Eubacterial HslV and HslU Subunits Homologs in Primordial Eukaryotes

Bernard Couvreur,‡ Rudy Wattiez,† Alex Bollen,‡ Paul Falmagne,† Dominique Le Ray,* and Jean-Claude Dujardin**

*Laboratory for Protozoology, Prince Leopold Institute for Tropical Medicine, Antwerpen, Belgium; †Laboratoire de Chimie Biologique, Université de Mons-Hainaut, Belgium; and ‡Applied Genetics, IBMM, Université Libre de Bruxelles, Gosselies, Belgium

ATP-dependent protease complexes are present in all three kingdoms of life, where they rid the cell of misfolded or damaged proteins and control the level of certain regulatory proteins. They include the proteasome in Eukaryotes, Archea, and Actinomycetales and the HslVU (ClpQY) complex in other eubacteria. We showed that genes homologous to eubacterial *HslV* (*ClpQ*) and *HslU* (*ClpY*) are present in the genome of trypanosomatid protozoa and are expressed. The features of the cDNAs indicated that bona fide trypanosomatid messengers had been cloned and ruled out bacterial contamination as the source of the material. The N-terminal microsequence of HslV from *Leishmania infantum* (Protozoa: Kinetoplastida) permitted the identification of the propeptide cleavage site and indicated that an active protease is present. High similarities (≥57.5%) with the prototypical HslV and HslU from *Escherichia coli* and conservation of residues essential for biochemical activity suggested that a functional HslVU complex is present in trypanosomatid protozoa. The structure of the N-termini of HslV and HslU further suggested mitochondrial localization. Phylogenetic analysis indicated that HslV and HslU from trypanosomatids clustered with eubacterial homologs but did not point to any particular bacterial lineage. Because typical eukaryotic 20S proteasomes are present in trypanosomatids, we concluded that the eubacterial HslVU and the eukaryotic multicatalytic protease are simultaneously present in these organisms. To our knowledge this is the first report of a eubacterial HslVU complex in eukaryotes and, consequently, of the simultaneous occurrence of both a proteasome and HslVU in living cells.

Introduction

ATP-dependent proteolytic complexes are present in all three kingdoms of life. They are self-compartmentalizing, multicomponent proteases that assume the vital function of ridding the cell of damaged, misfolded or abnormal proteins as well as controlling the levels of regulatory proteins. In Archea, Actinomycetales, and eukaryotes they are known as proteasomes, whereas in eubacteria other than Actinomycetales they include the Lon, ClpAP, ClpXP, and HslVU (ClpQY) complexes (De Mot et al. 1999; Voges, Zwickl, and Baumeister 1999).

In eukaryotes the 26S proteasome degrades proteins that have been targeted for destruction by the ubiquitin pathway. The catalytic core of the proteasome or 20S particle is an ~700-kDa complex of 28 subunits ranging from 20 to 35 kDa in molecular weight and from 4.5 to 8.7 in pI. They show a remarkable structural similarity and extensive sequence homologies at the amino acid level. As viewed under the electron microscope, the 20S complex is a hollowed cylinder made of four stacked rings of seven subunits each. Alpha subunits (α1–7) form the two outer rings and beta subunits (β1–7), the inner two ones. The latter assume the actual proteolytic activity (Coux, Tanaka, and Goldberg 1996; Bochtler et al. 1999; Voges, Zwickl, and Baumeister 1999). In Archea and Actinomycetales one or two representatives of each alpha and beta families are present (Coux, Tanaka, and Goldberg 1996; De Mot et al. 1999; Voges, Zwickl, and Baumeister 1999). On its own the eukaryotic 20S proteasome degrades only small peptides, and it is only when associated with the 19S particle or regulatory complex that it is able to process proteins. The 19S complex contains 17 subunits, including six ATPase subunits forming a hetero-hexameric ring at the base connecting with the 20S proteasome. The 19S complex is believed to recognize polyubiquinated proteins and unfold and translocate them into the 20S particle (Coux, Tanaka, and Goldberg 1996; Voges, Zwickl, and Baumeister 1999; Ferrell et al. 2000). In Archea and Actinomycetales several ATPases that are candidates for activators of their respective proteasomes have been identified (De Mot et al. 1999; Voges, Zwickl, and Baumeister 1999).

In eubacteria other than Actinomycetales, ATP-dependent proteolysis is performed by simpler complexes in which ATPase subunits associate with proteolytic ones, i.e., ClpA and ClpX with ClpP, HslU with HslV (Gottesman 1996; Bochtler et al. 1999; De Mot et al. 1999). Although the ClpX and HslU ATPases are both members of the HSP100/Clp family, the HslV protease is not related to serine protease ClpP but rather to β subunits of the proteasome. Indeed, HslV and proteasome β subunits share ~20% similarity over ~200 amino acids, a similar fold and threonine catalytic mechanism (Coux, Tanaka, and Goldberg 1996; Gottesman 1996; Bochtler et al. 1999). Two hexameric (HslV) or heptameric (ClpP) rings of identical proteolytic subunits form the proteolytic complex, which can be capped on either or both ends by a ring made of six identical ATPase subunits. Like the 19S particle in eukaryotes, they are thought to recognize and unfold the protein substrates and translocate them into the proteolytic chamber (Horwich, Weber-ban, and Finley 1999; Voges,

Key words: trypanosomatids, prokaryotes, ATP-dependent protease, evolution.

