OPEN ACCESS Freely available online

PLOS ONE

Transcriptome Analysis of Encystation in *Entamoeba invadens*

Aleyla Escueta De Cádiz[1,2☯], Ghulam Jeelani[1☯], Kumiko Nakada-Tsukui[1], Elisabet Caler[3], Tomoyoshi Nozaki[1,4*]

1 Department of Parasitology, National Institute of Infectious Diseases, Tokyo, Japan, **2** Department of Biological Science and Environmental Studies, College of Science and Mathematics, University of the Philippines Mindanao, Davao, Philippines, **3** J. Craig Venter Institute, Rockville, Maryland, United States of America, **4** Graduate School of Life and Environmental Sciences, University of Tsukuba, Tsukuba, Ibaraki, Japan

Abstract

Encystation is an essential differentiation process for the completion of the life cycle of a group of intestinal protozoa including *Entamoeba histolytica*, the causative agent of intestinal and extraintestinal amebiasis. However, regulation of gene expression during encystation is poorly understood. To comprehensively understand the process at the molecular level, the transcriptomic profiles of *E. invadens*, which is a related reptilian species that causes an invasive disease similar to that of *E. histolytica*, was investigated during encystation. Using a custom-generated Affymetrix platform microarray, we performed time course (0.5, 2, 8, 24, 48, and 120 h) gene expression analysis of encysting *E. invadens*. ANOVA analysis revealed that a total of 1,528 genes showed ≥3 fold up-regulation at one or more time points, relative to the trophozoite stage. Of these modulated genes, 8% (116 genes) were up-regulated at the early time points (0.5, 2 and 8h), while 63% (962 genes) were up-regulated at the later time points (24, 48, and 120 h). Twenty nine percent (450 genes) are either up-regulated at 2 to 5 time points or constitutively up-regulated in both early and late stages. Among the up-regulated genes are the genes encoding transporters, cytoskeletal proteins, proteins involved in vesicular trafficking (small GTPases), Myb transcription factors, cysteine proteases, components of the proteasome, and enzymes for chitin biosynthesis. This study represents the first kinetic analysis of gene expression during differentiation from the invasive trophozoite to the dormant, infective cyst stage in *Entamoeba*. Functional analysis on individual genes and their encoded products that are modulated during encystation may lead to the discovery of targets for the development of new chemotherapeutics that interfere with stage conversion of the parasite.

Citation: De Cádiz AE, Jeelani G, Nakada-Tsukui K, Caler E, Nozaki T (2013) Transcriptome Analysis of Encystation in *Entamoeba invadens*. PLoS ONE 8(9): e74840. doi:10.1371/journal.pone.0074840

Editor: Matthew Bogyo, Stanford University, United States of America

Received February 26, 2012; **Accepted** August 8, 2013; **Published** September 11, 2013

Funding: This work was supported by a Grant-in-Aid for Scientific Research from the Ministry of Education, Culture, Sports, Science and Technology (MEXT) of Japan (18GS0314, 18050006, 18073001, 20390119, 23390099), a Grant-in-Aid for Scientific Research on Innovative Areas from MEXT of Japan ("Matryoshka-type evolution", 3308), a grant for research on emerging and re-emerging infectious diseases from the Ministry of Health, Labour and Welfare of Japan (H20-Shinkosaiko-ippan-016, H23-Shinkosaiko-ippan-014), a grant for research to promote the development of anti-AIDS pharmaceuticals from the Japan Health Sciences Foundation (KAA1551, KHA1101), and by Global COE Program (Global COE for Human Metabolomic Systems Biology) from MEXT, Japan. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

* E-mail: nozaki@nih.go.jp

☯ These authors contributed equally to this work.

Introduction

Amebiasis is common among individuals exposed to unsanitary health conditions in developing countries. Amebiasis is also seen in developed countries among men who have sex with men and mentally handicapped people [1,2]. In both cases, the infection is established through ingestion of the cysts in feces, or fecal contaminated food and water [1]. Although *in vitro* cultivation and *in vivo* passage of the reference strains and recent clinical isolates of *E. histolytica* led to identification and characterization of the virulence mechanisms associated with amebiasis [3], the molecular mechanisms of differentiation from the invasive trophozoite to the dormant, infective cyst stage, called encystation, remains largely unknown. This is in part due to the lack of *in vitro* or *in vivo* systems that allow differentiation of *E. histolytica* [4]. To overcome this, *E. invadens*, which is a related reptilian species that causes an invasive disease similar to that of *E. histolytica*, has been used as a model system for encystation as *E. invadens* trophozoites can be induced to encyst in axenic

conditions [5–7]. The morphology, the life cycle consisting of binary stages, the sites of encystation, invasiveness to the colonic epithelium, and potential dissemination from the intestine into other organs through the portal vein are similar between the two species.

Several studies focused on identifying genes involved in the stage conversion of *Entamoeba*. It has been shown that galactose/N-acetylgalactosamine, proteasome, beta-adrenergic components, and transcription factors Myb affect stage switching [4,8–12]. Protein kinase C inhibitors and short chain fatty acids have also been linked to encystation [13–15]. In *Giardia lamblia*, cysteine proteases (CP) and UDP-N-acetylglucosamine pyrophosphorylase have been shown to be key enzymes during encystation [16,17]. Similarly, an *E. invadens* CP isotype was found elevated in encysting cells [18]. In the social ameba *Dictyostelium discoideum*, cyclic AMP is used as an autocrine factor for sporulation [19].

Availability of the whole genome sequence of *E. histolytica* has facilitated production of custom-made DNA microarray necessary for identification and classification of genes related to virulence [20,21], the response against oxidative and nitrosative stresses [22], and stage conversion [12]. A Myb transcription factor in *E. histolytica* was also found to regulate transcription of stage-specific genes [23]. Here, we present the whole genome transcriptional profiling of *E. invadens* during encystation. Genes modulated during encystation and their patterns are examined to identify genes and pathways that are involved in encystation.

Materials and Methods

Cultivation and encystation of *E. invadens*

Axenic cultures of *E. invadens* strain IP-1 trophozoites were maintained in BI-S-33 medium at 26 °C. To induce encystation, in two independent experiments in triplicate, trophozoites were harvested in the late logarithmic phase and the cells were transferred to a seven 36 ml flasks with 47% LG medium at a final concentration of 5 x 10^5/ml [8]. Cells were collected at seven time points: 0, 0.5, 2, 8, 24, 48, and 120 h after exposure to the encystation medium. After incubation in 47% LG medium, total numbers of cells were counted under a microscope. One portion of the cells was saved for RNA extraction and another portion was used for the differentiation of trophozoites and cysts. For the determination of cysts, the cells were resuspended in PBS containing 0.05% sarkosyl, and allowed to sit for 20 min at room temperature [24,25]. After lysed cells were stained with 0.22% trypan blue (Wako Pure Chemical Industries Ltd., Japan), intact cysts were counted and the encystation efficiency was measured by dividing the number of cysts resistant to 0.05% sarkosyl (Sigma-Aldrich, St. Louis MO, USA) with the total number of cells suspended in PBS without sarkosyl, in two independent experiments performed in triplicate.

RNA extraction

For isolation of RNA, the cells, harvested at various time points and tested for the sarkosyl sensitivity as described above, were washed three times with 1X PBS and collected by centrifugation at 1, 500 rpm for 5 minutes after induction to wash off the encystation medium. The collected cell pellets were resuspended in 1 ml of Trizol reagent (Invitrogen, Carlsbad, CA, USA) and lysed using a Dounce homogenizer (approximately 300 strokes) until majority of the cysts were lysed as previously described [39]. The RNA concentration for each sample was measured using a Nanodrop Spectrophotometer 1000 (Thermo Scientific, Wilmington, DE, USA). RNA integrity was checked using Bio-Rad's Experion Automated Electrophoresis System (RNA StdSens analysis kit).

Affymetrix Microarray Hybridization

All reagents and protocols used in this study were as described in GeneChip® Expression Analysis Technical Manual (Affymetrix, Inc. Santa Clara, CA, USA). Five arrays were used for five independently isolated RNA samples corresponding to 2 biological replicates (3 arrays for the first set and 2 arrays for the second set; two sets of encystation experiments were carried out > 1 year apart in-between), were used for each time point. Using the One-Cycle cDNA synthesis kit, 5 g of total RNA was reverse transcribed using T7-Oligo (dT) primer in the first strand cDNA synthesis. After the second strand synthesis, the double-stranded cDNA template was used for *in vitro* transcription (IVT), in the presence of biotinylated nucleotides (GeneChip IVT labeling kit) to produce Biotin-labeled cRNA. Unincorporated NTPs were removed from the biotinylated cRNA (GeneChip® sample cleanup module) and then purified, quantified and fragmented. Hybridization cocktail of eukaryotic hybridization controls and fragmented, labeled cRNA (GeneChip® Hybridization, Wash and Stain Kit) were hybridized for 16 hours at 45 °C (Hybridization Oven 640, Affymetrix) to custom-generated Affymetrix platform microarray (49-7875) with probe sets consisting of 11 probe pairs each representing 12, 384 *E. invadens* open reading frames (Eh_Eia520620F_Ei) and 9, 327 *E. histolytica* (Eh_Eia520620F_Eh). The array chips were washed and stained (GeneChip® Hybridization, Wash and Stain Kit) with Streptavidin–phycoerythrin Biotinylated anti-streptavidin antibody using a GeneChip® Fluidics Station 450 (Affymetrix) for 1.5 hours. After washing and staining, the GeneChip® arrays were scanned using the Hewlett-Packard Affymetrix Scanner 3000.

Analysis of microarray data

Raw probe intensities were generated by the GeneChip Operating Software (GCOS) and GeneTitan Instrument from Affymetrix. Normalized expression values for each probe set were obtained from R 2.7.0 downloaded from the BioConductor project (http://www.bioconductor.org) using robust multiarray averaging with correction for oligosequence (gcRMA). Standard correlation coefficients were calculated using GeneSpring GX 10.0.2. Reproducibility of the experiments was determined by Pearson's correlation coefficient and confirmed by principal component analysis. Only genes that were considered 'present' by GCOS at least one of three arrays at any time points were used in further analysis. One-way ANOVA analysis with Tukey's Post Hoc test was performed to extract

differentially expressed genes. Gene probe sets were considered differentially expressed between time points if they had at least a 3 fold change compared against the value at 0 hour and a p-value < 0.05, calculated using Welch's t-test, after multiple test correction by the Benjamini–Hochberg method. A post-hoc test using Tukey's Honestly Significant Difference test was conducted to determine significant differences between samples. The data presented in this publication have been deposited in NCBIs Gene Expression Omnibus (GEO, http://www.ncbi.nlm.nih.gov/geo/) and are accessible through GEO Series accession number GSE33312.

Annotation of Rab small GTPases and cysteine proteases

We searched the *Entamoeba* genome database using *E. histolytica* Rab small GTPases and cysteine proteases as query [3,26]. One hundred twenty one Rab and 64 EiCPs were annotated after CLUSTAL W alignment, manually edited using BioEdit and phylogenetic tree created using MEGA4 software [27]. For details, see references [28,29].

Results and Discussion

Kinetics of morphological differentiation

In order to identify and characterize genes and gene cascades involved in encystation, we examined the transcriptional profiles at 7 time points (0, 0.5, 2, 8, 24, 48, and 120 h) during encystation of the reptilian ameba *E. invadens*. At 8 h post-induction of encystation by transferring trophozoites to the differentiation medium of low osmolarity containing no glucose, trophozoites became highly motile as compared with those maintained in the BI-S-33 medium, and only 0.9-2.8% of cysts were formed (Figure 1A). At 24 h after induction, the trophozoites rounded up, became immobile, and formed clusters, and the percentage of cysts increased to 19.7%. At 48 h of encystation, large multicellular aggregates were formed, and 50.7% of the total cells transformed into cysts. At 120h post induction 91.5% of cells transformed into cysts (Figure 1A).

Overview of transcriptional changes during encystation

Among 12, 384 probe sets (Table S1) corresponding to *E. invadens* open reading frames (including 1, 272 probe sets that had been removed from NCBI), approximately 6, 014 genes were found to be expressed, i.e. had a "present" call in at least one of the five experimental replicates, at least one time point. We did not only choose genes that were found "present" in all five replicates because this will narrow down the size of genes to be analyzed during the course of encystation (Figure 1B and Table S2). These genes were filtered to extract genes whose probe sets represent a single gene (a probe set name contains suffix "_at") or that were so highly similar in sequence to other genes as to make it impossible to design a unique probe set (a probe set with a suffix "s_at"). We set significant levels of changes to 3 fold, similar to that used in the previous work [12], where cyst-specific genes were identified in recent clinical isolates of *E. histolytica*. Furthermore, much higher numbers of

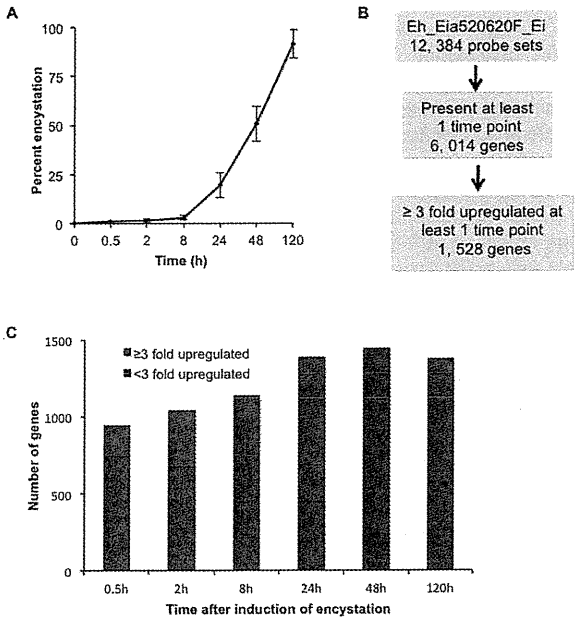

Figure 1. Overview of transcriptomic analysis. (**A**) **Kinetics of differentiation**. The percentages of the amoebae resistant to 0.05% sarkosyl during encystation are shown. Values are presented as % encystation and represent the mean ± S.D. of two independent experiments conducted in triplicate. (**B**) **Flow of analysis**. Microarray data were obtained in triplicates from *E. invadens* exposed to 47% LG medium for 0, 0.5, 2, 8, 24, 48, or 120 h, and genes expressed in at least one time point were selected for further analysis. The second data set of two biological replicates are shown as a representative. (**C**) The number of genes that were proven to be statistically significantly up-regulated by ≥3 or <3 fold at each time point of encystation.

doi: 10.1371/journal.pone.0074840.g001

genes were selected when lower fold (e.g., two fold) changes were used, which made description of modulated genes very lengthy. To validate the reproducibility of the results, we compared the transcriptomic data from the two biological replicates at different time points during encystation (Figure S1). The two data sets showed reasonable Pearson correlation coefficients (R values ranging from 0.7451 to 0.9021). We selected for the further analysis only genes that were modulated by ≥3 fold in both sets of biological replicates.

