J. Euk. Microbiol., 45(3), 1998 pp. 307–313

Functional Conservation of Calreticulin in *Euglena gracilis*

LORELLA NAVAZIO,* MARIA C. NARDI,* SIMONETTA PANCALDI,** PAOLA DAINESE,*** BARBARA BALDAN,* ANNE-CATHERINE FITCHETTE-LAINÉ,**** LOÏC FAYE,**** FLAVIO MEGGIO,***** WILLIAM MARTIN****** and PAOLA MARIANI*,[1]

*Dipartimento di Biologia, Università di Padova, Via U. Bassi 58/B, 35131 Padova, Italy, and
**Dipartimento di Biologia, Università di Ferrara, 44100 Ferrara, Italy, and
***Institut für Biochemie III, Eldgenössische Technische Hochschule Zentrum, CH-8092 Zürich, Switzerland, and
****Laboratoire des Transports Intracellulaires, CNRS UPRESA 6037, IFR 23, Université de Rouen, 76821 Mont-Saint Aignan cédex, France, and
*****Dipartimento di Chimica Biologica, Università di Padova, 35131 Padova, Italy, and
******Institut für Genetik, Technische Universität Braunschweig, D-38023 Braunschweig, Federal Republic of Germany

ABSTRACT. Calreticulin is the major high capacity, low affinity Ca^{2+} binding protein localized within the endoplasmic reticulum. It functions as a reservoir for triggered release of Ca^{2+} by the endoplasmic reticulum and is thus integral to eukaryotic signal transduction pathways involving Ca^{2+} as a second messenger. The early branching photosynthetic protist *Euglena gracilis* is shown to possess calreticulin as its major high capacity Ca^{2+} binding protein. The protein was purified, microsequenced and cloned. Like its homologues from higher eukaryotes, calreticulin from *Euglena* possesses a short signal peptide for endoplasmic reticulum import and the C-terminal retention signal KDEL, indicating that these components of the eukaryotic protein routing apparatus were functional in their present form prior to divergence of the euglenozoan lineage. A gene phylogeny for calreticulin and calnexin sequences in the context of eukaryotic homologues indicates i) that these Ca^{2+} binding endoplasmic reticulum proteins descend from a gene duplication that occurred in the earliest stages of eukaryotic evolution and furthermore ii) that Euglenozoa express the calreticulin protein of the kinetoplastid (trypanosomes and their relatives) lineage, rather than that of the eukaryotic chlorophyte which gave rise to *Euglena*'s plastids. Evidence for conservation of endoplasmic reticulum routing and Ca^{2+} binding function of calreticulin from *Euglena* traces the functional history of Ca^{2+} second messenger signal transduction pathways deep into eukaryotic evolution.

Supplementary key words. Calcium binding proteins, endosymbiosis, gene phylogeny, intracellular calcium store, signal transduction.

SEVERAL signal transduction pathways of higher eukaryotic cells involve the triggered release of Ca^{2+} sequestered in the endoplasmic reticulum (ER) [46]. Within the ER, sequestered, releasable Ca^{2+} does not occur as a free ion, but rather is complexed by Ca^{2+} binding proteins that are resident in the lumen of the ER and provide an efficient buffer of intracellular Ca^{2+} levels. Of these, calreticulin is generally considered as the major contributor to Ca^{2+} storage capacity of the compartment, accounting for about half of the total Ca^{2+} stored [34]; exceptions can be found in specialized forms of the ER, such as the sarcoplasmic reticulum of muscle cells, where calsequestrin, rather than calreticulin, assumes the major Ca^{2+} binding and storage function [37].

Calreticulin is a low affinity, high capacity Ca^{2+} binding protein that binds ~20 moles of Ca^{2+} per mole of protein and has an apparent molecular mass of 50 to 60 kDa in various species. Its ubiquity and high level of conservation across higher eukaryotes reflect an important role of this protein as a central element of cellular Ca^{2+} homeostasis. Calreticulin shares extensive sequence similarity with the integral ER membrane Ca^{2+} binding protein, calnexin, which appears to function as a molecular chaperone. Chaperoning has also been attributed to calreticulin, and it has been suggested that calnexin and calreticulin may act in concert in the quality control system of the ER by assisting glycoproteins in their maturation [18]. Calreticulin has been extensively studied and cloned from several vertebrates and invertebrates (see [35] for a review) and from a few higher plants as well [6, 8, 39, 42]. No biochemical or molecular evidence—including that from completely sequenced prokaryotic genomes [13, 4, 23]—exists to suggest the presence of calreticulin-like proteins in either eubacteria or archaebacteria, as expected from the absence of internal homeostatic devices for Ca^{2+} in prokaryotes [58]. Calreticulin is thus a characteristically eukaryotic protein found in a uniquely eukaryotic compartment, the ER.

From an evolutionary standpoint, the presence of a cellular Ca^{2+} buffering system must have been prerequisite to the emergence of Ca^{2+} as a controlled second messenger during the history of nucleate cells and, accordingly, the components of this system likely arose very early in eukaryotic evolution [58]. From these considerations, a number of intriguing questions ensue concerning the evolution of ER-mediated Ca^{2+} signaling pathways. These include: i) How far back the components of Ca^{2+} homeostasis can be traced within the history of contemporary eukaryotes? ii) What was the evolutionary order of appearance of calreticulin and calnexin proteins? iii) In eukaryotes of secondary endosymbiotic origin, where a eukaryote has engulfed a photosynthetic eukaryote (see [31, 55] for brief overviews), have components of the eukaryotic symbiont been incorporated into the host's Ca^{2+} homeostasis machinery?

Ca^{2+} binding proteins have been little studied in early branching protists and no data is available from protists of secondary symbiotic origin, yet these are of particular interest due to the complexity of their membrane systems. The plastids of *Euglena gracilis*, for example, possess three outer membranes instead of two as in higher plants. This finding led inter alia to the suggestion that *Euglena*'s plastids might have arisen through engulfment and degeneration of a eukaryotic alga, rather than a cyanobacterium [15]. Cytological [51, 57] and molecular [20, 48] data indicate that the host cell of *Euglena*'s secondary symbiosis belonged to the kinetoplastids, a non-photosynthetic group of protists encompassing trypanosomes and their relatives [56], whereas pigment content (chlorophylls a and b) and molecular data suggest that the plastid descends from the chlorophyte lineage [28, 55]. The outermost of *Euglena*'s plastid membranes derives from the ER, and proteins targeted to the plastids are imported with the help of bipartite, N-terminal signal/transit peptide regions [50]. It is not known from which eukaryote involved in *Euglena*'s history (host or symbiont) this ER derives. Here we report the biochemical characterization, purification, ER-localization, cloning of the gene, and studies of the evolutionary relationships of the major Ca^{2+} binding protein present in the ER of *Euglena gracilis*.

