction (29) to produce cDNA clones corresponding to the 5′ end of the mRNA (results not shown). These allowed us to map the 5′ end precisely 16 nucleotides upstream of the start of translation (Fig. 2). It is noteworthy that the mRNA sequence corresponds exactly to the genomic DNA sequence, indicating the absence of transsplicing of mRNA such as occurs in kinetoplastid protists (e.g., refs. 30–32).

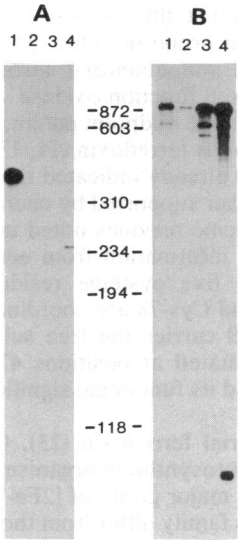

FIG. 6. Nuclease mapping of the ends of the *T. vaginalis* ferredoxin transcripts. (*A*) S1 nuclease mapping of the 5′ end. DNA fragment 1 (Fig. 1) was ^{32}P-labeled and annealed to total RNA. Lane 1, fragment 1 only; lane 2, fragment 1 plus S1 nuclease; lane 3, fragment 1 plus yeast tRNA and S1 nuclease; lane 4, fragment 1 plus *T. vaginalis* RNA and S1 nuclease. (*B*) Exonuclease VII mapping of the 3′ end. Lanes contain the same as in *A*, except that DNA fragment 2 was ^{32}P-labeled and used in annealing reactions, and exonuclease VII was used in the digestions instead of S1 nuclease. Sizes are given in nucleotides.

There are no typical transcription regulation motifs, such as TATA or CAAT boxes positioned at approximately −25 or −70, respectively, upstream of the ferredoxin gene. In the ≈150-bp-long sequence upstream of the gene only one TATA-like motif can be seen, at −97 relative to the start of transcription (TATAAA, underlined in Fig. 2). No CAAT-like box is present within the available sequence.

We have mapped the 3′ end of the ferredoxin mRNA using exonuclease VII and a 900-nucleotide fragment (fragment 2, Fig. 1) in RNA protection assays. A discrete band of ≈110 nucleotides was generated (Fig. 6*B*), mapping the 3′ end of the mRNA 18–20 nucleotides downstream of the stop codon (see Fig. 2). Analysis of 3′ sequences shows that the poly(A) site(s) is flanked by possible signals (underlined in Fig. 2) that comply, in part, with the consensus sequences AATAAA and YGTGTTYY, thought to direct polyadenylylation in higher eukaryotes (33, 34). Interestingly, the first A in the former of these sequences is also the last nucleotide of the stop codon TAA. Whether these 5′ or 3′ sequence elements are functional remains to be tested.

As this is, to our knowledge, the only gene to have been cloned and analyzed from a trichomonad, it is of interest to note that there are no introns in the gene (Fig. 2).

DISCUSSION

In the present study, we have determined the primary structure of the hydrogenosomal [2Fe-2S]ferredoxin of the anaerobic protist *T. vaginalis* by protein and DNA sequencing. In addition, we have characterized in detail the structures of the gene and of its transcript.

The primary structure of *T. vaginalis* ferredoxin was found to be most similar to two ferredoxins of different origin—on one hand, to putidaredoxin of the aerobic bacterium *Pseudomonas putida* (21) and, on the other, to the [2Fe-2S]ferredoxin II of a cyanobacterium, *Aphanothece sacrum* (23). Somewhat less similarity was noted to ferredoxins of vertebrate mitochondria (22, 25, 26) and of some photosynthetic and halophilic organisms.

These proteins belong, however, to two distinct families of [2Fe-2S]ferredoxin that differ in their metabolic function and in a number of their properties (10). One of these comprises putidaredoxin and mitochondrial ferredoxins, which are components of mixed function oxidase systems (10, 35, 36). Similarities noted in the midpoint potential, optical, and EPR spectra of *T. vaginalis* ferredoxin (11, 13) and of proteins of this family (35, 36) already indicated their possible relationship (11), now further supported by characteristic spacing of the first three cysteine residues noted in this study.

On the basis of information from adrenodoxin (37) it is likely that of the five cysteine residues found, Cys-38, Cys-44, Cys-47, and Cys-78 are coordinating the iron-sulfur center, and Cys-49 carries the free sulfhydryl group. The CXC sequence situated at positions 47–49 is unusual for ferredoxins (38) and its functional significance remains to be established.