In general, the number of up- and down-regulated genes increased as encystation proceeded. We mainly focused on the up-regulated genes during encystation. The number of genes that were up-regulated ≥3 fold at any time points was 1, 528 (Table 1 and Table S3). Among the up-regulated genes, the number and proportion of the genes up-regulated ≥3 fold, compared to up-regulated < 3 fold, tends to increase at the later time points of encystation (Figure 1C) with the highest number of up-regulated genes noted at 48 h of encystation. A total of 2841 genes were down regulated by ≥3 fold at one or

Table 1. Grouping and distribution of 1,528 genes that were up-regulated ≥3 fold at least one time points during encystation.

Category	0.5h	2h	8h	24h	48h	120h	Number of genes
1	+						11
2		+					30
3			+				10
4	+	+					26
5	+		+				1
6		+	+				15
7	+	+	+				23
8				+			85
9					+		4
10						+	57
11				+	+		89
12					+	+	358
13				+		+	7
14				+	+	+	362
15	+	+	+	+	+	+	54
16		+		+			3
17	+	+	+	+			12
18	+	+	+	+	+		13
19	+			+			3
20		+	+	+			18
21		+	+	+	+		19
22		+	+	+	+	+	81
23		+				+	2
24			+	+			16
25			+	+	+		23
26			+			+	2
27			+	+	+	+	107
28	+		+	+	+	+	4
29	+	+		+	+	+	4
30	+	+	+		+	+	1
31	+	+			+		1
32	+	+		+			3
33	+	+				+	3
34	+				+	+	10
35		+			+	+	17
36			+		+	+	6
37	+			+	+		1
38		+		+	+	+	20
39	+			+	+	+	12
40		+	+		+	+	3
41	+	+			+	+	11
42	+	+		+	+		1
			Total				1528

doi: 10.1371/journal.pone.0074840.t001

more time points during encystation (Table S1). At each time point, 263, 528, 543, 1325, 1835, and 1998 genes were downregulated by ≥3 fold at 0.5, 2, 8, 24, 48, and 120 h, respectively.

To identify genes that are modulated at specific time points during differentiation, we further grouped the genes into 42 categories based on expression profiles (Table 1). About 37% of genes (566) are up-regulated at 0 to 8 h of encystation; 20% (116 genes, categories 1-7) of which are exclusively up-regulated at 0 to 8 h while 80% (450 genes, categories 15-42) of those genes were also up-regulated at later time points (Table 1). List of genes up-regulated at 0.5 and/or 2 h are presented in Table S4. Genes up-regulated at 8 h are listed in Table S5.

For the 962 genes up-regulated at later time points (categories 8-14 Table 1), 9%, 0.4%, and 6% of the genes specifically peak at 24, 48 or 120 h respectively, whereas 37% (358 genes) peak at two time points (48 and 120h), 9% (89 genes) and 0.7% (7 genes) peak at 24/48 h and 24/120 h, respectively. Thirty eight percent of genes (362 genes) are continuously up-regulated at 24 to 120 h. List of genes up-regulated only at 24 h are listed in Table S6 and genes up-regulated at 48 and/or 120 h are listed in Table S7.

Only 31% (469 genes) of the up-regulated genes were annotated (Table S3). Of the 1,059 genes encoding for hypothetical proteins, 18% (187 genes) have orthologs in other organsisms. In the following sections, we summarize the modulated annotated genes based on functional classes.

A: Bacterial surface protein A (BspA) family

Leucine-rich repeat (LRR)-containing proteins, which were initially identified in *Bacteroides forsythus* (BspA), are one of the most abundant multicopy genes in the *E. invadens* genome representing about 1.4% of the total *E. invadens* sequence reads [30]. They are annotated in AmoebaDB as hypothetical proteins with conserved regions. Similarly, 114 genes encoding for BspA-like proteins were identified in *E. histolytica* genome [31]. Homology searches using these *E. histolytica* BspA-like proteins revealed that *E. invadens* contain 149 BspA-like proteins (data not shown). Our transcriptome analysis showed that 26 out of 149 *E. invadens BspA* genes were up-regulated during encystation (Figure 2 and Table S3). About half (11) the genes were up-regulated at 48 and/or 120 h (Table S7), while 5 *BspA* genes were up-regulated at 0.5 and/or 2 h (Table S4), and the other 10 genes were up-regulated at different time points (Figure 2, Table S3). Time-dependent up-regulation of specific subsets of BspA-like genes are intriguing, as BspA was implicated in the attachment and invasion to host cells in *Treponema denticola* and *Tannerella forsythiae* [32,33]. One of *E. histolytica* BspA-like protein, EhLRRP1, has been shown to be localized on the cell surface, but its possible role in interaction with host ligands is not yet established [34]. As proposed in *Trichomonas vaginalis*, where BspA-like proteins might be involved in cell-cell adhesion when *T. vaginalis* forms large aggregates [35], BspA-like proteins may also be involved in a similar phenomenon in *E. invadens* during the early stage of encystation. Despite the similarity in the LRR repeats among *E. invadens*, bacteria, and trichomonads, the lack of the amino-terminal sequence and the transmembrane domain of *E. invadens* LRR suggests that its function is divergent [30]. Thus, up-regulation of *E. invadens* BspA at the late stage of encystation is highly remarkable, as BspA-like proteins were not shown associated with cell differentiation in other organisms [32,33]. It was, however, shown in *T. vaginalis* that transcript level of some BspA proteins change upon exposure

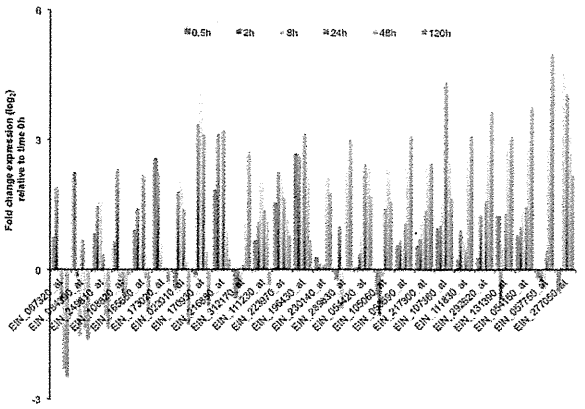

Figure 2. Modulation of the transcript level of the *E. invadens* BspA- like genes during encystation. Values are expressed as log_2 fold change of expression relative to time 0 h.

doi: 10.1371/journal.pone.0074840.g002

in high and low iron concentrations [35]. In *E. histolytica*, iron and serum starvation resulted in the trafficking of a cytoplasmic EhRab11A protein to the cell periphery and the development of detergent resistance, similar to the cyst stage [36]. It would be interesting to show the localization of *E. invadens* BspA in encysting cells to determine its possible function during encystation.

B: Cytoskeletal proteins

The Rho/Rac family of small GTP binding proteins is known to be involved in cytoskeleton regulation [37]. Two *E. invadens*-specific (i.e., no homolog in *E. histolytica*) *Rac* genes (EIN_166990 and EIN_017340) were modulated at early time points (Table S1), while one *RacJ* (EIN_243630) and one *RacD* (EIN_137540) genes, which also have *E. histolytica* counterparts, were up-regulated at late time points together with several Rho/Rac effectors [GTPase activating proteins (GAP) and guanine nucleotide exchange factors (GEF)]. In *E. histolytica*, it was previously shown that only a gene encoding for RacH, whose physiological role has not yet been established, was up-regulated specifically in cysts [12]. A homolog of *RacH* gene (EIN_105260) showed slight change (2 fold) in gene expression at 0.5 h of encystation (Table S1). These data are consistent with the notion that regulation of cytoskeletal rearrangement is essential at the early phase of encystation when trophozoites rearrange its surface for aggregation. It was also shown that cytochalasin D, a potent inhibitor of actin polymerization, inhibits encystation [38].

Phosphoinositides such as phosphatidylinositol 4,5-bisphosphate [PtdIns(4,5) P2] are important secondary messengers in cell surface receptor-mediated signal transduction and participate in actin cytoskeleton rearrangement [37]. The enzyme phosphatidyl 3-kinase (PI3K), which phosphorylates PtdIns, was previously shown to participate in the encystation process [13,14,39]. Our transcriptome data showed the mRNA level of one *PI3K* genes (EIN_083000) increased >3 fold at 24 to 120 h of encystation (Table S6). Furthermore, genes encoding PI(4,5) P2,3-kinase (EIN_310690) and diacylglycerol kinase (EIN_196180) were also up-regulated at later time points (Tables S6 and S7). The involvement of the pathway in encystation and excystation was previously suggested [39,40]. Recently, phosphoinositides, particularly PtdIns3P and PtdIns4P, were shown to participate in cytoskeletal rearrangement during phagocytosis of *E. histolytica* [41].

C: Kinases and phosphatases

Tyrosine kinases play a pivotal role in sensing changes in the environment. It has been previously shown by analyzing transcriptome of recent clinical *E. histolytica* isolates that at least 14 transmembrane kinases (TMKs) are developmentally regulated [12]. Modulation of 8 TMKs were also observed *in vivo* [20]. Recently, analysis of *E. histolytica* TMK 39, 54, and 96 have shown to be involved in phagocytosis and growth [42,43]. Similarly, we identified 16 up-regulated *E. invadens* genes that showed significant similarity to 11 genes encoding *E. histolytica* TMKs (Table S8). Six of them were up-regulated at the early time points (*TMK 87*, Tables S4 and S5), while ten were up-regulated at the later time points (Tables S6 and S7). Forty-six *E. histolytica TMK* genes were not detected to be transcribed in trophozoites, similar to the three major cyst-specific Jacob proteins [44]. Among the *E. invadens* homologs corresponding to these 46 *E. histolytica TMK* genes, five *E. invadens TMK* gene homologs (8, 40, 38, 73, and 82) were up-regulated in later time points of encystation. However, in contrast to the previous finding, which suggested that TMK54 is involved in growth and surface expression of Gal/GalNAc lectin in trophozoites of *E. histolytica* [43], *E. invadens TMK 54* gene expression was not significantly modulated at early time point of encystation. These data suggest that TMK54 may have divergent functions in two species. *E. invadens TMK87* gene was shown to be up-regulated in *E. histolytica* in recent clinical isolates [12].

Genes encoding for serine/threonine protein phosphatases and dual specificity phosphatases were also increased during the late phase of encystation (Tables S6 and S7). In particular, 4 genes (EIN_221990, EIN_105320, EIN_020140, EIN_230540) encoding for the serine threonine phosphatases 2C (PP2C) were up-regulated during the late phase of encystation (Table S7). In yeast, PP2C phosphatases are implicated in attenuating phosphorylation during heat and osmotic shock [45]. Thus, up-regulation of PP2C might reflect anti-stress responses of *E. invadens* during encystation, which was reported to be induced by glucose starvation and hypo-osmotic shock [8]. However, the *E. invadens* PP2C homologs (Table S7) in *E. histolytica* (EHI_194220 and EHI_092510) were not shown to be cyst specific [12].

D: Metabolism

A majority of metabolic genes involved in central energy metabolism in general were repressed. However, despite its dormant nature, genes encoding several metabolic enzymes involved in nucleotide metabolism, energy, lipids, and

sphingolipids metabolism were still transcribed during the late phase of encystation (Tables S6 and S7). In our previous study we discussed detailed analysis of metabolisms of glycolysis, amino acid, cyst wall biosynthesis [46]. Briefly, among genes involved in chitin biosynthesis, the transcript level of a gene encoding for glucosamine-fructose-6-phosphate aminotransferase (GFAT, EIN_136750), which is the first and rate limiting enzyme of the chitin biosynthetic pathway, was increased at 24 h of encystation (Table S3). Chitin is the major components of the cyst wall and a homopolymer of β-1, 4-linked *N*-acetyl-glucosamine (GlcNAc) [47]. It was shown in *Giardia* that UDP-GlcNAc pyrophosphorylase (UAP) promotes the synthesis of GlcNAc and cyst wall filaments [16]. In addition, genes encoding for UDP-glucose 4-epimerase (UAE), glucosamine 6-phosphate N-acetyltransferases (GNA), phosphoglucosamine mutase (AGM), and glucosamine-6-phosphate isomerase (GNP) were also shown to be increased at mRNA and protein levels during encystation in *Giardia* [48]. However, in contrast to the findings in *Giardia*, genes for only GNA (EIN_036890) and one of UAPs (EIN_224560) were found significantly up-regulated during encystation of *E. invadens* (Table S3).