MATERIALS AND METHODS

Cell cultures. *Euglena gracilis* Klebs, strain Z, was grown to midexponential phase in Hutner and Provasoli's growth liq-

[1] To whom correspondence should be addressed. Telephone: 39-49-8276277; Fax: 39-49-8276280; Email: marianip@civ.bio.unipd.it

uid medium [59] at 26° C on a rotary shaker under 16/8 light/dark photoperiod.

Cell disruption techniques. For protein purification, *Euglena* cells were broken in a French Press under a pressure of 10,000 psi for 10 s. For isolation of intact organelles, cells were gently ground for 2.5 min with acid-washed glass beads (212–300 μm, Sigma, St. Louis, MO) as described by Sulli and Schwartzbach [49].

Protein purification and microsequencing. Acidic Ca^{2+} binding proteins were isolated by using an ammonium sulphate precipitation procedure, followed by DEAE-Cellulose (DE52, Whatman, Maidstone, UK) column chromatography as described by Milner et al. [38]. Protein concentration of each fraction was determined by the method of Bradford [3] using bovine serum albumine as a standard. For N-terminal amino acid sequencing the gel band corresponding to the 56 kDa protein was excised from multiple lanes, loaded onto a funnel-shaped concentration gel, electroblotted onto a polyvinylidene difluoride membrane and directly analyzed on a Applied Biosystems 476A protein Sequencer.

Cell fractionation. Organelles were separated by sucrose density gradient centrifugation. Cell extracts (1.5 ml) were layered onto sucrose gradients, which consisted of 9.5-ml sucrose solution increasingly linearly in concentration from 15 to 35% (w/w) over a 0.5 ml cushion of 55% sucrose (w/w). All sucrose solutions contained 25 mM Hepes-KOH, pH 7.4 and 1 mM EDTA. Gradients were centrifuged for 16 h at 100,000 *g* in a SW 41 rotor on a Kontron (Zurich, Switzerland) T-2060 centrifuge and fractionated from the top into 1-ml fractions. Sucrose concentration was determined by refractometry (N1 and N2 refractometers Atago, Tokyo, Japan).

Biochemical techniques. One-dimensional SDS-PAGE was carried out according to Laemmli [26] on 7.5–10% linear gradient gels. Two-dimensional electrophoresis was performed according to O'Farrell [43], with isoelectrofocusing in the first dimension (pH gradient 4–6.5). Gels were stained first with Coomassie blue and then with Stains All (Bio-Rad, Richmond, CA), according to Campbell et al. [5]. Protein transfer onto nitrocellulose membranes was carried out as described by Towbin et al. [54]. $^{45}Ca^{2+}$-binding to the blotted proteins was determined by $^{45}Ca^{2+}$-overlay technique [29]. Affinodetection with Concanavalin A on blots was performed according to Faye and Chrispeels [10]. For immunoblot analysis, blots were incubated with: a) affinity-purified polyclonal antibodies against rabbit calreticulin, 1:2,000 diluted (from Dr. Michalak, Edmonton, Canada); b) antiserum against spinach calreticulin, 1:2,000 diluted [40]; c) antiserum against recombinant tobacco binding protein (BiP), 1:1,000 diluted (from Dr. Vitale, Milano, Italy); d) antibodies against plant complex N-linked glycans [11]. Calreticulin proteins isolated from spinach leaves and rabbit liver [41] were used as controls in immunoblot experiments. Phosphorylation assays were carried out as previously described [2]. In some experiments, samples were pretreated with 200 mM alkaline phosphatase (Sigma) for 30 min, heated for 3 min at 50° C to inactivate the phosphatase and then phosphorylated by protein kinase CK2, either in the absence or presence of 420 nM polylysine. DEAE-Cellulose fractions eluted between 0.1 and 0.4 M NaCl were pooled and subjected to phenyl-Sepharose (CL-4B, Pharmacia, Uppsala, Sweden) column chromatography followed by elution with 10 mM $CaCl_2$ [41].

Molecular techniques. For amplification of the homologous calreticulin probe, the degenerate oligonucleotides 5'TAYTAYAARGARACNTTYGARCC3' and 5'RTANCCNCCNCCRCARTC3' were used in a PCR reaction containing 10 mM Tris-HCl, pH 8.3, 50 mM KCl, 1 mM $MgCl_2$, 50 μM of each dNTP, 1.6 μM of each primer, 0.02 U/μl Taq polymerase (Perkin-Elmer, Foster City, CA) and 0.4 ng/μl of double stranded *Euglena* cDNA in a total volume of 25 μl. Forty cycles of 1 min at 93° C, 1 min at 50° C and 2 min at 72° C were performed. The 260-bp amplification product was used as a substrate for reamplification under the same conditions, the product of which was purified by ultrafiltration on Microcon 30 devices (Amicon, Beverly, MA), filled in with Klenow polymerase, phosphorylated with T4 kinase, again purified by ultrafiltration on Microcon 30 devices, ligated into Hinc II-cut p-BluescriptSK+ (Stratagene, La Jolla, CA) and transformed into E. coli nm522 to yield the plasmid pPCRcalrH. The sequence of pPCRcalrH was determined by the didedoxy method to confirm the identity of the amplified product. The 279-bp Cla I—XhoI fragment of pPCRcalrH was isolated by gel electrophoresis and DE52 ion-exchange chromatography as a hybridization probe for cDNA screening. Messenger RNA isolation from *Euglena*, cDNA synthesis, cDNA cloning as well as standard molecular techniques were performed as previously described [19, 20]. The sequences reported in this paper have been deposited with GenBank under the accession number EMBL Y09816.

Data handling. Sequences were extracted from GenBank and aligned with ClustalW [53]. Regions of uncertain alignment were omitted from the analysis leaving 275 amino acid positions for analysis. Numbers of substitutions per site were estimated using the Dayhoff matrix option of PHYLIP [12] and used to construct the neighbor joining tree [47].