Address for correspondence and reprints: Bernard Couvreur, Virology Department, Veterinary and Agrochemical Research Centre (VAR), Groeselenberg 99, B-1180 Bruxelles, Belgium.
E-mail: becou@var.fgov.be.

Mol. Biol. Evol. 19(12):2110–2117. 2002
© 2002 by the Society for Molecular Biology and Evolution. ISSN: 0737-4038

Zwickl, and Baumeister 1999). Finally, in addition to activating proteolytic complexes, ATPases homopolymeric complexes have a chaperone capability of their own (Wickner, Maurizi, and Gottesman 1999). So far, the presence of a proteasome or an HslVU complex was considered mutually exclusive (De Mot et al. 1999). Here, we describe the presence of HslV and HslU subunits coding sequences and their expression in trypanosomatid protozoa. To our knowledge this is the first report of an HslVU complex in a eukaryote and, consequently, the first report of simultaneous occurrence of both a proteasome and an HslVU complex in living organisms.

Materials and Methods

Molecular Biology Techniques

Genomic DNA and polyA+ RNA were extracted from promastigotes of *Leishmania infantum* strain MHOM/67/MA(BE)/ITMAP 263, using standard procedures (Sambrook, Fritsch, and Maniatis 1989). PolyA+ RNA was reverse transcribed with the Superscript II system (Life Technologies) and a *Not*1oligod(T)18 primer (Pharmacia). All PCR amplification experiments were performed with the Expand HF PCR system (Roche Molecular Biomedicals) in 100-μl reaction volumes containing 1.5 mM $MgCl_2$. The TC1 Thermal Cycler (Perkin Elmer) was used. Cycling conditions were 2 min denaturation at 94°C, then 30 cycles of 1 min denaturation at 94°C, 1 min annealing at 50–65°C (according to the primer sets), 3 min elongation at 72°C, then a final elongation of 7 min at 72°C. PCR products were inserted into pCRII by the TOPO TA cloning method (Invitrogen). Nucleotide sequencing was performed on purified PCR products and on double-stranded plasmid DNA by the Dye Terminator Cycle Sequencing method, and the reaction products were analyzed on a 373 DNA sequencer (Applied Biosystems). New sequences received the following accession numbers: AJ298867 (*LiHsl*V); AJ421946, AJ428522, and AJ428521 (*LiHsl*U); AJ298868 (*TcHsl*V); and BN000068 (*TbHsl*U).

Computer-Assisted Analysis of Sequences and Database Searches

Most nucleotide and amino acid sequence analyses as well as searches in GENBANK and EMBL databases were performed using programs from the Genetic Computer Group (GCG) suite version 10. Prediction of leader peptides and subcellular localizations were performed with the SignalP 1.1 (Nielsen et al. 1997) and TargetP 1.0 (Emmanuelson et al. 2000) programs, respectively, at the Danish Center for Biological Sequence Analysis (CBS) server (http://www.cbs.dtu.dk/services/). Keyword interrogations of the expressed sequences tags database (DBEST) and of the genome database (entrez: genome) were performed at the National Center for Biotechnology Information server (http://www.ncbi.nlm.nih.gov). Finally, genome survey sequences (gss) from the *Trypanosoma brucei* genome project were interrogated by TBLASTN, using the prototypical *E. coli* HslV (accession number P31059) and HslU (accession number P32168) amino acid sequences as queries. The first hits obtained with the *E. coli* queries were used in turn to interrogate the databases.

Sequence Alignment and Phylogenetic Analysis

Amino acid sequences were aligned using the Clustal program (Thompson et al. 1997) as implemented in the Clustal X version 1.8 package (http://www-igbmc.u-strasbg.fr/BioInfo/). Alignments were visualized and edited using the Genedoc package version 2.6 (http://www.psc.edu/biomed/genedoc/). Phylogenetic analysis by the distance method, using neighbor-joining for the construction of dendrograms, was performed using programs in the Clustal package. The options ‘‘exclude positions with gaps’’ and ‘‘correct for multiple substitutions’’ were invoked as indicated in the text. Bootstrap values were calculated on 1,000 repeats of the initial data set. The analysis was repeated using programs from the Phylip package version 3.5c (http://evolution.genetics.washington.edu/phylip.html). One thousand repeats of the Clustal alignment were obtained by SEQBOOT. Distances were computed by PROTDIST, using Dayhoff PAM matrix. Nonrooted trees were constructed by neighbor-joining with NEIGHBOR and a consensus tree was obtained with CONSENSE.

For the study of HslV, representative proteasome β subunit amino acid sequences from *Arabidopsis thaliana*, *Homo sapiens*, and *Saccharomyces cerevisiae* as well as available ones from the trypanosomatids *Leishmania* and *Trypanosoma* were retrieved from public databases and aligned with Clustal. Profile alignment was used to align available HslVs to proteasome β subunits. First, an unrooted distance dendrogram was constructed by neighbor-joining from the alignment of the mature proteins only (i.e., omitting propeptides). Next, phylogenetic analysis was performed on the HslVs aligned with either the archetypal β subunit of *Thermoplasma acidophilum* (acc. P28061), or subunit β5 from *T. brucei* (acc. AJ132959) or *S. cerevisiae* (acc. P30656) as an out-group.

Similarly, available eubacterial HslU sequences were retrieved from the databases and aligned. Selected members of the ClpX were included too. Alignments of the entire proteins or only regions homologous to the N-terminal domain and the C-terminal domain of HslU were used for phylogenetic analysis.