Three genes encoding for β-1,3-N-acetylglucosaminyltransferase, involved in glycosphingolipid and glycan biosynthesis, were up-regulated at either early (EIN_068160; Table S4) or late phase (EIN_112490 and EIN_200230; Table S6) of encystation. These enzymes participate in the transfer of GlcNAc from UDP-GlcNAc onto Gal ↓β-3 (GlcNAc β-6) GalNAc-mucin and are therefore important in chitin biosynthesis.

E: Proteasome components, ubiquitin, and SUMO

It has been shown that expression of ubiquitin (*Ub*) gene is co-up-regulated with known cyst-specific genes during encystation. In addition, encystation was inhibited by proteasome inhibitors, suggesting that ubiquitin-proteasome activity is essential for encystation [9]. In agreement to the premise, transcription of major components of the Ub pathway such as the anaphase-promoting complex (EIN_034040), cell cycle division (EIN_192160), E2 Ub conjugating enzymes (EIN_101850), and E4 ubiquitination factor (EIN_107750) genes were up-regulated in the late phase of encystation (Tables S6 and S7). However, only *E. invadens Ub* gene (EIN_063840) in the *E. invadens* genome database was not significantly modulated during encystation process, which seems to contradict with the previous finding [9]. However, one should note that the *E. invadens Ub* gene previously shown to be up-regulated (AF016643 [9]) was only 52% identical to EIN_063840. The genes encoding for E1 Ub activating enzymes, 26S proteasome regulatory and core particle subunits (Table S1) were, though highly expressed, not significantly up-regulated during encystation. The gene encoding for ubiquitin carboxy-terminal hydrolase (EIN_243050) with a peptidase C19 motif was up-regulated at 8 to 120 h with a 67-fold peak expression at 24 h. A gene encoding for Ub-specific protease (EIN_107760) was also up-regulated at 24 to 48 h of encystation (Table S3). These de-ubiquitinating enzymes are likely required to process Ub-conjugated products, negatively regulate ubiquitination, and regenerate free Ub [49].

An antagonistic relationship has been established between the Ub system and sentrin/small ubiquitin-related modifier (SUMO) [50]. It has been shown that Ub and SUMO compete for a single modification site of an inhibitory protein involved in the signaling of the transcriptional activator nuclear factor–κB. SUMO-specific E2 (conjugating enzyme) Ubc9 inhibits NF-κB-dependent transcription in response to a variety of signals [50]. Our transcriptome data also showed an up-regulation of genes encoding SUMO-specific proteases (EIN_157340 and EIN_200450), SUMO ligases (EIN_168610 and EIN_081680), and Ubc9 (EIN_220240) on the late phase of encystation (Tables S6 and S7). It needs to be further determined whether Ub/SUMO antagonistic system is operated in *Entamoeba*, and the target substrates for ubiquitylation or sumoylation need to be identified [9].

F: Protein transporters

Genes encoding for major facilitator superfamily (MFS) transporter proteins were up-regulated at different time points: EIN_257160 and EIN_040590 at 0.5 to 2 h (Table S4), EIN_054130 at 8 h (Table S5), EIN_035840 at 24 h (Table S6) and EIN_059680 at 48 to 120 h (Table S7). In general, MFS proteins facilitate the transport across the cytoplasmic or internal membranes of a variety of substrates including ions, sugar phosphates, drugs, neurotransmitters, nucleosides, amino acids, and peptides [51]. Genes encoding for two MFS (EIN_035840 and EIN_059680) that were up-regulated at the late time points were predicted to be involved in multidrug efflux as predicted by TransportDB (http://membranetransport.org/), whereas the substrates of the MFS that was expressed at the early time points (EIN_257160 and EIN_040590) were not predicted. Two genes encoding for CorA (EIN_053430 and EIN_222130) metal ion transporters (MIT), which transport magnesium/cobalt ions, were also up-regulated at 2 h (Table S4). It was previously shown that supplementation of varying concentrations and mixtures of Mg^{2+}, Mn^{2+}, and Co^{2+} ions to PEHPS culture medium was essential for the production of "cyst-like" structure in *E. histolytica* trophozoites [52,53]. The presence of bivalent metal ions Mn^{2+} and Co^{2+} was also shown to be necessary for augmenting chitin synthase activity in encysting *E. invadens* [54] and recognized as co-factors in the synthesis of the cyst wall chitin [52]. Genes encoding chitin synthases (EIN_040930 and EIN_168780) were up-regulated >3 fold starting from 2 and 8 h, respectively, and remained up-regulated up to 120 h of encystation (Table S3). The simultaneous up-regulation of metal ion transporters and chitin synthases during encystation likely supports the previous report showing that *E. histolytica* generated chitin-like material during axenic cultivation upon supplementation of these metal ions [55].

Genes encoding other ion transporters including voltage ion superfamily (EIN_036050), P-type ATPase (EIN_153520 and EIN_051610), and resistance nodulation cell division (RND) transporter (EIN_016330) were also up-regulated at the early time points (Table S4). In contrast, six transporter genes encoding the ATP-binding cassette (ABC) superfamily

(EIN_015980, EIN_135600, EIN_103360, EIN_167910, EIN_146950, and EIN_059680) and four *importin* genes (EIN_219050, EIN_093910, EIN_040110, and EIN_069500) were up-regulated at the later time points (Table S7). A gene encoding for an ABC transporter (EIN_103360), previously reported to be expressed in a cyst-specific manner (EHI_178050), was also up-regulated [12]. Importin α subunit is known to bind to the nuclear localization signal of the proteins to be imported, whereas importin β subunit facilitates the docking of the importin-protein complex to the nuclear pore, respectively.

It remains still uncharacterized how the cyst wall proteins are transported in encysting trophozoites [11]. UDP-GlcNAc is the end product of the hexosamine biosynthesis pathway and the essential precursor of chitin. It was shown that accumulation of UDP-GlcNAc precedes chitin formation [46]. Two genes encoding for the UDP-GlcNAc transporter (EIN_294920 and EIN_248420) were not up-regulated with statistical significance, but gene expression slightly increased at 24 to 120 h of encystation (Table S1). These genes are known to be mainly involved in the transmembrane transport of nucleotides and sugars in the Golgi apparatus, which is the site of glycosylation, sulfation, and phosphorylation of proteoglycans and sphingolipids [56]. Up-regulation of genes encoding for UDP-GlcNAc transporters and chitinase genes (EIN_239240, EIN_053310, EIN_059870) (see below) simultaneously occurred in the late phase of encystation.

G: Vesicular trafficking: small GTPases and their effectors

E. invadens possesses 121 *Rab* genes, which were previously designated [28]. Of these 121 *Rab* genes, 14 genes (Figure 3) including 7 genes encoding RabX isotypes, which have corresponding homologs in *E. histolytica*, and 5 genes encoding *E. invadens*-specific (i.e., no homolog in *E. histolytica*) RabZ were up-regulated during encystation. Among the *EiRabX* isotype genes, three genes were up-regulated at later time points, one gene was up-regulated at early time points, while three other *EiRabX* isotype genes were intermittently modulated during the entire encystation process. Variation in the expression pattern of *E. invadens* specific *RabZ* isotype genes was also observed.

Three *E. histolytica* Rabs were previously suggested to be involved in encystation. A gene encoding for EhRab11A (previously named as EhRab11, and re-designated in reference 26) was up-regulated in a "cyst-like" form formed in a serum-deprived medium [36], while *EhRabM1* and *EhRabN1* genes were found to be highly expressed in recent clinical isolates that retained encystation ability, compared to laboratory strain, HM-1 [12]. Up-regulation of *EhRab7D* gene of an avirulent HM-1 strain was reported previously [57], but this gene was also shown to be down-regulated in recent clinical isolates [12]. Our transcriptome data showed that among two Rab subfamily (Rab7 and RabN), *EiRab7D* gene expression was up-regulated at 2 to 120 h of encystation while up-regulation of *EiRab7l* gene expression started earlier (0.5 h) and remained up-regulated up to 48 h of encystation. *EiRabN1* gene expression was upregulated at the 120 h of encystation (Figure 3).

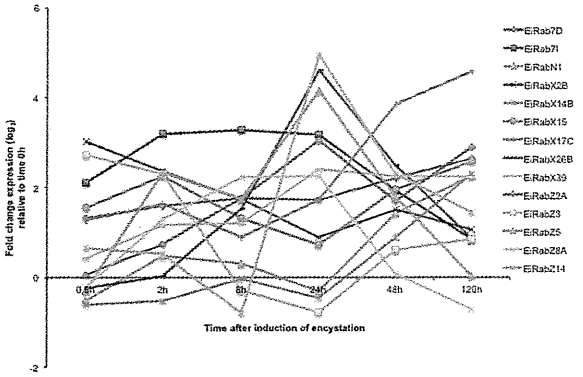

Figure 3. Modulation of the transcripts of 14 *E. invadens* Rab genes during encystation (0.5-120 h). Values are expressed as $\log_2$ fold change of expression relative to time 0 h. Gene IDs: EiRab7D, EIN_133760; EiRab7l, EIN_196420; EiRabN1, EIN_136950; EiRabX2B, EIN_099000; EiRabX14B, EIN_147580; EiRabX15, EIN_238750; EiRabX17C, EIN_107380; EiRabX26B, EIN_060100; EiRabX39, EIN_238590; EiRabZ2A, EIN_289320; EiRabZ3, EIN_192430; EiRabZ5, EIN_039070; EiRabZ8A, EIN_270650; EiRabZ14, EIN_061010.

doi: 10.1371/journal.pone.0074840.g003

H: Cyst wall components

The major components of *E. invadens* cyst wall are the Jacob, Jessie lectins, and chitinase [58]. It has been proposed that these components are assembled in the cyst wall in a "wattle and daub" model, in which Jacob lectins form the wattle, chitinases cross-link the microfibrils, and Jessie lectins form the mortar or daub [59]. Chitin deacetylase acts by cleaving chitin to form chitosan on the surface of the cyst wall [47]. Based on our transcriptomic data (Figure 4), the expression patterns of each lectin and chitin subtype varied. Two genes encoding for EiJacob 1 (EIN_050710_s_at) and 4 (EIN_294450_at) were up-regulated, but the upregulation was not statistically significant. However, expression of *EiJacob1* gene increased at 0.2-120 h, while expression of *EiJacob4* gene increased from 0.5 to 48 h (Table S1), expression of genes encoding EiJacob 2, 3, and chitin synthases were up-regulated at 2 h, followed by those encoding EiJacob 5, 6, 7, chitin deacetylase 2, and chitinase 2 at 8 h. Expression of *EiJessie1c* gene was up-regulated at 8-48 h. Expression of *EiJessie3a, 3b*, and *chitinase 1* genes were upregulated at 24 h (Figure 4). Chitinase 3 was constitutively expressed. Therefore, the expression profiles of these components do not support the proposed "wattle and daub" model, and may suggest post-transcriptional regulation of these proteins.

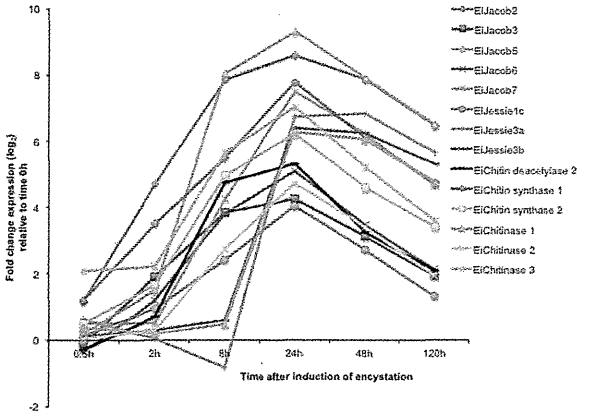

Figure 4. Modulation of the transcript level of the *E. invadens* cyst wall components during encystation (0.5-120 h). Values are expressed as $\log_2$ fold change of expression relative to time 0 h. Gene IDs: EiJacob 2, EIN_137570; EiJacob 3, EIN_016240; EiJacob 5, EIN_104770; EiJacob 6, EIN_015880; EiJacob 7, EIN_186850; EiJessie 1c, EIN_243430; EiJessie 3a, EIN_040990; EiJessie 3b, EIN_058620; EiChitinase 1, EIN_239240; Eichitinase 2, EIN_053310; Eichitinase 3, EIN_059870; EiChitin deacetylase 2, EIN_058630; EiChitin synthase 1, EIN_040930; EiChitin synthase 2, EIN_168780.

doi: 10.1371/journal.pone.0074840.g004

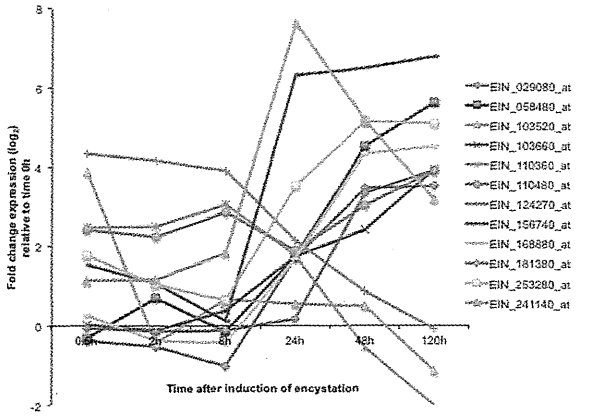

Figure 5. Modulation of the transcript level of the *E. invadens* Myb transcription factors during encystation (0.5-120 h). Values are expressed as $\log_2$ fold change of expression relative to time 0 h after induction of encystation.

doi: 10.1371/journal.pone.0074840.g005

I. Myb transcription factors with R2R3 and $^{T}/_{S}$HAQK$^{Y}/_{F}$ motifs

The Myb family of transcription factors is essential in regulating cell differentiation, cell proliferation, and cell cycle [60]. Recent *in silico* analysis of the *E. histolytica* genome showed 34 proteins with Myb DNA-binding domains [61]. A gene encoding for one of these Mybs with a $^{T}/_{S}$HAQK$^{Y}/_{F}$ motif was shown to be developmentally regulated [12]. It has also been shown that this Myb transcription factor regulates expression of a subset of stage-specific genes in *E. histolytica* [23]. Using the Pathema *Entamoeba* genome database, we searched for *E. invadens* Mybs, and identified a total of 37 *E. invadens* Myb genes encoding Myb proteins with R2R3 and $^{T}/_{S}$HAQK$^{Y}/_{F}$ conserved motifs. Twenty-eight putative *E. invadens* Myb proteins (25 annotated and 3 hypothetical proteins) possess conserved R2R3 repeats, eleven of these were differentially expressed (Figure 5). The $^{T}/_{S}$HAQK$^{Y}/_{F}$ motif was found in 9 hypothetical proteins; and only one gene (EIN_241140_at) of them was up-regulated during encystation (Figure 5). The expression profiles of most of these *Myb* genes were different, which is consistent with a notion that each Myb transcription factor controls expression of specific subsets of genes on a specific phase of encystation, as previously suggested [23].