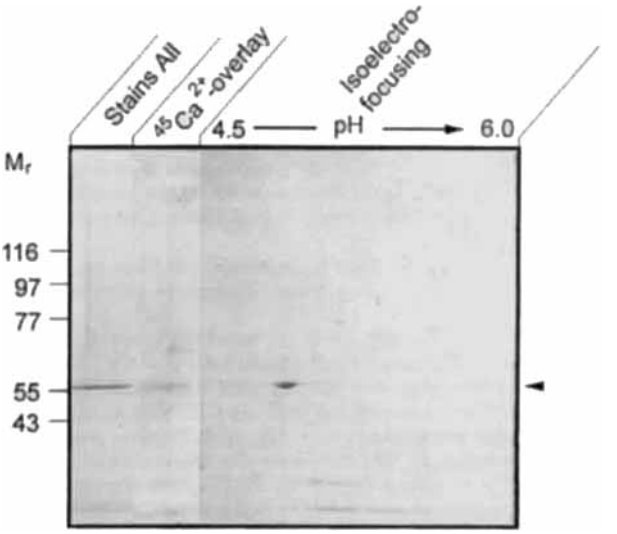

Fig. 1. Identification of *Euglena* Ca^{2+}-binding proteins. The DEAE-Cellulose fraction, eluted at 0.15 M NaCl, was analyzed by Stains All staining of 14 μg of protein after SDS-PAGE, $^{45}Ca^{2+}$-ligand overlay of 40 μg of protein blotted to nitrocellulose (autoradiography exposed for 7 days), isoelectrofocusing of 20 μg of protein (only the region between pH 4.5 and 6.0 is shown). Positions of the size markers in kDa are indicated on the left. Arrowhead indicates the 56 kDa blue-staining, Ca^{2+}-binding protein.

RESULTS

Isolation and properties of the main Ca^{2+} buffering protein. Ca^{2+} binding proteins from *Euglena* were isolated by using a selective ammonium sulphate precipitation procedure, followed by purification on DEAE-Cellulose chromatography. Electrophoretic analysis of protein fractions showed in the fraction eluted at 0.15 M NaCl the presence of a single major Stains All-blue staining band, migrating in SDS-PAGE as a 56 kDa protein (Fig. 1). The same protein was found to bind $^{45}Ca^{2+}$

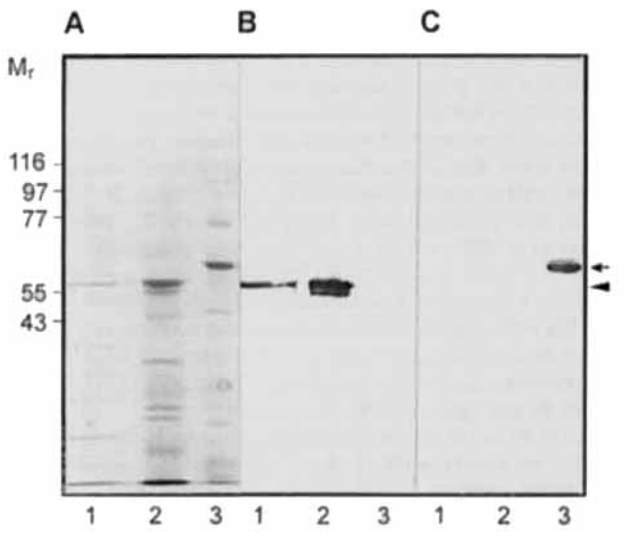

Fig. 2. Immunological cross-reactivity of the *Euglena* 56 kDa protein with antibodies against spinach calreticulin and rabbit calreticulin. **A**. Stains All staining of DEAE-Cellulose fractions electrophoresed on SDS-PAGE. **B**. Western blot with anti-spinach calreticulin antiserum. **C**. Western blot with anti-rabbit calreticulin antibodies. Lane 1: *Euglena* 56 kDa Ca^{2+} binding protein (A, B and C, 14 μg each). Lane 2: 56 kDa and 54 kDa spinach calreticulins (A, 8 μg; B and C, 2 μg each). Lane 3: 61 kDa and 59 kDa rabbit calreticulins (A, 4 μg; B and C, 2 μg each). Positions of the size markers in kDa are indicated on the left. Arrow indicates 59 kDa rabbit calreticulin, arrowhead indicates 56 kDa *Euglena* and spinach calreticulins.

when blotted onto nitrocellulose and subjected to $^{45}Ca^{2+}$-overlay technique. When analyzed by two-dimensional electrophoresis after isoelectrofocusing in the first dimension, a single 56 kDa protein was observed in the pH range of approximately 4.9 (Fig. 1). Blue-staining with Stains All, $^{45}Ca^{2+}$ binding ability, and a highly acidic isoelectric point are diagnostic properties of high capacity, low affinity Ca^{2+} binding proteins [36]. But these criteria do not establish the identity of the isolated protein. Therefore, the N-terminal amino acid sequence of the purified protein was determined. It revealed detectable similarity to the N-terminal sequence of spinach calreticulin [41] and rabbit calreticulin [14], suggesting that the 56 kDa protein isolated from *Euglena* belongs to the calreticulin family of Ca^{2+} binding proteins.

The isolated 56 kDa protein cross-reacted with antibodies raised against purified spinach calreticulin (Fig. 2B), but did not show any cross-reaction with polyclonal antibodies raised against rabbit calreticulin (Fig. 2C). The presence of N-linked glycans on the *Euglena* 56 kDa protein was investigated by affinity and immunodetection using Concanavalin A, a lectin specific for high mannose N-glycans, and antibodies directed against plant complex N-linked glycans [11], respectively. Calreticulins from higher plants are often glycosylated, possessing N-linked carbohydrate chains of the high-mannose type [40]. By contrast, the 56 kDa *Euglena* Ca^{2+}-binding protein is apparently not glycosylated either with high-mannose or complex type glycans, since no positive reaction was observed in either case (data not shown).

Furthermore, the 56 kDa *Euglena* protein failed to undergo phosphorylation by protein kinase CK2 (data not shown), whereas spinach calreticulin is an excellent substrate for this enzyme [2]. Neither previous treatment with alkaline phosphatase nor the addition of polylysine, a known stimulator of CK2 which is required for the phosphorylation of calmodulin [32], rescued significant phosphorylation of the 56 kDa *Euglena* Ca^{2+}-binding protein by CK2. With respect to the latter two properties, the biochemical behavior of the 56 kDa *Euglena* protein more closely resembled that of rabbit calreticulin, which is neither glycosylated [24], nor phosphorylated by CK2 [2].

We examined the possibility that calsequestrin, a functional analogue of calreticulin in muscle cells, might be co-expressed with the 56 kDa protein in *Euglena*. To test this, we employed Ca^{2+}-dependent elution from a hydrophobic matrix (phenyl-Sepharose), which is a characteristic property of calsequestrin proteins that is not shared by calreticulins [7]. No *Euglena* proteins were found to elute from phenyl-Sepharose in the presence of Ca^{2+} (data not shown), providing further evidence that *Euglena* expresses a single high capacity, low affinity Ca^{2+} binding protein.

***Euglena* possesses a classical 56 kDa calreticulin.** Degenerate primers were constructed against the N-terminal sequence of the purified protein and against the hexapeptide motif "DCGGGY" conserved among all the calreticulin proteins from higher eukaryotes studied to date. PCR with these primers using *Euglena* cDNA as a substrate yielded an amplification product of the expected length (260 bp). This was shown by sequencing to encode the expected N-terminal amino acid sequence and a peptide with high similarity to known calreticulin sequences. From 60,000 independent recombinants of the *Euglena* cDNA library, all five positives obtained were identical in sequence. The sequence of pCRET1, pCRET5, pCRET9, pCRET18 and pCRET22 and the protein encoded (accession number Y09816) are shown in Fig. 3.

