Analysis of a Novel Mitosomal β-Barrel Outer Membrane Protein in *Entamoeba histolytica*

A Dissertation Submitted to

the Graduate School of Life and Environmental Sciences,

the University of Tsukuba

in Partial Fulfillment of the Requirements

for the Degree of Doctor of Philosophy in Biological Sciences

(Doctoral Program in Biological Sciences)

Herbert J. Santos

Table of Contents

Abstract

Entamoeba possesses a highly divergent mitochondrion-related organelle known as the mitosome, which has an outer membrane containing β-barrel Outer Membrane Proteins (BOMPs). Mitosomes are highly degenerate organelles, lacking many components and functions associated with the canonical aerobic mitochondria. In fact, the lone established role of *Entamoeba* mitosomes currently known is the compartmentalization of the sulfate activation pathway. Thus, further searches for functional homologs and unique proteins associated with sulfate activation, or even mitochondrial processes in the mitosomes of *Entamoeba* is warranted to promote better understanding of the biology and evolution of this organelle and MROs in general. Here, I report the discovery of a novel protein in the mitosome of *Entamoeba*, which is named Mitosomal β-barrel Outer Membrane Protein of 30 kDa (MBOMP30). Initially identified through *in silico* analysis, I experimentally confirmed that MBOMP30 is indeed a β-barrel protein. The data from circular dichroism analysis indicated that MBOMP30 has a predominant β-sheet structure. Localization to *Entamoeba histolytica* mitosomes was observed through immunofluorescence assay and Percoll-gradient fractionation. Mitosomal membrane integration was demonstrated by carbonate fractionation, proteinase K digestion, and immunoelectron microscopy. Interestingly, MBOMP30 also possesses a putative β-signal, a sequence believed to guide β-barrel outer membrane protein (BOMP) assembly. The deletion of the β-signal did not affect membrane integration, but abolished the formation of a ~240 kDa complex. MBOMP30 represents only the seventh subclass of BOMPs in mitochondria and related organelles,

discovered to date. Three of them (Tom40, Sam50, and VDAC) are found in diverse branches of the Eukaryota, while the others have only been clearly confirmed in specific lineages, namely Mdm10 in yeasts and ATOM and Tac40 in trypanosomatids. Interestingly, results of searches in current public databases indicate that only *Entamoeba* species have detectable homologs to MBOMP30, suggesting that this protein may be unique to *Entamoeba* mitosomes.

Introduction

One of the major highlights in the evolution of eukaryotes throughout its biological history is the endosymbiotic event that paved the way to the formation of present organelles namely mitochondria and chloroplasts. Mitochondria, as originally proposed by Margulis in 1970, are believed to have arisen from a α-proteobacterium that was engulfed by a methanogenic archaea or an ancient eukaryote, and established a successful endosymbiotic relationship. As eukaryotic lineages begin to diverge, in response to changing environments and establishment of niches, organelles have undergone certain modifications that provide better adaptation for survival.

Mitochondria and related organelles

Mitochondria can possess highly divergent and often degenerate morphology, function, and components in eukaryotes that are adapted to anoxic or hypoxic environments. In cases in which morphology is drastically changed and some hallmark mitochondrial processes such as oxidative phosphorylation, the TCA cycle, and â-oxidation are lost, these organelles are called Mitochondrion-Related Organelles (MROs), or specifically, hydrogenosomes and mitosomes. Mitosomes represent the simplest form of a mitochondrion, and being particularly degenerate organelles, they lack cristae structure, and even the ability to synthesize ATP. It is believed that mitosomes, as well as hydrogenosomes, have occurred multiple times during eukaryotic evolution because organisms that possess mitosomes do not cluster together in eukaryote phylogenies, and the size, function, and content of mitosomes differ amongst organisms (Heinz and Lithgow, 2012; Lithgow

and Schneider, 2010; Makiuchi and Nozaki, 2014).

β-barrel Outer Membrane Proteins

Like Gram-negative bacteria and chloroplasts in land plants, green algae, red algae and glaucophytes, mitochondria and MROs possess two membranes. Transport of proteins and metabolites across the outer membrane is mediated by pore-formingβ-barrel Outer Membrane Proteins. (Hereafter "BOMP" is used to denote any β-barrel outer membrane protein and "MBOMP" to denote BOMPs from mitochondria and MROs).

In mitochondria, six subclasses of MBOMPs have been previously identified: Tom40, Sam50, VDAC, Mdm10, ATOM and Tac40. Tom40 is the core pore component of the Translocase of the Outer Membrane (TOM) complex required for the import of mitochondrial precursor proteins into mitochondria (Baker et al., 1990; Hill et al., 1998). Sam50 is the central component of the Sorting and Assembly Machinery (SAM) complex and promotes the integration of MBOMPs (Kozjak et al., 2003; Paschen et al., 2003; Wiedemann et al., 2003) to the outer membrane. Both Tom40 and Sam50 are essential for yeast viability (Baker et al., 1990; Hill et al., 1998; Kozjak et al., 2003). VDAC (Voltage-Dependent Anion Channel) primarily serves as a non-specific diffusion pore for small molecules entering or leaving the mitochondria (Colombini, 2004). Mdm10 (Mitochondrial Dynamics and Morphology 10) has only been clearly identified in fungi, and is involved in mitochondrial morphogenesis and dynamics (Sogo and Yaffe, 1994), as well as in the biogenesis of mitochondrial BOMPs, as it was reported to be a part of the SAM complex (Yamano et al., 2010). It is also a member of the ER-

mitochondria tethering complex known as the ER-Mitochondria Encounter Structure, ERMES (Kornmann et al., 2009). Trypanosomatids lack Tom40, and instead have a unique translocase called ATOM (Pusnik et al., 2011) (Archaic Translocase of the Outer Mitochondrial membrane). Recently another trypanosome-specific MBOMP Tac40, a member of the Tripartite Attachment Complex, was identified. Tac40 also belongs to the mitochondrial porin family, and is essential to mitochondrial DNA inheritance, as it physically links the mitochondrial genome to cytoskeletal components of both the mitochondrion and flagellum of *Trypanosoma brucei* (Schnarwiler et al., 2014). Among all MBOMPs, only Sam50 and ATOM have recognizable homologs outside of the Eukaryota. Sam50 has a bacterial homolog, Omp85/BamA, a 16-stranded bacterial BOMP (Gentle et al., 2005), which also integrates and assembles BOMPs (Gentle et al., 2004); and the homolog of ATOM belongs to the Omp85 subgroup called YtfM (Pusnik et al., 2011), which is a BOMP required for normal growth in *Escherichia coli* (Stegmeier et al., 2007).

BOMP assembly and integration

Precursor BOMPs synthesized in the cytosol are assembled and integrated in the outer membrane of endosymbiotic organelles namely chloroplasts, mitochondria and MROs, and the evolutionarily related Gram-negative bacteria. Some conserved features and mechanisms exist between the machineries required for BOMP assembly in eukaryotic and prokaryotic systems. However, there are also some notable differences, such as the presence of conserved signal sequences, and the exactly opposite direction by which the manner of protein integration to the outer membrane proceeds.

In Gram-negative bacteria, precursor BOMPs expressed in the cytoplasm has to pass through the inner membrane using the channel forming Sec translocase. Once in the periplasmic space, chaperone molecules assist the nascent polypeptide to the BAM complex, specifically to BamA/Omp85, also a BOMP, which is responsible for properly folding and integrating the BOMP to the outer membrane (Tamm et al., 2001;Walther et al., 2009).In the case of the mitochondria, the nuclear-encoded BOMP gene is expressed and the protein synthesized in the cytosol. The precursor protein goes to the mitochondria through the TOM complex, specifically the Tom40 channel which is a BOMP itself. Upon entry to the intermembrane space, small Translocase of the Inner Membrane (small TIM) complexes Tim9-Tim10 and Tim8-Tim13 interact with the precursor BOMP to direct it to the SAM complex (Paschen et al., 2003; Wiedemann et al., 2003). Sam50, which is also a BOMP, works in the same way as its bacterial homolog BamA/Omp85, by assembling and integrating the protein to the outer membrane of the organelle (Taylor and Pfanner, 2004; Chacinska et al., 2009; Schmidt et al., 2010). In yeasts in particular, a short amino acid sequence at the last β-strand near the carboxyl end of the protein called β-signal is involved in BOMP recognition by Sam35, another component of the SAM complex, prior to its integration to the outer membrane (Kutik et al., 2008).

Entamoeba possesses mitosomes

Entamoeba histolytica is an anaerobic unicellular parasite that causes dysentery and extra-intestinal abscesses that are responsible for an estimated 100,000 deaths annually. This organism possesses highly

divergent mitosomes, as predicted by a recent proteomic study (Mi-ichi et al., 2009). It appears that the mitosome proteome in *E. histolytica* is far less complex than mitochondria (e.g., yeast mitochondria are believed to harbor around 1000 proteins (Meisinger et al., 2008; Reinders et al., 2006)), and remarkably different from other MROs. Indeed, even Fe-S cluster biogenesis, which is the only known common function of mitochondria and MROs, is uncertain in *E. histolytica* mitosomes (Ali et al., 2004; van der Giezen et al., 2004). It was reported that the iron-sulfur (Fe-S) cluster assembly genes *iscS* and *iscU* of *E.histolytica* were acquired by horizontal gene transfer (Maralikova et al., 2010; van der Giezen et al., 2004). Unlike other organisms which use ISC system for Fe-S cluster biogenesis, *E. histolytica* (Ali et al., 2004; van der Giezen et al., 2004) and the distantly related *Mastigamoeba balamuthi* (Gill et al., 2007) use a NIF (nitrogen fixation)-like system for Fe-S cluster biogenesis; which is primarily, if not exclusively, cytosolic and appears to have been obtained by horizontal transfer from ε-proteobacteria. Thus far, the only established role of the *Entamoeba* mitosome is sulfate activation (Mi-ichi et al., 2009; Mi-ichi et al., 2011) (Figure 1).

MBOMPs in *Entamoeba*

Only two homologs of MBOMP subclasses have been found in *Entamoeba* in the recent years, namely Tom40 and Sam50 (Dolezal et al., 2010) (Figure 1). Among the many “missing links” in understanding mitosomal biology is the conspicuous absence of detectable homologs to VDAC, an MBOMP which is the usual channel for metabolites in mitochondrial outer membrane. Intrigued by this, we endeavored to collaborate and

combine bioinformatics and experimental techniques to search for novel *Entamoeba* BOMPs. It was also motivating that the previously developed method by collaborators (Imai et al., 2008), (Imai et al., 2011), used for predicting MBOMPs from amino acid sequence had already been able to predict a candidate novel plastid BOMP TGD4 (At3g06960), which subsequently confirmed to localize to the plastid outer membrane (Wang et al., 2012); These efforts led to the discovery of *E. histolytica* MBOMP30 (Mitosomal β-barrel Outer Membrane Protein of 30kDa), a novel subclass of mitosomal BOMP unique to *Entamoeba* species. The objectives of this study were to, analyze and characterize MBOMP30 to give concrete support to the *in silico* prediction that the protein is indeed a β-barrel protein, by showing that it possesses high amounts of β-strands, and that it is localized and integrated to the mitosome membrane. In addition, I also explored on the possible role of the β-signal in MBOMP30.

Materials and Methods

MBOMP prediction pipeline

For MBOMP prediction, our collaborators from the Computational Biology Research Center (CBRC), National Institute of Advanced Industrial Science and Technology (AIST), Japan, prepared a modified version of their Support Vector Machine (SVM)-based predictor, developed for a previous study (Imai et al., 2011). Like other BOMP predictors (Berven et al., 2004; Park et al., 2005), this predictor considers physicochemical features reflecting BOMP structural motifs (such as β-strands with alternating hydrophobic residues) expected to be found in the β-barrel region. Unique to this predictor is the consideration of sorting signals; e.g. the β-signal, which MBOMPs might be expected to have, and signal peptides which MBOMPs are expected *not* to have (Figure 2). For this study the predictor was trained on a dataset including mitosomal and hydrogenosomal BOMP sequences, to improve its ability to detect novel BOMPs in MROs.

The first step in the pipeline requires a relatively high MBOMP probability score (probability > 0.7) from the MBOMP classifier. Additionally it requires that the secondary structure composition as predicted by PSI-PRED (McGuffin et al., 2000) to be at least 25% β-strand and no more than 25% α-helix.

Training Dataset

A dataset of 81 MBOMP sequences (including presumed MBOMPs inferred by sequence similarity) was prepared, consisting of 71 mitochondrial BOMPs and 10 MRO-BOMPs: Tom40 and Sam50 from *E. histolytica, E. invadens, Giardia intestinalis, Encephalitozoon cuniculi, Cryptosporidium*

parvum, Trichomonas vaginalis and *Blastocystis hominis*. No pair of positive examples shared more than 40% identity. For negative examples, a total of 2464 non-MBOMP yeast proteins with clear Uniprot annotation and less than 20% mutual sequence identity was used. The sequence data were obtained from Uniprot and EuPathDB (Aurrecoechea et al., 2010; The Uniprot Consortium, 2012).

Newly Integrated β-Signal and Physicochemical Sequence Features

As shown in Figure 3, most MRO-BOMPs in the training set contained matches to the β-signal, therefore our collaborators decided to add β-signal-inspired features to the MBOMP predictor via two representations: Position Weight Matrix (PWM) and regular expression. To define the PWM based β-signal inspired features, the 81 training MBOMPs were divided by protein family and MAFFT (Katoh et al., 2002) was used to obtain four multiple alignments (one for each MBOMP family). Then a total of 81 aligned regions (octomers) matching the β-signal from the multiple alignments were extracted and the duplicate octomers were discarded to obtain 71β-signal examples. Next a PWM based on those 71 examples and background frequencies as described in their previous study (Imai et al., 2011) was defined. Finally, the two features derived from this PWM were defined; its maximum value over the C-terminal 50 residues and over entire sequence.

Also five binary features based on the occurrence of matches in the C-terminal 50 residues to regular expressions representing the β-signal motif were defined. For these regular expressions the original pattern $P_oxGxxH_yxH_y$, as proposed by Kutik et al. (2008) and four variants: $P_oH_yGh_yxH_yxH_y$, $P_oH_yGh_y\hat{H}_yH_yxH_y$, $\hat{H}_yH_yGh_yxH_yxH_y$ and $\hat{H}_yH_yGh_y\hat{H}_yH_yxH_y$, proposed in their

previous work (Imai et al., 2011) were used. In these expressions x is any residue, H_y, h_y and P_o denote [VLIMFYW], [ACVLIMFYW], and [KRHSTNQ], respectively, and $\hat{H}_y$ matches any residue not included in H_y.

To integrate physicochemical sequence features, an easily implemented method was used, potentially enabling the SVM can automatically learn the characteristics of various sorting signals based on the physicochemical properties of sequence segments. The amino- and carboxyl-terminal 90 residues into 6 blocks of 15 residues were partitioned, and then the average hydrophobicity, α-helical periodicity score (Imai, 2005), and the density of positive, negative, and aromatic residues for each block were all computed. These features are also relevant to structural motifs.

MBOMP30 homolog search, sequence comparison and topology prediction

Search for MBOMP30 homologs in other organisms was performed using the JackHMMR and HHblits algorithms (Finn et al., 2011; Pearson, 1991). Pairwise alignments and E-values of MBOMP30 homologs among *E. histolytica*, *E. dispar* and *E. invadens* were computed with the SSEARCH program (Pearson, 1991), and multiple alignments were built with Clustal omega (Sievers et al., 2011). For topology prediction of MBOMP30 and its homologs, BOCTOPUS and TMBETAPRED-RBF were used (Ou et al., 2010).

Cell-free synthesis of EhMBOMP30 proteoliposomes

The open reading frame of the *E. histolytica* MBOMP30 gene EHI_178630 was codon-optimized for wheat-germ expression, and cloned to

pY08 vector using *Spe*I and *Sal*I restriction sites. Messenger RNA was synthesized *in vitro* as previously described (Nozawa et al., 2007). Briefly, EhMBOMP30 was synthesized by wheat-germ expression system (Nozawa et al., 2007), in the presence of 0.5 mg/mL 1,2-diphytaynyl-sn-glycero-3-phosphocholine (DPhPc) liposomes (Avanti Polar Lipids, Alabaster, AL).

Far-UV and circular dichroism spectroscopy analysis of MBOMP30 proteoliposome

Synthesized *E. histolytica* MBOMP30 proteoliposomes were pelleted, then washed twice with CD buffer (25 mM sodium phosphate buffer, pH 7.0) by centrifugation at 15,000 rpm at 4°C for 20 min. The pellet was resuspended in 150 µL CD buffer and sonicated briefly until translucent. Recombinant His-tagged GFP reconstituted in DPhPc liposomes was used as a positive control. CD measurements were performed using a Jasco J-820 Spectropolarimeter, and the samples were loaded onto a quartz cuvette with 0.1-cm path length (Starna Cells, Inc, Atascadero, CA). The far-UV spectra (185 to 250 nm) were obtained by taking the average of 9 consecutive scans with a resolution of 1 nm, a scanning speed of 200 nm/min and a response time of 2 s. The secondary structure ratios were estimated by the CONTIN algorithm (Provencher and Glockner, 1981; van Stokkum et al., 1990) using a 42-protein set database available through the Dichroweb online tool (Whitmore and Wallace, 2004, 2008).

Plasmid construction

Total RNA was isolated from *E. histolytica* trophozoites by TRIZOL®

reagent (Invitrogen, Carlsbad,San Diego, CA). Poly-A RNAs were purified using GenElute^{TM} mRNA Miniprep Kits (Sigma-Aldrich, Japan). cDNA was synthesized from poly-A RNAs using SuperScript^{TM} III RNase H^{-} reverse transcriptase (Invitrogen), and oligo(dT)$_{20}$ primer (Invitrogen). The *E. histolytica* gene EHI_178630 was PCR-amplified from cDNA using Phusion DNA polymerase (New England Biolabs, Beverly, MA) using the following primer sets (sense primer; 5'- GTTCCCGGGATGTTGGGTAAAACTGC -3' and antisense primers; 5'- GAACTCGAGTTAAAGTGATAAATCAATTCCA -3' or

5'- GAACTCGAGTTAGAGTTGATTTTTCTGGTCTT -3') for full length (HA-MBOMP30) or β-signal-truncated (HA-MBOMP30$^{\Delta 275\text{-}282}$) inserts respectively. After digestion by *Xma*I and *Xho*I, the restriction fragments were ligated into pEhEx-HA (Nakada-Tsukui et al., 2009) using Ligation-Convenience Kit (Nippongene, Tokyo, Japan).