J: Cysteine proteases

Fifty cysteine proteases (CPs) have been identified in the *E. histolytica* genome [24]. Ten of these CPs were detected in trophozoites, most of which appear to be linked with virulence [3,26,62–65]. However, the role of CPs in encystation has not been elucidated. Recently, 8 CP genes have shown to be differentially expressed in cysts using xenic *E. histolytica* clinical isolates [12].

Cysteine proteases in *E. invadens* were previously identified and annotated [29]. We grouped 64 *E. invadens* CPs into three categories: cyst specific CPs (11 EiCPs), expression of which increased at 24 to 120 h; trophozoite specific CPs (19 EiCPs), expression of which was higher at 0 to 8 h of encystation compared to 24 to 120 h; and constitutively expressed CPs (34 EiCPs) (Figure 6A). Among the modulated CPs (Figure 6B), six belongs to C1 papain superfamily clades A and B (EiCP-A2c, EiCP-A3e, EiCP-BA, EiCP-BB, EiCP-B6, and EiCP-B9), one calpain-like protease (EiCalp2b), two ubiquitin carboxy-terminal hydrolases (EiUCHa and EiUCHc), and one Ulp protease (EiUlpC). Ulp proteases are a group of peptidases that control the function of SUMO. Expression of *EiCP-A2c* and *EiCP-BB* genes was up-regulated at early time points, while that of *EiCP-B9* and *EiUCHa* genes were up-regulated at 24 and 120 h, respectively. Six EiCPs were up-regulated at 24 h or later of encystation. The most striking result was the 61-fold up-regulation of *EiCP-BA* transcript at 24h of encystation. The closest homolog of EiCP-BA is EhCP-B6, although EiCP-BA has a transmembrane domain and an ERFNIN motif similar with a cathepsin L-like enzyme [3,65]. Additionally, expression levels of *EiCP-B6* and *EiCP-UCHa* genes, which were not expressed at 0-0.5 h of encystation, dramatically increased by 99 and 68 fold at 24 h, respectively. *E. histolytica* genes homologous to *EiCP-A3e* and *EiCP-B9* genes in were also up-regulated in cysts [12]. Further studies on the cellular localization of these newly-identified stage-regulated CPs are required to determine whether *E. invadens* CPs may be involved in encystation, as demonstrated in *Giardia lamblia*, where cysteine protease 2 was co-transported with cyst wall

protein in encystation-specific vesicles and plays a central role during encystation [17].

In addition, the EhCP-B9/EhCP112 homolog in *E. invadens* [29] was shown to be accumulated near the cyst wall of immature cysts and further evenly distributed in the cytosol of mature cysts. Transcription of EiCP-B9 also increased 126 fold at 24 h of encystation, which coincide the initiation of the cyst wall formation [18].

K: Heat shock proteins

Microarray analysis of heat shock induced *E. histolytica* showed up-regulation of Gal/GalNAc lectin, cysteine proteases, and heat shock proteins such as Hsp70 (EHI_197860, EHI_199590) and Hsp90 (EHI_102270, EHI_163480) [66]. Exposure of *E. invadens* trophozoites to similar conditions also increased the mRNA expression of *BiP* gene (GenBank AAF64243.1) and was suggested to be partially linked with encystation based on an increased expression of *Jacob* and *chitinase* genes in *E. invadens*, although heat shock per se did not result in the formation of the cyst wall [67]. The closest homolog of this *E. invadens BiP* gene on our array is luminal binding protein 4 precursor (EIN_105260, 58% identity), which also contains the ER-retention signal motif (KDEL) required for its proper targeting to the endoplasmic reticulum [67,68]. However, the mRNA level of EIN_105260 was unchanged at 0-24 h and only slightly increased (1.4 fold) at 120 h of encystation (Table S1). Our transcriptome data did not support the premise that the expression of *BiP* and *chitinase* genes is linked or coincides in *E. invadens*. Up-regulation of *chitinase 1* gene expression peaked at 24 h (78-fold up-regulated). Expression of *chitinase 2* genes increased by 7 fold at as early as 8 h, and peaked at 24 h (26 fold up-regulation) of encystation. In contrast, *chitinase 3* gene was constitutively expressed (Figure 4).

Conclusions

Our transcriptomic analysis of *E. invadens* revealed global changes of gene expression during encystation, and should help to identify key regulatory genes that are essential for the process. Further studies on individual genes and their encoded products that are modulated during encystation may lead to the discovery of targets for the development of new chemotherapeutics that interfere with stage conversion of the parasite.

Supporting Information

Figure S1. Correlation between two biological replicates. The correlation levels of transcripts in DNA microarray analysis between first and second biological replicates at different time points during encystation is shown. The Pearson correlation coefficients were calculated using Excel (2011) workbook.
(TIF)

Table S1. List of all probe sets representing *E. invadens* open reading frames. The *E. invadens* probeset ID, fold change and regulation relative to time 0 h (trophozoite stage), average hybridization signal intensities (raw expression data) of five arrays at each time point, normalized $\log_2$ transformed value, common names, gene ID and predicted GO function are shown.
(XLSX)

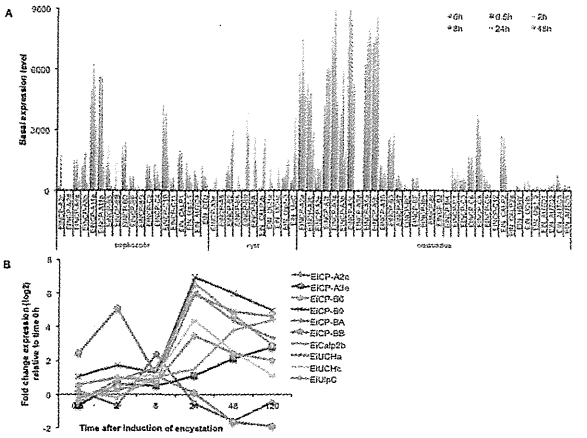

Figure 6. Modulation of the transcript level of the *E. invadens* cysteine proteases during encystation (0.5-120 h). (**A**) Sixty four CPs were grouped into trophozoite-, cyst-predominant, and constitutively expressed CPs based on the transcriptome profiles. **B**) Line graphs showing the fold change expression ($\log_2$) relative to time 0h of ten EiCPs whose expression were significantly modulated during encystation. Gene IDs: EiCP-A2c, EIN_168460; EiCP-A3e, EIN_192250; EiCP-B6, EIN_292720; EiCP-B9, EIN_152250; EiCP-BA, EIN_184830; EiCP-BB, EIN_199850; EiCalp2b, EIN_187000; EiUCHa, EIN_243050; EiUCHc, EIN_107760; EiUlpC, EIN_200450.
doi: 10.1371/journal.pone.0074840.g006

Table S2. Normalized transcriptome data of genes present at least one time points during encystation. The *E. invadens* probe set ID, p-value and corrected p-value of ANOVA, post-hoc test, fold change and regulation relative to time 0 h (trophozoite stage), normalized expression levels in $\log_2$ scale, Pathema/AmoebaDB gene ID, common names, predicted GO function of two biological replicate are shown.
(XLSX)

Table S3. List of genes which were up-regulated ≥ 3 fold at one or more time points during encystation. The *E. invadens* probeset ID, fold change and regulation relative to time 0 h (trophozoite stage), average hybridization signal intensities (raw expression data) of five arrays at each time point, normalized $\log_2$ transformed value, common names, gene id and predicted GO function are shown.
(XLSX)

Table S4. List of genes which were up-regulated ≥ 3 fold at 0.5 and 2 h of encystation. The *E. invadens* probeset ID, fold change and regulation relative to time 0 h (trophozoite stage), average hybridization signal intensities (raw expression data) of five arrays at each time point, normalized $\log_2$ transformed value, common names, gene ID and predicted GO function are shown.
(XLSX)

Table S5. List of genes which were up-regulated ≥ 3 at 8 h of encystation. The *E. invadens* probeset ID, fold change and regulation relative to time 0h (trophozoite stage), average hybridization signal intensities (raw expression data) of five arrays at each time point, normalized $\log_2$ transformed value, common names, gene ID and predicted GO function are shown.
(XLSX)

Table S6. List of genes which were up-regulated ≥ 3 fold at 24 h of encystation. The *E. invadens* probeset ID, fold change and regulation relative to time 0 h (trophozoite stage), average hybridization signal intensities (raw expression data) of five arrays at each time point, normalized $\log_2$ transformed value, common names, gene ID and predicted GO function are shown.
(XLSX)

Table S7. List of genes which were up-regulated ≥ 3 fold at 48 and 120h of encystation. The *E. invadens* probeset ID, fold change and regulation relative to time 0h (trophozoite stage), average hybridization signal intensities (raw expression data) of five arrays at each time point, normalized $\log_2$ transformed value, common names, gene ID and predicted GO function are shown.
(XLSX)

Table S8. List of *E. invadens* transmembrane kinases genes induced ≥ 3 fold at one or more time points during encystation. The *E. invadens* probeset ID, annotation, closest homolog in *E. histolytica* TMKs (number and group), expression in previous studies [12,20,43] predicted number of transmembrane domains (TMHMM) and conserved sequence motifs [43] are shown.
(XLSX)

Author Contributions

Conceived and designed the experiments: AE GJ KN TN. Performed the experiments: AE GJ. Analyzed the data: AE GJ KN TN. Contributed reagents/materials/analysis tools: AE KN EC TN. Wrote the manuscript: AE GJ TN.