The encoded protein shares an average of about 50% amino acid identity in comparison with calreticulin sequences from higher eukaryotes and an average 40% identity with the related Ca^{2+} binding protein calnexin from various sources. These data, together with the biochemical results described above, indicate that the 56 kDa protein purified from *Euglena* is calreticulin. The cleavage site of the N-terminal signal peptide for *Euglena* calreticulin was determined to be after serine (-1) by comparison to the N-terminal sequence of the directly sequenced purified protein (Fig. 3). The mature protein of 383 residues has a predicted Mr of 43,884 and a pI of 4.56. Acidity is likely to be responsible for the anomalous migration in gels of all calreticulin proteins so far investigated [1]. The sequence of *Euglena* calreticulin possesses the typical tripartite organization in the N-, P- and C-domains recognized in higher eukaryotic calreticulin proteins [1]. No potential sites for N-glycosylation were observed in either the N-domain (in the standard numbering position 32, common in plant calreticulins [40] and position 162, in bovine brain [30], *Schistosoma* [17] and *Leishmania* [22]) calreticulins], nor in the C-domain (position 327, typical of animal calreticulins [1]), nor in any other position, in agreement with our biochemical results described above concerning the absence of protein glycosylation.

Euglena calreticulin lacks the putative nuclear targeting signal located in the central proline-rich region (P-domain) of some animal and plant calreticulins [1]. Variants of the amino acid sequence repeats (Repeat A and Repeat B) found in the P-domain of calreticulin and calnexin, were also conserved in the *Euglena* calreticulin. The C-domain is very rich in acidic amino acids (38% aspartate and glutamate residues), as expected for the low affinity, high capacity Ca^{2+}-binding domain. *Euglena* calreticulin reveals six threonines (T47, T173, T185, T227, T304 and T362) and one serine (S314) as potential phosphorylation sites for CK2 (S/T-X-X-E/D). These are however predicted to be poor substrates for CK2 since they meet just the minimum requirements (one acidic residue at position +3) for

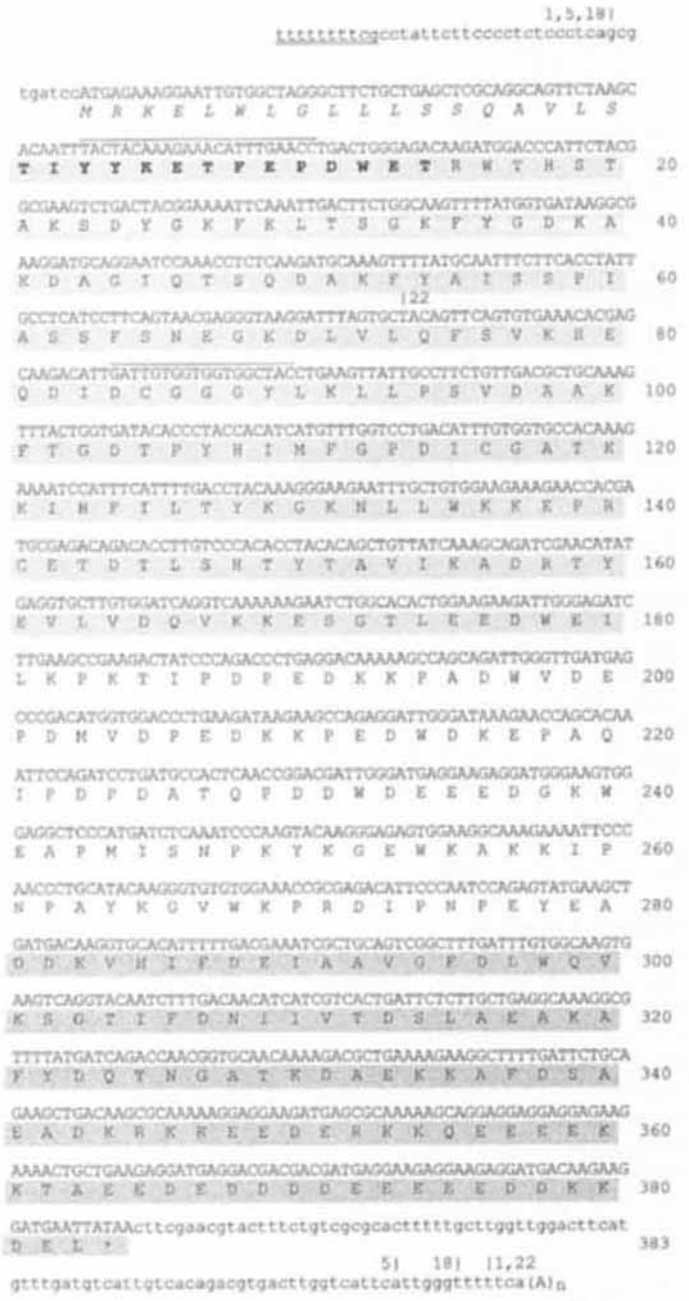

Fig. 3. Sequence of pCRET9 encoding *Euglena gracilis* calreticulin. The signal peptide is shown in italics, noncoding regions are shown in lower case. The directly sequenced N-terminus of the purified protein is given in bold face type. The 5′ termini and polyadenylation sites of the independent cDNA clones pCRET1, pCRET5, pCRET18 and pCRET22 are indicated with vertical bars. The 5′ sequence corresponding to the spliced leader of other *Euglena* mRNAs is double underlined. Amino acid motifs against which primers were designed for isolation of the homologous probe are overlined. Numbering refers to amino acid positions of the mature *Euglena* protein. The N, P and C regions are distinguished by shading.

phosphorylation by CK2 [33], in agreement with the lack of phosphorylation by CK2 of *Euglena* calreticulin found in vitro.

Localization of calreticulin in *Euglena*'s ER. The C-terminus of *Euglena* calreticulin possesses the characteristic ER-retrieval sequence KDEL, suggesting an ER-localization of the calreticulin protein, as demonstrated for all calreticulins investigated so far [25]. In higher eukaryotes calreticulin is a marker protein of the ER [25], but in early branching protists, the localization of the protein has not been demonstrated. Therefore, to determine whether this ER-retrieval sequence is functional and the localization of *Euglena* calreticulin, we prepared cell fractions from *Euglena* cultures and tested for immunoreaction with anti-calreticulin serum. The result is shown in Fig. 4. Calreticulin was predominantly localized in the 28–30% sucrose fractions (Fig. 4B), which were previously suggested to contain ER-vesicles from *Euglena* [9]. This was confirmed by immunoreaction with antibodies against BiP, a second marker protein of the ER [45], which localized in the same fractions (Fig. 4C).