Cell culture and amoeba transformation

Trophozoites of *Entamoeba histolytica* HM-1:IMSS cl6 (Diamond et al., 1978) and G3 (Bracha et al., 2006) strains were cultivated axenically in Diamond BI-S-33 medium (Diamond et al., 1978). Lipofection of amoebic trophozoites, selection, and maintenance of transformants were performed as previously described (Mi-ichi et al., 2009). Briefly trophozoites from semi-confluent culture were harvested and washed thrice using ice-cold 2% glucose/PBS. Approximately 5 x 10^5 cells were seeded on each well of a 12-well plate, and 400 mL BIS was added before incubating at 35.5°C under microaerophilic conditions to allow the cells to attach to the surface of the

wells. Then, 5 µg of the plasmid was mixed with 10 µL of Plus reagent (Invitrogen) and transfection medium [1X Opti-MEM (Invitrogen), 1% ascorbic acid and 0.5% L-cysteine with final pH of 6.8] to a total volume of 50 µL. The solution was incubated for 10 min at room temperature before adding 20 µL Lipofectamine reagent (Life Technologies) and 30 µL transfection medium. After 10 minutes, the plate containing amoeba cells were removed from the incubator. Excess medium was discarded, and the plasmid mixture was then added to individual wells to commence transfection. Then, 400 µL of transfection medium was added to each well before a five-hour incubation at 35.5°C microaerophilic conditions. Then, the plate was put on ice to allow cell detachment. Transfected cells were seeded onto amoeba tubes containing BIS medium and were incubated at 35.5°C. Selection of overexpressing or gene silencing mutants was performed by daily replacement of medium supplemented with Geneticin/G418 (Gibco) at increasing concentration, until control cells (transfected without plasmid) fail to survive the antibiotic challenge.

Immunoflourescence assay (IFA)

IFA was performed as previously described (Mi-ichi et al., 2009; Nakada-Tsukui et al., 2005). Amoeba trophozoites were fixed onto glass slides using 3.7% paraformaldehyde in PBS for 10 min at room temperature. Then, the slides were washed thrice with PBS, and the cells were made permeable, and blocked using 0.2% saponin in 1% bovine serum albumin (BSA) in PBS for 10 min at room temperature. Double staining of the cells was performed using mouse anti-HA antibody and rabbit anti-adenosine-5'-

phosphokinase [(APSK; XP_656278; mitosomal matrix marker (Mi-ichi et al., 2009)] antisera for 1 hr at room temperature. After washing the slides thrice using 0.1% BSA in PBS, the cells were reacted with goat anti-mouse IgG-conjugated to Alexa Fluor 488 and goat anti-rabbit IgG-conjugated to Alefa Fluor 568 (Invitrogen) for 1 hr at room temperature. The slides were washed thrice with 0.1% BSA in PBS, then each well was overlayed with 2 μL mounting medium, before viewing using Carl Zeiss LSM 510 Meta laser-scanning confocal microscope.

Preparation of organelle fraction

Amoeba strains that expressed HA-MBOMP30, Tom40-HA (Makiuchi et al., 2013) , AAC-HA (Mi-ichi et al., 2009), and APSK-HA (Mi-ichi et al., 2009) proteins, and also the mock transformant, pEhEx-HA, were washed three times with 2% glucose/PBS. After resuspension in lysis buffer (10 mM MOPS-KOH, pH 7.2, 250 mM sucrose, protease inhibitors), the cells were disrupted mechanically by a Dounce homogenizer. Unbroken cells were removed by centrifugation at 5,000 *g* for 10 min, and the supernatant centrifuged at 100,000 *g* for 60 min to separate the organelle and cytosolic fractions. The 100,000 *g* organelle fractions were resuspended with lysis buffer, and were recollected by centrifugation at 100,000 *g* for60 min.

Percoll-gradient Ultracentrifugation

Sequential Percoll-gradient fractionation was performed following the protocol previously reported (Mi-ichi et al., 2009). Briefly, organelle fractions of HA-MBOMP30 overexpressing trophozoites were collected and overlayed

onto a 30% Percoll in lysis buffer, before centrifugation at 100,000 *g* for 1 hr at 4°C. Fractions of 200 µL each were collected, meanwhile 150 µL of the last 8 fractions were pooled and placed on top of a 70% Percoll-lysis buffer solution. Then, 15% Percoll-lysis buffer solution was placed on top of the pooled fractions from the first ultracentrifugation. It was subjected to a second ultracentrifugation step following the same conditions, after which 200 µL fractions were collected as well. The samples from both first and second ultracentrifugation steps were run in denaturing sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) and then blotted onto nitrocellulose membranes reacted with anti-HA and anti-Chaperonin 60 (Cpn60; a canonical soluble mitosomal matrix marker) antibodies respectively.

Sodium carbonate mitosomal membrane fractionation

Na_2CO_3 treatment of mitosomal membrane fractions was performed as previously described (Makiuchi et al., 2013) but using Cpn60 as a soluble mitosome protein marker. Organelle fractions were washed once in lysis buffer. A total of 1 mg of washed organelle fraction was diluted 20-fold with 100 mM Na_2CO_3 and further reacted with 150 mM NaCl. To break the organellar membranes, the suspension was mixed well, and then incubated on ice for 30 minutes. The soluble organelle proteins were separated from the membrane-integrated proteins by ultracentrifugation at 100,000 *g*, 4°C for 1 hr. The supernate (soluble organelle fraction) was collected while the particulate (membrane) fraction was washed once with 100 mM Na_2CO_3 and 150 mM NaCl. Finally, the four fractions containing cytosolic, organelle, soluble organelle, and organellar membrane proteins were subjected to denaturing

SDS-PAGE, followed by immunoblotting using anti-HA and anti-Cpn60 antibodies.

Proteinase K protection assay

Proteinase K assay was performed following the previous protocol with some modifications (Makiuchi et al., 2013). Briefly, 100 µg protein from organelle fractions of amoebic trophozoites were prepared, then treated with proteinase K (the final concentration was 3.125 µg/ml) (Roche) for 15 min at 4°C. Several control strains were used to estimate the topology of HA-MBOMP30. Tom40-HA, AAC-HA, and APSK-HA were used as controls for outer membrane, inner membrane, and soluble matrix mitosomal proteins respectively. Control samples without proteinase K were likewise prepared for all overexpressing strains. Samples were then run on SDS-PAGE followed by immunoblot analysis using anti-HA mouse monoclonal antibody and HRP-conjugated anti-mouse antibody (Thermo Scientific) as primary and secondary antibodies respectively. Quantitation of band intensities was performed using the Analysis Toolbox in ImageQuant TL software (GE Healthcare).

Immunoelectron microscopy

Sample preparation was carried out as previously described(Mi-ichi et al., 2011) with some modifications. Trophozoites overexpressing HA-MBOMP30, HA-MBOMP30$^{\Delta 275\text{-}282}$ and mock transformant (pEhEx-HA) in BIS medium were incubated with gold disks at 35.5°C for 15 minutes to allow for attachment. The disks carrying amoebae were then frozen in liquid propane at

-175°C. Once frozen, the samples were freeze-substituted with 0.2% glutaraldehyde in ethanol and 3% distilled water at -80°C overnight. The samples were infiltrated with a 50:50 mixture of ethanol and resin (LR White: London Resin Co., Ltd., Berkshire, UK) at 4°C for 30 min. Afterwards, three changes of 100% LR white at 4°C for 30 min each were performed. The samples were then transferred to a fresh 100% resin, followed by an overnight polymerization at 50°C. Then, the blocks were ultra-thin sectioned at 80 nm using a diamond knife of an ultramicrotome (ULTRACUT UCT, Leica) and the resulting sections were placed on nickel grids. The disks were then double-stained using anti-HA mouse and anti-Cpn60 rabbit antibodies in 1% BSA-PBS for 90 min at room temperature, then rinsed with 1% BSA-PBS three times for 1 min each. Then, they were floated on drops of the secondary anti-mouse and anti-rabbit IgG antibody conjugated to gold particles for 1 hr at room temperature. They were rinsed and placed in 2% glutaraldehyde in 0.1 M phosphate buffer before drying. Then, the grids were stained with 2% uranyl acetate for 15 min, and secondary-stained with lead stain solution at room temperature for 3 min. The grids were observed at Tokai Microscopy Inc. (Nagoya, Japan), using a transmission electron microscope (JEM-1400 Plus, JEOL Ltd.,) at an acceleration voltage of 80kV. Digital images (2048 x 2048 pixels) were taken with a CCD camera (VELETA, Olympus Soft Imaging Solution GmbH).

Immunoprecipitation of the EhMBOMP30 complex

Organelle fractions of HA-MBOMP30 and HA-MBOMP30$^{\Delta 275-282}$overexpressing strains, and pEhEx-HA control were prepared and

solubilized in 2% digitonin in IP Buffer containing 50 mM BisTris-HCl, pH 7.2, 50 mM NaCl, 0.001 % Ponceau S, and 10 % w/v glycerol. The solubilizing condition is 2 µg of proteins per uL of reaction mix and at 4^{o}C for 30 min. Then, the reaction mix was centrifuged at 20,000 *g* for 30 min at 4^{o}C and the resulting supernatant containing solubilized proteins were transferred to a new tube. Immunoprecipitation was performed as previously described (Makiuchi et al., 2013).The solubilized organelle fraction was reacted with Protein G-Sepharose 4 Fast Flow (GE Healthcare) to remove proteins with non-specific interactions with the target proteins. The mixture was rotated at 4^{o}C for 60 min. Then, the Sepharose beads were removed by centrifugation at 16,100 *g* at 4^{o}C for 1 min. The pre-cleared solubilized organelle fractions were then reacted with 30 µL anti-HA mouse monoclonal antibody conjugated with agarose (Sigma-Aldrich Japan) at 4°C for 3 h with rotation. Then, the anti-HA agarose beads were dropped by centrifugation at 16,100 *g* at 4^{o}C for 1 min. The anti-HA agarose beads were washed three times with IP buffer containing 1% digitonin. Afterwards, the anti-HA agarose beads were dropped by centrifugation at 16,000 *g* at 4^{o}C for 1 min. Bound proteins were eluted overnight, while rotating at 4^{o}C in IP buffer containing 1% digitonin and 600 µg/ml HA peptide (Sigma-Aldrich Japan). The samples were loaded on both blue-native polyacrylamide gel electrophoresis (BN-PAGE) and SDS-PAGE gels, followed by immunoblotting and silver staining. The blots were immunodecorated with mouse anti-HA and AP-conjugated anti-mouse IgG antibodies.

Results

In silico screening of novel MBOMP candidates in *E. histolytica*

Novel MBOMP candidates in *E. histolytica* were identified by the customized MBOMP predictor (Imai et al., 2011) for MRO's, using a generally refined screening method (Figure 2) conducted by our collaborators from AIST-Japan. A total of 8,306 proteins in the *E. histolytica* genome database was screened, yielding six MBOMP candidates: EHI_178630, EHI_007460, EHI_163510, EHI_050690, EHI_068370, and EHI_104420 (Table 1). Among the six predicted MBOMPs, EHI_104420 was the top hit, and incidentally is a homolog of Tom40, which has been previously characterized (Dolezal et al., 2010). Four of the other five proteins are annotated as non-mitochondrial proteins and have sequence similarity to known functional domains. Thus, in this study, I focused to analyze the remaining protein EHI_178630 (Uniprot accession C4LUS8), which attained the second highest predictor score, second only to the known MBOMP Tom40, and is now named MBOMP30.

MBOMP30 has no detectable homologs outside of *Entamoeba*

To examine the phylogenetic distribution of MBOMP30across the three domains of life, the highly sensitive iterative profile HMM comparison search methods JackHMMR and HHblits (Finn et al., 2011; Remmert et al., 2012)were used to search for homologs of *E. histolytica* MBOMP30 in other organisms. Interestingly, no significant hits (E-value < 1e-5) were found in either bacteria or archaea. In addition, homologs of MBOMP30 were detected only in genus *Entamoeba* among all eukaryotes, namely, *E. nuttalli*, *E. dispar* and *E. invadens*. Both *E. nuttalli* and *E. dispar* are closely related to *E.*

histolytica. The former, a pathogenic parasite of non-human primates, and the latter, a non-pathogenic species infecting humans, possess very similar sequences ENU1_140620 and EDI_035580, sharing 97.2% and 86.5% identity with the *E. histolytica* homologue (although EDI_035580 is truncated, E-value 1.2e-66), respectively. On the other hand, the reptilian parasite *E. invadens*, which is distantly related to *E. histolytica*, yielded a diverged homologue EIN_041060 (with an identical sequence assigned as EIN_066350 in NCBI), sharing a 32.8% identity with the *E. histolytica* homologue (E-value 1.8e-27) (Table 2). Notably, there were no hits in the relatively well-characterized free-living aerobic amoebozoan *Dictyostelium discoideum,* or in unicellular organisms similarly possessing mitosomes such as *Giardia lamblia, Encephalitozoon cuniculi, Cryptosporidium parvum,* and *Mastigamoeba balamuthi*. Furthermore, homolog search in other representative organisms with hydrogenosomes (*Trichomonas vaginalis*), hydrogen-producing mitochondria (*Blastocystis hominis*), or aerobic mitochondria (*Trypanosoma brucei, Saccharomyces cerevisiae*) detected non-significant hits.

MBOMP30 is predicted to contain transmembrane β-strands

The MBOMP predictor gives a high probability score for all four *Entamoeba* MBOMP30 homologs. To assure that this result was not a quirk of the predictor, the *E. histolytica* MBOMP30 sequences were analyzed using two tools designed for topology prediction of bacterial BOMPs, which may be expected to share some structural properties with MBOMPs. The topology prediction tools BOCTOPUS and TMBETAPRED-RBF (Hayat and Elofsson, 2012; Ou et al., 2010) predicted the *Entamoeba* MBOMP30 to contain

multiple transmembrane β-strand regions (Figure 4, Figure 5a), consistent with the premise that MBOMP30 is a mitosomal BOMP. The outside surface of MBOMPs faces a lipid environment and is expected to display a relatively high hydrophobicity. The MBOMP30 sequences indeed share this property, and their predicted β-strand regions and hydrophobic stretches align well based on its hydropathy profile (Figure 5b).

EhMBOMP30 synthesized by cell-free system is integrated to DPhPC proteoliposomes

I synthesized *E. histolytica* MBOMP30 by wheat germ cell-free expression system in the presence of DPhPC liposomes. To ensure that MBOMP30 was integrated to the DPhPC liposomes, Accudenz-flotation assay was performed (Figure 6a). Integration of EhMBOMP30 to DPhPC liposomes was clearly observed following the autoradiographic analysis of SDS-PAGE blots of fractions collected from Accudenz-density gradient ultracentrifugation of proteoliposomes. In this assay, liposomes float due to their relatively low density. EhMBOMP30 bands were observed to be concentrated on the top fractions, suggesting it is associated to the liposome fraction (Figure 6b). On the other hand, GFP, a soluble β-barrel protein was used as negative control and expectedly concentrated in the bottom fractions, implying its non-integration to or non-association with liposomes (Figure 6b).

Far-UV spectroscopy and circular dichroism (CD) suggests high β-strand content of *E. histolytica* MBOMP30

The secondary structure composition of MBOMP30 was estimated by

performing CD spectroscopy of cell-free synthesized EhMBOMP30-DPhPC proteoliposomes. The far-UV spectra of EhMBOMP30 indicated a pattern representative of β-strand-rich proteins, exhibiting minimum and maximum ellipticities near 220 and 195 nm, respectively (Figure 7a, Table 3). This was similar to the spectra of GFP (Figure 7b), a soluble β-barrel control (Greenfield and Fasman, 1969). In addition, using the CONTIN algorithm accessed through the Dichroweb webpage, the deconvoluted CD spectra predicted that *E. histolytica* MBOMP30 has a high β-strand-content estimated at 30.5 ± 2.9%, while having only 14.1 ± 1.7% α-helices. It was also estimated that the protein has 14.9 ± 4.9% β-turns, and 40.5 ± 6.6% random/unordered secondary structure (Table 3). The estimatedβ-strand andα-helix ratios obtained from the CD analysis are comparable to the prediction made *in silico* using PSI-PRED (Table 1). Likewise, a previous report on the CD spectra of human VDAC (hVDAC1) in DPhPC liposome estimated the β-strand, α-helix, β-turn, and random coil contents of hVDAC1 as 37.3%, 7.7%, 22.6%, and 32.4%, respectively (Shanmugavadivu et al., 2007). The association of MBOMP30 in DPhPc liposomes and the far-UV and CD data further support the *in silico* prediction that MBOMP30 is a β-barrel protein.

Localization of MBOMP30 to the *Entamoeba* mitosome

Next, I experimentally verified the localization of MBOMP30 in *E. histolytica* trophozoites using an *E. histolytica* cell line expressing MBOMP30 tagged with the hemagglutinin epitope (HA) at the amino terminus (HA-MBOMP30). I avoided potential interference with the putative β-signal, which is located at the carboxyl terminus, by using amino-terminal HA tagging. I

confirmed the molecular mass of the expressed product was as expected using whole amoebic lysates loaded onto SDS-PAGE followed by blotting and immunodecoration with anti-HA antibody (Figure 8). I then performed an immunofluorescence assay by double-staining HA-MBOMP30 overexpressing strain, with anti-HA antibody and anti-APSK antisera. The IFA results showed colocalization of the EhMBOMP30 with APSK (Figure 9), suggesting mitosome localization of the protein. In addition, I fractionated lysates from HA-MBOMP30-expressing trophozoites by two consecutive rounds of Percoll-gradient ultracentrifugation, followed by subjecting collected fractions to SDS-PAGE, blotting and immunodecoration with anti-HA antibody and antiserum raised against Cpn60, a canonical mitosomal protein. The distribution of the band corresponding to HA-MBOMP30 throughout the fractions was similar to that of Cpn60 (Figure10), suggesting that MBOMP30 localizes to mitosomes. These independent results confirm the mitosomal localization of MBOMP30, which was likewise detected in a previously conducted mitosome proteomic survey (Mi-ichi et al., 2009).