The cyanobacterial ferredoxin (23), together with ferredoxins of other photosynthetic organisms and halobacteria, belongs to another major group of [2Fe-2S]ferredoxins (10). Ferredoxins of this family differ from those of the first group in the spacing of the cysteine residues coordinating the iron-sulfur center (10) and in their midpoint potentials and EPR spectra (39).

T. vaginalis ferredoxin is the major electron transport component of hydrogenosomes. It links pyruvate:ferredoxin oxidoreductase to hydrogenase (1, 2, 40). It is not surprising that an organelle that produces hydrogen contains a ferredoxin. This protein is the electron donor for hydrogenase in many organisms (9). What is peculiar, however, is that this ferredoxin is not a 2[4Fe-4S]ferredoxin, which is the usual electron carrier in hydrogen-producing bacteria, but is related to ferredoxins that are components of mixed function oxidase systems (35, 36).

Ferredoxins of other anaerobic protists have not been characterized save the ferredoxin of *E. histolytica*, an organism which contains neither mitochondria nor hydrogenosomes. This is a 2[4Fe-4S]ferredoxin (27, 41) with marked similarities in its amino acid sequence to *Clostridium pasteurianum* ferredoxin (28). *E. histolytica* ferredoxin also serves as electron acceptor for pyruvate:ferredoxin oxidoreductase (41). It is noteworthy that two main groups of anaerobic protists contain distinct types of ferredoxins, although both are acceptors for similar enzymes (2).

The DNA sequence reveals the presence of an amino-terminal octapeptide that is absent from the purified protein. The simplest interpretation of this finding is that the additional amino acids comprise a leader sequence that is proteolytically removed upon translocation into hydrogenosomes. Although we cannot rigorously exclude the possibility of proteolytic degradation of ferredoxin during its purification, we regard it unlikely. In three independently purified samples, only glycine was detected as the amino-terminal residue, indicating the absence of random degradation. The purified protein was functionally active in a number of assays (11, 40). Unfortunately, our attempts to confirm this assumption *in vivo* by pulse–chase experiments have been thwarted by the apparent rapid processing of the protein.

Although the presence of a leader on this hydrogenosomal protein remains to be proven, the similarity of the overall features of this octapeptide to those of mitochondrial leader peptides (presequences) is striking (e.g., refs. 19, 20). For example, this short peptide includes all 3 amino acids that are typically overrepresented in mitochondrial leader sequences (leucine, serine, arginine). Furthermore, there is an arginine at −2 from the predicted cleavage site (20) and the peptide is hydrophobic and devoid of negatively charged amino acids. Finally, it appears to have the potential to form an amphiphilic α-helix, a characteristic of mitochondrial presequences that is regarded to be essential for function. Whether such a short α-helix (two turns) could function is unclear, however. It should be noted that mitochondrial presequences as short as 12 amino acids (three helical turns) have been shown to be translocationally effective (42). Mitochondrial and chloroplast proapoferredoxins have leader sequences. However, these are significantly longer than the octapeptide noted here (26, 43, 44).

In addition to providing a definitive primary structure of ferredoxin, our characterization of the *T. vaginalis* ferredoxin gene provides information on the structure and expression of what is to our knowledge the first gene to be cloned from a trichomonad flagellate. Analysis of the transcript of this single-copy gene reveals the presence of unusually short 5′ and 3′ untranslated regions (16 and 18 nucleotides, respectively). It is noteworthy that very short untranslated regions have invariably been observed in the limited number of genes that have been analyzed from other anaerobic protists, *E. histolytica* (26, 45) and *Giardia lamblia* (46). Furthermore, the only other *T. vaginalis* protein-coding gene for which the 3′ end has been mapped likewise has a short untranslated region of only 20 nucleotides (C. Lahti and P.J.J., unpublished observations). The possibility that this is the consequence of a limited genome size can be ruled out, as genomes of these organisms are as large as those of other protists that are not characterized by this feature. It may, however, reflect some unusual mechanism(s) used by these organisms for the formation of mRNAs.

The lack of TATA and CAAT sequence motifs situated at approximately 25 and 70 nucleotides 5′ of the start of the mRNA is not surprising, as transcription units from lower eukaryotes often lack these regulatory signals. Interestingly,

the only TATA box identified 5′ of the *T. vaginalis* ferredoxin transcript, situated at −97, has considerable homology to TATA boxes situated at −80 in the *E. histolytica* actin gene (45), −102 in a *G. lamblia* β-tubulin gene (46), and −121 in the *E. histolytica* ferredoxin gene (27). Examination of these sequences reveals a consensus sequence of TATAAAA. However, the significance of this consensus is still uncertain in view of the small number of genes analyzed so far and the general AT-richness of examined 5′ sequences. Nonetheless, there is precedence for distal TATA boxes, as those of yeast genes are typically found at −40 to −120 5′ of the cap site of the mRNA (47). Whether this 5′ distal sequence motif as well as those identified in the 3′ untranslated region (underlined in Fig. 2) of the *T. vaginalis* gene are functional awaits further analysis.