Results

An HslV Homolog is Expressed in *L. infantum*, and Its Features Are Consistent with Its Being a Mitochondrial Protease

During a search for antigenic proteins in *L. infantum* promastigotes, analysis of a series of 21- to 31-kDa protein bands yielded a 27-residues microsequence (PEP1) that had a high similarity to the amino-terminal extremity of eubacterial HslV mature proteins and mature catalytically active proteasome β subunits (fig. 1*a*). The corresponding gene was then cloned from promastigotes using a two-step RT-PCR methodology. The 3′ end of the cDNA was amplified from cDNA by

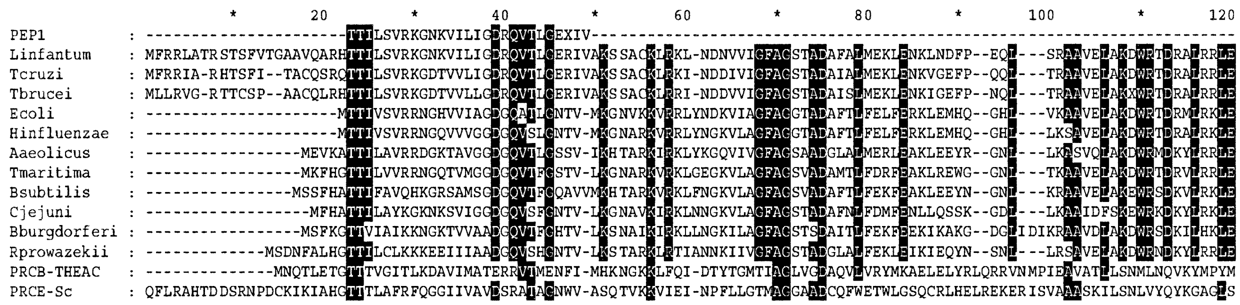

FIG. 1.—Alignment of HslV proteins from kinetoplastid protozoa and eubacteria. The Clustal X algorithm was used to align the N-terminal microsequence from an *L. infantum* promastigote protein (PEP1), with sequences calculated from full-length *L. infantum* cDNA, and partial cDNAs from *T. brucei* and *T. cruzi* with eubacterial HslVs (named after the species from which they come). The archetypal β proteasome subunit of the archea *Thermoplasma acidophilum* (PRCB-Theac) and a catalytically active β subunit from the yeast (PRCE-Sc) were included too. For brevity, only the first 120 positions of the alignment are shown.

using a degenerate upstream primer corresponding to the PEP1 micropeptide in association with an arbitrary downstream primer corresponding to the 5′ adapter of the reverse transcription primer. From the resulting partial nucleotide sequence, another downstream primer was designed and used to amplify the 5′ part of the cDNA in association with an upstream primer colinear to the *Leishmania* spliced leader. The sequences of the 5′ and 3′ parts of the gene overlapped perfectly over 300 nt; we thus reconstituted the full-length sequence (1,137 bp) of the mRNA coding for the HslV-like protein (assigned EMBL accession number AJ298867). The presence of a 5′ spliced leader, together with polypyrimidine stretches in the 3′ UTR, demonstrated the kinetoplastic nature of the cDNA. The sequence contained an open reading frame of 663 bp, of which 59% were G's or C's, which is close to the average 63% GC content of *Leishmania* coding sequences (Alonso, Guevara, and Ramirez 1992). The deduced 220 amino acid protein had a 24,124-Da calculated molecular weight and pI = 6.16. Perfect alignment was observed between the conceptual translation and the amino-terminal peptide PEP1 (fig. 1*a*). Also, this pointed to a cleavage site between His21 and Thr22, yielding a 21,678-Da mature protein with pI = 4.92. Moreover, the conservation of residues Thr1, Asp17, and Lys33 (fig. 1*a*) as well as Ser129 and Ser 169 (not shown) on the mature protein was consistent with its being an active threonine protease (Bochtler et al. 1999). No glycosylation sites were detected, and the typical motif for proteasome β subunits (PS00854 in Prosite) was not found. SignalP prediction also pointed to a signal peptide ending at residue 22, thus yielding a predicted mature protein in accordance with the observed N-terminal microsequence. Furthermore, the TargetP algorithm indicated a high probability of mitochondrial localization for the protein. This was further supported by (1) the prediction of an amphipatic helix structure for the N-terminal terminus using the program HelicalWheel, and (2) the presence of N-terminal residues that have been experimentally shown to direct a target protein in the yeast and in the *T. brucei* mitochondria (Häusler et al. 1997).

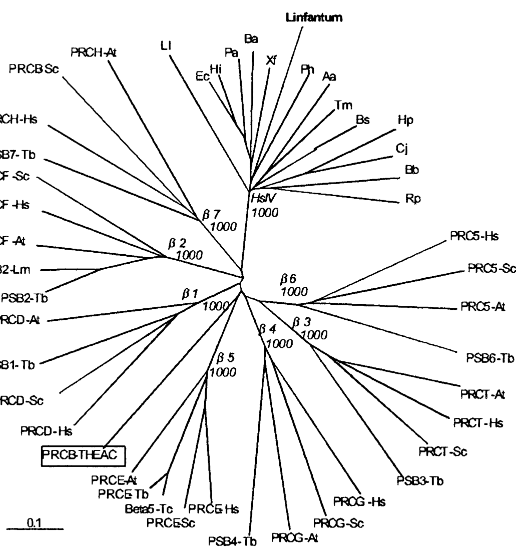

FIG. 2.—Phylogenetic analysis of HslV homologs and proteasome β subunits. The seven β subunits from *H. sapiens* (Hs), *A. thaliana* (At), and *S. cerevisiae* (Sc) as well as available ones from *T. brucei* (Tb), *T. cruzi* (Tc), and *L. major* (Lm) were aligned with PRCB-Theac (boxed), *L. infantum* (in bold), and representative eubacterial HslVs. An unrooted dendrogram was then constructed from the uncorrected distance matrix by neighbour-joining. Bootstrap values were calculated on 1,000 repeats of the initial alignment using a seed value of 111. Only values over 950 were reported near the relevant nodes. The scale bar represents the number of substitutions per site. All algorithms were part of the Clustal X package version 1.8.