References

1. Stanley SL (2003) Amebiasis. Lancet 361: 1025-1034. doi:10.1016/S0140-6736(03)12830-9. PubMed: 12660071.
2. Nozaki T, Kobayashi S, Takeuchi T, Haghighi A (2005) The diversity of clinical isolates of *Entamoeba histolytica* in Japan. Arch Med Res 37: 276-278.
3. Clark CG, Alsmark UC, Tazreiter M, Saito-Nakano Y, Ali V et al. (2007) Structure and content of the *Entamoeba histolytica* genome. Adv Parasitol 65: 51–190. doi:10.1016/S0065-308X(07)65002-7. PubMed: 18063096.
4. Eichinger D (1997) Encystation of *Entamoeba* parasites. Bioessays 19: 633-638. doi:10.1002/bies.950190714. PubMed: 9230696.
5. Donaldson M, Heyneman D, Dempster R, Garcia L (1975) Epizootic of fatal amebiasis among exhibited snakes: epidemiologic, pathologic, and chemotherapeutic considerations. Am J Vet Res 36: 807-817. PubMed: 1147335.
6. Kojimoto A, Uchida K, Horii Y, Okumura S, Yamaguch R et al. (2001) Amebiasis in four ball pythons, Python reginus. J Vet Med Sci 63: 1365-1368. doi:10.1292/jvms.63.1365. PubMed: 11789622.
7. Chia MY, Jeng CR, Hsiao SH, Lee AH, Chen CY et al. (2009) *Entamoeba invadens* myositis in a common water monitor lizard (Varanus salvator). Vet Pathol 46: 673–676. doi:10.1354/vp.08-VP-0224-P-CR. PubMed: 19276058.
8. Sanchez L, Enea V, Eichinger D (1994) Identification of a developmentally regulated transcript expressed during encystation of *Entamoeba invadens*. Mol Biochem Parasitol 67: 125–135. doi:10.1016/0166-6851(94)90102-3. PubMed: 7838173.
9. Gonzalez J, Bai G, Frevert U, Corey EJ, Eichinger D (1999) Proteasome-dependent cyst formation and stage-specific ubiquitin mRNA accumulation in *Entamoeba invadens*. Eur J Biochem 264: 897-904. doi:10.1046/j.1432-1327.1999.00682.x. PubMed: 10491138.
10. Coppi A, Eichinger D (1999) Regulation of *Entamoeba invadens* encystation and gene expression with galactose and N-acetylglucosamine. Mol Biochem Parasitol 102: 67-77. doi:10.1016/S0166-6851(99)00085-7. PubMed: 10477177.
11. Eichinger D (2001) Encystation in parasitic protozoa. Curr Opin Microbiol 4: 421-426. doi:10.1016/S1369-5274(00)00229-0. PubMed: 11495805.
12. Ehrenkaufer GM, Haque R, Hackney JA, Eichinger DJ, Singh U (2007) Identification of developmentally regulated genes in *Entamoeba histolytica*: insights into mechanisms of stage conversion in a protozoan parasite. Cell Microbiol 9: 1426–1444. doi:10.1111/j.1462-5822.2006.00882.x. PubMed: 17250591.
13. Makioka A, Kumagai M, Ohtomo H, Kobayashi S, Takeuchi T (2000) *Entamoeba invadens*: Protein kinase C inhibitors block the growth and encystation. Exp Parasitol 95: 288-290. doi:10.1006/expr.2000.4538. PubMed: 11038313.
14. Makioka A, Kumagai M, Ohtomo H, Kobayashi S, Takeuchi T (2001) Inhibition of encystation of *Entamoeba invadens* by wortmannin. Parasitol Res 87: 371-375. doi:10.1007/s004360000339. PubMed: 11403379.
15. Byers J, Faigle W, Eichinger D (2005) Colonic short-chain fatty acids inhibit encystation of *Entamoeba invadens*. Cell Microbiol 7: 269-279. PubMed: 15659070.
16. Bulik DA, va Ophem P, Manning JM, Shen Z, Newburg DS et al. (2000) UDP-N-acetylglucosamine pyrophosphorylase, a key enzyme in encysting *Giardia*, is allosterically regulated. J Biol Chem 275: 14722-14728. PubMed: 10799561.
17. DuBois KN, Abodeely M, Sakanari J, Craik CS, Lee M et al. (2008) Identification of the major cysteine protease of *Giardia* and its role in encystation. J Biol Chem 283: 18024–18031. doi:10.1074/jbc.M802133200. PubMed: 18445589.
18. Ebert F, Bachmann A, Nakada-Tsukui K, Hennings I, Drescher B et al. (2008) An *Entamoeba* cysteine peptidase specifically expressed during encystation. Parasitol Int 57: 521–524. PubMed: 18723116.
19. Kriebel PW, Parent CA (2004) Adenylyl cyclase expression and regulation during the differentiation of *Dictyostelium discoideum*. IUBMB Life 56: 541–546. doi:10.1080/15216540400013887. PubMed: 15590560.
20. Gilchrist CA, Houpt E, Trapaidze N, Fei Z, Crasta O et al. (2006) Impact of intestinal colonization and invasion on the *Entamoeba histolytica* transcriptome. Mol Biochem Parasitol 147: 163–176. doi:10.1016/j.molbiopara.2006.02.007. PubMed: 16569449.
21. MacFarlane RC, Singh U (2006) Identification of differentially expressed genes in virulent and nonvirulent *Entamoeba* species: potential implications for amebic pathogenesis. Infect Immun 74: 340-351. doi:10.1128/IAI.74.1.340-351.2006. PubMed: 16368989.
22. Vicente JB, Ehrenkaufer GM, Saraiva LM, Teixeira M, Singh U (2009) *Entamoeba histolytica* modulates a complex repertoire of novel genes

in response to oxidative and nitrosative stresses: implications for amebic pathogenesis. Cell Microbiol 11: 51-69. doi:10.1111/j.1462-5822.2008.01236.x. PubMed: 18778413.
23. Ehrenkaufer GM, Hackney JA, Singh U (2009) A developmentally regulated Myb domain protein regulates expression of a subset of stage-specific genes in *Entamoeba histolytica*. Cell Microbiol 11: 898-910. doi:10.1111/j.1462-5822.2009.01300.x. PubMed: 19239479.
24. González-Salazar F, Viader-Salvadó JM, Martınez-Rodrıguez HG, Campos-Góngora E, Mata-Cárdenas BD et al. (2000) Identification of seven chemical factors that favor high-quality *Entamoeba histolytica* cyst-like structure formation under axenic conditions. Arch Med Res 31 (Suppl): 192–193. doi:10.1016/S0188-4409(00)00193-4. PubMed: 11070279.
25. Barrón-González MP, Villarreal-Treviño L, Verduzco-Martínez JA, Mata-Cárdenas BD, Morales-Vallarta MR (2005) *Entamoeba invadens*: in vitro axenic encystation with a serum substitute. Exp Parasitol 110: 318–321. doi:10.1016/j.exppara.2005.03.021. PubMed: 15955331.
26. Tillack M, Biller L, Irmer H, Freitas M, Gomes MA et al. (2007) The *Entamoeba histolytica* genome: primary structure and expression of proteolytic enzymes. BMC Genomics 8: 170. doi: 10.1186/1471-2164-8-170. PubMed: 17567921.
27. Kumar S, Nei M, Dudley J, Tamura K (2008) MEGA: A biologist-centric software for evolutionary analysis of DNA and protein sequences. Brief Bioinform 9: 299-306. doi:10.1093/bib/bbn017. PubMed: 18417537.
28. Nakada-Tsukui K, Saito-Nakano Y, Husain A, Nozaki T (2010) Conservation and function of Rab small GTPases in *Entamoeba*: annotation of E. invadens Rab and its use for the understanding of *Entamoeba* biology. Exp Parasitol 126: 337-347. doi:10.1016/j.exppara.2010.04.014. PubMed: 20434444.
29. Nakada-Tsukui K, Nozaki T (2010) Genomic and post-genomic approaches to understand the pathogenesis of the enteric protozoan parasite *Entamoeba histolytica*. In: P FratamicoY LiuS Kathariou. Genomes of Food and Waterborne Pathogens. Washington, DC: ASM Press. pp 321-341.
30. Wang Z, Samuelson J, Clark CG, Eichinger D, Paul J et al. (2003) Gene discovery in the *Entamoeba invadens* genome. Mol Biochem Parasitol 129: 23–31. doi:10.1016/S0166-6851(03)00073-2. PubMed: 12798503.
31. Lorenzi HA, Puiu D, Miller JR, Brinkac LM, Amedeo P et al. (2010) New assembly, reannotation and analysis of the *Entamoeba histolytica* genome reveal new genomic features and protein content information. PLOS Negl Trop Dis 4(6): e716. doi:10.1371/journal.pntd.0000716. PubMed: 20559563.
32. Ikegami A, Honma K, Sharma A, Kuramitsu HK (2004) Multiple functions of the leucine-rich repeat protein LrrA of *Treponema denticola*. Infect Immun 72: 4619–4627. doi:10.1128/IAI.72.8.4619-4627.2004. PubMed: 15271922.
33. Inagaki S, Onishi S, Kuramitsu HK, Sharma A (2006) *Porphyromonas gingivalis* vesicles enhance attachment, and the leucine-rich repeat BspA protein is required for invasion of epithelial cells by "Tannerella forsythiae". Infect Immun 74: 5023-5028. doi:10.1128/IAI.00062-06. PubMed: 16926393.
34. Davis PH, Zhang Z, Chen M, Zhang X, Chakraborty S et al. (2006) Identification of a family of BspA like surface proteins of *Entamoeba histolytica* with novel leucine rich repeats. Mol Biochem Parasitol 145: 111–116. doi:10.1016/j.molbiopara.2005.08.017. PubMed: 16199101.
35. Noël CJ, Diaz N, Sicheritz-Ponten T, Safarikova L, Tachezy J et al. (2010) *Trichomonas vaginalis* vast BspA-like gene family: evidence for functional diversity from structural organisation and transcriptomics. BMC Genomics 11: 99. doi:10.1186/1471-2164-11-99. PubMed: 20144183.
36. McGugan GC, Temesvari LA (2003) Characterization of a Rab11-like GTPase, EhRab11 of *Entamoeba histolytica*. Mol Biochem Parasitol 129: 137–146. doi:10.1016/S0166-6851(03)00115-4. PubMed: 12850258.
37. Hon CC, Nakada-Tsukui K, Nozaki T, Guillen N (2010) Dissecting the actin skeleton from a genomic prespective. In: CG ClarkPJ JohnsonRD Adam. Anaerobic Parasitic Protozoa Genomics and Molecular Biology. Norfolk, UK: Caister Academic Press. pp 81-155.
38. Makioka A, Kumagai M, Ohtomo H, Kobayashi S, Takeuchi T (2000) Effect of cytochalasin D on the growth, encystation, and multinucleation of *Entamoeba invadens*. Parasitol Res 86: 599-602. doi:10.1007/PL00008536. PubMed: 10935912.
39. Picazarri K, Nakada-Tsukui K, Nozaki T (2008) Autophagy during proliferation and encystation in the protozoan parasite *Entamoeba invadens*. Infect Immun 76: 278–288. doi:10.1128/IAI.00636-07. PubMed: 17923513.
40. Makioka A, Kumagai M, Kobayashi S, Takeuchi T (2003) Involvement of signaling through protein kinase C and phosphatidylinositol 3-kinase in the excystation and metacystic development of *Entamoeba invadens*. Parasitol Res 91: 204–208. doi:10.1007/s00436-003-0955-x. PubMed: 12923632.
41. Nakada-Tsukui K, Okada H, Mitra BN, 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. doi:10.1111/j.1462-5822.2009.01341.x. PubMed: 19496789.
42. Boettner DR, Huston CD, Linford AS, Buss SN, Houpt E (2008) *Entamoeba histolytica* phagocytosis of human erythrocytes involves PATMK, a member of the transmembrane kinase family. PLOS Pathog 4(1): e8. doi:10.1371/journal.ppat.0040008. PubMed: 18208324.
43. Buss SN, Hamano S, Vidrich A, Evans C, Zhang Y et al. (2010) Members of the *Entamoeba histolytica* transmembrane kinase family play non-redundant roles in growth and phagocytosis. Int J Parasitol 40: 833-843. doi:10.1016/j.ijpara.2009.12.007. PubMed: 20083116.
44. Beck DL, Boettner DR, Dragulev B, Ready K, Nozaki T et al. (2005) Identification and gene expression analysis of a large family of transmembrane kinases related to the Gal/GalNAc lectin in *Entamoeba histolytica*. Eukaryot Cell 4: 722-732. doi:10.1128/EC.4.4.722-732.2005. PubMed: 15821132.
45. Shiozaki K, Russell P (1994) Cellular function of protein phosphatase 2C in yeast. Cell Mol Biol Res 40: 241-243. PubMed: 7874201.
46. Jeelani G, Sato D, Husain A, Escueta-de Cadiz A, Sugimoto M, Soga T, Suematsu M, Nozaki T (2012) Metabolic profiling of the protozoan parasite *Entamoeba invadens* revealed activation of unpredicted pathway during encystation. PLOS ONE 7: e37740. doi:10.1371/journal.pone.0037740. PubMed: 22662204.
47. Das S, Van Dellen K, Bulik D, Magnelli P, Cui J et al. (2006) The cyst wall of *Entamoeba invadens* contains chitosan (deacetylated chitin). Mol Biochem Parasitol 148: 86–92. PubMed: 16621070.
48. Lopez AB, Sener K, Jarroll EL, Van Keulen H (2003) Transcription regulation is demonstrated for five key enzymes in Giardia intestinalis cyst wall polysaccharide biosynthesis. Mol Biochem Parasitol 128: 51-57. doi:10.1016/S0166-6851(03)00049-5. PubMed: 12706796.
49. Gallego E, Alvarado M, Wasserman M (2007) Identification and expression of the protein ubiquitination system in *Giardia intestinalis*. Parasitol Res 101: 1-7. doi:10.1007/s00436-007-0605-9. PubMed: 17252268.
50. Ulrich HD (2005) Mutual interactions between the SUMO and ubiquitin systems: A plea of no contest. Trends Cell Biol 15: 525–532. doi: 10.1016/j.tcb.2005.08.002. PubMed: 16125934.
51. Pao SS, Paulsen IT, Saier MH Jr (1998) Major Facilitator Superfamily. Microbiol Mol Biol Rev 62: 1-34. PubMed: 9529885.
52. Campos-Góngora E, Viader-Salvadó JM, Martínez-Rodríguez HM, Zuñiga-Charles MA (2000) Mg, Mn, and Co Ions enhance the formation of *Entamoeba histolytica* cyst-Like structures resistant to sodium dodecyl sulfate. Arch Med Res 31: 162–168. doi:10.1016/S0188-4409(00)00054-0. PubMed: 10880721.
53. González-Salazar F, Viader-Salvadó JM, Martínez-Rodríguez HG, Campos-Góngora E, Mata-Cárdenas BD et al. (2000) Identification of seven chemical factors that favor high-quality *Entamoeba histolytica* cyst-like structure formation under axenic conditions. Arch Med Res 31: S192-S193. doi:10.1016/S0188-4409(00)00193-4. PubMed: 11070279.
54. Das S, Gillin FD (1991) Chitin synthase in encysting *Entamoeba invadens*. Biochem J 20: 641-647. PubMed: 1764027.
55. Said-Fernández S, Campos-Góngora E, González-Salazar F, Martínez-Rodríguez HG, Vargas-Villarreal J (2001) Mg^{2+}, Mn^{2+}, and Co^{2+} stimulate *Entamoeba histolytica* to produce chitin-like material. J Parasitol 87: 919-923. doi: 10.1645/0022-3395(2001)087[0919:MMACSE]2.0.CO;2. PubMed: 11534662.
56. Abeijon C, Mandon EC, Hirschberg CB (1997) Transporters of nucleotide sugars, nucleotide sulfate and ATP in the Golgi apparatus. Trends Biochem Sci 22: 203-207. doi:10.1016/S0968-0004(97)01053-0. PubMed: 9204706.
57. Biller L, Davis PH, Tillack M, Matthiesen J, Lotter H, Stanley SL, Tannich E, Bruchhaus I (2010) Differences in the transcriptome signatures of two genetically related *Entamoeba histolytica* cell lines derived from the same isolate with different pathogenic properties. BMC Genomics 11: 63. doi:10.1186/1471-2164-11-63. PubMed: 20102605.
58. Van Dellen KL, Chatterjee A, Ratner DM, Magnelli PE, Cipollo JF et al. (2006) Unique posttranslational modifications of chitin-binding lectins of *Entamoeba invadens* cyst walls. Eukaryot Cell 5: 836-848. doi: 10.1128/EC.5.5.836-848.2006. PubMed: 16682461.