Evolution of *Euglena*'s calreticulin gene. A BLAST search with *Euglena* calreticulin identified numerous calreticulin and calnexin homologues. Calnexin possesses two ~20-residue insertions near the A-repeats of calreticulin, an N-terminal extension of ~50 amino acids, and a C-terminal extension of ~130 amino acids relative to calreticulin. After purging the data of such regions and of regions of uncertain alignment (i.e. the extremely acidic region of the C-domain), 275 positions remained for phylogenetic analysis. The neighbor joining tree (Fig. 5) reveals that calreticulin of *Euglena* branches with its recently characterized homologue from the kinetoplastid *Leishmania donovani* [22]. The relatively weak bootstrap support (63%) for that branch may relate to difficulties imposed by the elevated substitution rate in *Leishmania* calreticulin as reflected in its long terminal branch. Fig. 5 clearly shows that *Euglena* 56 kDa protein is a member of the calreticulin, not the calnexin, family. The finding that *Euglena* calreticulin branches with its homologue from kinetoplastids, but not with those from higher plants, suggests that in *Euglena* this component of the Ca^{2+} homeostasis apparatus was derived from the eukaryotic host rather than the eukaryotic symbiont.

DISCUSSION

We have shown that the early-branching photosynthetic protist, *Euglena gracilis*, possesses the high-capacity low-affinity Ca^{2+} binding protein calreticulin, an integral component of intracellular Ca^{2+} buffering systems in higher eukaryotes [25]. The results of cell fractionation and immunodetection furthermore demonstrate that calreticulin in *Euglena* is localized in the ER. Previous data indicated that *Euglena* arose through secondary symbiosis [15] in which a member of the kinetoplastid lineage engulfed a chlorophyte. Indeed, expressed genes of both kinetoplastid [20, 27] and chlorophyte [20, 28] descent are found in *Euglena*'s nucleus. Calreticulin from *Euglena* clearly reflects the kinetoplastid ancestry of its nucleocytoplasmatic component, which represents one of the earliest branching lineages of mitochondriate eukaryotes known. Since the function (high-capacity low-affinity Ca^{2+} binding) and localization of calreticulin is conserved across the euglenozoan-higher eukaryote boundary, these findings suggest that the ER-localized Ca^{2+} buffering machinery necessary for employment of Ca^{2+} as a second-messenger in signal transduction was already in place very early in eukaryotic evolution.

By sequencing the N-terminus of the purified protein, we deduced the sequence of the 18-amino acid signal peptide from the cDNA as "MRKELWLGLLLSSQAVLS." This is one of the few [16] ER-signal sequences to be directly determined from the early branching *Euglena* lineage and, to our knowledge, it is the most ancient experimentally verified ER-specific signal peptide known to date. Its sequence characteristics are similar to those of signal peptides from higher eukaryotes. The C-terminal KDEL-motif is also similar to higher eukaryotic ER-retrieval signals, and since calreticulin in *Euglena* is localized in the ER, it appears that the KDEL ER-retrieval signal is con-

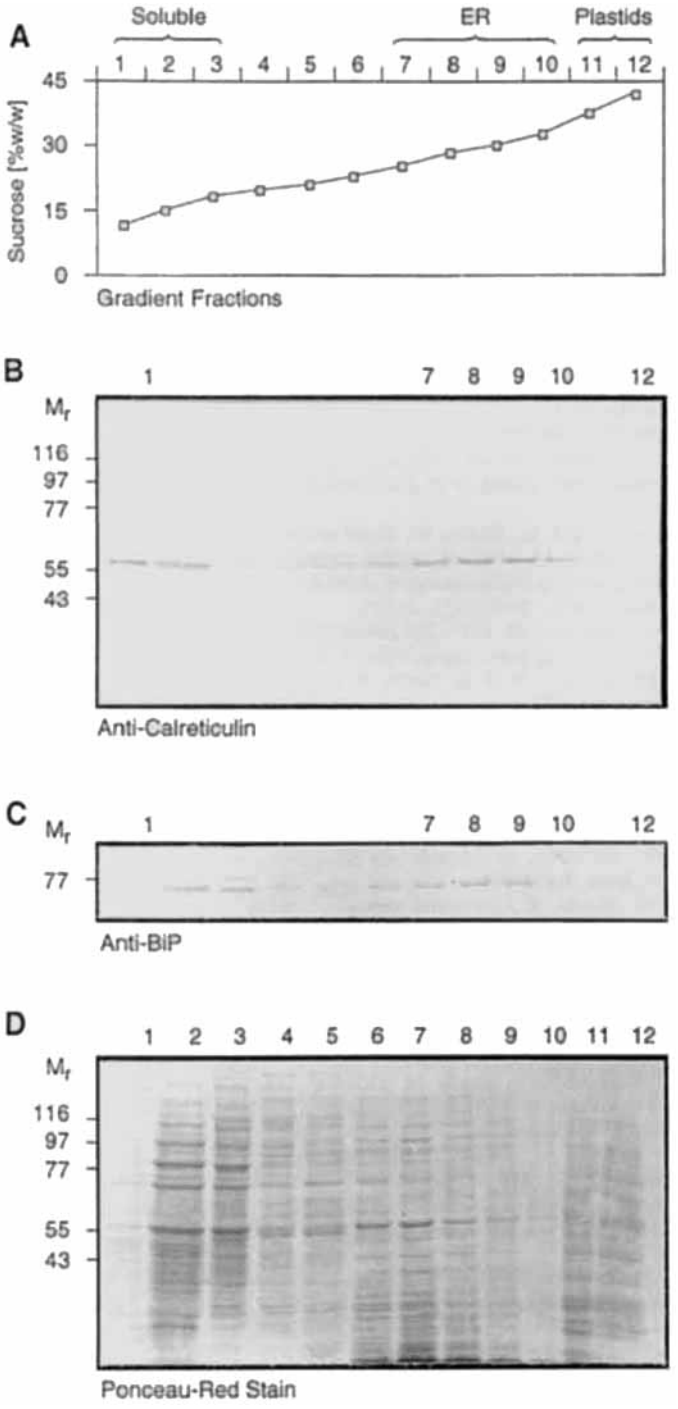

Fig. 4. Subcellular localization of *Euglena* calreticulin. Organelles were separated by sucrose gradient centrifugation of *Euglena* homogenates. **A**. Sucrose concentration of fractions collected from the top to the bottom of the gradients. **B**. Western blot of fractions decorated with anti-spinach calreticulin antiserum. **C**. Western blot of fractions decorated with anti-tobacco BiP antiserum. **D**. Ponceau Red staining of blot (100 μl each fraction). Soluble fractions at the top of the gradients contain lumenal proteins released from vesicles during cell disrupture. Positions of the size markers in kDa are indicated on the left.