Interrogation of the EMBL and GENBANK databanks showed that the calculated protein sequence had the highest similarity to eubacterial heat shock locus V protein (HslV) (e.g., 60% to the HslV of *E. coli*). Lower similarities were observed with archeal and eukaryotic proteasome subunits (e.g., 36% to *Thermoplasma acidophilum* PRCB and 31% to subunit PRCE of *S. cerevisiae*). Phylogenetic analysis by the distance method revealed that the protozoan HslV joined with eubacterial HslVs and away from trypanosomatid proteasome β subunits that clustered into the expected seven eukaryotic subfamilies (fig. 2). Bootstrap analysis confidently supported these observations (values >950/1,000). However, within the HslV family, the *L. infantum* protein could not be assigned to any particular bacterial lineage because the bootstrap values were lower than 750/1,000.

Therefore, the data indicated the presence of a genuine *HslV* homolog in the genome of *L. infantum*, ex-

pressing and processing into a mature threonine protease in the promastigote stage of the parasite, and suggested localization in the mitochondria.

HslV Homologs Are Also Expressed in the Related Trypanosomatids *T.* (*Trypanozoon*) *brucei* and *T.* (*Schizotrypanum*) *cruzi*

A search in the database of expressed sequence tags (DBEST, National Center for Biotechnology Information) pointed to possible bacterial HslV in *T.* (*S.*) *cruzi* epimastigotes (accession numbers AA952689, AA890822, AA952660) and *T.* (*T.*) *brucei rhodesiense* bloodstream forms (AA689168). The sequences were retrieved from the databases and analyzed. Moreover, one of the *T.* (*S.*) *cruzi* clones (clone TENS 1873) was kindly provided by the authors and resequenced (AJ298868). The typical spliced leader sequence was observed at the 5′ end of the *T.* (*S.*) *cruzi* cDNA, confirming the kinetoplastic origin of the material. Both *T.* (*S.*) *cruzi* and *T.* (*T.*) *brucei rhodesiense* partial calculated proteins presented high similarity with eubacterial HslVs (table 1), absence of the β subunit proteasome signature, possible cleavage site yielding a N-terminal threonine, and conservation of residues Asp17 and Lys33 in the mature protein (fig. 1*a*). Mitochondrial localization was suggested by the same arguments as described above (table 1). The observations thus indicated the presence of *HslV* homologs in members of the genus *Trypanosoma* and their transcription into mature messengers, at least in the developmental stages used for constructing the cDNA libraries.

ATPase Subunit HslU Homologs in *L. infantum*, *L. major*, and *T.* (*T.*) *brucei*

In *E. coli*, proteolytic HslV subunits associate with ~50-kDa ATPase HslU subunits to constitute the HslVU (ClpQY) proteolytic complex (Bochtler et al. 1999; De Mot et al. 1999). We found a kinetoplastid homolog of *HslU* in a preliminary release of the *L. major* genome (AL132763). Using a primer colinear to the 3′ end of the coding sequence and a miniexon primer, we isolated the corresponding cDNA from *L. infantum*. Two cDNA species of 1,700 and 2,100 bp were amplified (acc. AJ428522 and AJ428521). Sequence analysis indicated the presence of a spliced leader in each of the sequences at the 5′ end, followed by 5′ UTRs of 104 and 538 nt, respectively. The larger and smaller cDNAs differed only in the size of the 5′ UTR. The coding sequence was also amplified from genomic DNA (acc. AJ421946). Finally, nine overlapping partial DNA sequences (acc. AQ639522, AQ639633, AQ642698, AQ647404, AQ648527, AQ655149, AQ661372, AQ942802, AQ951737) coding for peptides highly similar to the HslU of *E. coli* were found in preliminary data of the *T. brucei* genome. We compiled them and reconstructed a 1,533-nt open reading frame (acc. BN000068).

The properties of the coding sequences and calculated HslUs from *L. major*, *L. infantum*, and *T. brucei* are presented in table 1 and figure 3. Features supporting

Table 1
Properties of *HslV* and *HslU* from Trypanosomatids Compared with the *E. coli* Homologs

Gene	Accession Number	Species	Coding Sequence Size (nt)	Coding Sequence G or C (%)	Calculated Protein Size (aa)	MW (Da)	pI Value	Putative Localization[a]	Similarity (%)
HslV (ClpQ)	AE000467	*E. coli*	531	53	175	18,961	6.37	NA	100
LiHslV	AJ298867	*L. infantum*	663	59	220 (199)[b]	24,124 (21,678)[b]	6.16 (4.92)[b]	mTP (0.849)	61
TcHslV	AJ298868	*T. cruzi*	NA	58	NA	NA	NA	mTP (0.805)	62
TbHslV	AA689168	*T. brucei*	NA	56	NA	NA	NA	mTP (0.770)	61
HslU (ClpY)	AE000467	*E. coli*	1,329	53	443	49,593	5.11	NA	100
LiHslU	AJ421946	*L. infantum*	1,512	60	504	55,673	5.12	mTP (0.434) or SP (0.556)	57.50
LmHslU	AL132763	*L. major*	1,515	61	505	55,808	5.07	mTP (0.419) or SP (0.426)	58
TbHslU.	BN000068	*T. brucei*	1,533	50	511	56,296	7.85	mTP (0.860)	58

Note.—MW, molecular weight; NA, not available; mTP, mitochondrial targeting peptide; SP, secretory pathway targeting peptide.
[a] Scores are given in parentheses.
[b] Values for the mature protein.