MBOMP30 is integrated to the mitosomal membrane

To verify the prediction that MBOMP30 is an integral outer membrane protein, I subjected the particulate (membrane) fraction to 100,000 *g* ultracentrifugation with Na_2CO_3, which is known to liberate soluble matrix and peripheral membrane proteins from organelles (Fujiki et al., 1982). Immunoblot analysis of the fractions showed that HA-MBOMP30 was only detected in the pellet fraction after Na_2CO_3 treatment (Figure 11 a,b), similar to the positive control, mitosomal BOMP EhTom40-HA (Figure 11 a,c),

suggesting membrane integration of HA-MBOMP30. In contrast, the soluble mitosomal matrix protein Cpn60 was observed only in the supernatant fraction after Na_2CO_3 treatment (Figure11 a,c).

Topology of MBOMP30 based on Proteinase K protection assay

I also performed proteinase K protection assay to ascertain the membrane topology of MBOMP30 in amoeba trophozoites. Immunoblots revealed the sensitivity of HA-MBOMP30 and control proteins to proteinase K degradation. Quantitation of band signal intensity and comparing to untreated fractions revealed that MBOMP30 was 55.2% degraded, having a profile intermediate to that of an outer membrane control, Tom40-HA which was 71.2% digested, and an inner membrane control, AAC (ATP/ADP Carrier (Chan et al., 2005; Mi-ichi et al., 2011))-HA, which was less susceptible with a 42.5% degradation ratio. Meanwhile as expected, the matrix marker APSK-HA, showed the least susceptibility to proteinase K digestion with only 8.98% degraded by the treatment (Figure 12).

Mitosomal membrane localization of MBOMP30

HA-MBOMP30 trophozoites were subjected to immunoelectron microscopy to visualize and confirm mitosomal membrane localization of the protein. Interestingly, the electron micrographs of the mitosomes of HA-MBOMP30 overexpressors showed protein localization on both the outer and inner membranes of *E. histolytica* mitosomes (Figure 13). The micrographs reveal mitosomes having an electron-dense region, with a diameter ranging from 100-600 nm, enclosed by a double membrane, and marked by anti-

Cpn60 staining (5 nm gold particles). Notably, abundant anti-HA staining (15 nm gold particles) of mitosomal membranes was observed in HA-MBOMP30-expressing transformants. In addition, the micrographs of randomly selected mitosomes (n=11) in HA-MBOMP30 trophozoites, revealed that the gold anti-HA particle distribution of 25.2 ± 15.7 gold/μm^2 in the mitosomal membranes was significantly higher compared to just 0.18 ± 0.22 gold/μm^2 in the cytoplasm ($p<0.001$, using Student's *t*-test). BOMPs are known to be present exclusively in the outer membrane. However, there are micrographs showing anti-HA staining of the inner membrane of the mitosomes of HA-MBOMP30 over-expressing cells. It is possible that the over-expression of MBOMP30 could have caused the mislocalization to the inner membrane, or that fixation and other technical issues concerning preparation of samples for immunoelectron microscopy, might be potential reasons for such observations.

MBOMP30 has a putative β-signal

It was previously reported that a short, carboxyl-terminal sequence known as the β-signal (Kutik et al., 2008)plays an important role in MBOMP integration and/or assembly in the mitochondrial outer membrane of *Saccharomyces cerevisiae,* and by a comparative sequence analysis is inferred to do so in a wide range of eukaryotes. Based on a comparison of numerous putative MBOMP sequences, Imai et al. (2011) have proposed a slightly refined β-signal consensus sequence: $P_o x G h_y x H_y x H_y$ (P_o: non-negatively charged polar residue, G: glycine, H_y: large hydrophobic residue, h_y: loosely defined hydrophobic residue including alanine and cysteine, x: any residue). Notably, the β-signal almost never contains the secondary structure

breaker proline in any position. As shown in Figure 4, *E. histolytica* MBOMP30 has an appropriately placed match to the β-signal, but in *E. invadens* MBOMP30 the match is not perfect, because an alanine is aligned where a polar residue should occur. However, a few MBOMPs with imperfect matches to the β-signal are known, for example the *Saccharomyces pombe* Mdm10 β-signal (FFGVHFEY) has a phenylalanine residue in place of a polar residue in the first position (Imai et al., 2008).

I also subjected HA-MBOMP30$^{\Delta 275\text{-}282}$ (HA-MBOMP30 lacking the β-signal) over-expressing amoeba trophozoites to immunoelectron microscopy. Results of mitosome staining show localization in or near the mitosomal outer membrane (Figure 14). Moreover, the gold anti-HA particle distribution in 13 randomly selected micrographs of HA-MBOMP30$^{\Delta 275\text{-}282}$ mitosomes is higher in the mitosomal membranes 13.4 ± 11.9 gold/μm^2 compared to 0.03 ± 0.07 gold/μm^2 in the cytoplasm ($p=0.0016$, using Student's *t*-test). This data clearly suggests that membrane integration was unaffected even with the truncation of the putative β-signal, and therefore this sequence is not required for outer membrane targeting per se. This observation is supported by the immunoblot of HA-MBOMP30$^{\Delta 275\text{-}282}$ cells fractions, showing that the protein was detected in the organelle fraction and was retained in the organellar membrane fraction even after treatment with Na_2CO_3 (Figure 11b). This strongly suggests that the protein is integrated to the membrane of the mitosome, similar to HA-MBOMP30 as well as to the Tom40-HA control, and not like Cpn60. Double staining of HA-MBOMP30$^{\Delta 275\text{-}282}$ trophozoites with both anti-HA and anti-APSK antibodies by immunofluorescence assay also show clear colocalization between the two signals (Figure 15), further supporting the

observation that the β-signal-truncated MBOMP30 remains targeted to the mitosome despite such deletion.

MBOMP30 forms a ~240 kDa protein complex

I investigated whether MBOMP30 forms a complex, by immunoprecipitation of digitonin-solubilized organelle fraction of HA-MBOMP30. By performing BN-PAGE of immunoprecipitated samples, followed by immunoblot analysis using anti-HA antibody, I detected a band suggestive of an MBOMP30-containing complex of approximately 240 kDa (Figure 16 a,b). I also tested if the complex is formed without the putative β-signal sequence. Interestingly, the ~240 kDa complex was not observed in either amino- or carboxy-terminal HA-tagged MBOMP30$^{\Delta 275\text{-}282}$ trophozoites. Furthermore, I also observed a similar phenomenon with immunoprecipitated HA-Tom40 and HA-Tom40$^{\Delta 275\text{-}284}$ from solubilized organelle fractions (Figure 16 b,c). This suggests that the β-signal may have a role in establishing stable MBOMP complex formation in *E. histolytica*.

Discussion

MBOMP30 is a novel mitosomal outer membrane integral β-barrel protein

By a collaborative effort of combining computational and experimental work, I have analyzed an *in silico*-predicted novel eukaryotic subclass of mitochondrion-related-organelle β-barrel proteins, found exclusively in the genus *Entamoeba*. MBOMP30 is the seventh MBOMP subclass, lacking any recognizable sequence homology to any of the six previously identified MBOMP subclasses. Although I have not determined the structure of MBOMP30 (e.g. via protein NMR or X-ray crystallography) and therefore such conclusion may not be considered as proven definitively, I have presented a considerable amount of experimental evidence to support this case. First, the *in silico* prediction of the secondary structure ratios by PSI-PRED (Table 1), was strongly corroborated by the deconvolution of the CD and far-UV spectroscopic data of MBOMP30 in DPhPC (Table 3), revealing a high ratio of β-sheet, consistent with the protein being a β-barrel. In a previous report, the CD spectral deconvolution analysis of human MBOMP hVDAC1 in DPhPC was also found to have high β-strand content (Shanmugavadivu et al., 2007), similar to what I observed in MBOMP30 and GFP. Second, Percoll gradient ultracentrifugation and Na_2CO_3 fractionation experiments showed MBOMP30 behaved like the representative *Entamoeba* MBOMP Tom40 (Makiuchi et al., 2013). Third, imaging data, provided by immunofluorescence and immunoelectron microscopy, strongly indicate mitosomal, as well as outer membrane localization of MBOMP30. Fourth, the membrane integration of MBOMP30 does not appear to be mediated by lipid modifications. It lacks a

canonical amino-terminal secretory signal peptide, thus it is unlikely to go through the ER. Potential lipid modification sites including GPI attachment or isoprenylation are absent as well. Highly reliable and well-benchmarked predictors of α-helical transmembrane regions such as Phobius (Kall et al., 2004) also predicted no such region in MBOMP30. Moreover, the synthesis of MBOMP30 by an *in vitro* wheat germ translation system demonstrated that it is spontaneously integrated into lipid bilayers (Figure 6), which was similarly observed in other integral mitochondrial membrane proteins such as the VDAC1 in *Homo sapiens* (Shanmugavadivu et al., 2007) and the dicarboxylate–tricarboxylate carrier in *Arabidopsis thaliana* and *Plasmodium falciparum* (Nozawa et al., 2011).

BOMP assembly and integration into *Entamoeba* mitosomal outer membranes appears to be highly divergent from mitochondrial models

The mitosome of *Entamoeba* lacks most of the canonical processes and components existing in the mitochondrion and even the hydrogenosome (Makiuchi and Nozaki, 2014). In particular, the mitosomal outer and inner membranes appear almost bare, lacking homologs for most of the components associated with protein import. The assembly of BOMPs typically requires the outer membrane complexes TOM and SAM, containing the β-barrel proteins Tom40 and Sam50 respectively (Wiedemann et al., 2003). Recently, Tom60 (Makiuchi et al., 2013), a component unique to the *Entamoeba* TOM complex, was discovered. This essential tetratricopeptide repeat-containing protein acts as both a cytoplasmic carrier of soluble and membrane pre-mitosomal proteins, as well as a lone structural component of

the *Entamoeba* TOM complex. Aside from Tom60, the *Entamoeba* genome contains no other detectable homologs to the non-β-barrel proteins which usually work in concert with Tom40 and Sam50 (Dolezal et al., 2010; Makiuchi and Nozaki, 2014), including the small TIM complexes, Tim9-Tim10 and Tim8-Tim13 required for translocation of precursor BOMPs from the TOM to the SAM complex (Paschen et al., 2003; Wiedemann et al., 2003). Also, there appears to be no homolog of Sam35 in *E. histolytica*, the non-BOMP subunit of the SAM complex which recognizes the β-signal in yeast prior to BOMP assembly via Sam50 (Kutik et al., 2008).

Interestingly, the *E. histolytica* MBOMP30 and Tom40 match the slightly refined β-signal consensus sequence (Figure 3). However, in the case of the *E. histolytica* Sam50 homolog, it does not have anything close to a β-signal match, but instead has a terminal phenylalanine, which (although possibly coincidental) matches the membrane integration/assembly signal for bacterial BOMPs, simply consisting of a large aromatic residue [FYW] as the final residue of the protein (Struyve et al., 1991) (Figure 3). This observation is intriguing in its own right, as Sam50 is an MBOMP with a recognizable bacterial BOMP homolog (Omp85/BamA), which itself typically ends in [FYW]. Given that the *Entamoeba* proteome seems to exhibit an unusual mixture of systems of diverse phylogenetic origins (Mi-ichi et al., 2009), it is conceivable that their MBOMPs utilize both eukaryotic and bacterial mechanisms for membrane insertion.

To investigate the role of the putative *Entamoeba* β-signal in BOMP assembly I transfected amoebic trophozoites to overexpress MBOMP30 lacking the β-signal. Interestingly, the data from sodium carbonate

fractionation, immunofluorescence assay, immunoelectron microscopy, and Blue-Native PAGE of immunoprecipitated soluble organelle fractions, suggest that the putative β-signal in MBOMP30 is not essential for the integration to the outer membrane, but could be required for sorting, formation, and/or stability of the ~240 kDa complex. Similarly, the abolishment of the formation of the 600 kDa TOM complex was observed in *E. histolytica* Tom40 lacking the β-signal (Figure 6b). Although an exact role of the β-signal in the MBOMP30 complex formation remains unknown, different mechanisms for recognition, sorting, and integration of β-barrel proteins possibly exist in the mitosomes of *Entamoeba*. It is plausible that MBOMP30 release from either Tom40 or Sam50 is blocked by the deletion of its β-signal, preventing it to properly forming the ~240 kDa complex on the outer membrane. A similar phenomenon was observed when truncation of some residues of the β-signal in yeast Tom40, Mdm10, and porin, did not impair binding to Sam50, but prevented formation of the TOM complex on the outer membrane (Kutik et al., 2008). It is tempting to speculate that, with the minimal components for protein import, especially BOMP assembly, the *Entamoeba* mitosome has developed simplistic alternative mechanisms of membrane protein translocation and integration, and/or that possible functional homologs of "missing" components associated with protein import, are yet to be discovered.

Physiological role of MBOMP30

Entamoeba mitosomes compartmentalize enzymes required for sulfate activation. Since *Entamoeba* mitosomes transport substrates and produce intermediary metabolites of sulfate activation, such as sulfate,

adenosine phosphates, phosphate, and 3'-phosphoadenosine-5'-phosphosulfate (PAPS), it is expected that they possess suitable transporters in their membranes. However, a sensitive hidden Markov model-based similarity search failed to find an *Entamoeba* homolog of VDAC, the usual channel for metabolites in the outer membrane of mitochondria (Dolezal et al., 2010). Without VDAC, it is unclear how metabolites are transported across the outer membrane in this organism. This study is partially motivated by the speculation that novel transporters may serve this role, although it is also possible that Tom40 serves as the transporter of both proteins and metabolites, as it is known to transport some small molecules in yeast (Budzinska et al., 2009; Kmita and Budzinska, 2000). In the same manner, there is also a possibility that MBOMP30 has roles similar to those of the MBOMPs, Mdm10 or Tac40. In yeasts, Mdm10 tethers the mitochondria to the ER and is involved in the ERMES (Kornmann et al., 2009). Several cellular processes are associated with this BOMP, including the maintenance of mitochondrial morphology, and Ca^{2+} transport and lipid exchange between the ER and mitochondria. Although there are no detectable *Entamoeba* homologs of the proteins comprising ERMES (Wideman et al., 2013), such as the Maintenance of Mitochondrial Morphology protein 1 (Mmm1), Mdm12, and Mdm 34 (Kornmann et al., 2009), it is conceivable that MBOMP30, like Mdm10, may be involved in linking the mitosome to the ER via unknown functional homologs, to other *Entamoeba* organelles, or even to cytoskeletal proteins, as Mdm10 binds to actin (Boldogh et al., 1998) and Tac40 attaches to the basal body of the *T. brucei* flagellum (Schnarwiler et al., 2014).

My repeated attempts at knocking down expression of MBOMP30 by

gene silencing have failed, suggesting that the protein is central and essential to the survival of *E. histolytica* trophozoites, and in good contrast to the fact that the knockdown of Cpn60, AAC, and three enzymes involved in sulfate activation was not lethal(Mi-ichi et al., 2011). In fact, gene silencing of ATP sulfurylase, Cpn60, and AAC caused slight upregulation of MBOMP30 expression based on unpublished microarray data. I have also attempted to identify the components of the ~240 kDa complex. Several ER and cytoplasmic proteins were detected in the MS-MS analysis. Converse immunoprecipitation of one candidate binding partner (data not shown), yielded a complex with almost identical size, however MS sequencing did not detect the presence of MBOMP30. There is a possibility that the protein exists as a homo-oligomer, like bacterial BOMPs such as MspA, a porin uniquely found in *Mycobacterium smegmatis* that forms a homo-octameric complex (Faller et al., 2004).

In conclusion, the discovery of MBOMP30 represents only the seventh class of eukaryotic MBOMPs and therefore significantly increases the understanding of the range of sequence, function, and phylogenetic distribution possible for this structural class of eukaryotic proteins. Although BOMPs are numerous and diverse in bacteria (Wimley, 2003), systematic attempts to find them in eukaryotic genomes have not yielded novel MBOMPs so far (Imai et al., 2011; Imai et al., 2008; Remmert et al., 2009). However, it is important to remember that the diversity of eukaryotes, and consequently the diversity of mitochondria and MROs, remains hidden in organisms whose genomes are yet to be completely sequenced. This study could potentially aid

and stimulate further searches for novel MBOMPs. In addition to shedding light on MBOMPs in general, the knowledge on a lineage-specific MBOMP30 will help guide experiments to elucidate its function and better understand the biology and evolution of *Entamoeba*.

Acknowledgements

First, I would like to express my deepest gratitude to my academic supervisor and adviser, Professor Tomoyoshi Nozaki for accepting me as a PhD student in his laboratory, and for his constant guidance, unending support, and reassuring advice throughout the conduct of this research. I also thank Dr. Takashi Makiuchi of the Department of Infectious Diseases, Tokai University School of Medicine, Japan, for mentoring me in the early stages of my PhD studies in Japan when he was still part of our laboratory. Professor Nozaki and Dr. Makiuchi also provided many insights on the approaches and strategies that were undertaken in this project, as well as in the interpretation of the data and results. I also would like to express my appreciation to the members of the *Entamoeba* group (both current and former) of the Nozaki laboratory namely Dr. Kumiko Tsukui, Dr. Yumiko Nakano, Dr. Yoko Chiba, Dr. Ghullam Jeelani, Emi Sato, Yuki Hanadate, Konomi Marumo, Natsuki Watanabe, Mihoko Imada, Yoshitaka Murakami, Nozomi Mase, and Saki Arakawa, for the critical comments, fruitful discussions and exchanges of ideas. I also thank Dr. Jesus Flores of the Center for Research and Advanced Studies of the National Polytechnic Institute (CINVESTAV), Mexico, for his advices during some of our discussions.