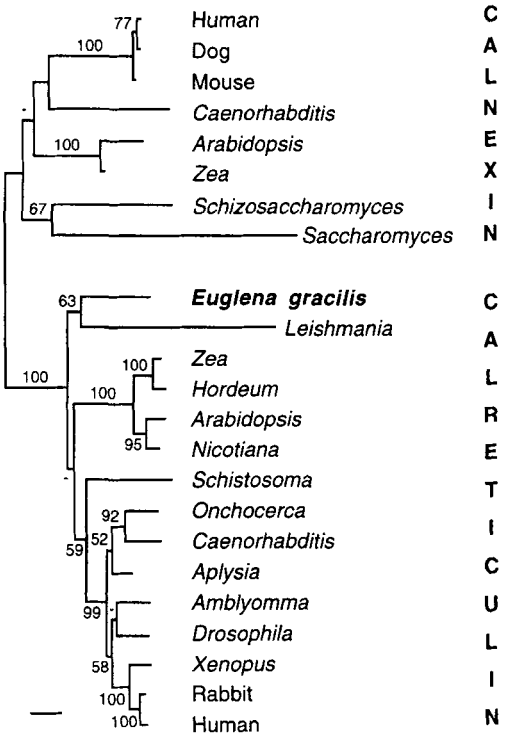

Fig. 5. Phylogenetic tree for calreticulin and calnexin sequences. The scale bar indicates 0.1 substitutions per site. Bootstrap proportions greater than 50/100 are indicated at branches, bootstrap proportions less than 50/100 are indicated as "-".

served by virtue of its function. Taken together, these findings suggest that the basic machinery of both ER-targeting and ER-retention of proteins was in existence at the time when Euglenozoa and higher eukaryotes diverged.

Euglena is known to attach spliced leader (SL) sequences to the 5′ ends of the majority of nuclear encoded mRNAs [52]. Such SL sequences are probably involved in the cleavage of polycistronic mRNAs in *Caenorhabditis* into monocistronic mRNAs for translation [21]. In *Euglena,* the SL in several mRNAs ends with the motif 5′GTGTCT(A/T)TTTTTTTCG3′ [52]. The 5′ terminus of pCRET9 is 5′TTTTTTTCG3′, suggesting that the transcript for calreticulin is processed with the help of spliced leader addition.

The primary sequence reveals properties specific to calreticulin function, i.e. a highly acidic C-terminal domain involved in high-capacity low-affinity Ca^{2+} binding. The C-domain consists of 38% negatively charged amino acid residues, comparable to 40% in rabbit calreticulin [14] and 38% in maize calreticulin [39]. N-linked carbohydrates are not a widespread feature of calreticulin proteins, being more common in plants [40] rather than animals [1]. The amino acid sequence of *Euglena* calreticulin contains no consensus site for N-linked glycosylation, consistent with our biochemical assays that revealed no evidence for glycosylation of the purified protein. Furthermore, like animal calreticulin, but unlike higher plant calreticulin [2], the *Euglena* protein lacks good potential sites of phosphorylation by protein kinase CK2 and does not undergo phosphorylation with this enzyme. Thus, glycosylation and phosphorylation of calreticulin do not appear to be conserved properties necessary for calreticulin function within the ER, rather they appear to have undergone lineage-specific modifications during eukaryotic evolution.

We found no evidence for the presence of other high capac-

ity, low affinity Ca^{2+} binding proteins in *Euglena*. The occurrence of both calreticulin and calsequestrin has recently been reported in *Paramecium* cells [44].

The gene phylogeny indicates that the Ca^{2+} binding proteins calreticulin and calnexin share a common ancestor and that they are related via a gene duplication that occurred early in the evolution of eukaryotic cells, prior to the separation of the euglenozoan and higher eukaryotic lineages. Our finding of a specifically calreticulin-related Ca^{2+} binding protein in the early-branching eukaryote *Euglena gracilis* indicates that the origin of this protein is indeed ancient, but clearly more recent than the calreticulin-calnexin gene duplication, suggesting that the earliest eukaryotes may have utilized an ancestral Ca^{2+} binding protein from which both calnexin and calreticulin descend. Finally, homologues of calreticulin have not been found in eubacteria or archaebacteria, either as proteins or in completely sequenced genomes, consistent with the view [58] that the evolution of this integral component of the Ca^{2+} homeostatic machinery was a specifically eukaryotic invention, and a critical step in allowing the Ca^{2+} ion to become a universal second messenger for eukaryotic signal transduction pathways.

ACKNOWLEDGMENTS

The excellent technical assistance of Christiane Lindemann is gratefully acknowledged. We would like to thank Dr. A. Vitale (Milano, Italy) and Dr. M. Michalak (Edmonton, Canada) for kindly providing anti-tobacco BiP and anti-rabbit calreticulin antibodies, respectively. L. N. was recipient of a short term fellowship from the European Commission (Project of Technological Priority, PTP). This work was supported by MURST (to P. M.) and PTP (to W. M.) grants.

LITERATURE CITED

1. Baksh, S. & Michalak, M. 1996. Basic characteristics and ion binding to calreticulin. In: Michalak, M. (ed.). Calreticulin. Springer-Verlag, Heidelberg. Pp. 11–30.

2. Baldan, B., Navazio, L., Friso, A., Mariani, P. & Meggio, F. 1996. Plant calreticulin is specifically and efficiently phosphorylated by protein chinase CK2. *Biochem. Biophys. Res. Commun.*, **221**:498–502.

3. Bradford, M. M. 1976. A rapid and sensitive method for the quantitation of micrograms quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.*, **72**:248–254.

4. Bult, C. J., White, O., Olsen, G. J., Zhou, L., Fleischmann, R. D., Sutton, G. G., Blake, J. A., FitzGerald, L. M. , Clayton, R. A., Gocayne, J. D., Kerlavage A. R., Dougherty, B. A., Tomb J-F., Adams, M. D., Reich, C. I., Overbeek, R., Kirkness, E. F., Weinstock, K. G., Merrick, J. M., Glodek, A., Scott J. L., Geoghagen, N. S. M., Weidman, J. F., Fuhrmann, J. L., Nguyen, D., Utterback, T. R., Kelley, J. M., Peterson, J. D., Sadow, P. W., Hanna, M. C., Cotton, M. D., Roberts, K. M., Hurst, M. A., Kaine, B. P., Borodovsky, M., Klenk, H-P., Fraser, C. M., Smith, H. O., Woese, C. R. & Venter, J. C. 1996 . Complete genome sequence of the methanogenic archaeon, *Methanococcus jannaschii*. *Science*, **273**:1058–1073.