59. Chatterjee A, Ghosh SK, Jang K, Bullitt E, Moore L et al. (2009) Evidence for a "wattle and daub" model of the cyst wall of *Entamoeba*. PLOS Pathog 5(7): e1000498.
60. Rosinski JA, Atchley WR (1998) Molecular evolution of the Myb family of transcription factors: evidence for polyphyletic origin. J Mol Evol 46: 74–83. doi:10.1007/PL00006285. PubMed: 9419227.
61. Meneses E, Cárdenas H, Zárate S, Brieba LG, Orozco E et al. (2010) The R2R3 Myb protein family in *Entamoeba histolytica*. Gene 455: 32-42. doi:10.1016/j.gene.2010.02.004. PubMed: 20156532.
62. Que X, Reed SL (2000) Cysteine proteinases and the pathogenesis of amebiasis. Clin Microbiol Rev 13: 196-206. doi:10.1128/CMR.13.2.196-206.2000. PubMed: 10755997.
63. Moncada D, Keller K, Chadee K (2003) *Entamoeba histolytica* cysteine proteinases disrupt the polymeric structure of colonic mucin and alter its protective function. Infect Immun 71: 838–844. doi:10.1128/IAI.71.2.838-844.2003. PubMed: 12540564.
64. Bruchhaus I, Loftus BJ, Hall N, Tannich E (2003) The intestinal protozoan parasite *Entamoeba histolytica* contains 20 cysteine protease genes, of which only a small subset is expressed during in vitro cultivation. Eukaryot Cell 2: 501–509. doi:10.1128/EC.2.3.501-509.2003. PubMed: 12796295.
65. He C, Nora GP, Schneider EL, Kerr ID, Hansell K et al. (2010) A novel *Entamoeba histolytica* cysteine proteinase, EhCP4, is key for invasive amebiasis and a therapeutic target. J Biol Chem 285: 18516–18527.
66. Weber C, Guigon G, Bouchier C, Frangeul L, Moreira S et al. (2006) Stress by heat shock induces massive down regulation of genes and allows differential allelic expression of the Gal/GalNAc lectin in *Entamoeba histolytica*. Eukaryot Cell 5: 871-875. doi:10.1128/EC.5.5.871-875.2006. PubMed: 16682464.
67. Field J, Van Dellen K, Ghosh SK, Samuelson J (2000) Responses of *Entamoeba invadens* to heat shock and encystation are related. J Eukaryot Microbiol 47: 511-514. doi:10.1111/j.1550-7408.2000.tb00083.x. PubMed: 11001149.
68. Ghosh SK, Field J, Frisardi M, Rosenthal B, Mai Z et al. (1999) Chitinase secretion by encysting *Entamoeba invadens* and transfected *Entamoeba histolytica* trophozoites: localization of secretory vesicles, endoplasmic reticulum and golgi apparatus. Infect Immun 67: 3073-3081. PubMed: 10338523.

Author's Choice

The Cell Surface Proteome of *Entamoeba histolytica*

Laura Biller‡‡‡, Jenny Matthiesen‡‡‡, Vera Kühne‡, Hannelore Lotter‡, Ghassan Handal‡, Tomoyoshi Nozaki§, Yumiko Saito-Nakano§, Michael Schümann¶, Thomas Roeder‖, Egbert Tannich‡, Eberhard Krause¶, and Iris Bruchhaus‡**

Surface molecules are of major importance for host-parasite interactions. During *Entamoeba histolytica* infections, these interactions are predicted to be of prime importance for tissue invasion, induction of colitis and liver abscess formation. To date, however, little is known about the molecules involved in these processes, with only about 20 proteins or protein families found exposed on the *E. histolytica* surface. We have therefore analyzed the complete surface proteome of *E. histolytica*. Using cell surface biotinylation and mass spectrometry, 693 putative surface-associated proteins were identified. *In silico* analysis predicted that ~26% of these proteins are membrane-associated, as they contain transmembrane domains and/or signal sequences, as well as sites of palmitoylation, myristoylation, or prenylation. An additional 25% of the identified proteins likely represent non-classical secreted proteins. Surprisingly, no membrane-association sites could be predicted for the remaining 49% of the identified proteins. To verify surface localization, 23 proteins were randomly selected and analyzed by immunofluorescence microscopy. Of these 23 proteins, 20 (87%) showed definite surface localization. These findings indicate that a far greater number of *E. histolytica* proteins than previously supposed are surface-associated, a phenomenon that may be based on the high membrane turnover of *E. histolytica*. *Molecular & Cellular Proteomics 13: 10.1074/mcp.M113.031393, 132–144, 2014.*

The intestinal protozoan *Entamoeba histolytica* is an important human parasite. Its life cycle is relatively simple, consisting of infectious cysts that can survive outside the host and vegetative trophozoites that proliferate in the human gut. After infection, *E. histolytica* trophozoites are normally present in the intestine where they asymptomatically persist for months in the lumen. *E. histolytica* can become a pathogen by penetrating the intestinal mucosa and inducing colitis, or by disseminating to other organs, most commonly to the liver, where it induces abscess formation.

The factors that determine the clinical outcomes of *E. histolytica* infections have not been well defined. Decisive factors may include genetic aspects of the host and/or parasite, the type of immune response mounted by the host, the presence of concomitant infections, and host diet. *E. histolytica* surface proteins are regarded to be of prime importance for host-parasite interactions. Members of the galactose/N-acetyl D-galactosamine-inhibitable (Gal/GalNAc) lectin family exposed on the surface of the parasite are considered important for adherence to target cells (1, 2), with adherence necessary for killing and/or phagocytosis. In addition to their involvement in adhesion and phagocytosis, the surface molecules of *E. histolytica* are exposed to the host's immune system. To date, only about 20 proteins or protein families have been identified as exposed on the plasma membrane of the parasite. These proteins include EhADH112 and the cysteine peptidase EhCP112 (EhCP-B9), which form a 112 kDa adhesion protein (3, 4); the serine-rich *E. histolytica* protein (SREHP) (5); a calreticulin (6); an as-yet unidentified mannose binding lectin (7); transmembrane kinases, including phagosome-associated TMK96 (PATMK), TMK39 and TMK54 (8, 9); a family of Bsp-A-like molecules (10); a rhomboid protease (11); an EhRab7 molecule (12); an actinin-like protein (AAF20148) (13); the lysine (K) and glutamic acid (E) enriched proteins KERP-1 and KERP-2 (13); the cysteine peptidases EhCP-A2 and EhCP-A5 (14, 15); a peroxiredoxin (29 kDa thiol-dependent peroxidase) (16); the transcription factor URE3-BP, which localizes to the cytoplasm and inner surface of the plasma membrane (17); the ARIEL antigen (18, 19); a LIM protein (EhLimA) associated with lipid rafts in the plasma membrane (20); the M8 family surface metalloprotease (EhMSP-1) (21); alcohol dehydrogenase 3 (22); and syntaxin 1 and SNAP-25 (23).

To identify the complete set of membrane proteins that are thought to be of prime importance for host-parasite interactions, we chose a combined approach, consisting of an *in silico* analysis of predictable membrane association and bio-

From the ‡Bernhard Nocht Institute for Tropical Medicine, Bernhard-Nocht-Str. 74, 20359 Hamburg, Germany; §Department of Parasitology, National Institute of Infectious Diseases, 1-23-1 Toyama, Shinjuku-ku, Tokyo 162-8640, Japan; ¶Leibniz Institute for Molecular Pharmacology, Robert-Rössle-Strasse 10, 13125 Berlin, Germany; ‖Zoological Institute, Christian-Albrechts-University, Olshausenstraße 40, 24098 Kiel, Germany

Author's Choice—Final version full access.
Received May 29, 2013, and in revised form, September 24, 2013
Published, MCP Papers in Press, October 17, 2013, DOI 10.1074/mcp.M113031393

tinylation of surface proteins followed by mass spectrometric analysis. *In silico* analysis of the 8306 predicted *E. histolytica* proteins (AmoebaDB, version 1.7, http://amoebadb.org/amoeba/), using a program that predicts transmembrane protein topology, identified 1326 proteins with one or more transmembrane domains and 1079 proteins with a signal peptide with an overlap of 561 proteins. These findings support the hypothesis that the vast majority of *Entamoeba* surface proteins await identification.

Biotinylation of surface-exposed proteins using a nonpermeable reagent combined with subsequent purification and sequencing is a well-established method to characterize the cell surface proteome of various cell types and organisms (24–26). So far, the surface proteome of only one protozoan, *Trichomonas vaginalis,* was analyzed by de Miguel and colleagues using the cell surface biotinylation approach. They identified more than 400 proteins to be associated with the surface (27). Here, the surface-exposed proteins of *E. histolytica* were biotinylated, purified by affinity chromatography on Avidin agarose resin and analyzed by SDS-PAGE followed by liquid-chromatography mass spectrometry (LC-MS/MS). Of the total of 693 identified proteins, ~50% showed no predictable membrane association. Nevertheless, localization analysis indicated that about 85% of these proteins were found on the plasma membrane surface.

EXPERIMENTAL PROCEDURES

E. histolytica Cell Culture—*E. histolytica* trophozoites of strain HM-1:IMSS, obtained in 2001 from the American Type Culture Collection (ATCC; Manassas, VA, USA; Catalogue No 30459), were cultured axenically in TYI-S-33 medium supplemented with 10% adult bovine serum (ABS) at 36 °C in plastic tissue culture flasks (28).

Biotinylation of Cell Surface Molecules and Purification of Biotinylated Proteins—The experiments were performed three times independently. Cell surface proteins were biotinylated with Thermo Scientific EZ-Link Sulfo-NHS-SS-Biotin (Pierce Biotechnology, Rockford, IL) according to the manufacturer's instructions with some modifications. Briefly, trophozoites were grown in six T75 culture flasks until cells reached 90–95% confluence ($\leq 4 \times 10^7$ trophozoites/flask). The culture medium was removed and the trophozoites were washed carefully with 8 ml phosphate-buffered saline (PBS)[1] (6.7 mM $NaHPO_4$, 3.3 mM NaH_2PO_4, 140 mM NaCl, pH 7.2) at room temperature. One aliquot of EZ-Link Sulfo-NHS-SS-Biotin (12 mg) was dissolved in 48 ml of ice-cold PBS, and 10 ml of biotin solution was added to each flask. The solutions were incubated for 30 min at 4 °C using an orbital shaker, and 500 μl of quenching solution was added to each flask. The cells from each flask were transferred to a 50 ml reaction tube and centrifuged (400 × *g*, 3 min, 4 °C), and the trophozoites were resuspended in 5 ml TBS, centrifuged again, and lysed by resuspension in 500 μl lysis buffer containing 10 μM cysteine peptidase inhibitor E-64 and 0.5% SDS. Before lysis of the cells the viability of the labeled trophozoites was quantified by trypan blue exclusion. The viability has to be >98% to proceed in the experiments. After sonication of the cell lysate for 10 s, samples were incubated for 30 min on ice, vortexed every 5 min and sonicated twice more every 10 min. The lysed cells were centrifuged at 10000 × *g* for 2 min at 4 °C and the supernatants were transferred to new tubes. The biotin-labeled proteins were isolated using NeutrAvidin Agarose and eluted as described by the manufacturer (Pierce Cell surface Protein Isolation Kit, Instructions, Thermo Scientific). Summarized, the cell lysate was incubated in a column containing NeutrAvidin Agarose for 60 min at RT using an end-over-end rotator. Before elution the agarose was washed four times using Wash Buffer. Next 400 μl of SDS-PAGE sample buffer, containing 50 mM dithiothreitol, was added to the column, incubated for 60 min at room temperature with an end-over-end rotator. Afterward the proteins were eluted by centrifugation.

As a negative control, trophozoites were treated in the same manner as described, without the addition of biotin.

One-dimensional SDS-PAGE and In-gel Digestion—The eluted protein fractions were separated on 12% SDS-PAGE gels, which were stained with Coomassie brilliant blue R-250. Each lane was cut into 36–40 slices, and each slice was washed with 50% (v/v) acetonitrile in 50 mM ammonium bicarbonate, shrunk by dehydration in acetonitrile and dried in a vacuum centrifuge. Each gel piece was reswollen in 10 μl of 50 mM ammonium bicarbonate containing 50 ng trypsin (sequencing grade modified, Promega). After incubation for 17 h at 37 °C, the enzymatic reaction was terminated by addition of 10 μl of 0.5% (v/v) trifluoroacetic acid in acetonitrile, and the separated liquid was dried under vacuum. Each sample was reconstituted in 6 μl of 0.1% (v/v) trifluoroacetic acid and 5% (v/v) acetonitrile in water.