Second, I would like to acknowledge all the collaborators that provided significant contribution for the completion of this study. First, I thank Dr. Kenichiro Imai, Dr. Kentaro Tomii, and Professor Paul Horton, from the Computational Biology Research Center (CBRC), National Institute of Advanced Industrial Science and Technology (AIST), Japan for the *in silico* prediction and analyses. Also, I would like to thank Dr. Akira Nozawa of the

Proteo-Science Center, Ehime University, and Professor Yuzuru Tozawa of the Graduate School of Science and Engineering, Saitama University for helping me in the experiments involving MBOMP30 expressed using cell-free expression system such as the CD spectroscopy analysis. I also thank Dr. Mohamed Ibrahim of the Botany Department, Ain Shams University, Egypt for sharing his expertise in CD spectroscopy.

Third, I would like to acknowledge the financial support provided for this work, namely the Grants-in-Aid for Scientific Research on Innovative Areas (“Matryoshka-type evolution”, No. 3308) and the Platform for Drug Discovery, Informatics, Structural Life Science from Ministry of Education, Culture, Sports, Science and Technology (MEXT), Japan. I also thank MEXT for providing me the scholarship throughout the three years of my PhD education in Japan.

I also thank the professors from the University of Tsukuba namely, Professor Tetsuo Hashimoto, Professor Yuji Inagaki, and Professor Kisaburo Nagamune, for their unending support and encouraging comments and advice. I also thank the entire Institute of Biology, College of Science, University of the Philippines Diliman, for their support and the training I gained throughout my stay as student and Instructor.

Finally, I would like to express my gratitude and love to my family and friends in the Philippines. I thank my parents, Benjamin and Patricia, my sisters Cherry Lou and Ma. Cecilia, my aunt Elvira, and all my relatives for their unconditional love and unending support. I also thank all my friends, former colleagues and students, and newfound friends in Japan, for all the warm acts of kindness and encouragement. Thank you all very much!

References

Ali, V., Shigeta, Y., Tokumoto, U., Takahashi, Y., and Nozaki, T. (2004). An intestinal parasitic protist, *Entamoeba histolytica*, possesses a non-redundant nitrogen fixation-like system for iron-sulfur cluster assembly under anaerobic conditions. J Biol Chem *279*, 16863-16874.

Aurrecoechea, C., Brestelli, J., Brunk, B.P., Fischer, S., Gajria, B., Gao, X., Gingle, A., Grant, G., Harb, O.S., Heiges, M.*, et al.* (2010). EuPathDB: a portal to eukaryotic pathogen databases. Nucleic Acids Res *38*, D415-419.

Baker, K.P., Schaniel, A., Vestweber, D., and Schatz, G. (1990). A yeast mitochondrial outer membrane protein essential for protein import and cell viability. Nature *348*, 605-609.

Berven, F.S., Flikka, K., Jensen, H.B., and Eidhammer, I. (2004). BOMP: a program to predict integral beta-barrel outer membrane proteins encoded within genomes of Gram-negative bacteria. Nucleic Acids Res *32*, W394-399.

Boldogh, I., Vojtov, N., Karmon, S., and Pon, L.A. (1998). Interaction between mitochondria and the actin cytoskeleton in budding yeast requires two integral mitochondrial outer membrane proteins, Mmm1p and Mdm10p. J Cell Biol *141*, 1371-1381.

Bracha, R., Nuchamowitz, Y., Anbar, M., and Mirelman, D. (2006). Transcriptional silencing of multiple genes in trophozoites of *Entamoeba histolytica*. PLoS Pathog *2*, e48.

Budzinska, M., Galganska, H., Karachitos, A., Wojtkowska, M., and Kmita, H. (2009). The TOM complex is involved in the release of superoxide anion from mitochondria. J Bioenerg Biomembr *41*, 361-367.

Chacinska, A., Koehler, C., Milenkovic, D., Lithgow, T., Pfanner, N. (2009). Importing mitochondrial proteins: machineries and mechanisms. Cell *138*, 628-644.

Chan, K.W., Slotboom, D.J., Cox, S., Embley, T.M., Fabre, O., van der Giezen, M., Harding, M., Horner, D.S., Kunji, E.R., Leon-Avila, G.*, et al.* (2005). A novel ADP/ATP transporter in the mitosome of the microaerophilic human parasite *Entamoeba histolytica*. Curr Biol. *15*, 737-742.

Colombini, M. (2004). VDAC: the channel at the interface between mitochondria and the cytosol. Mol Cell Biochem *256-257*, 107-115.

The Uniprot Consortium (2012). Reorganizing the protein space at the Universal Protein Resource (UniProt). Nucleic Acids Res, *40 (Database issue)*, D71-75.

Diamond, L.S., Harlow, D.R., and Cunnick, C.C. (1978). A new medium for the axenic cultivation of *Entamoeba histolytica* and other *Entamoeba*. Trans R Soc Trop Med Hyg *72*, 431-432.

Dolezal, P., Dagley, M.J., Kono, M., Wolynec, P., Likic, V.A., Foo, J.H., Sedinova, M., Tachezy, J., Bachmann, A., Bruchhaus, I.*, et al.* (2010a). The essentials of protein import in the degenerate mitochondrion of *Entamoeba histolytica*. PLoS Pathog *6*, e1000812.

Faller, M., Niederweis, M., and Schulz, G.E. (2004). The structure of a mycobacterial outer-membrane channel. Science *303*, 1189-1192.

Finn, R.D., Clements, J., and Eddy, S.R. (2011). HMMER web server: interactive sequence similarity searching. Nucleic Acids Res *39*, W29-37.

Fujiki, Y., Fowler, S., Shio, H., Hubbard, A.L., and Lazarow, P.B. (1982). Polypeptide and phospholipid composition of the membrane of rat liver peroxisomes: comparison with endoplasmic reticulum and mitochondrial membranes. J Cell Biol *93*, 103-110.

Gentle, I., Gabriel, K., Beech, P., Waller, R., and Lithgow, T. (2004). The Omp85 family of proteins is essential for outer membrane biogenesis in mitochondria and bacteria. J Cell Biol *164*, 19-24.

Gentle, I.E., Burri, L., and Lithgow, T. (2005). Molecular architecture and function of the Omp85 family of proteins. Mol Microbiol *58*, 1216-1225.

Gill, E.E., Diaz-Trivino, S., Barbera, M.J., Silberman, J.D., Stechmann, A., Gaston, D., Tamas, I., and Roger, A.J. (2007). Novel mitochondrion-related organelles in the anaerobic amoeba *Mastigamoeba balamuthi*. Mol Microbiol *66*, 1306-1320.

Greenfield, N., and Fasman, G.D. (1969). Computed circular dichroism spectra for the evaluation of protein conformation. Biochemistry US *8*, 4108-4116.

Hayat, S., and Elofsson, A. (2012). BOCTOPUS: improved topology prediction of transmembrane beta barrel proteins. Bioinformatics *28*, 516-522.

Heinz, E., and Lithgow, T. (2012). Back to basics: A revealing secondary reduction of the mitochondrial protein import pathway in diverse intracellular parasites. Biochim Biophys Acta *1833,* 295-303.

Hill, K., Model, K., Ryan, M.T., Dietmeier, K., Martin, F., Wagner, R., and Pfanner, N. (1998). Tom40 forms the hydrophilic channel of the mitochondrial import pore for preproteins. Nature *395*, 516-521.

Imai, K., Fujita, N., Gromiha, M.M., and Horton, P. (2011). Eukaryote-wide sequence analysis of mitochondrial beta-barrel outer membrane proteins. BMC Genomics *12*, 79.

Imai, K., Gromiha, M.M., and Horton, P. (2008). Mitochondrial beta-barrel proteins, an exclusive club? Cell *135*, 1158-1159; author reply 1159-1160.

Imai, K.a.M., S (2005). Mechanisms of secondary structure breakers in soluble proteins. Biophysics*1*, 55-65.

Kall, L., Krogh, A., and Sonnhammer, E.L. (2004). A combined transmembrane topology and signal peptide prediction method. J Mol Biol *338*, 1027-1036.

Katoh, K., Misawa, K., Kuma, K., and Miyata, T. (2002). MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. Nucleic Acids Res *30*, 3059-3066.

Kmita, H., and Budzinska, M. (2000). Involvement of the TOM complex in external NADH transport into yeast mitochondria depleted of mitochondrial porin1. Biochim Biophys Acta *1509*, 86-94.

Kornmann, B., Currie, E., Collins, S.R., Schuldiner, M., Nunnari, J., Weissman, J.S., and Walter, P. (2009). An ER-mitochondria tethering complex revealed by a synthetic biology screen. Science *325*, 477-481.

Kozjak, V., Wiedemann, N., Milenkovic, D., Lohaus, C., Meyer, H.E., Guiard, B., Meisinger, C., and Pfanner, N. (2003). An essential role of Sam50 in the protein sorting and assembly machinery of the mitochondrial outer membrane. J Biol Chem *278*, 48520-48523.

Kutik, S., Stojanovski, D., Becker, L., Becker, T., Meinecke, M., Kruger, V., Prinz, C., Meisinger, C., Guiard, B., Wagner, R.*, et al.* (2008). Dissecting membrane insertion of mitochondrial beta-barrel proteins. Cell *132*, 1011-1024.

Kyte, J., and Doolittle, R.F. (1982). A simple method for displaying the hydropathic character of a protein. J Mol Biol *157*, 105-132.

Lithgow, T., and Schneider, A. (2010). Evolution of macromolecular import pathways in mitochondria, hydrogenosomes and mitosomes. Philos Trans R Soc Lond B Biol Sci *365*, 799-817.

Makiuchi, T., Mi-ichi, F., Nakada-Tsukui, K., and Nozaki, T. (2013). Novel TPR-containing subunit of TOM complex functions as cytosolic receptor for *Entamoeba mitosomal* transport. Sci Rep *3*, 1129.

Makiuchi, T., and Nozaki, T. (2014). Highly divergent mitochondrion-related organelles in anaerobic parasitic protozoa. Biochimie *100*, 3-17.

Maralikova, B., Ali, V., Nakada-Tsukui, K., Nozaki, T., van der Giezen, M., Henze, K., and Tovar, J. (2010). Bacterial-type oxygen detoxification and iron-sulfur cluster assembly in amoebal relict mitochondria. Cell Microbiol *12*, 331-342.

Margulis, L. (1970). Origin of eukaryotic cells evidence and research implications for a theory of the origin and evolution of microbial, plant, and animal cells on the precambrian earth vol. xxii., Yale University Press, New Haven, 1970

McGuffin, L.J., Bryson, K., and Jones, D.T. (2000). The PSIPRED protein structure prediction server. Bioinformatics *16*, 404-405.

Meisinger, C., Sickmann, A., and Pfanner, N. (2008). The mitochondrial proteome: from inventory to function. Cell *134*, 22-24.

Mi-ichi, F., Abu Yousuf, M., Nakada-Tsukui, K., and Nozaki, T. (2009). Mitosomes in *Entamoeba histolytica* contain a sulfate activation pathway. PNatl Acad SciUSA *106*, 21731-21736.

Mi-ichi, F., Makiuchi, T., Furukawa, A., Sato, D., and Nozaki, T. (2011). Sulfate activation in mitosomes plays an important role in the proliferation of *Entamoeba histolytica*. PLoS NeglectTrop D *5*, e1263.

Nakada-Tsukui, K., Okada, H., Mitra, B.N., and Nozaki, T. (2009). Phosphatidylinositol-phosphates mediate cytoskeletal reorganization during phagocytosis via a unique modular protein consisting of RhoGEF/DH and FYVE domains in the parasitic protozoon *Entamoeba histolytica*. Cell Microbiol *11*, 1471-1491.

Nakada-Tsukui, K., Saito-Nakano, Y., Ali, V., and Nozaki, T. (2005). A retromerlike complex is a novel Rab7 effector that is involved in the transport of the virulence factor cysteine protease in the enteric protozoan parasite *Entamoeba histolytica*. Mol Biol Cell *16*, 5294-5303.

Nozawa, A., Fujimoto, R., Matsuoka, H., Tsuboi, T., and Tozawa, Y. (2011). Cell-free synthesis, reconstitution, and characterization of a mitochondrial dicarboxylate-tricarboxylate carrier of *Plasmodium falciparum*. Biochem Bioph Res Co *414*, 612-617.

Nozawa, A., Nanamiya, H., Miyata, T., Linka, N., Endo, Y., Weber, A.P., and Tozawa, Y. (2007). A cell-free translation and proteoliposome reconstitution system for functional analysis of plant solute transporters. Plant Cell Physiol *48*, 1815-1820.

Ou, Y.Y., Chen, S.A., and Gromiha, M.M. (2010). Prediction of membrane spanning segments and topology in beta-barrel membrane proteins at better accuracy. J Comput Chem *31*, 217-223.

Park, K.J., Gromiha, M.M., Horton, P., and Suwa, M. (2005). Discrimination of outer membrane proteins using support vector machines. Bioinformatics *21*, 4223-4229.

Paschen, S.A., Waizenegger, T., Stan, T., Preuss, M., Cyrklaff, M., Hell, K., Rapaport, D., and Neupert, W. (2003). Evolutionary conservation of biogenesis of beta-barrel membrane proteins. Nature *426*, 862-866.

Pearson, W.R. (1991). Searching protein sequence libraries: comparison of the sensitivity and selectivity of the Smith-Waterman and FASTA algorithms. Genomics *11*, 635-650.

Provencher, S.W., and Glockner, J. (1981). Estimation of globular protein secondary structure from circular dichroism. Biochemistry US *20*, 33-37.

Pusnik, M., Schmidt, O., Perry, A.J., Oeljeklaus, S., Niemann, M., Warscheid, B., Lithgow, T., Meisinger, C., and Schneider, A. (2011). Mitochondrial preprotein translocase of trypanosomatids has a bacterial origin. Curr Biology *21*, 1738-1743.

Reinders, J., Zahedi, R.P., Pfanner, N., Meisinger, C., and Sickmann, A. (2006). Toward the complete yeast mitochondrial proteome: multidimensional separation techniques for mitochondrial proteomics. J Proteome Res *5*, 1543-1554.

Remmert, M., Biegert, A., Hauser, A., and Soding, J. (2012). HHblits: lightning-fast iterative protein sequence searching by HMM-HMM alignment. Nat Methods *9*, 173-175.

Remmert, M., Linke, D., Lupas, A.N., and Soding, J. (2009). HHomp--prediction and classification of outer membrane proteins. Nucleic Acids Res *37*, W446-451.

Schmidt, O., Pfanner, N., and Meisinger, C. (2010). Mitochondrial protein import: from proteomics to functional mechanisms. Nat Rev Mol Cell Biol *11*, 655-667.

Schnarwiler, F., Niemann, M., Doiron, N., Harsman, A., Kaser, S., Mani, J., Chanfon, A., Dewar, C.E., Oeljeklaus, S., Jackson, C.B.*, et al.* (2014). Trypanosomal TAC40 constitutes a novel subclass of mitochondrial beta-barrel proteins specialized in mitochondrial genome inheritance. PNatl Acad Sci USA *111*, 7624-7629.

Shanmugavadivu, B., Apell, H.J., Meins, T., Zeth, K., and Kleinschmidt, J.H. (2007). Correct folding of the beta-barrel of the human membrane protein VDAC requires a lipid bilayer. J Mol Biol *368*, 66-78.

Sievers, F., Wilm, A., Dineen, D., Gibson, T.J., Karplus, K., Li, W., Lopez, R., McWilliam, H., Remmert, M., Soding, J.*, et al.* (2011). Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Mol Syst Biol *7*, 539.

Sogo, L.F., and Yaffe, M.P. (1994). Regulation of mitochondrial morphology and inheritance by Mdm10p, a protein of the mitochondrial outer membrane. J Cell Biol *126*, 1361-1373.

Stegmeier, J.F., Gluck, A., Sukumaran, S., Mantele, W., and Andersen, C. (2007). Characterisation of YtfM, a second member of the Omp85 family in *Escherichia coli*. Biol Chem *388*, 37-46.

Struyve, M., Moons, M., and Tommassen, J. (1991). Carboxy-terminal phenylalanine is essential for the correct assembly of a bacterial outer membrane protein. J Mol Biol *218*, 141-148.

Tamm, L. Arora, A., and Kleinschmidt, J. (2001). Structure and assembly of β-barrel membrane proteins. J Biol Chem *276,* 32399-32402.

Taylor R. and Pfanner, N. (2004). The protein import and assembly machinery of the outer membrane. Biochim Biophys Act *1658*, 37-43.

van der Giezen, M., Cox, S., and Tovar, J. (2004). The iron-sulfur cluster assembly genes iscS and iscU of *Entamoeba histolytica* were acquired by horizontal gene transfer. BMC Evol Biol *4*, 7.

van Stokkum, I.H., Spoelder, H.J., Bloemendal, M., van Grondelle, R., and Groen, F.C. (1990). Estimation of protein secondary structure and error analysis from circular dichroism spectra. Anal Biochem *191*, 110-118.

Walther, D.M., Rapaport, D., and Tommassen, J. (2009). Biogenesis of beta-barrel membrane proteins in bacteria and eukaryotes: evolutionary conservation and divergence. Cell Mol Life Sci *66*, 2789-2804.

Wang, Z., Xu, C., and Benning, C. (2012). TGD4 involved in endoplasmic reticulum-to-chloroplast lipid trafficking is a phosphatidic acid binding protein. Plant J *70*, 614-623.

Whitmore, L., and Wallace, B.A. (2004). DICHROWEB, an online server for protein secondary structure analyses from circular dichroism spectroscopic data. Nucleic Acids Res *32*, W668-673.

Whitmore, L., and Wallace, B.A. (2008). Protein secondary structure analyses from circular dichroism spectroscopy: methods and reference databases. Biopolymers *89*, 392-400.

Wideman, J.G., Gawryluk, R.M., Gray, M.W., and Dacks, J.B. (2013). The ancient and widespread nature of the ER-mitochondria encounter structure. Mol Biol Evol *30*, 2044-2049.

Wiedemann, N., Kozjak, V., Chacinska, A., Schonfisch, B., Rospert, S., Ryan, M.T., Pfanner, N., and Meisinger, C. (2003). Machinery for protein sorting and assembly in the mitochondrial outer membrane. Nature *424*, 565-571.

Wimley, W.C. (2003). The versatile beta-barrel membrane protein. Curr Opin Struc Biol*13*, 404-411.

Yamano, K., Tanaka-Yamano, S., and Endo, T. (2010). Mdm10 as a dynamic constituent of the TOB/SAM complex directs coordinated assembly of Tom40. EMBO Rep *11*, 187-193.