5. Campbell, K. P., MacLennan, D. H. & Jorgensen, A. O. 1983. Staining of the Ca^{2+}-binding proteins, calsequestrin, calmodulin, troponin C, and S-100, with the cationic carbocyanine dye "Stains-all". *J. Biol. Chem.*, **258**:11267–11273.

6. Chen, F., Hayes, P. M., Mulrooney, D. M. & Pan, A. 1994. Identification and characterization of cDNA clones encoding plant calreticulin in barley. *Plant Cell*, **6**:835–843.

7. Damiani, E., Heilmann, C., Salvatori, S. & Margreth, A. 1989. Characterization of high-capacity low-affinity calcium binding protein of liver endoplasmic reticulum: calsequestrin-like and divergent properties. *Biochem. Biophys. Res. Commun.*, **165**:973–980.

8. Denecke, J., Carlsson, L. E., Vidal, S., Hoglund, A. S., Ek, B., van Zeijl, M. J., Sinjorgo, K. M. C. & Palva, E. T. 1995. The tobacco homolog of mammalian calreticulin is present in protein complexes in vivo. *Plant Cell*, **7**:391–406.

9. Dockerty, A. & Merret, M. J. 1979. Isolation and enzymic characterization of *Euglena* proplastids. *Plant Physiol.*, **63**:468–473.

10. Faye, L. & Chrispeels, M. J. 1985. Characterization of N-linked oligosaccharides by affinoblotting with concanavalin A-peroxidase and treatment of the blots with glycosidases. *Anal. Biochem.*, **149**:218–224.

11. Faye, L., Gomord, V., Fitchette-Lainé, A-C. & Chrispeels, M. J. 1993. Affinity purification of antibodies specific for Asn-linked glycans containing α 1→3 fucose or β 1→2 xylose. *Anal. Biochem.*, **209**:104–108.

12. Felsenstein, J. 1993. Phylip (Phylogeny inference package) manual, Version 3.5c. Distributed by the author.—University of Washington, Seattle: Department of Genetics.

13. Fleischmann, R. D., Adams, M. D., White, O., Clayton, R. A., Kirkness, E. F., Kerlavage, A. R., Bult, C. J., Tomb, J-F., Dougherty, B. A., Merrick, J. M., McKenney, K., Sutton, G., FitzHugh, W., Fields, C., Gocayne, J. D., Scott, J., Shirley, R., Liu, L-I., Glodek, A., Kelley, J. M., Weidman, J. F., Phillips, C. A., Spriggs, T., Hedblom, E., Cotton, M. D., Utterback, T. R., Hanna, M. C., Nguyen, D. T., Saudek, D. M., Brandon, R. C., Fine, L. D., Fritchman, J. L., Fuhrmann, J. L., Geoghagen, N. S. M., Gnehm, C. L., McDonald, L. A., Small, K. V., Fraser, C. M., Smith, H. O. & Venter, J. C. 1995. Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd. *Science*, **269**: 496–512.

14. Fliegel, L., Burns, K., MacLennan, D. H., Reithmeier, R. A. F. & Michalak, M. 1989. Molecular cloning of the high affinity calcium-binding protein (calreticulin) of skeletal muscle sarcoplasmic reticulum. *J. Biol. Chem.*, **264**:21522–21528.

15. Gibbs, S. H. 1978. The chloroplast of *Euglena* may have evolved from symbiotic green algae. *Can. J. Bot.*, **56**:2883–2889.

16. Gumpel, N. J. & Smith, A. G. 1992. A novel calcium-binding protein from *Euglena gracilis*. *Eur. J. Biochem.*, **210**:721–727.

17. Hawn, T. R., Tom, T. D. & Strand, M. 1993. Molecular cloning and expression of SmIrV1, a *Schistosoma mansoni* antigen with similarity to calnexin, calreticulin, and OvRal1. *J. Biol. Chem.*, **268**:7692–7698.

18. Helenius, A., Trombetta, E. S., Herbert, D. N. & Simons, J. F. 1997. Calnexin, calreticulin and the folding of glycoproteins. *Trends Cell Biol.*, **7**:193–200.

19. Henze, K., Schnarrenberger, C., Kellermann, J. & Martin, W. 1994. Chloroplast and cytosolic triosephosphate isomerases from spinach: purification, microsequencing and cDNA cloning of the chloroplast enzyme. *Plant Mol. Biol.*, **26**:1961–1973.

20. Henze, K., Badr, A., Wettern, M., Cerff, R. & Martin, W. 1995. A nuclear gene of eubacterial origin in *Euglena gracilis* reflects cryptic endosymbioses during protist evolution. *Proc. Natl. Acad. Sci.* USA, **92**:9122–9126.

21. Hirsch, D. 1994. Operons in eukaryotes follow the spliced leader. *Nature*, **372**:222–223.

22. Joshi, M., Pogue, G. P., Duncan, N. S., Lee, N. K., Singh, H., Atreya, C. D., Dwyer, D. M. & Nakhasi, H. L. 1996. Isolation and characterization of *Leishmania donovani* calreticulin gene and its conservation of the RNA binding activity. *Mol. Biochem. Parasitol.*, **81**: 53–64.

23. Kaneko, T., Sato, S., Kotani, H., Tanaka, A., Asamizu, E., Nakamura, Y., Miyajima, N., Hirosawa, M., Sugiura, M., Sasamoto, S., Kimura, T., Hosouchi, T., Matsuno, A., Muraki, A., Nakazaki, N., Naruo, K., Okumura, S., Shimpo, S., Takeuchi, C., Wada, T., Watanabe, A., Yamada, M., Yasuda, M., & Tabata, S. 1996. Sequence analysis of the genome of the unicellular cyanobacterium *Synecocystis* sp. Strain PCC6803. II. Sequence determination of the entire genome and assignment of potential protein-coding regions. *DNA Res.*, **3**:109–136.

24. Khanna, N. C., Tokuda, M. & Waisman, D. M. 1987. Comparison of calregulins from vertebrate livers. *Biochem. J.*, **242**:245–251.

25. Krause, K.-H. & Michalak, M. 1997. Calreticulin. *Cell*, **88**:439–443.

26. Laemmli, U. K. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature*, **227**:680–685.

27. Levasseur, P. J., Meng, K. & Bouck, B. 1994. Tubulin genes in the algal protist *Euglena gracilis*. J. *Euk. Microbiol.*, **41**:468–477.

28. Martin, W., Somerville, C. C. & Loiseaux—de Goër, S. 1992. Molecular phylogenies of plastid origins and algal evolution. *J. Mol. Evol.*, **35**:385–403.

29. Maruyama, K., Mikawa, T. & Ebashi, S. 1984. Detection of calcium binding proteins by ^{45}Ca autoradiography on nitrocellulose mem-

brane after sodium dodecyl sulfate gel electrophoresis. *J. Biochem.*, **95**: 511–519.