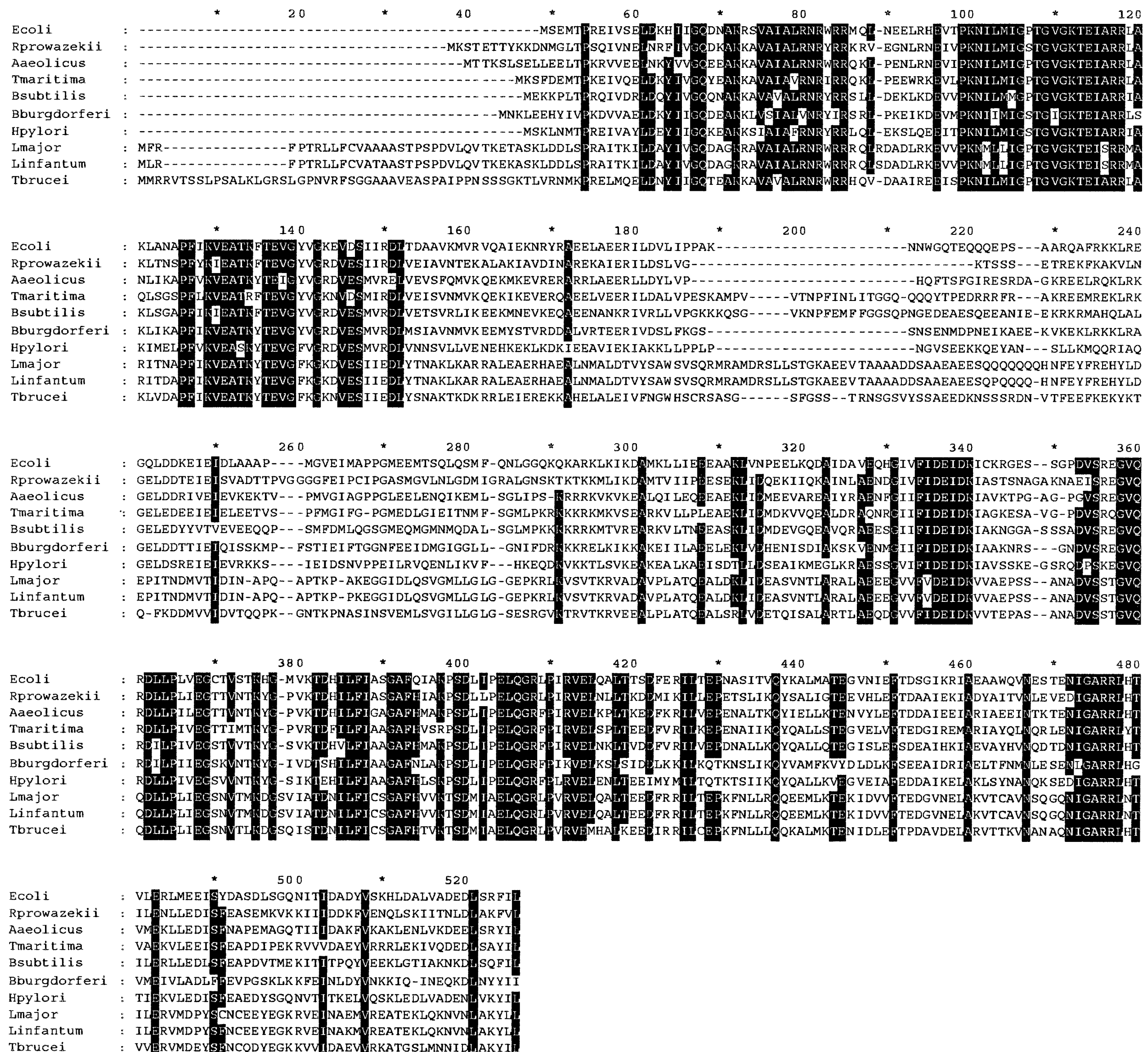

FIG. 3.—Alignment of eubacterial and trypanosomatid HslU. For brevity, only representatives of the main eubacterial groups were aligned to *L. infantum*, *L. major*, and *T. brucei* HslUs. Programs from the ClustalX package version 1.8 were used for the alignment, which was then visualized by Genedoc 2.

their identification with genuine HslU homologs are ~55-kDa molecular weight, presence of an ATP-GTP binding site (P-loop, Prosite P00017) near the N-terminus (positions 101–118 in fig. 3), and ~60% similarity to the prototypical HslU of *E. coli*. Furthermore, residues found to be essential to activities of the HslVU complex in mutagenesis studies of *E. coli* HslU (Song et al. 2000) were conserved. They were Arg325, Arg393, Glu321, Lys80, and Glu286. Residue Tyr91, which is central to the translocation pore (residues 87–95) and a key element in the model of Wang et al. (2001), was replaced by similar Phe91 as in the related FtsH protein (Wang et al. 2001) and in the *Helicobacter pylori* sequence (fig. 3). Throughout the bacterial and trypanosomatid sequences, high similarity was observed in regions corresponding to the highly structured N-terminal (N, residues 2–109 and 244–332 on the *E. coli* sequence) and C-terminal (C, residues 333–443) domains as defined by Bochtler et al. (2000) from crystal structures of HslU in *E. coli*. On the contrary, the intermediate domain (I, residues 110–243) was less conserved in bacteria and trypanosomatids alike and had several insertions and deletions. In mutagenesis studies the intermediate region supports some degree of deletions and insertions without loss of activity (Song et al. 2000).