Mass Spectrometric Identification of Proteins—NanoLC-MS/MS experiments were performed using a reversed-phase capillary liquid chromatography system (Eksigent 2D nanoflow LC; Axel Semrau GmbH, Germany) connected to an LTQ-Orbitrap XL mass spectrometer (Thermo Scientific, Bremen, Germany). Peptides were separated on a capillary column (Atlantis dC18, 3 μm, 100 Å, 150 mm × 75 μm i.d., Waters) at an eluent flow rate of 250 nL/min using a linear gradient of 2–50% B in 60 min. Mobile phase A was 0.1% formic acid (v/v) in water; mobile phase B was 0.1% formic acid in acetonitrile. Mass spectra were acquired in a data-dependent mode with one MS survey scan in the Orbitrap and MS/MS scans of the five most intense precursor ions in the LTQ. The MS survey range was *m/z* 350–1500. The dynamic exclusion time (for precursor ions) was set at 120 s, and automatic gain control was set at 3×10^6 and 10,000 for Orbitrap-MS and LTQ-MS/MS scans, respectively.

For database searching, raw data files were processed with the Mascot Distiller (version 2.2.1; Matrix Science) to generate Mascot generic files (mgf files). The MASCOT server (version 2.2; Matrix Science, London, UK) was used to search in-house against the NCBI nonredundant database (NCBInr, date 02/02/2009, 7787617 sequences, 2685418921 residues). A maximum of two missed cleavages was allowed, and the mass tolerances of precursor and sequence ions were set at 10 ppm and 0.35 Da, respectively. Acrylamide modifications of cysteine and methionine oxidation were considered possible modifications. Scaffold (version 2.01; Proteome Software, Inc., Portland, OR) was used to validate MS/MS based peptide and protein identifications. Peptide identifications were accepted if their probability was established at >70.0%, as specified by the Peptide Prophet algorithm. Protein identifications were accepted if their probability was established at >99% and if they contained at least two identified tryptic peptides. If the coverage of the protein is ≥ 5%, one identified tryptic peptide is sufficient.

In silico Analysis—The following programs were used for functional annotations as well as for identification of membrane-association sites or secretion: AmoebaDB, version 1.7, http://amoebadb.org/amoeba/; SignalP 3.0 Server, http://www.cbs.dtu.dk/services/SignalP/; NMT - The MYR Predictor, http://mendel.imp.ac.at/myristate/SUPLpredictor.htm;

[1] The abbreviations used are: PBS, phosphate-buffered saline; TBS, Tris-buffered saline; TFA trifluoacetic acid; IPTG, isopropyl-beta-D-thiogalactoside; GST, glutathione-S-transferase; ECL, enhanced chemiluminescence.

SecretomeP 2.0 Server, http://www.cbs.dtu.dk/services/SecretomeP/; TMHMM Server v. 2.0, http://www.cbs.dtu.dk/services/TMHMM-2.0/; Predotor, a prediction service for identifying putative N-terminal targeting sequences, http://urgi.versailles.inra.fr/predotar/predotar.html; PrePS - Prenylation Prediction Suite, http://mendel.imp.ac.at/sat/PrePS/index.html; Big-PI Prediction, http://mendel.imp.ac.at/gpi/cgi-bin/gpi_pred.cgi.

Isolation of Genomic DNA and RNA and cDNA Synthesis—*E. histolytica* trophozoites (1×10^6) were cultivated in 75 ml culture flasks for 24 h, harvested by chilling on ice for 5 min, and sedimented at $200 \times g$ for 5 min at 4 °C. The cell pellet was washed twice with PBS. Genomic DNA was isolated using the Easy-DNA™ kit (Invitrogen, Life Technologies, Carlsbad, CA) according to the manufacturer's instructions.

Expression and Purification of Recombinant Proteins—Fragments of genes encoding putative surface localized proteins of *E. histolytica* were amplified by PCR. The sequences of the oligonucleotide primers as well as the respective restriction sites allowed the amplified DNA to be cloned in a predicted orientation into the prokaryotic expression plasmid pJC45 (supplemental Table S1), a derivative of pJC40 that also encodes a histidine leader of ten residues attached to the N terminus of the gene product (29). The following plasmids were generated: pJCrEhAct1 (Activator 1 40 kDa subunit, XM_646064.1), pJCrEhC2–1 (C2 domain-containing protein, XM_650207), pJCrEhC2–3 (C2 domain protein, XM_649407), pJCrEhDnaJ (DnaJ family protein, XM_648397), pJCrEhFeHyd (Fe-hydrogenase, XM_647747), pJCrEhGrainin1 (grainin 1, XM_645280), pJCrEhGNBP (guanine nucleotide-binding protein subunit beta 2-like 1, XM_651958), pJCrEhHypP (hypothetical protein, XM_647328), pJCrEhP120 (proliferating-cell nucleolar antigen p120, XM_649901.1), pJCrEhHydA (putative iron hydrogenase HydA, AF262400), pJCrEhRho (Rho family GTPase, XM_649396), pJCrEhRub (rubrerythrin, XM_647039.2), and pJCrEhURE3-BP (URE3-BP sequence specific DNA binding protein, AF291721). After transformation of *E. coli* BL21(DE3) pAPlacI^Q with each recombinant plasmid, the bacteria were transferred into 25 ml LB medium supplemented with 100 μg/ml ampicillin and 50 μl kanamycin, and grown overnight at 37 °C. An aliquot was transferred to 500 ml LB medium containing ampicillin and kanamycin, and grown at 37 °C until OD_{600} reached 0.6. Recombinant expression was induced by addition of isopropyl-beta-D- thiogalactoside (IPTG) at a final concentration of 0.1 mM, and culturing was continued for 3 h. Subsequently, bacteria were sedimented (6000 × *g*, 30 min), and the pellet was resuspended in purification buffer A (10 mM Tris-HCl, 100 mM Na_3PO_4, 4 M guanidinium hydrochloride, pH 8.0), treated by ultrasonification, and centrifuged for 30 min at 10000 × *g*. The supernatant was applied to a Ni^{2+}-NTA-column in binding buffer (8 M urea, 5 mM imidazole, 0.5 M NaCl, 20 mM Tris-HCl, pH 7.9), and eluted with 0.5 M imidazole in the same buffer. PCR-amplified fragments encoding EhRab7D (XM_646823.1), EhRab7E (XM_646110.2), EhRab7G (XM_651385), EhRab7H (XM_648322), or EhRab11A (XM_642856.2) were introduced into BamHI and XhoI restriction enzyme sites of pGEX6P-1 vector (GE Healthcare) to produce N-terminal GST-fusion recombinant proteins. Expression and purification of recombinant proteins of EhRab GTPases were performed according to the manufacturer's instructions.

Generation of Polyclonal Antibodies—Antibodies were generated by injecting 100 μg of recombinant protein in complete Freund's adjuvant into BALB/c mice, followed by two booster injections, at 2 week intervals, with 100 μg of recombinant protein in incomplete Freund's adjuvant. Thirteen antisera were generated: to αAct1, αEhC2–1, αEhC2–3, αFhDnaj, αEhFeHyd, αEhGrainin1, αEhGNBP, αEhHypP, αEhP120, αEhHydA, αEhRho, αEhRub and αURE3-BP. αEhRab7D, αEhRab7E, αEhRab7G, αEhRab7H, and αEhRab11A antibodies were raised in rabbit. Additional antibodies were generated in rabbits against αEhCP-A1 (XP_650156), αEhCP-A2 (XP_650642) (30, 31), αEhCoronin (XP_654419, provided by Frank Ebert), αEhFeSOD (XP_648827) (32), αEhPrx (XP_647907) (33), αEhTR (XP_655748), αEhLectin (XP_655415) (34), whereas an antibody against αEhCP-A5 (XP_650937, (30)) was generated in chickens.

Protein Analyses—To prepare amoebic extracts, trophozoites were washed twice with PBS (6.7 mM $NaHPO_4$, 3.3 mM NaH_2PO_4, 140 mM NaCl, pH 6.8) and sedimented by centrifugation at 400 × *g* for 2 min at 4 °C. To minimize proteolysis, 20 μM trans-epoxysuccinyl-L-leucylamino-(4-guanodino)butane (E64, Sigma-Aldrich) was added. Cells were alternately flash frozen in liquid nitrogen, thawed at room temperature, and vortexed five times. Lysates were centrifuged at 40000 × *g* for 1 h at 4 °C. The supernatants contained PBS-soluble proteins. The pellets were washed twice in ice-cold PBS and solubilized in PBS supplemented with 1% Triton X-100. Extracts (50 μg/lane) were separated under reducing conditions on 12% SDS-PAGE gels. The proteins were transferred to nitrocellulose membranes using the wet blotting technique, with 25 mM Tris-HCl, 192 mM glycine, 1.3 mM SDS, pH 8.3, and 20% methanol as the transfer buffer. For Western blot analyses, primary antibodies were used at 1:500 dilution and secondary antibodies (anti-mouse HRP and anti-rabbit HRP, DAKO A/S; Glostrup, Denmark; anti-chicken IgY HRP, Sigma) at 1:5000 dilution. Blots were developed using ECL (Amersham Biosciences, ECL plus Western blotting detection reagents; GE Healthcare).

Immunofluorescence Assays (IFA)—Freshly harvested and washed amoebae (5×10^5) were fixed in 3% paraformaldehyde in PBS. Cells were permeabilized with PBS containing 0.2% saponin. Free aldehyde groups were blocked by subsequent incubation with 50 mM ammonium chloride in PBS (± saponin), followed by incubation with blocking buffer (PBS supplemented with 2% fetal calf serum) for 10 min. The trophozoites were subsequently incubated with specific antiserum (diluted 1:200 in PBS ± saponin), washed three times with PBS, and finally incubated in the dark with 1:400 dilutions of Alexa Fluor®488 goat anti-mouse or anti-rabbit antibody or Alexa Fluor®598 goat anti-chicken (Invitrogen, Molecular Probes, Eugene, OR, USA). After additional three washes with PBS, the amoebae were analyzed by fluorescence microscopy (Leica, DM BR, Wetzlar, Germany). Confocal fluorescence microscopy was performed using the Olympus IX81 microscope with FluoView Version 1.7b software (Olympus, Hamburg, Germany).

Biotin-surface labeling was validated using a 1:100 dilution of mouse monoclonal anti-biotin antibody (Sigma-Aldrich, München, Germany).

Live immunofluorescence staining was performed as described (22), with slight modifications. Briefly, amoebae (~3×10^5) were sedimented by centrifugation at 400 × *g* for 2 min at 4 °C and resuspended in 2 ml resuspension/blocking buffer, consisting of 50% v/v TYI-S-33 (without serum or antibiotics), 50% v/v PBS with 20% heat-inactivated fetal calf serum (Sigma-Aldrich) and 20 μM E64 for 10 min at 4 °C. The amoebae were again sedimented, resuspended in 400 μl resuspension buffer containing a 1:100 dilution of primary antibody, and incubated for 20 min at 4 °C. The cells were washed three times, incubated with fluorescence-labeled secondary antibody for 20 min at 4 °C, washed three times in resuspension buffer and once in PBS, and fixed in 4% paraformaldehyde for 30 min at room temperature, followed by Hoechst staining of the nuclei.

RESULTS

Preparation of E. histolytica Surface Protein-Enriched Fractions—To identify the surface-associated proteins of *E. histolytica*, live trophozoites were labeled with sulfo-NHS-SS-biotin, a thiol-cleavable amine-reactive biotinylation reagent. To reduce the possibility that this reagent would be internalized

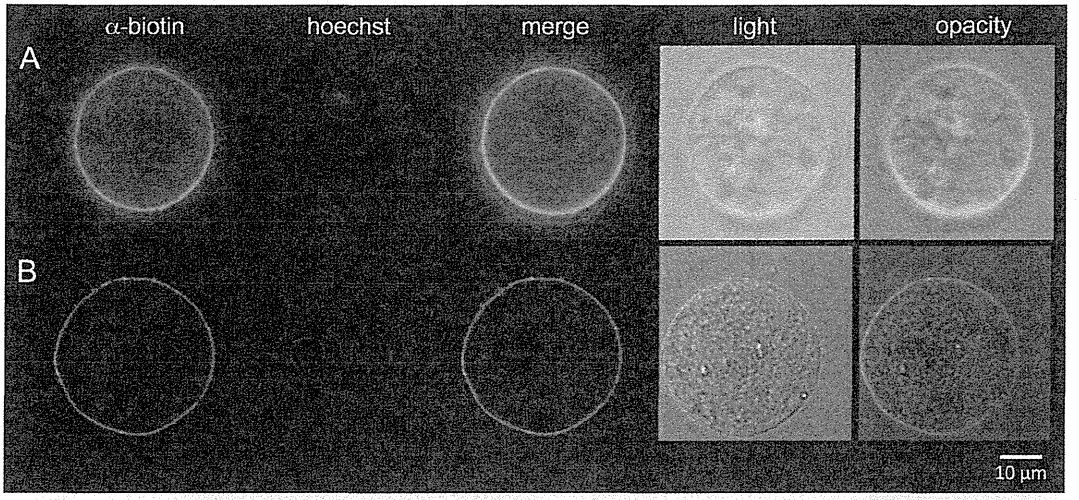

FIG. 1. **Biotinylation of *E. histolytica* surface proteins using sulfo-NHS-SS-biotin.** *E. histolytica* trophozoites were incubated with sulfo-NHS-SS-Biotin, fixed, and treated with saponin. Afterward, the biotin labeling of surface molecules was visualized with a monoclonal anti-biotin antibody and an Alexa488 conjugated secondary antibody, followed by Hoechst staining of nuclei. Fluorescence microscopy was performed using (*A*) a fluorescence microscope (Leica, DM BR, Wetzlar, Germany) or (*B*) a confocal microscope (Olympus IX81 microscope with the FluoView Version 1.7b software; Olympus, Hamburg, Germany).

by phagocytosis, labeling was performed at 4 °C for 30 min. As a control, amoebae were biotinylated and exposed to an anti-biotin antibody for immunofluorescence microscopy. The results clearly showed that biotin labeling was confined to the trophozoite surface (Figs. 1*A*, 1*B*). Nevertheless, confocal microscopy indicated that the surface of the trophozoites is not always uniformly labeled. Instead some bead-like protrusions are visible (Fig. 1*B*). The viability of the amoebae after the labeling procedure was always greater than 98%.