Table

Table 1. List of predicted MBOMP candidates in *E. histolytica*.

JackHMMER searches were performed against NCBI non-redundant protein database with the best E-value amongst annotated proteins shown. The β-signal motif (P_oxGxxH_yxH_y) was matched against the β-strand closest to the C-terminus, as predicted by BOCTOPUS.

Protein ID	Length	MBOMP prediction probability	Predicted β-strand ratio	Predicted α-helix ratio	Annotation inferred from homology	JackHMMER Hit (E-value)	Pfam domain (E-value)	β-signal motif hit
EHI_104420	284	0.961	0.29	0.21	Tom40	*C. parvum*, Tom40p like translocase (0.026)	Eukaryotic porin (2.9e-09)	Yes
EHI_178630	282	0.954	0.34	0.23	Unknown	No significant hit	No significant hit	Yes
EHI_007460	304	0.871	0.27	0.20	Phospholipase C	*T. thermophile*, Endonuclease/ Exonuclease/ phosphatase family protein (1.2e-76)	Endonuclease/ Exonuclease/ phosphatase family (3.5e-15)	Yes
EHI_163510	263	0.853	0.31	0.18	Exosome complex exonuclease RRP4	*D. rerio*, Exosome complex exonuclease RRP4 (8.9e-100)	S1 RNA binding domain (0.19)	No
EHI_050690	321	0.749	0.43	0.09	Apyrase	*P. sergenti*, salivary apyrase (3.1e-79)	Apyrase (5.4e-09)	Yes
EHI_068370	466	0.734	0.56	0.00	Cysteine protease binding protein family like protein	*E.histolytica*, cysteine protease binding protein family 4 (8.5e-178)	No significant hit	No

Table 2. Results of EhMBOMP30 homolog search. Results of SSEARCH analysis highlight that the candidate homologs of EhMBOMP30 are limited to the genus *Entamoeba*. Furthermore, the results of SSEARCH targeting specific organisms clearly indicate the lack of homologs in representative eukaryotes having MROs or mitochondria.

Organisms	Organelle type	Best Hit (length)	E-value	Annotation
Entamoeba nuttalli	Mitosome	ENU1_140620 (283)	3.7e-99	Uncharacterized protein
Entamoeba dispar	Mitosome	EDI_035580 (239)	1.2e-66	Uncharacterized protein
Entamoeba invadens	Mitosome	EIN_041060/ EIN_066350 (286)	1.8e-27	Uncharacterized protein
Giardia lamblia	Mitosome	GSB_150221 (418)	0.46	Uncharacterized protein
Encephalitozoon cuniculi	Mitosome	ECU01_0380 (310)	0.3	Uncharacterized protein
Cryptosporidium parvum	Mitosome	cgd6_5020 (216)	1.4	Protein with WD40 repeats
Trichomonas vaginalis	Hydrogenosome	TVAG_193490 (552)	2.3	Uncharacterized protein
Blastocystis hominis	Hydrogen -producing mitochondria	GSBLH_T00006135001 (329)	0.5	Uncharacterized protein
Dictyostelium discoideum	Mitochondria	DDB_0184270 (187)	1.7	Uncharacterized protein
Trypanosoma brucei	Mitochondria	Tb927.3.5550 (183)	0.54	Small GTP-binding protein, putative
Saccharomyces cerevisiae	Mitochondria	GYP1 (637)	0.96	GTPase-activating protein GYP1

Table 3.Deconvolution of EhMBOMP30 CD spectra. Estimation of MBOMP30 secondary structure organization based on CD spectra analysis using CONTIN software.

Trial	Predicted β-strand %	Predicted α-helix %	Predicted β-turns %	Predicted random coil %	NMRSD	λ_{min}, nm	λ_{max}, nm
1	33.5	12.2	19.2	35.1	0.059	221.5	196.0
2	27.8	14.8	9.5	47.9	0.088	221.9	192.5
3	30.3	15.6	15.9	38.5	0.094	220.9	192.8
Avg±SD	30.5±2.9	14.1±1.7	14.9±4.9	40.5±6.6			

NMRSD is the normalized root mean square deviation from the CONTIN analysis and should be <0.1.

Figures

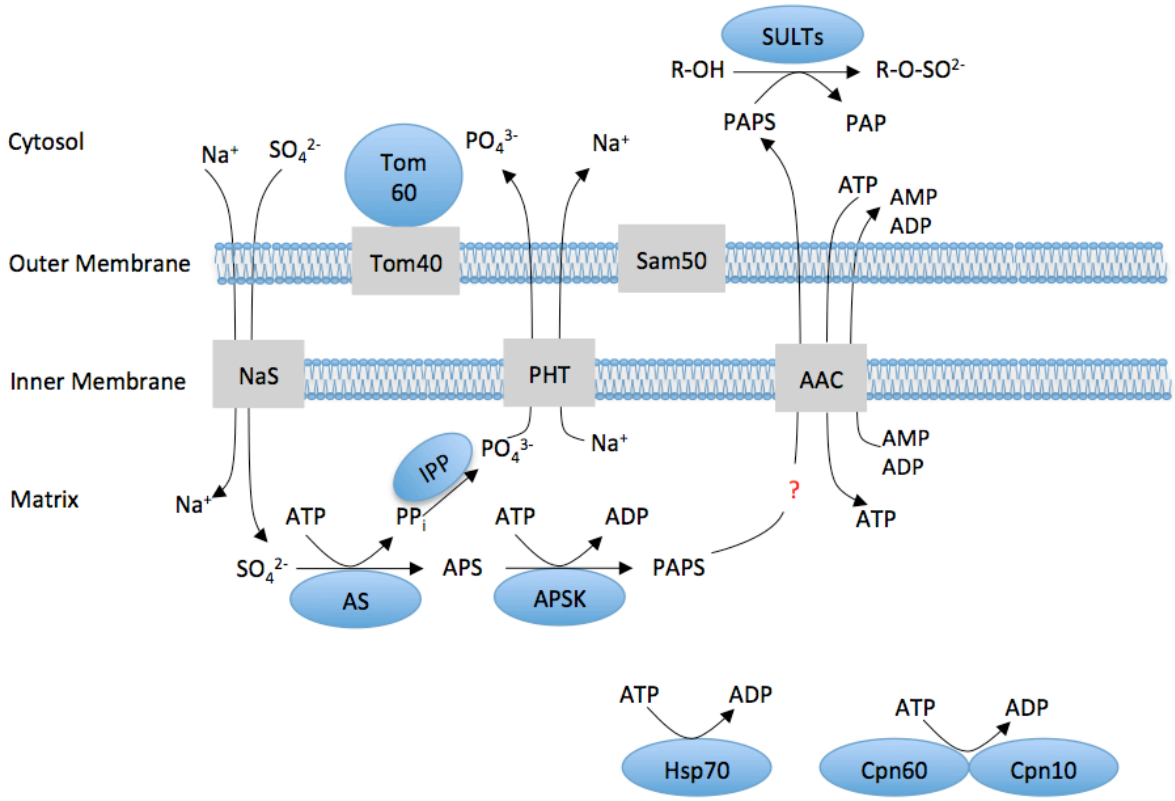


Figure 1. Schematic diagram of the *Entamoeba histolytica* mitosome.

The *E. histolytica* mitosome lacks most of the proteins associated with the canonical mitochondria. Compartmentalization of the sulfate activation pathway is the only known role of this organelle. Enzymes required for sulfate activation pathway were identified previously (Mi-ichi et al., 2009).Tom40 and Sam50 were previously identified by Dolezal et al., (2010) while the discovery of Tom60 was recently reported by Makiuchi et al., (2013).

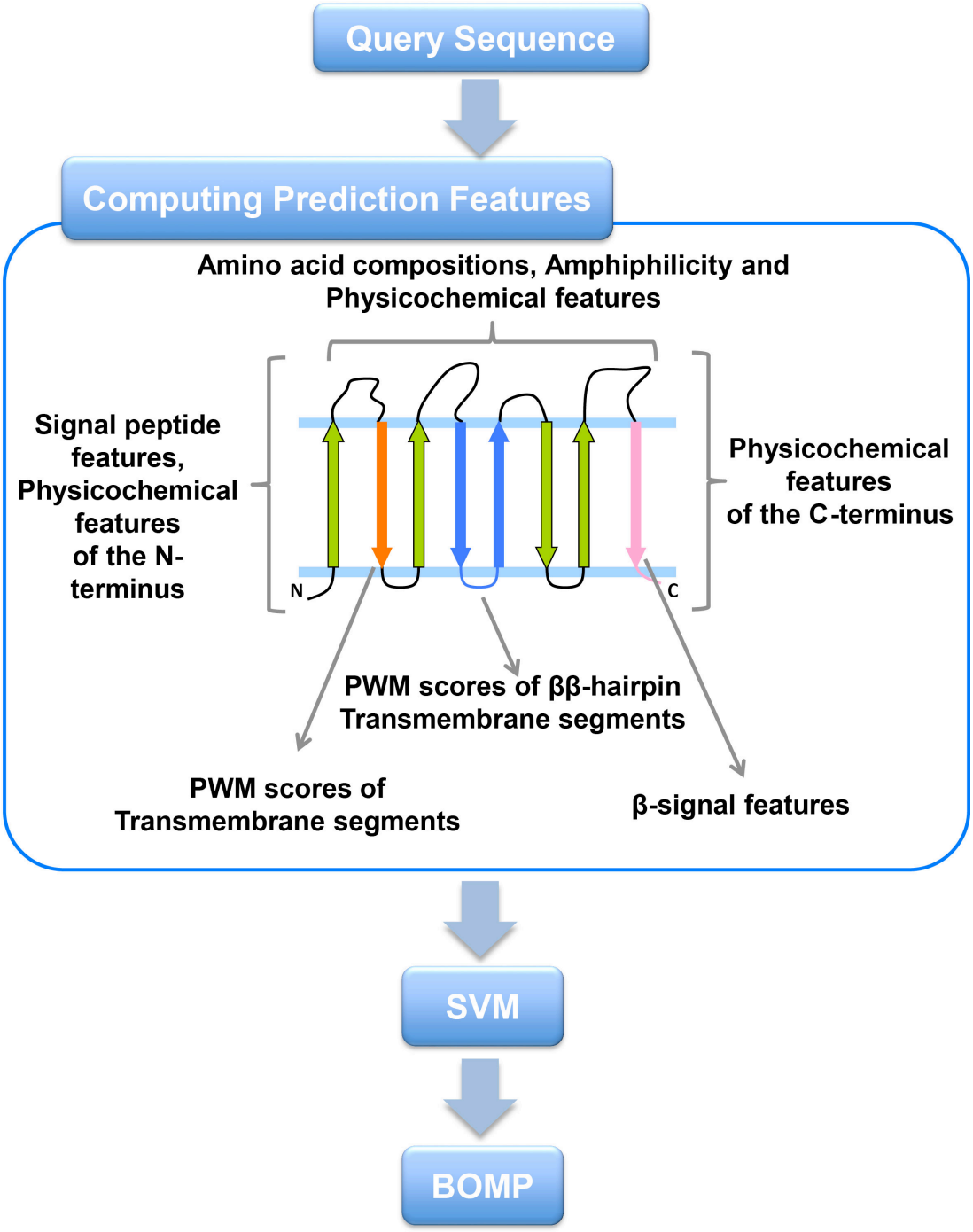


Figure 2. The architecture of the MBOMP prediction system. Sequence features used by the MBOMP predictor of Dr. Imai, Dr. Tomii, and Dr. Horton of AIST-Japan, are listed in relation to a schematic depiction of MBOMP structure.

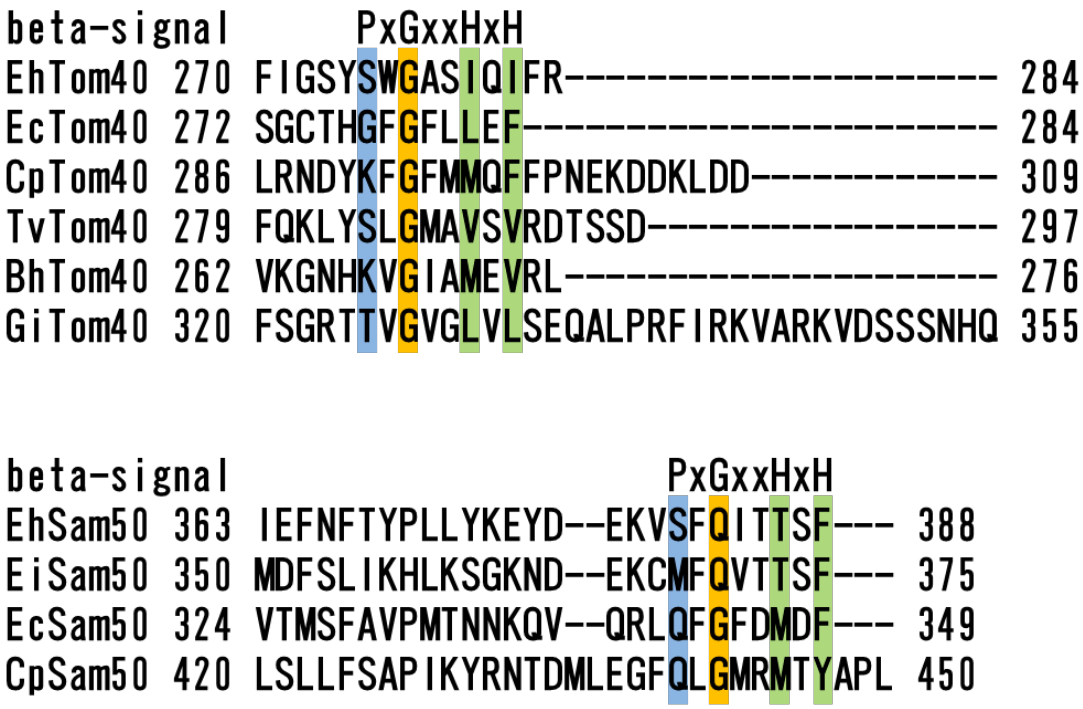


Figure 3. Multiple alignments of potential mitosomal β-signals. Multiple alignments of the vicinity the β-signal motif matches of MRO orthologs of Tom40 and Sam50 are shown. In EcTom40, EhSam50 and EiSam50, the motif match is not perfect. (Abbreviations: Eh – *Entamoeba histolytica,* Ec – *Encephalitozoon cuniculi*, Ei – *Entamoeba invadens*, Cp – *Cryptosporidium parvum*, Tv – *Trichomonas vaginalis*, Bh – *Blastocystis hominis*, Gi – *Giardia intestinalis*)

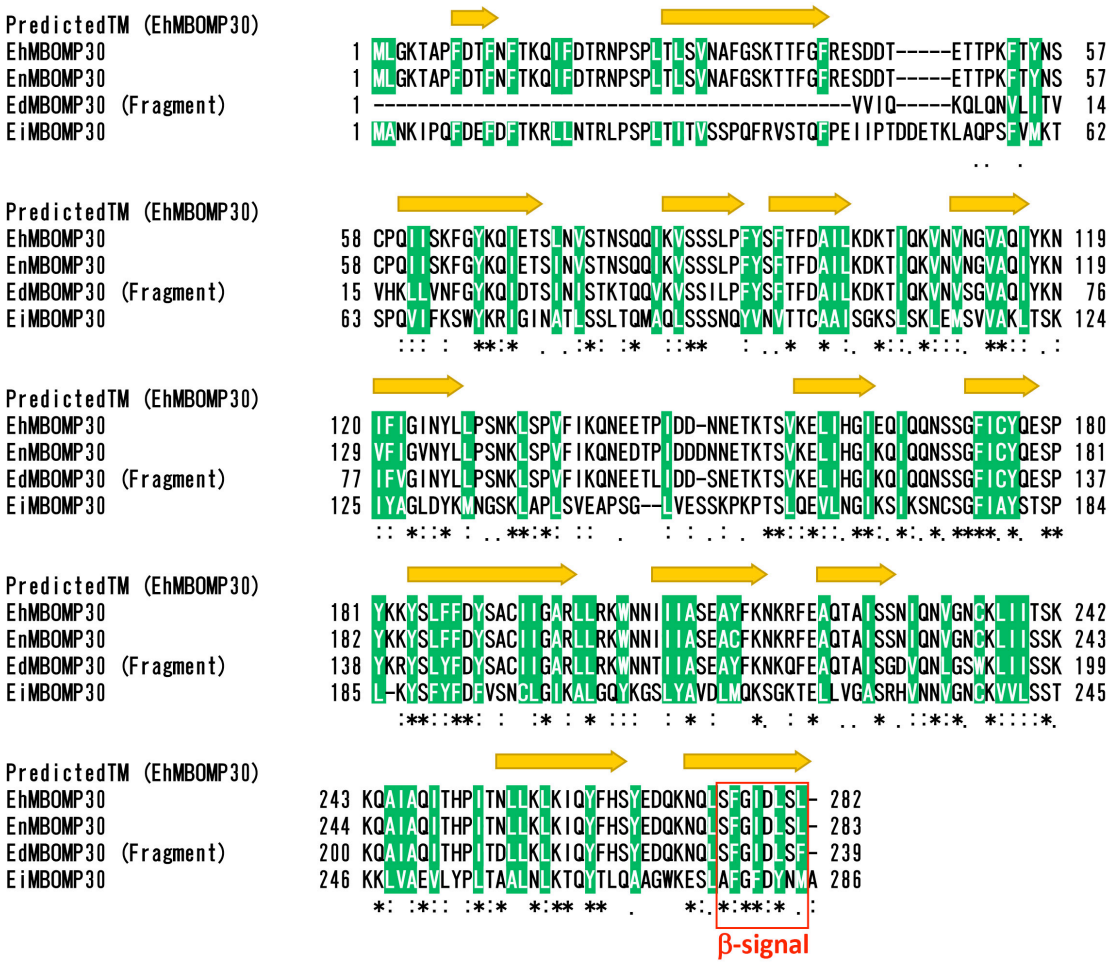


Figure 4. Prediction of transmembrane β-strandsof EhMBOMP30 and multiple alignment with its homologs. Multiple alignment of EhMBOMP30, EnMBOMP30, EdMBOMP30, and EiMBOMP30, built with Clustal omega (Sievers et al., 2011) is shown in CLUSTAL format (Eh-*Entamoeba histolytica*, En-*Entamoeba nuttalli*, Ed-*Entamoeba dispar*, Ei-*Entamoeba invadens*). Predicted transmembrane β-strands of EhMBOMP30 are indicated with yellow arrows. The arrows cover all positions predicted as part of a β-strand by either of two BOMP topology predictors: BOCTOPUS and TMBETAPRED-RBF (Hayat and Elofsson, 2012; Ou et al., 2010). Conserved hydrophobic residues are highlighted in green. A red dashed box indicates the region corresponding to the β-signal motif in the last predicted transmembrane β-strand.