In the coding sequences the G-C contents were close to the average values for the respective species (table 1; Alonso, Guevara, and Ramirez 1992). The N-termini showed mitochondrion-targeting conserved residues and amphipatic helix structures. TargetP predicted mitochondrial localization for the *T. brucei* protein and mitochondrial or secretory pathway for the *Leishmania* homologs (table 1). Phylogenetic analysis indicated that

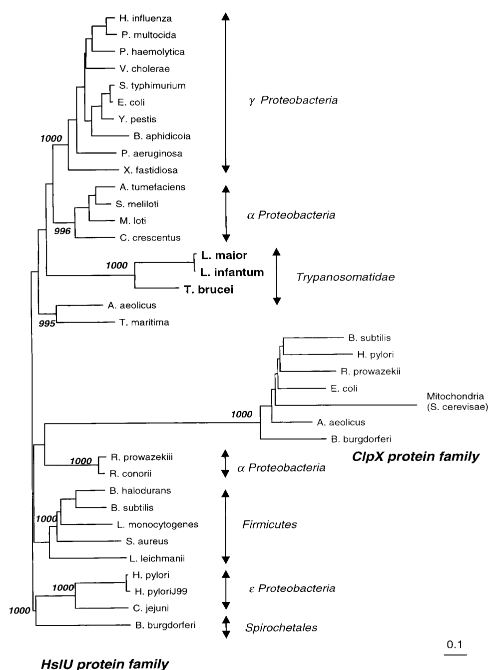

FIG. 4.—Phylogenetic analysis of HslU amino acid sequences from eubacteria and trypanosomatid protozoa. Available HslU peptide sequences were recovered from databases and aligned to the *L. infantum*, *L. major*, and *T. brucei* HslU peptide sequences and selected members of the ClpX family using Clustal X version 1.8. A neighbour-joining distance dendrogram was constructed using programs in the Phylip package. Bootstrap values over 950/1,000 are indicated near the corresponding nodes. The scale bar represents the number of substitutions per site.

trypanosomatid sequences clustered together with bacterial HslUs and away from members of the ClpX family (fig. 4). This was supported by high bootstrap values (>950/100). However, they could not be confidently assigned or joined to any particular bacterial clade (bootstrap values <750/1,000). Altogether, these data indicated the presence of *HslU* homologs in the genome of trypanosomatid protozoa and suggested conservation of fold and biochemical activity as well as possible mitochondrial targeting.

Discussion

We describe eubacterial HslV and HslU homologs in trypanosomatid (kinetoplastid) protozoa from the genera *Leishmania* and *Trypanosoma*. The structure of cDNA sequences indicated that the genes were transcribed and correctly processed into mature messengers. Furthermore, the presence of the typical spliced leader sequence at the 5′ end of each clone indicated that they originated from bona fide kinetoplastid messengers and ruled out contaminating bacterial material as the source of the sequences. In the case of HslV, a protein microsequence from *Leishmania* promastigotes confirmed that the protein is expressed. The micropeptide corresponded to the amino-terminal extremity of a mature HslV, implying that the polypeptide had been processed into a mature protein. For HslV and HslU, calculated proteins in *L. infantum* and other trypanosomatids had the expected molecular weight (~22 and 55 kDa, respectively), and presented specific motifs (like the ATP-GTP binding site for HslU) and conservation of residues essential for the activity of the complex. Finally, the structure of the N-termini of trypanosomatid HslV and HslU suggested a mitochondrial localization for the two proteins. Therefore, the observations supported the notion that a functional HslVU complex was present in trypanosomatid protozoa, probably in the mitochondrion.

In parasitic protozoa, including *L. mexicana* (Robertson 1999), *T. cruzi* (Gonzalez et al. 1996), and *T. brucei* (Hua et al. 1996), a typical eukaryotic proteasome has been biochemically purified, and its role has been explored. Trypanosomatids would therefore be endowed with both a cytoplasmic-eukaryotic and a mitochondrial-eubacterial machinery for degradation of proteins. This is reminiscent of the compartmentalization of enzymes of the glycolytic pathway, where cytoplasmic and glycosomal isoforms of glyceraldehyde-3-phosphate dehydrogenase glycosomal (GAPDH) are coded for by different genes and have a different evolutionary origin. Glycosomal GAPDH is present in the common ancestor to kinetoplastids, whereas the cytoplasmic isoform was acquired at a later stage by horizontal transfer (Hannaert, Opperdoes, and Michels 1998).

In *E. coli*, HslV and HslU are heat shock proteins involved in the proteolysis of misfolded proteins (Missiakas et al. 1996). HslVU and the HslU homopolymeric complexes are also implicated in the regulation of cell division in *E. coli* through regulation of cell division inhibitor SulA (Kanemori, Yanagi, and Yura 1999; Seong et al. 1999, 2000).