30. Matsuoka, K. M., Seta, K., Yamakawa, Y., Okuyama, T., Shinoda, T. & Ysobe, T. 1994. Covalent structure of bovine brain calreticulin. *Biochem J.*, **298**:435–442.

31. McFadden, G. & Gilson, P. R. 1995. Something borrowed, someting green: lateral transfer of chloroplasts by secondary endosymbiosis. *Trends Ecol. and Evol.*, **10**:12–17.

32. Meggio, F., Boldyreff, B., Marin, O., Marchiori, F., Perich, J. W., Issinger, O-G. & Pinna, L. A. 1992. The effect of polylysine on casein-kinase-2 activity is influenced by both the structure of the protein/peptide substrates and the subunit composition of the enzyme. *Eur. J. Biochem.*, **205**:939–945.

33. Meggio, F., Marin, O. & Pinna, L. A. 1994. Substrate specificity of protein kinase CK2. *Cell Mol. Biol. Res.*, **40**:401–409.

34. Meldolesi, J., Krause, K.-H. & Michalak, M. 1996. Calreticulin: how many functions in how many cellular compartments? *Cell Calcium*, **20**:83–86.

35. Michalak, M. 1996. Calreticulin. Springer-Verlag, Heidelberg. Pp.1–211.

36. Michalak, M., Milner, R. E., Burns, K. & Opas, M. 1992. Calreticulin. *Biochem. J.*, **285**:681–692.

37. Milner, R. E., Famulski, K. F. & Michalak, M. 1992. Calcium binding proteins in the sarcoplasmic/endoplasmic reticulum of muscle and nonmuscle cells. *Mol. Cell. Biochem.*, **112**:1–13.

38. Milner, R. E., Baksh, S., Shemanko, C., Carpenter, M. R., Smillie, L., Vance, J. E., Opas, M. & Michalak, M. 1991. Calreticulin, and not calsequestrin, is the major calcium binding protein of smooth muscle sarcoplasmic reticulum and liver endoplasmic reticulum. *J. Biol. Chem.*, **266**:7155–7165.

39. Napier, R. M., Trueman, S., Henderson, J., Boyce, J. M., Hawes, C., Fricker, M. D. & Venis, M. A. 1995. Purification, sequencing and functions of calreticulin from maize. *J. Exp. Bot.*, **46**:1603–1613.

40. Navazio, L., Baldan, B., Mariani, P., Gerwig, G. J. & Vliegenthart, J. F. G. 1996. Primary structure of the N-linked carbohydrate chains of calreticulin from spinach leaves. *Glycoconjugate J.*, **13**:977–983.

41. Navazio, L., Baldan, B., Dainese, P., James, P., Damiani, E., Margreth, A. & Mariani, P. 1995. Evidence that spinach leaves express calreticulin but not calsequestrin. *Plant Physiol.*, **109**:983–990.

42. Nelson, D. E., Glaunsinger, B. & Bohnert, H. J. 1997. Abundant accumulation of the calcium-binding molecular chaperone caleticulin in specific floral tissues of *Arabidopsis thaliana*. *Plant Physiol.*, **114**:29–37.

43. O'Farrell, P. H. 1975. High resolution two-dimensional electrophoresis of proteins. *J. Biol. Chem.*, **250**:4007–4021.

44. Plattner, H., Habermann, A., Kissmehl, R., Klauke, N., Majoul, I. & Söling, H-D. 1997. Differential distribution of calcium stores in *Paramecium* cells. Occurrence of a subplasmalemmal store with a calsequestrin-like protein. *Eur. J. Cell Biol.*, **72**:297–306.

45. Pedrazzini, E. & Vitale, A. 1996. The binding protein (BiP) and the synthesis of secretory proteins. *Plant Physiol. Biochem.*, **34**:207–216.

46. Pozzan, T., Rizzuto, R., Volpe, P. & Meldolesi, J. 1994. Molecular and cellular physiology of intracellular calcium stores. *Physiol. Rev.*, **74**:595–636.

47. Saitou, N. & Nei, M. 1987. The neighbor-joining method: A new method for the reconstruction of phylogenetic trees. *Mol. Biol. Evol.*, **4**:406–425.

48. Sogin, M., Gunderson, J., Elwood, H., Alonso, R. & Peattie, D. 1989. Phylogenetic meaning of the kingdom concept: an unusual ribosomial RNA from *Giardia lamblia*. *Science*, **243**:75–77.

49. Sulli, C. & Schwartzbach, S. D. 1995. The polyproteine precursor to the *Euglena* light-harvesting chlorophyll a/b-binding protein is transported to the Golgi apparatus prior to chloroplast import and polyprotein processing. *J. Biol. Chem.*, **270**:13084–13090.

50. Sulli, C. & Schwartzbach, S. D. 1996. A soluble protein is imported into *Euglena* chloroplasts as a membrane-bound precursor. *Plant Cell*, **8**:34–53.

51. Surek, B. & Melkonian, M. 1986. A cryptic cytostome is present in *Euglena*. *Protoplasma*, **133**:39–49.

52. Tessier, L. H., Keller, M., Chan, R. L., Fournier, R., Weil, J-H. & Imbault, P. 1991. Short leader sequences may be transferred from small RNAs to pre-mature mRNAs by trans-splicing in *Euglena*. *EMBO J.*, **10**:2621–2625.

53. Thompson, J. D., Higgins, D. G. & Gibson, T. J. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucl. Acids Res.*, **22**:4673–4680.

54. Towbin, H., Stahelin, T. & Gordon, J. 1979. Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. *Proc. Natl. Acad. Sci.* USA, **76**:4350–4354.

55. Van de Peer, Y., Rensing, S., Maier, U-G. & De Wachter, R. 1996. Substitution rate calibration of small subunit ribosomal RNA identifies chlorarachniophyte endosymbionts as remnants of green algae. *Proc. Natl. Acad. Sci.* USA, **93**:7732–7736.

56. Vickerman, K. 1990. Phylum Zoomastigina class Kinetoplastida. In: Margulis, L., Corliss, J. O., Melkonian, M. & Chapman, D. J. (ed.), Handbook of Protoctista. Jones & Bartlett, Boston. Pp. 215–238.

57. Walne, P. L. & Kivic, P. A. 1990. Phylum Euglenida. In: Margulis, L., Corliss, J. O., Melkonian, M. & Chapman, D. J. (ed.), Handbook of Protoctista. Jones & Bartlett, Boston. Pp. 270–287.

58. Williams, R. J. P. 1996. Calcium binding proteins in normal and transformed cells. *Cell Calcium*, **20**:87–93.

59. Wolken, J. J. 1967. *Euglena*. Appleton-Century Crofts, New York. Pp. 174–180.

Received 9-16-97, 12-5-97; accepted 1-7-98.