a

```
SEQ             MLGKTAPFDTFNFTKQIFDTRNPSPLTLSVNAFGSKTTFGFRESDDTETTPKFTYNSCPQIISKFGYKQIETSLNVSTNSQQ
BOCTOPUS        -------------------------BBBBBBBBBB-----------------------------------------------
TMBETAPRED-RBF  -------BBBB-----------------BBBBBBBBBBBB-----------------BBBBBBBBBBBB----------

SEQ             IKVSSSLPFYSFTFDAILKDKTIQKVNVNGVAQIYKNIFIGINYLLPSNKLSPVFIKQNEETPIDDNNETKTSVKELIHGIE
BOCTOPUS        -BBBBBBB---BBBBBBB---------BBBBBBB---BBBBBBBB---------------------------BBBBBBB-
TMBETAPRED-RBF  -BBBBBBBBBBBBBBBBB----------BBBBBBBBBBBBBBBBB---------------------------------

SEQ             QIQQNSSGFICYQESPYKKYSLFFDYSACIIGARLLRKWNNIIIASEAYFKNKRFEAQTAISSNIQNVGNCKLIITSKKQAI
BOCTOPUS        -------BBBBBBB-----BBBBBBBBBBBBBB-------BBBBBBB--------BBBBBBB-----------------
TMBETAPRED-RBF  ---------BBBBB-BBBBBBBBBBBBBBBBB--------BBBBBBBBB-----------------------------

SEQ             AQITHPITNLLKLKIQYFHSYEDQKNQLSFGIDLSL
BOCTOPUS        ----------BBBBBBB---------BBBBBBBBB
TMBETAPRED-RBF  --------BBBBBBBBBBB-----BBBBBBBBB--
```

b

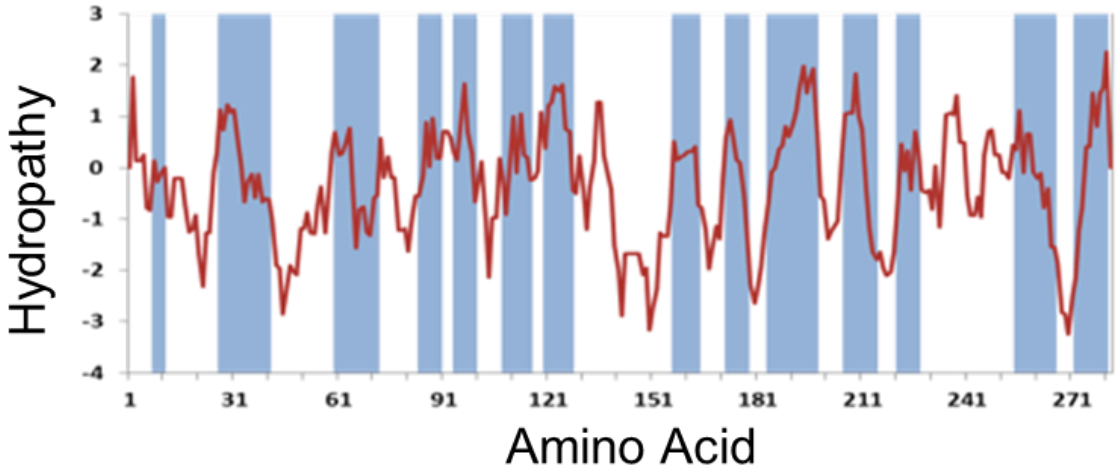
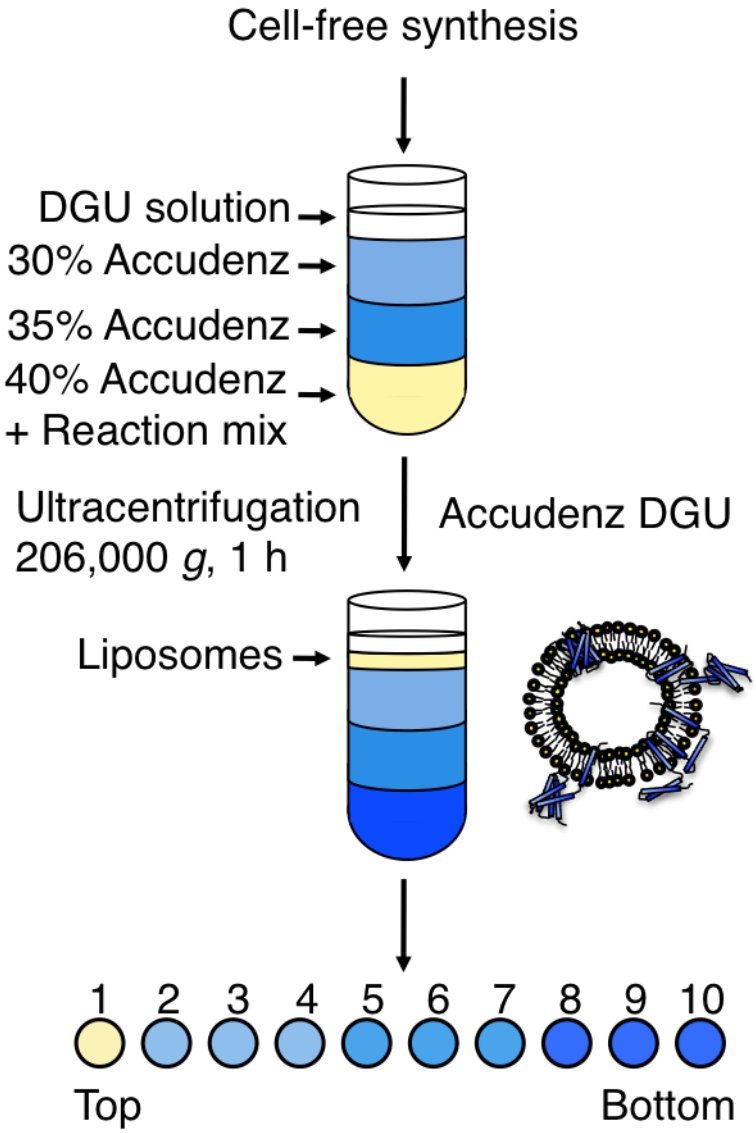


Figure 5. Prediction of transmembrane β-strands and hydrophobicity profile of EhMBOMP30. (a) Residues predicted to be part of a transmembrane β-strand are indicated by a "B" in the track labeled by the prediction method, BOCTOPUS (Hayat and Elofsson, 2012) or TMBETAPRED-RBF (Ou et al., 2010). (b) Hydropathy profile of EhMBOMP30. The horizontal axis shows the position in the sequence and the vertical axis shows the average hydrophobicity (Kyte and Doolittle, 1982) of the 7 residues centered on that position. Blue boxes indicate predicted transmembrane β-strands.

a


Cell-free synthesis
DGU solution
30% Accudenz
35% Accudenz
40% Accudenz
+ Reaction mix
Ultracentrifugation
206,000 g, 1 h
Accudenz DGU
Liposomes
1 2 3 4 5 6 7 8 9 10
Top
Bottom

b

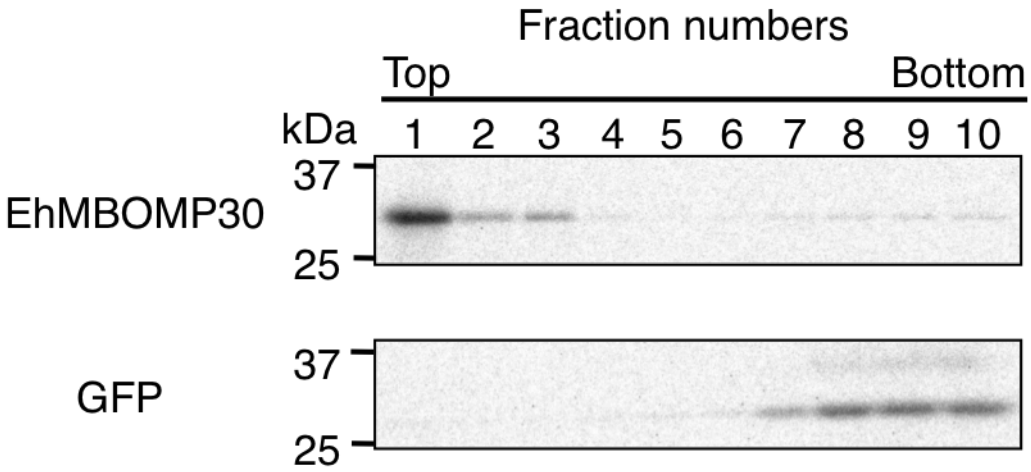
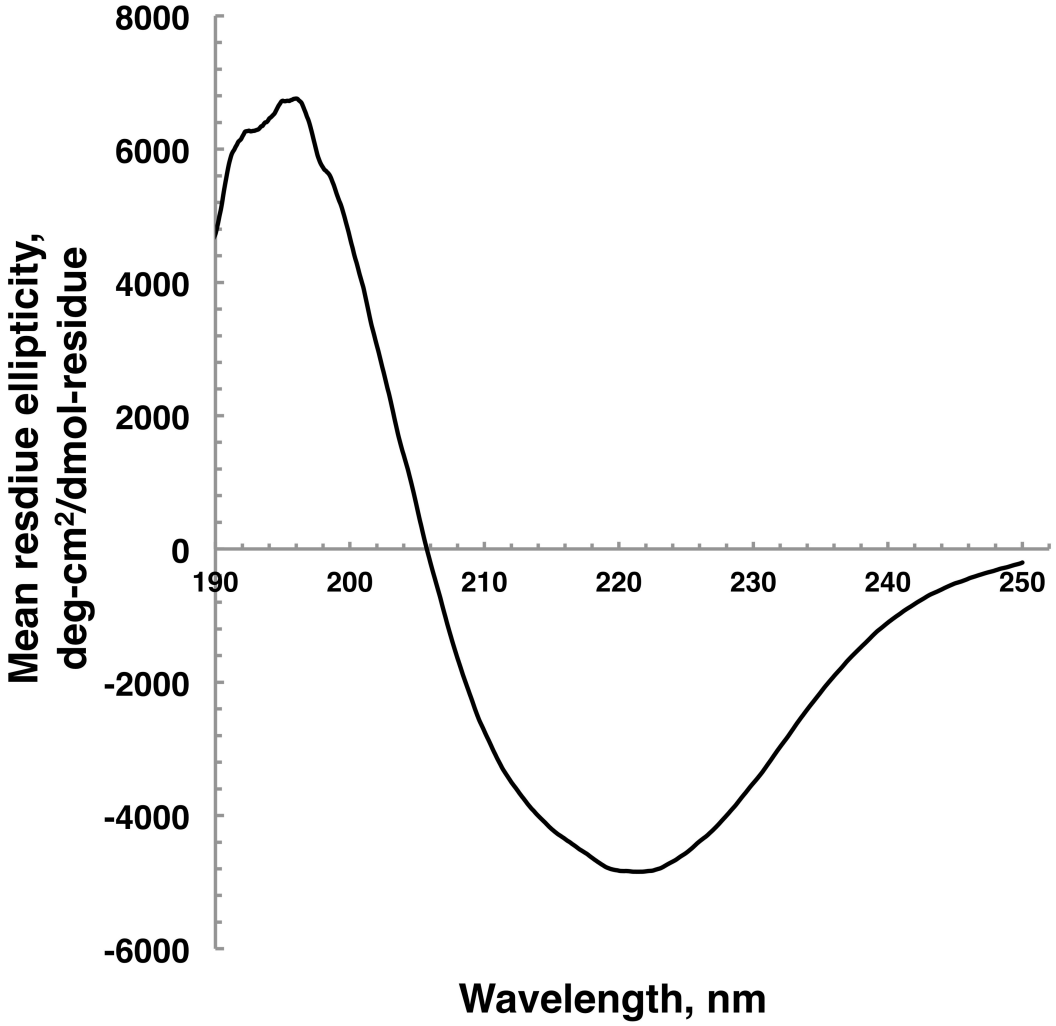


Figure 6. **Accudenz flotation assay of EhMBOMP30 proteoliposome**. (a) Density gradient ultracentrifugation (DGU) of EhMBOMP30 proteoliposomes, synthesized by cell-free system, in DGU solution and Accudenz (Accurate Chemical and Scientific, Westbury, NY) was performed as previously described (Nozawa et al., 2011). (b) Fractions were collected from the top of the tube and analyzed by SDS–PAGE and autoradiography.

a


8000
6000
4000
2000
0
-2000
-4000
-6000
190
200
210
220
230
240
250
Mean resdiue ellipticity, deg-cm²/dmol-residue
Wavelength, nm

b

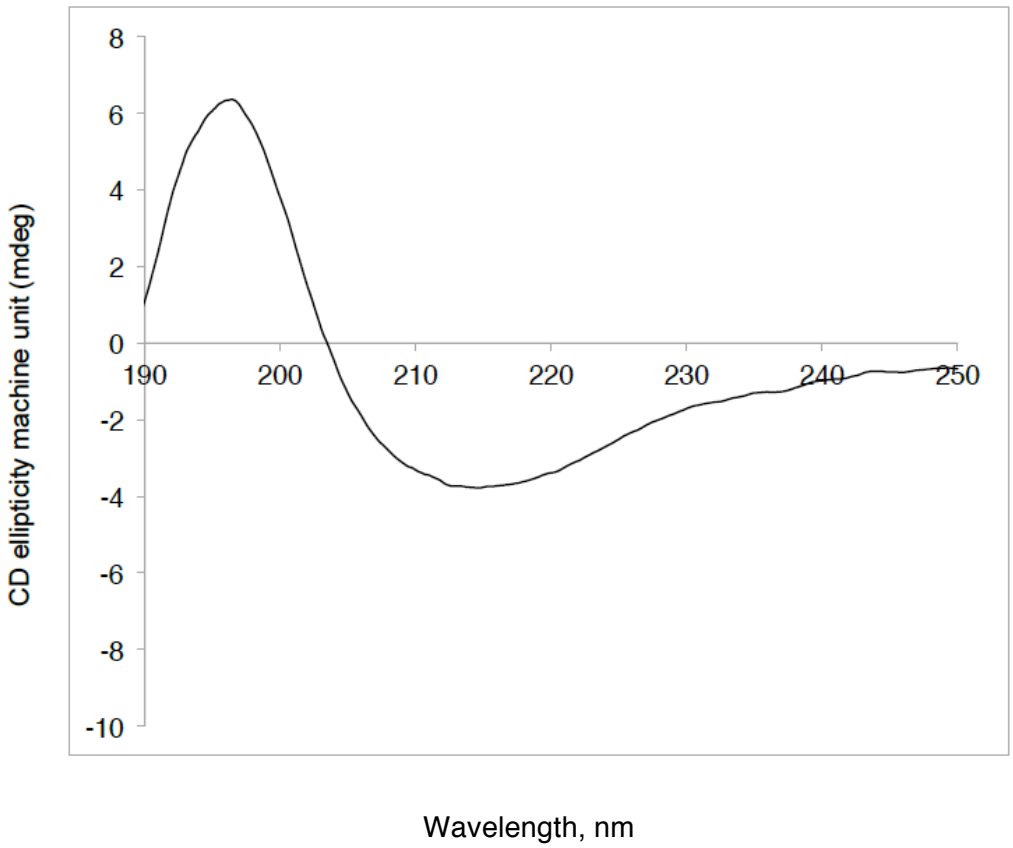


Figure 7. Circular dichroism spectroscopy. The far-ultraviolet CD spectra of (a) MBOMP30 in DPhPc liposomes and (b) GFP, expressed in a cell-free system, were taken from the accumulation of 9 consecutive scans.

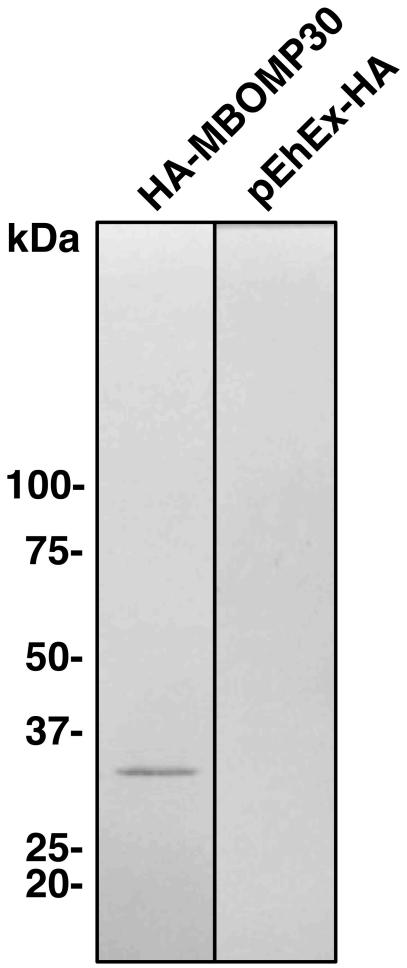


Figure 8. Overexpression of HA-MBOMP30 in amoeba. Ectopic expression of HA-MBOMP30 in *E. histolytica* trophozoites. Approximately 20 μg of cell lysates from HA-MBOMP30 and control strain was fractionated on SDS-PAGE and subjected to immunoblot analysis using anti-HA antibody.