In eukaryotes, homologs of other bacterial ATP-dependent proteases such as Lon and ClpXP (but not HslVU so far) have been found in the mitochondria. They seem essential for biogenesis and maintenance of the mitochondria through degradation of nonassembled or misfolded polypeptides and control of the steady-state levels of regulatory proteins (Käser and Langer 2000). Similarly, HslVU might complement or supplement other complexes in protein turnover of the mitochondrion in trypanosomatids. It may also have specific functions related to the peculiar biology of digenetic parasites which undergo dramatic morphological and, at least for *T. brucei*, metabolic remodeling (Vickerman 1994), including the mitochondrion, as they change host as well as differentiate from one stage to the next in the life cycle.

The observation of eubacterial genes in eukaryotic nuclear genomes is usually seen as evidence of migration of genes from the early endosymbiotic bacteria at the origin of the mitochondria or even evidence of ancient genome fusion at the very origin of the eukaryotic lineage (Gray, Burger, and Lang 1999). Accordingly, kinetoplastid *HslV* and *HslU* might have been present in early eukaryotes and retained in the kinetoplastids but might have been lost in more recent lineages. Indeed, kinetoplastids have been assigned a primordial status and in numerous phylogenetic reconstructions they branch earlier than most lineages of eukaryotes (Vickerman 1994; Stevens et al. 2001). Alternatively, a hor-

izontal transfer event in the last common ancestor of trypanosomatids would account for the presence of the gene in the three species. These organisms have ample opportunities of contact with bacteria in the gut of insect vectors. This is best illustrated by the observation of an endosymbiotic β-proteobacteria in some species of monogenic genera *Crithidia*, *Blastocrithidia*, and *Herpetomonas* (de Souza and Machado Motta 1999). A search for HslV and HslU homologs in free-living kinetoplastids or in related primitive eukaryotes such as the euglenids could help resolve these questions.

Our phylogenetic analyses indicated that trypanosomatid HslV and HslU, despite clustering with eubacterial sequences, could not be confidently assigned to any particular bacterial clade. Although there is now a large scientific support for the notion that mitochondria evolved once from an endosymbiotic a-proteobacterium, only ~200 out of ~400 proteins in the yeast mitochondrial proteome can be related to bacteria, and only ~50 of them can be closely related to α-proteobacteria (Kurland and Anderson 2000). Therefore, the lack of safe clustering with the a-proteobacteria sequences does not exclude early mitochondrial origin for trypanosomatid *HslV* and *Hsl*U.

In summary, we present evidence that the eubacterial HslVU and the eukaryotic multicatalytic protease are simultaneously present in trypanosomatid protozoa and expressed, at least in the developmental stages studied. We believe that exploring the role of HslVU in trypanosomatid protozoa could further our knowledge of the metabolic and developmental processes in these important human pathogens. A search for *HslV* and *HslU* homologs in free-living kinetoplastids and in other early-branching eukaryotes would help improve our understanding of aspects of mitochondria origins and eukaryotes evolution.

Supplementary Material

Accession numbers for the new sequences are AJ298867 (*LiHsl*V); AJ421946, AJ428522, and AJ428521 (*LiHsl*U); AJ298868 (*TcHsl*V); and BN000068 (*TbHsl*U).

Acknowledgments

The authors thank Prof. Daniel O. Sanchez of the Instituto de Investigaciones Biotecnologicas, Universidad Nacional de General San Martin, Buenos Aires, Argentina, for the gift of cDNA clone TENS 1873 from *T. cruzi*. Sequence databases and the GCG suite are maintained and updated at regular intervals at the Belgian EMBnet node facility in Brussels. The skilful technical assistance of Mr. Sylvain Lestrade and Mr. Ary Van Elsen at Applied Genetics is gratefully acknowledged. R.W. is Research Associate at the Belgian Fond National de la Recherche Scientifique. This work was supported in part by contract FNRS 1.5.043.98 from the Belgian Fond National de la Recherche Scientifique.

LITERATURE CITED

ALONSO, G., P. GUEVARA, and J. L. RAMIREZ. 1992. Trypanosomatidae codon usage and GC distribution. Mem. Inst. Oswaldo Cruz Rio de Janeiro **87**:517–523.

BOCHTLER, M., L. DITZEL, M. GROLL, C. HARTMANN, and R. HUBER. 1999. The proteasome. Annu. Rev. Biophys. Biomol. Struct. **28**:295–317.

BOCHTLER, M., C. HARTMANN, H. K. SONG, G. P. BOURENKOV, H. D. BARTUNIK, and R. HUBER. 2000. The structures of HslU and the ATP-dependent protease HslU-HslV. Nature **403**:800–805.

COUX, O., K. TANAKA, and A. L. GOLDBERG. 1996. Structure and functions of the 20S and 26S proteasomes. Annu. Rev. Biochem. **65**:801–47.

DE MOT, R., I. NAGY, J. WALZ, and W. BAUMEISTER. 1999. Proteasomes and other self-compartmentalizing proteases in prokaryotes. Trends Microbiol. **7**:88–92.

DE SOUZA, W., and M. C. MACHADO MOTTA. 1999. Endosymbiosis in protozoa of the *Trypanosomatidae* family. FEMS Microbiol. Lett. **173**:1–8.

EMMANUELSON, O., H. NIELSEN, S. BRUNAK, and G. VON HEIJNE. 2000. Predicting subcellular localization of proteins based on their N-terminal amino-acid sequence. J. Mol. Biol. **300**:1005–1016.