The similarity of the primary structure of *T. vaginalis* ferredoxin to functionally different ferredoxins present in organisms of different systematic positions is a challenging observation. Data on other hydrogenosomal ferredoxins will be needed to assess the phylogenetic significance of this finding. The similarities, in amino acid sequence and in the environment of the iron-sulfur center, between *T. vaginalis* ferredoxin and ferredoxins of an aerobic bacterium and of mitochondria are intriguing. Taken by themselves, these similarities and the presence of an amino-terminal extension comparable to mitochondrial presequence could support the idea that hydrogenosomes are derived from mitochondria or from a common progenitor of these two organelles (4, 5). Nevertheless, one should not lose sight of the fundamental differences in enzymatic composition between mitochondria and hydrogenosomes, notably the presence of a pyruvate dehydrogenase complex and the lack of hydrogenase in mitochondria as against the presence of pyruvate:ferredoxin oxidoreductase and hydrogenase in hydrogenosomes (1, 2). These differences seem to be sufficiently large as to argue against a direct conversion of mitochondria into hydrogenosomes. In any case, the question of the origin of hydrogenosomes remains open at present. Further comparative data on different constituents of this organelle from various protists could shed light on this problem.

We thank Dr. Martin M. Weiss for providing the polyclonal antiserum, Audrey Sher and Deborah Mulready for technical assistance, Dr. Margaret Perkins, Dr. David Campbell, Carol Lahti, and Doris Quon for helpful discussions, and Dr. David Campbell and Dr. Alexander van der Bliek for their critical comments on the manuscript. The purified protein was sequenced by the Protein Sequencing Facility of The Rockefeller University. This work was supported by grants from the National Institute of Allergy and Infectious Diseases (AI 27857 to P.J.J. and AI 11942 to M.M.) and the National Science Foundation (PCM 8415003 to M.M.).