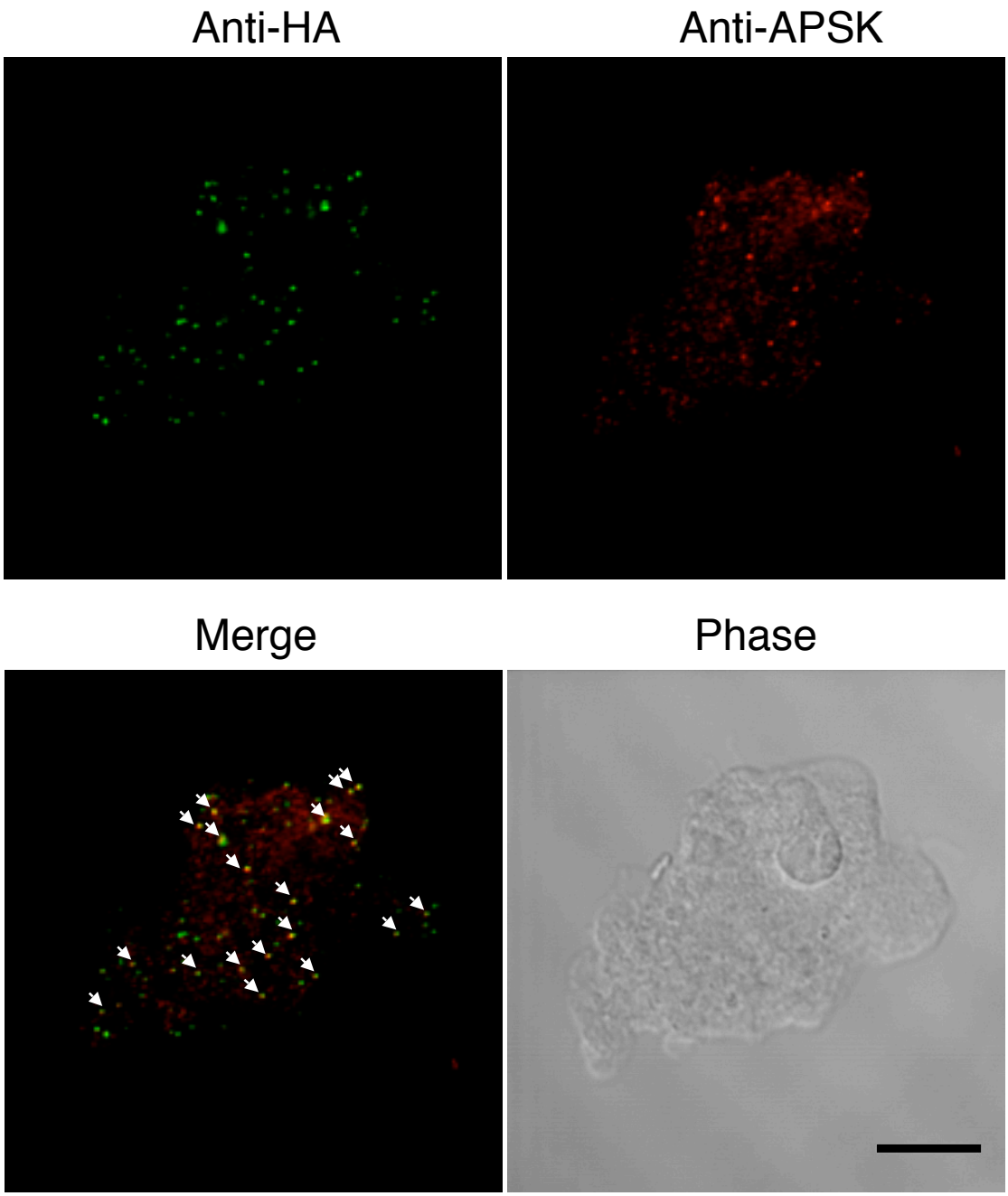


Figure 9. Immunofluorescence analysis of HA-MBOMP30 trophozoites.

White arrows indicate colocalization of punctate anti-HA (green) and anti-APSK (red; mitosomal marker) signals [scale bar, 10 μm].

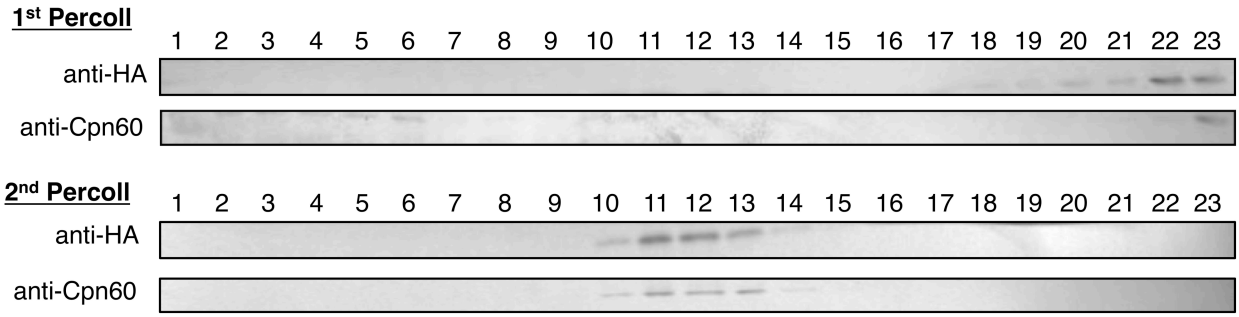
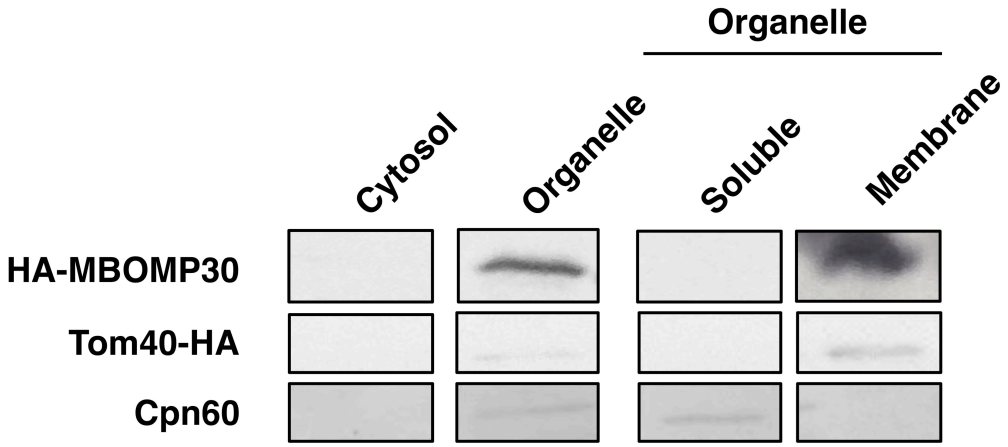
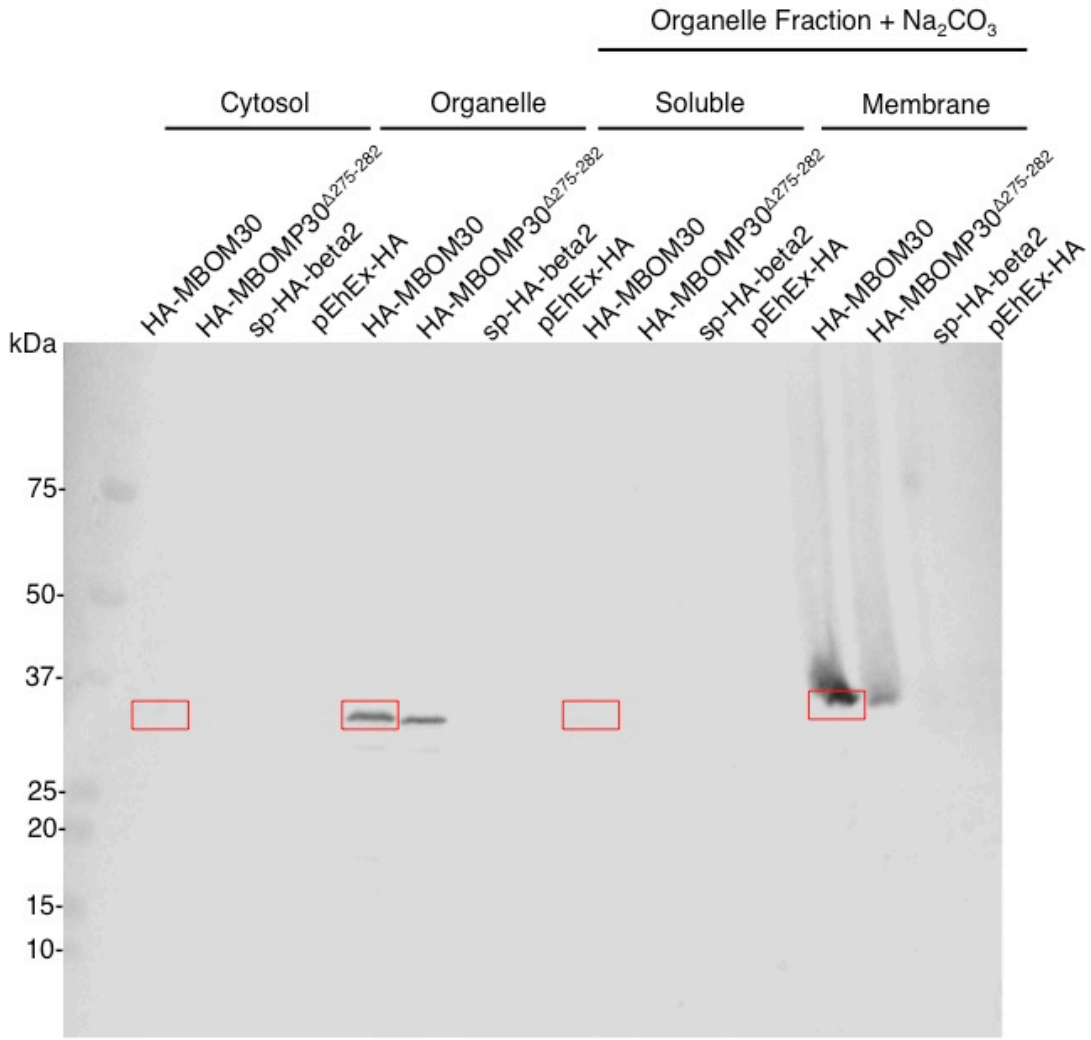


Figure 10. Immunoblot analysis of the fractions by two series of Percoll gradient ultracentrifugation. Approximately 20 µL of each fraction of the first and second ultracentrifugation was separated by SDS-PAGE and blotted to nitrocellulose membranes. The blots were cut into strips containing the region of the target proteins, before reacting with anti-HA and anti-Cpn60 antibodies.

a


Organelle
Cytosol
Organelle
Soluble
Membrane
HA-MBOMP30
Tom40-HA
Cpn60

b


Organelle Fraction + Na2CO3
Cytosol
Organelle
Soluble
Membrane
HA-MBOM30
HA-MBOMP30Δ275-282
sp-HA-beta2
pEhEx-HA
kDa
75-
50-
37-
25-
20-
15-
10-

c

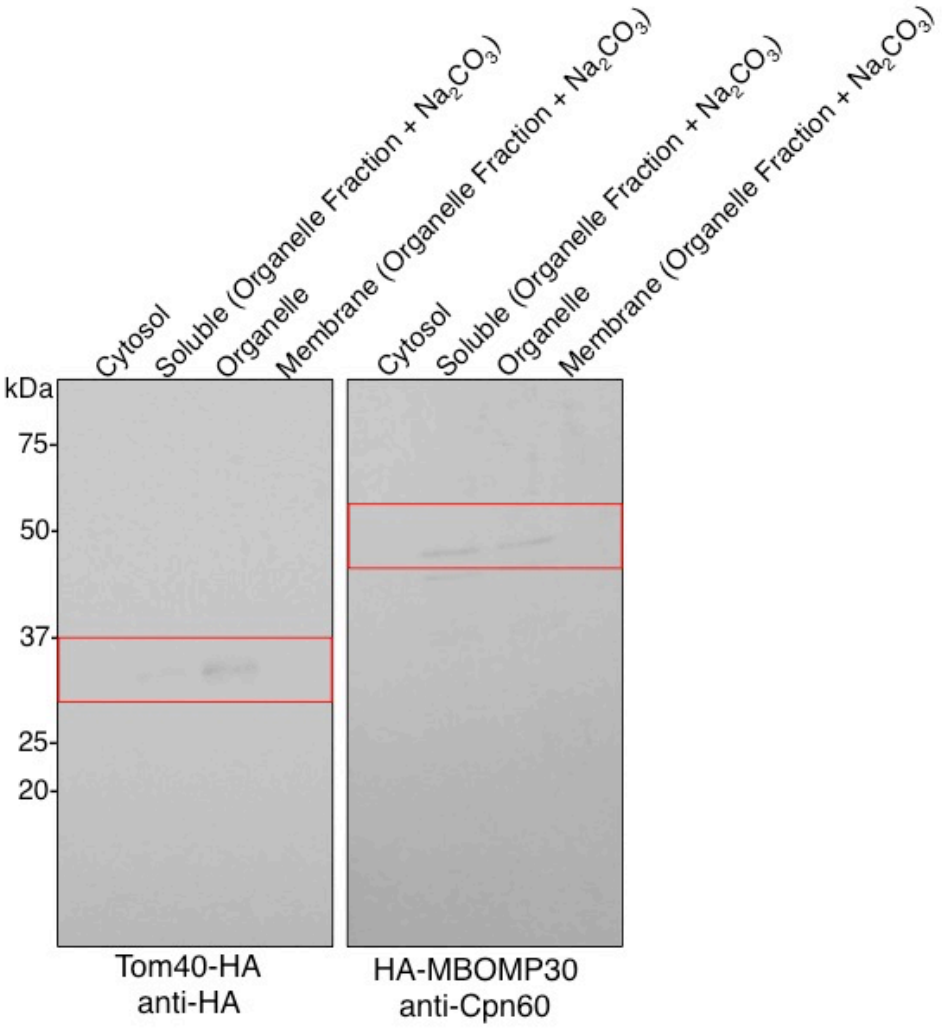
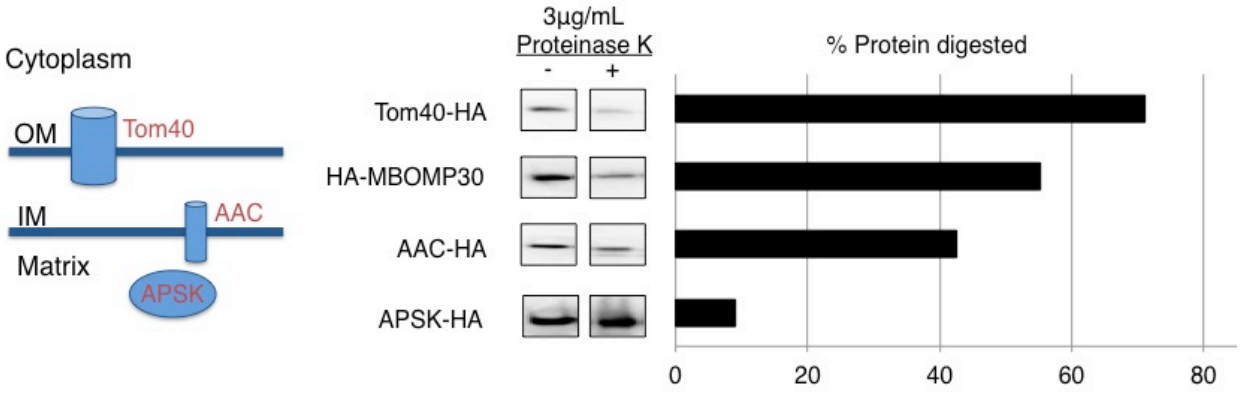
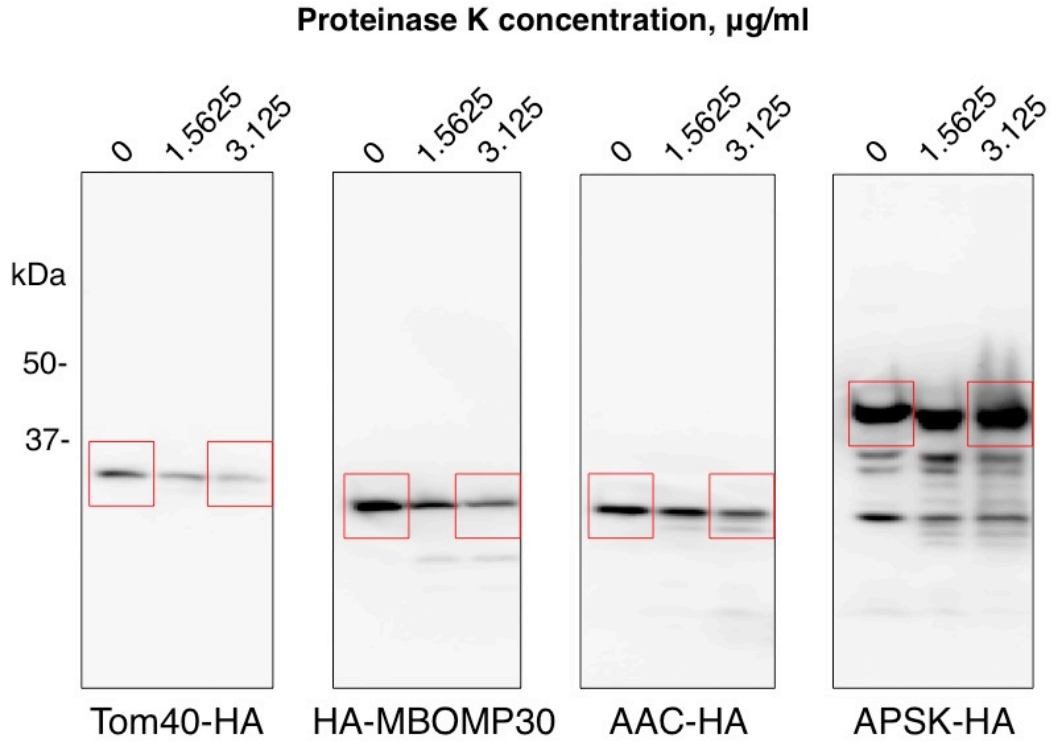