FERRELL, K., C. R. M. WILKINSON, W. DUBIEL, and C. GORDON. 2000. Regulatory subunit interactions of the 26S proteasome, a complex problem. Trends Biochem. Sci. **25**:83–88.

GONZALEZ, J., F. J. RAMALHO-PINTO, U. FREVERT, J. GHISO, S. TOMLINSON, J. SCHARFSTEIN, E. J. COREY, and V. NUSSENZWEIG. 1996. Proteasome activity is required for the stage-specific transformation of a protozoan parasite. J. Exp. Med. **184**:1909–1918.

GOTTESMAN, S. 1996. Proteases and their targets in *Escherichia coli*. Annu. Rev. Genet. **30**:465–506.

GRAY, M. W., G. BURGER, and B. F. LANG. 1999. Mitochondrial evolution. Science **283**:1476–1481.

HANNAERT, V., F. OPPERDOES, and P. A. M. MICHELS. 1998. Comparison and evolutionary analysis of the glycosomal glyceraldehyde-3-phosphate dehydrogenase from different kinetoplastida. J. Mol. Evol. **47**:728–738.

HÄUSLER, T., Y.-D. STIERHOF, J. BLATTNER, and C. CLAYTON. 1997. Conservation of mitochondrial targeting sequence function in mitochondrial and hydrogenosomal proteins from the early-branching eukaryotes *Crithidia*, *Trypanosoma* and *Trichomonas*. Eur. J. Cell. Biol. **73**:240–251.

HORWICH, A., E. U. WEBER-BAN, and D. FINLEY. 1999. Chaperone rings in protein folding and degradation. Proc. Natl. Acad. Sci. USA **96**:11033–11040.

HUA, S.-B., W.-Y. TO, T. T. NGUYEN, M.-L. WONG, and C. C. WANG. 1996. Purification and characterization of proteasomes from *Trypanosoma brucei*. Mol. Biochem. Parasitol. **78**:33–56.

KANEMORI, M., H. YANAGI, and T. YURA. 1999. The ATP-dependent HslVU/ClpQY protease participates in turnover of cell division inhibitor SulA in *Escherichia coli*. J. Bacteriol. **181**:3674–3680.

KÄSER, M., and T. LANGER. 2000. Protein degradation in mitochondria. Semin. Cell. Dev. Biol. **11**:181–190.

KURLAND, C. G., and S. G. E. ANDERSON. 2000. Origin and evolution of the mitochondrial proteome. Microbiol. Mol. Biol. **64**:786–820.

MISSIAKAS, D., F. SCHRAGER, J. M. BETTON, G. GEORGOPOULOS, and S. RAINA. 1996. Identification and characterization of HslV-HslU (ClpQ-ClpY) proteins involved in overall proteolysis of misfolded proteins in *Escherichia coli*. EMBO J. **15**:6899–6909.

NIELSEN, H., J. ENGELBRECHT, S. BRUNAK, and G. VON HEIJNE. 1997. Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. Protein Eng. **10**:1–6.

ROBERTSON, C. R. 1999. The *Leishmania mexicana* proteasome. Mol. Biochem. Parasitol. **103**:49–60.

SAMBROOK, J., E. F. FRITSCH, and T. MANIATIS. 1989. Molecular cloning: a laboratory manual. 2nd edition. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.

SEONG, I. S., J. Y. OH, J. W. LEE, K. TANAKA, and C. H. CHUNG. 2000. The HslU ATPase acts as a molecular chaperone in prevention of aggregation of SulA, an inhibitor of cell division in *Escherichia coli*. FEBS Lett. **477**:224–229.

SEONG, I. S., J. Y. OH, S. J. YOO, J. H. SEOL, and C. H. CHUNG. 1999. ATP-dependent degradation of SulA, a cell division inhibitor, by the HslVU protease in *Escherichia coli*. FEBS Lett. **456**:211–214.

SONG, H. K., C. HARTMANN, R. RAMACHANDRAN, M. BOCHTLER, R. BERHENDT, L. MORODER, and R. HUBER. 2000. Mutational studies on HslU and its docking mode with Hsl*V*. Proc. Natl. Acad. Sci. USA **97**:14103–14108.

STEVENS, J. R., H. A. NOYES, C. J. SCHOFIELD, and W. GIBSON. 2001. The molecular evolution of *Trypanosomatidae*. Adv. Parasitol. **48**:1–56.

THOMPSON, J. D., T. J. GIBSON, F. PLEWNIAK, F. JEANMOUGIN, and D. G. HIGGINS. 1997. The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res. **24**:4876–4882.

VICKERMAN, K. 1994. The evolutionary expansion of the trypanosomatid flagellates. Int. J. Parasitol. **24**:1317–1331.

VOGES, D., P. ZWICKL, and W. BAUMEISTER. 1999. The 26S proteasome: a molecular machine designed for controlled proteolysis. Annu. Rev. Biochem. **68**:1015–1068.

WANG, J., J. J. SONG, M. C. FRANKLIN, S. KAMTEKAR, Y. J. IM, S. H. RHO, I. S. SEONG, C. S. LEE, C. H. CHUNG, and S. H. EOM. 2001. Crystal structures of the HslVU peptidase-ATPase complex reveal an ATP-dependent proteolysis mechanism. Structure **9**:177–184.

WICKNER, S., M. R. MAURIZI, and S. GOTTESMAN. 1999. Post-translational quality control: folding, refolding, and degrading proteins. Science **286**:1888–1893.

ANTONY DEAN, reviewing editor

Accepted July 17, 2002