1. Müller, M. (1980) *Symp. Soc. Gen. Microbiol.* **30,** 127–142.
2. Müller, M. (1988) *Annu. Rev. Microbiol.* **42,** 465–488.
3. Finlay, B. & Fenchel, T. (1989) *New Scientist* **123** (1671), 66–69.
4. Finlay, B. J. & Fenchel, T. (1989) *FEMS Microbiol. Lett.* **65,** 311–314.
5. Cavalier-Smith, T. (1987) *Ann. N.Y. Acad. Sci.* **503,** 55–72.
6. Whatley, J. M., John, P. & Whatley, F. R. (1979) *Proc. R. Soc. London Ser. B* **204,** 165–187.
7. Cammack, R. (1983) *Chem. Scr.* **21,** 87–95.
8. George, D. G., Hunt, L. T., Yeh, L.-S. L. & Barker, W. C. (1985) *J. Mol. Evol.* **22,** 20–31.
9. Yasunobu, K. T. & Tanaka, M. (1973) in *Iron-Sulfur Proteins*, ed. Lovenberg, W. (Academic, New York), Vol. 1, pp. 27–130.
10. Meyer, J., Bruschi, M. H., Bonicel, J. J. & Bovier-Lapierre, G. E. (1986) *Biochemistry* **25,** 6054–6061.
11. Gorrell, T. E., Yarlett, N. & Müller, M. (1984) *Carlsberg Res. Commun.* **49,** 259–268.
12. Marczak, R., Gorrell, T. E. & Müller, M. (1983) *J. Biol. Chem.* **258,** 12427–12433.
13. Chapman, A., Cammack, R., Linstead, D. J. & Lloyd, D. (1986) *Eur. J. Biochem.* **156,** 193–198.
14. Acharya, A. S. & Manjula, B. N. (1987) *Biochemistry* **26,** 3524–3530.
15. Young, R. A. & Davis, R. W. (1983) *Proc. Natl. Acad. Sci. USA* **80,** 1194–1198.
16. Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74,** 5463–5467.
17. Wang, A. L. & Wang, C. C. (1985) *J. Biol. Chem.* **260,** 3697–3702.
18. Berk, A. J. & Sharp, P. A. (1977) *Cell* **12,** 721–732.
19. Roise, D. & Schatz, G. (1988) *J. Biol. Chem.* **263,** 4509–4511.
20. von Heijne, G., Steppuhn, J. & Herrmann, R. G. (1989) *Eur. J. Biochem.* **180,** 535–545.
21. Koga, H., Yamaguchi, E., Matsunaga, K., Aramaki, H. & Horiuchi, T. (1989) *J. Biochem.* **106,** 831–836.
22. Akhrem, A. A., Lapko, A. G., Lapko, V. N., Morozova, L. A., Repin, V. A., Tishchenko, I. V. & Chashchin, V. L. (1978) *Bioorg. Khim.* **4,** 462–475.
23. Hase, T., Wakabayashi, S., Wada, K. & Matsubara, H. (1978) *J. Biochem.* **83,** 761–770.
24. Pearson, W. R. & Lipman, D. J. (1988) *Proc. Natl. Acad. Sci. USA* **85,** 2444–2448.
25. Kagimoto, K., McCarthy, J. L., Waterman, M. R. & Kagimoto, M. (1988) *Biochem. Biophys. Res. Commun.* **155,** 379–383.
26. Picado-Leonard, J., Voutilainen, R., Kao, L.-c., Chung, B.-c., Strauss, J. F., III, & Miller, W. L. (1988) *J. Biol. Chem.* **263,** 3240–3244.
27. Huber, M., Garfinkel, L., Gitler, C., Mirelman, D., Revel, M. & Rozenblatt, S. (1989) *Mol. Biochem. Parasitol.* **31,** 27–34.
28. Graves, M. C., Mullenbach, G. T. & Rabinowitz, J. C. (1985) *Proc. Natl. Acad. Sci. USA* **82,** 1653–1657.
29. Frohman, M. A., Dush, M. K. & Martin, G. R. (1988) *Proc. Natl. Acad. Sci. USA* **85,** 8998–9002.
30. Murphy, W. J., Watkins, K. P. & Agabian, N. (1986) *Cell* **47,** 517–525.
31. Sutton, R. E. & Boothroyd, J. C. (1986) *Cell* **47,** 527–535.
32. Laird, P. W., Zomerdijk, J. C. B. M., de Korte, D. & Borst, P. (1987) *EMBO J.* **6,** 1055–1062.
33. Proudfoot, N. J. & Brownlee, G. G. (1976) *Nature (London)* **263,** 211–214.
34. Birnstiel, M. L., Busslinger, M. & Strub, K. (1985) *Cell* **41,** 349–359.
35. Gunsalus, I. C. & Lipscomb, J. D. (1973) in *Iron-Sulfur Proteins*, ed. Lovenberg, W. (Academic, New York), Vol. 1, pp. 151–171.
36. Estabrook, R. W., Suzuki, K., Mason, J. I., Baron, J., Taylor, W. E., Simpson, E. R., Purvis, J. & McCarthy, J. (1973) in *Iron-Sulfur Proteins*, ed. Lovenberg, W. (Academic, New York), Vol. 1, pp. 193–223.
37. Cupp, J. R. & Vickery, L. E. (1988) *J. Biol. Chem.* **263,** 17418–17421.
38. Stout, C. D. (1982) in *Iron-Sulfur Proteins*, ed. Spiro, T. G. (Wiley, New York), pp. 97–146.
39. Tsukihara, T., Fukuyama, K. & Katsube, Y. (1987) in *Iron-Sulfur Protein Research*, eds. Matsubara, H., Katsube, Y. & Wada, K. (Japan Sci. Soc., Tokyo), pp. 59–68.
40. Steinbüchel, A. & Müller, M. (1986) *Mol. Biochem. Parasitol.* **20,** 57–65.
41. Reeves, R. E., Guthrie, J. D. & Lobelle-Rich, P. (1980) *Exp. Parasitol.* **49,** 83–88.
42. Hurt, E. C. & van Loon, A. P. G. M. (1986) *Trends Biochem. Sci.* **11,** 204–207.
43. Smeekens, S., van Binsbergen, J. & Weisbeek, P. (1985) *Nucleic Acids Res.* **13,** 3179–3194.
44. Mittal, S., Zhu, Y.-Z. & Vickery, L. E. (1988) *Arch. Biochem. Biophys.* **264,** 383–391.
45. Edman, U., Meza, I. & Agabian, N. (1987) *Proc. Natl. Acad. Sci. USA* **84,** 3024–3028.
46. Kirk-Mason, K. E., Turner, M. J. & Chakraborty, P. R. (1989) *Mol. Biochem. Parasitol.* **36,** 87–100.
47. Struhl, K. (1989) *Annu. Rev. Biochem.* **58,** 1051–1077.