Figure 11. Mitosomal membrane localization of MBOMP30 demonstrated by Na_2CO_3 fractionation. (a) Homogenates from amoebae expressing HA-MBOMP30 and Tom40-HA were fractionated. The organelle fraction was treated with Na_2CO_3 and NaCl to lyse the organelle and liberate loosely bound proteins. The resulting fractions were separated on SDS-PAGE followed by immunoblotting. Parts of the full-length immunoblot for the anti-HA (first two rows) and anti-Cpn60 (last row) antibody reactions are shown respectively. Tom40-HA and Cpn60 serve as a control for mitosomal BOMP and soluble mitosomal proteins, respectively. (b-c) The original blots are shown with red boxes indicating cropped regions presented in (a).

a


Cytoplasm
OM
Tom40
IM
AAC
Matrix
APSK
3µg/mL
Proteinase K
-
+
Tom40-HA
HA-MBOMP30
AAC-HA
APSK-HA
% Protein digested
0
20
40
60
80

b


Proteinase K concentration, µg/ml
0
1.5625
3.125
kDa
50-
37-
Tom40-HA
HA-MBOMP30
AAC-HA
APSK-HA

c

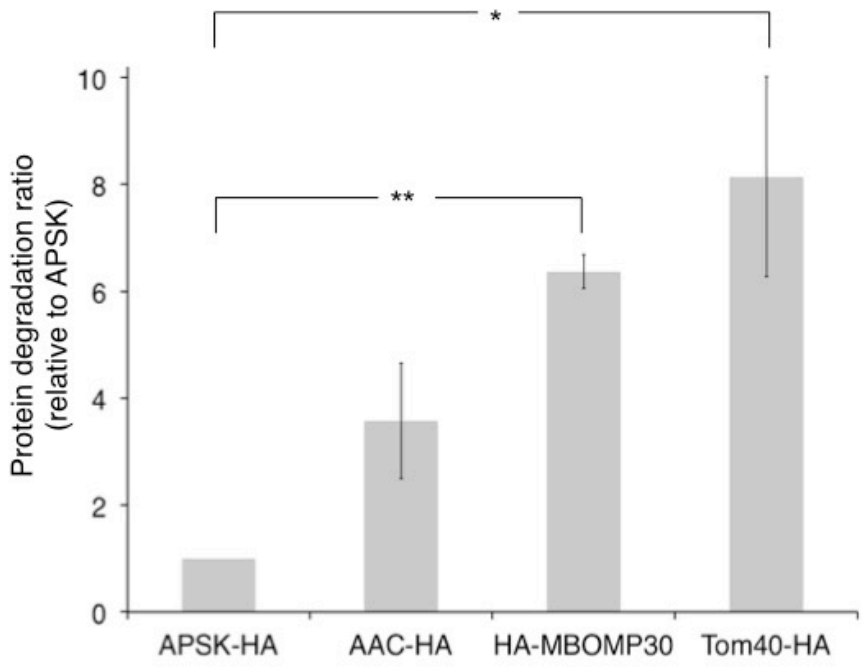


Figure 12. Proteinase K protection assay. (a) Left panel: Cartoon diagram presents the localization of three mitosomal proteins used as control. Middle panel: Cropped immunoblots of proteinase K-treated (+) and untreated (-) organelle fractions are shown. Right panel: The corresponding ratio of protein digestion are provided. (b) Full-length immunoblots with cropped regions highlighted by red boxes. (c) Protease K degradation ratio of HA-MBOMP30 and mitosomal membrane proteins relative to APSK. The percentages of Protease K digestion of HA-MBOMP30, outer membrane control Tom40-HA, and inner membrane control AAC-HA, were normalized against matrix control APSK-HA presented as mean ± standard deviation. The data was analyzed using Student's *t*-test, indicating significant difference between APSK and MBOMP30 (**$p<0.01$) as well as APSK and Tom40 (*$p<0.05$) degradation ratios, strongly suggesting mitosomal membrane and not matrix localization of MBOMP30 based on sensitivity to protease degradation (n=3).

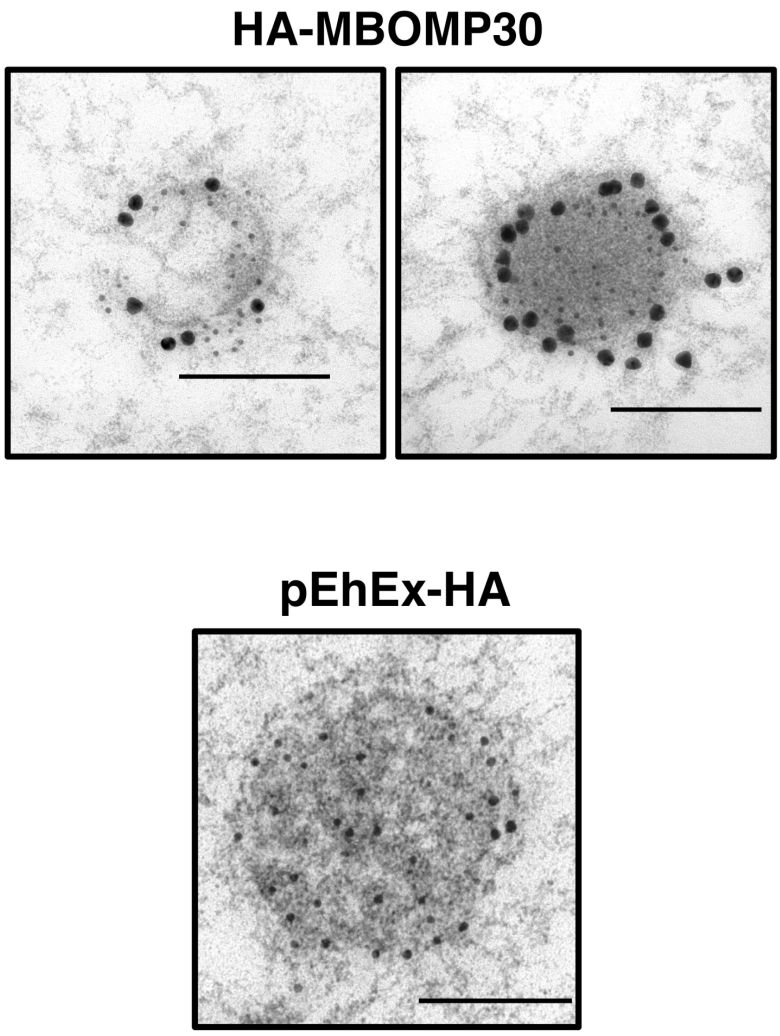


Figure 13. Mitosomal membrane localization of MBOMP30 as observed by immunoelectron microscopy. Immunodecoration of mitosomes of HA-MBOMP30 and control trophozoites with anti-HA (15 nm gold particles) and anti-Cpn60 antibodies (5 nm gold particles) showed electron-dense organelles surrounded by double membranes [scale bar, 200 nm].

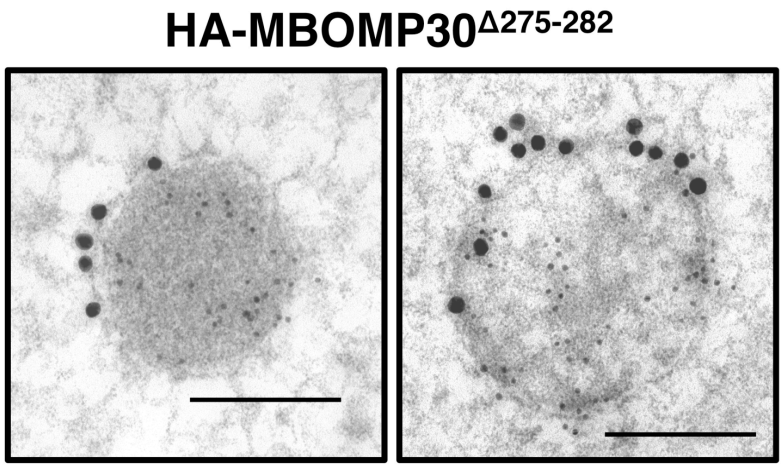


Figure 14. Immunoelectron microscopy of HA-MBOMP30$^{\Delta 275\text{-}282}$.

Mitosomes of trophozoites overexpressing MBOMP30$^{\Delta 275\text{-}282}$ were similarly immunodecorated with anti-HA (15 nm gold particles) and anti-Cpn60 antibodies (5 nm gold particles) [scale bar, 200 nm].

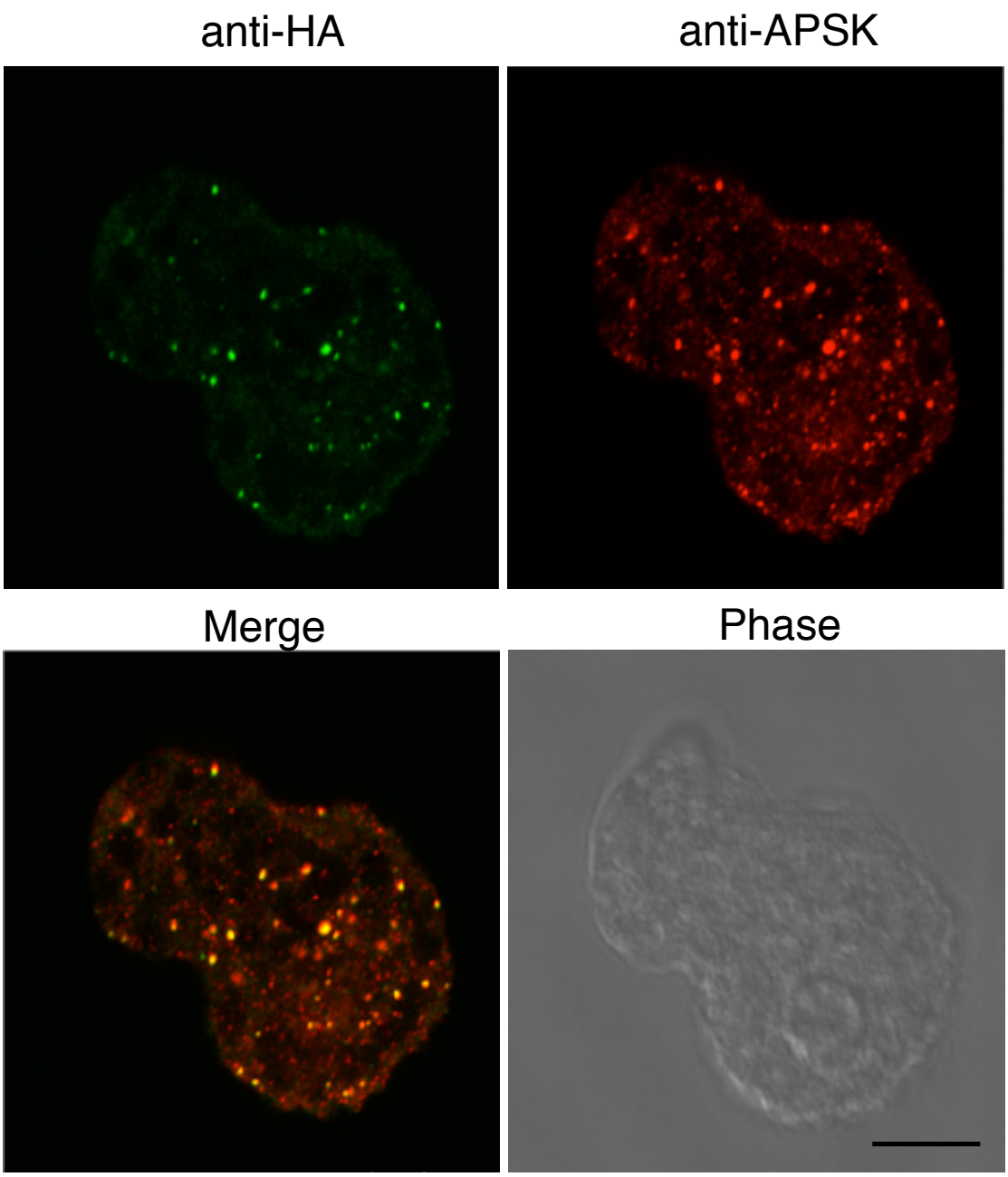
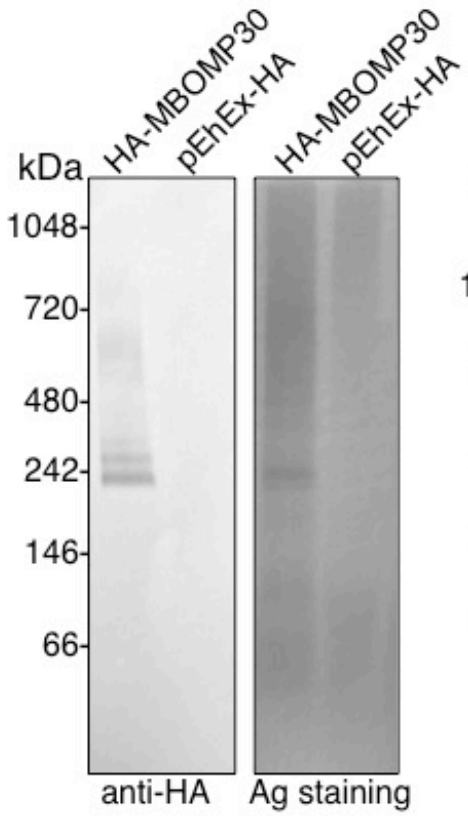
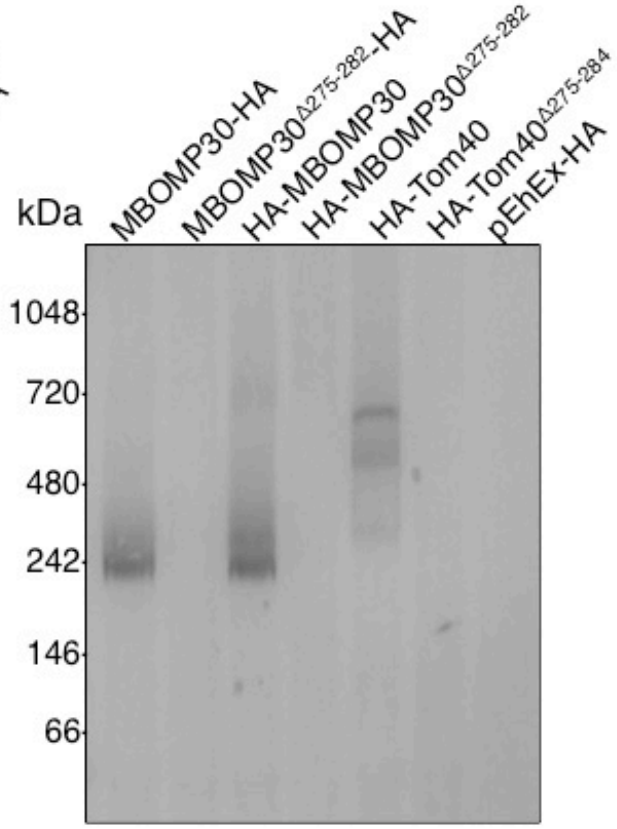


Figure 15. Immunofluorescence analysis of HA-MBOMP30$^{\Delta 275\text{-}282}$ trophozoites. Double staining reaction of HA-MBOMP30$^{\Delta 275\text{-}282}$ trophozoites clearly show colocalization of punctate anti-HA (green) and anti-APSK (red; mitosomal marker) signals [scale bar, 10μm].

a


HA-MBOMP30
pEhEx-HA
HA-MBOMP30
pEhEx-HA
kDa
1048
720
480
242
146
66
anti-HA
Ag staining

b


MBOMP30-HA
MBOMP30Δ275-282-HA
HA-MBOMP30
HA-MBOMP30Δ275-282
HA-Tom40
HA-Tom40Δ275-284
pEhEx-HA
kDa
1048
720
480
242
146
66

c

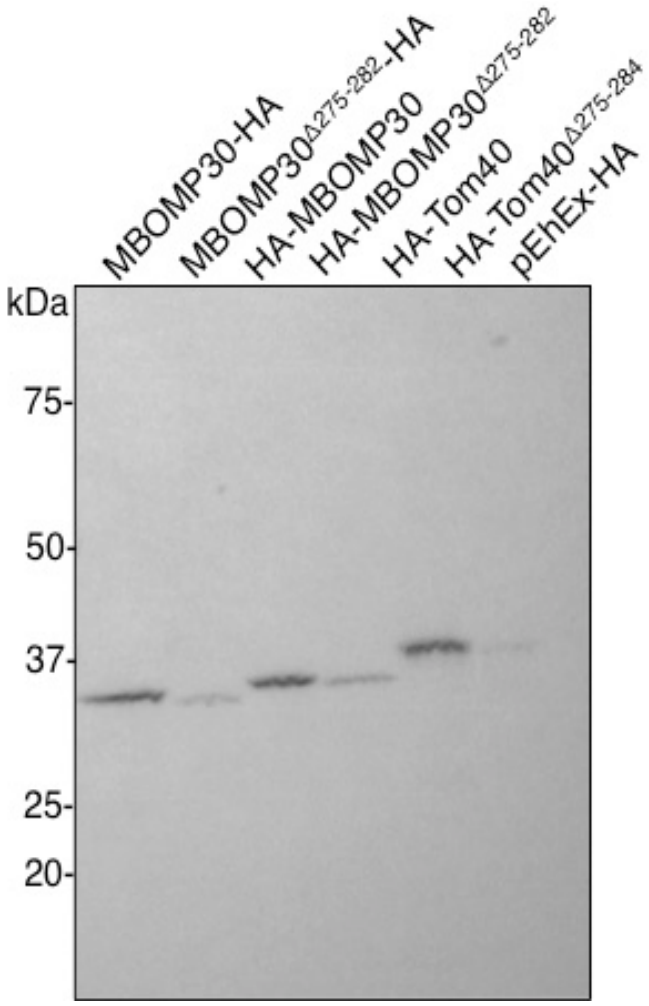


Figure 16. Detection of a 240 kDa EhMBOMP30 complex.

(a) Immunoprecipitation of organelle fraction solubilized in 2% digitonin, followed by BN-PAGE and anti-HA immunostaining (left panel) and silver staining (right panel). **(b)** Anti-HA immunoblots of BN-PAGE and **(c)** SDS-PAGE of subsequent immunoprecipitations using solubilized organelle fractions from amoeba overexpressing MBOMP30 with HA-tagging at either ends of the protein, with or without the putative β-signal. HA-Tom40 was used as a control.