Cells that had been washed to remove unbound biotin were lysed, and membrane proteins were solubilized and loaded onto a streptavidin column. The column was washed to remove contaminating proteins, and the biotin-labeled proteins remaining on the column were eluted with SDS-PAGE sample buffer. The biotin-labeled protein fractions from three independent experiments were separated on a 12% SDS-PAGE gel (Fig. 2). After Coomassie blue staining, each lane was cut into 1 mm slices and the separated proteins were analyzed by mass spectrometry.

Analysis of the Surface Protein-Enriched Fractions—Out of the total of 8306 proteins encoded by the *E. histolytica* genome, 1326 (16.0%) contain one or more transmembrane domains, 1079 (13.0%) contain a signal peptide, and 561 (6.75%) contain both (AmoebaDB, version 1.7). The supplemental Table S2 shows 693 putatively surface-associated proteins. 70% (488 proteins) of them could be detected in all three independent experiments performed and 30% of them (205 proteins) could be detected in two out of three experiments.

Of these 693 surface associated proteins, 38 (5.5%) contain one or more transmembrane domains, 42 (6%) contain a signal peptide, and 14 (2%) contain both (Figs. 3*A*, 3*B*). Therefore, only 2.9% of the 1326 proteins containing one or more transmembrane domains, 3.9% of the 1079 proteins containing a predicted signal peptide and 2.5% of the 561 proteins containing both domains were identified in the membrane fraction. However, out of the 693 putative surface-associated proteins, 175 (25.2%) were predicted by the SecretomeP 2.0 server for mammalian sequences (35) to be nonclassically secreted (*i.e.* not triggered by signal peptide) proteins (Figs. 3*A*, 3*B*). So far, it has not been determined whether this type of protein secretion occurs in *E. histolytica*. Furthermore, pro-

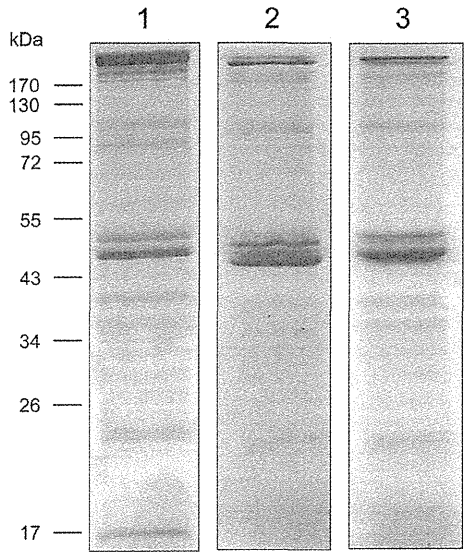

FIG. 2. **SDS-PAGE analysis of proteome preparations from three independent experiments.** Proteins were separated on a 12% SDS-PAGE gels and visualized by staining with Coomassie brilliant blue R-250.

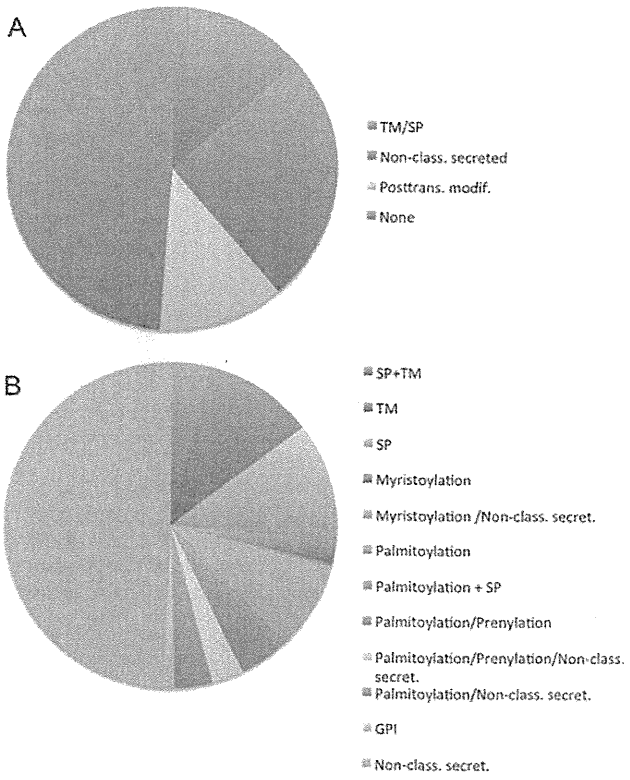

FIG. 3. **Proteins identified in the surface proteome of *E. histolytica*.** A total of 693 proteins were identified. *A*, *In silico* analysis showing that 13.5% of the identified proteins contain one or more transmembrane domains (TM) and/or a signal peptide (SP); that 12.4% contain various post-translational modifications (*e.g.* myristylation, palmitylation, farnesylation, or anchoring to GPI); and that 25.2% were nonclassically secreted. Membrane association could not be predicted for the remaining 48.8%. *B*, Pie chart showing the detailed analysis of the domains found in the 355 putative domain-containing proteins.

teins may be covalently modified post translationally by a variety of lipids, including myristate, palmitate, farnesyl, geranylgeranyl, and glycosylphosphatidylinositol (GPI). Of the 693 identified proteins, 86 (12.4%) contained at least one of these special post-translational modification sites (Figs. 3*A*, 3*B*). By contrast, none of the mentioned membrane-association domains were found in ~49% of the identified proteins (Fig. 3*A*).

BLAST analysis sorted the putative surface-associated proteins into several groups (Fig. 4). The largest group (159 of the 693 proteins, or 23%) consisted of hypothetical proteins. In addition, 11% of the identified proteins were found to be ribosomal proteins, 8.8% were nucleotide-binding/small GTPases, 7.3% were found to be involved in membrane trafficking, 3.9% were chaperones, 8.2% were involved in cytoskeletal organization, 5.5% were involved in protein processing and degradation, 12.2% were involved in general metabolic processes, 3.2% were classical membrane proteins involved in adhesion or as transporters, 4.9% were involved in nucleic acid metabolism, and 4% were involved in protein synthesis.

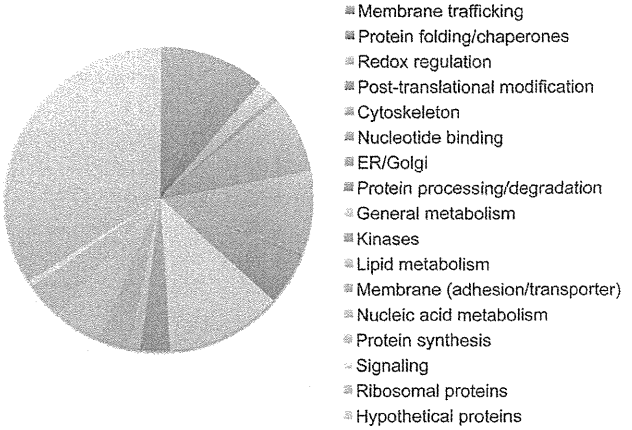

FIG. 4. **Distribution of functional annotations of the identified *E. histolytica* surface-associated proteins.** Functional groups were classified by BLAST homology analysis.

Only a few of the identified proteins (between 0.4 to 3%) were found to be involved in redox regulation, post-translational modifications, lipid metabolism, or signaling, or were ER/Golgi proteins or kinases (Fig. 4, supplemental Table S2).

Verification of Candidate Surface Proteins—To estimate the relative number of proteins that are actually surface-associated and not contaminants of membrane preparations, the location of 27 proteins was analyzed by Western blotting and/or immunofluorescence microscopy. Following random selection of a set of 23 putative surface-associated proteins, we generated antibodies against each. *In silico* analyses predicted that 11 of these proteins had no membrane-association domains, seven had a palmitoylation or prenylation domain, five had a signal peptide, one had a signal peptide and a transmembrane domain, and three were nonclassically secreted proteins. As control we also utilized antibodies against four proteins (Fe-hydrogenase, rubrerythrin, Rab7G, Rab7H) that were not detected in our proteomic study (Table I).

A solubility assay of 19 of these 27 proteins showed that DnaJ and Rab7E were present only in the membrane fraction, whereas EhC2–1, EhC2–3, EhCP-A1, EhCP-A2, EhCP-A5, EhFeSOD, EhGNBP, EhGrainin1, EhHydA, EhHypP, EhP120, EhPrx, EhRab7D, and EhRho were found in both the PBS-soluble and -insoluble fractions, and Fe-hydrogenase, rubrerythrin and EhURE3-BP were present only in the PBS-soluble fraction (Fig. 5, Table I).

To determine which of the proteins are associated with the surface live trophozoites were labeled using the antisera against these recombinant proteins. Twenty of these proteins were present on the surface, whereas seven were not, including EhAct1, EhDnaJ, EhFeHyd, EhGrainin1, EhHydA, EhHypP, and EhURE3-BP (Fig. 6). Nevertheless, immunofluorescence analyses of paraformaldehyde-fixed cells indicated that EhAct1, EhGrainin1, EhRub and EhURE3-BP were present both inside and on the surface of these cells (Fig. 7). The antibody against EhDNAJ was apparently inadequate for im-

TABLE I

Verification of candidate surface proteins. Nd: Not determined; IFA: Immunofluorescence analyses; Sap: Saponin; Pre: Prenylation Palm: Palmitoylation domain; SP: Signal peptide; TM: Transmembrane domain

Name abbr.	Gene name	Accession number	Domains/putative localization	Surface proteome	Western blot		IFA–live trophozoites	IFA–fixed trophozoites	
					PBS-soluble	PBS non-soluble		+ Sap.	– Sap.
EhAct1	Activator 1 40 kDa subunit	XP_651156, XM_646064.1	Palm	+	nd	nd	–	+	+
EhC2–1	C2 domain containing protein	XP_655299, XM_650207.2	None	+	+	+	+	+	+
EhC2–3	C2 domain protein	XP_654499, XM_649407.2	None	+	+	+	+	+	+
EhCoronin	Coronin	XP_654419, XM_649327.1	Non-class. secreted	+	nd	nd	+	nd	nd
EhCP-A1	Cysteine proteinase	XP_650156, XM_645064.2	SP	+	+	+	+	nd	nd
EhCP-A2	Cysteine proteinase 2	XP_650642, XM_645550.2	SP	+	+	+	+	nd	nd
EhCP-A5	Cysteine proteinase	XP_650937, XM_645845.2	SP	+	+	+	+	nd	nd
EhDnaJ	DnaJ family protein	XP_653489, XM_648397.1	SP	+	–	+	–	–	–
EhFeHyd	Fe hydrogenase	XP_652839, XM_647747.2	None	–	+	–	–	+	–
EhFeSOD	Iron-containing superoxide dismutase	XP_648827, XM_643735.2	None	+	+	+	+	+	+
EhGNBP	Guanine nucleotide-binding protein subunit beta 2-like 1	XP_657050, XM_651958	None	+	+	+	+	+	-
EhGrainin1	Grainin 1	XP_650372, XM_645280.2	SP	+	+	+	–	+	+
EhHydA	Putative iron hydrogenase	AAG31036, AF262400.1	None	+	+	+	–	nd	nd
EhHypP	Hypothetical protein	XP_652420, XM_647328.1	None	+	+	+	–	nd	nd
EhLectin	Galactose-specific adhesin 170kD subunit	XP_655415, XM_650323.2	SP/TM	+	nd	nd	+	+	+
EhP120	Proliferating-cell nucleolar antigen p120	XP_654993, XM_649901.1	None	+	+	+	+	+	+
EhPP2A	Serine/threonine-protein phosphatase 2A	XP_656214, XM_651122.2	Non-class. secreted	+	nd	nd	+	nd	nd
EhPrx	Peripheral membrane protein peroxiredoxin	XP_647907, XM_642815.2	Palm	+	+	+	+	+	+
EhRab7D	Rab family GTPase	XP_651915, XM_646823.1	Pre/Palm	+	+	+	+	nd	nd
EhRab7E	Rab family GTPase	XP_651202, XM_646110.2	Pre/Palm	+	–	+	+	+	+
EhRab7G	Rab family GTPase	XP_656477, XM_651385.1	Pre/Palm	–	nd	nd	+	+	+
EhRab7H	Rab family GTPase	XP_653414, XM_648322.1	Pre/Palm	–	nd	nd	+	nd	nd
EhRab11A	Rab family GTPase	XP_653051, XM_647959.2	Pre/Palm	+	nd	nd	+	nd	nd
EhRho	Rho family GTPase	XP_654488, XM_649396.2	None	+	+	+	+	+	+
EhRub	Rubrerythrin	XP_652131, XM_647039.2	None	–	+	–	+	+	+
EhTR	Thioredoxin reductase	XP_655748, XM_650656.2	None	+	nd	nd	+	+	+
EhURE3-BP	URE3-BP sequence specific DNA binding protein	AAG18423, AF291721.1	Non-class. secreted	+	+	–	–	+	+

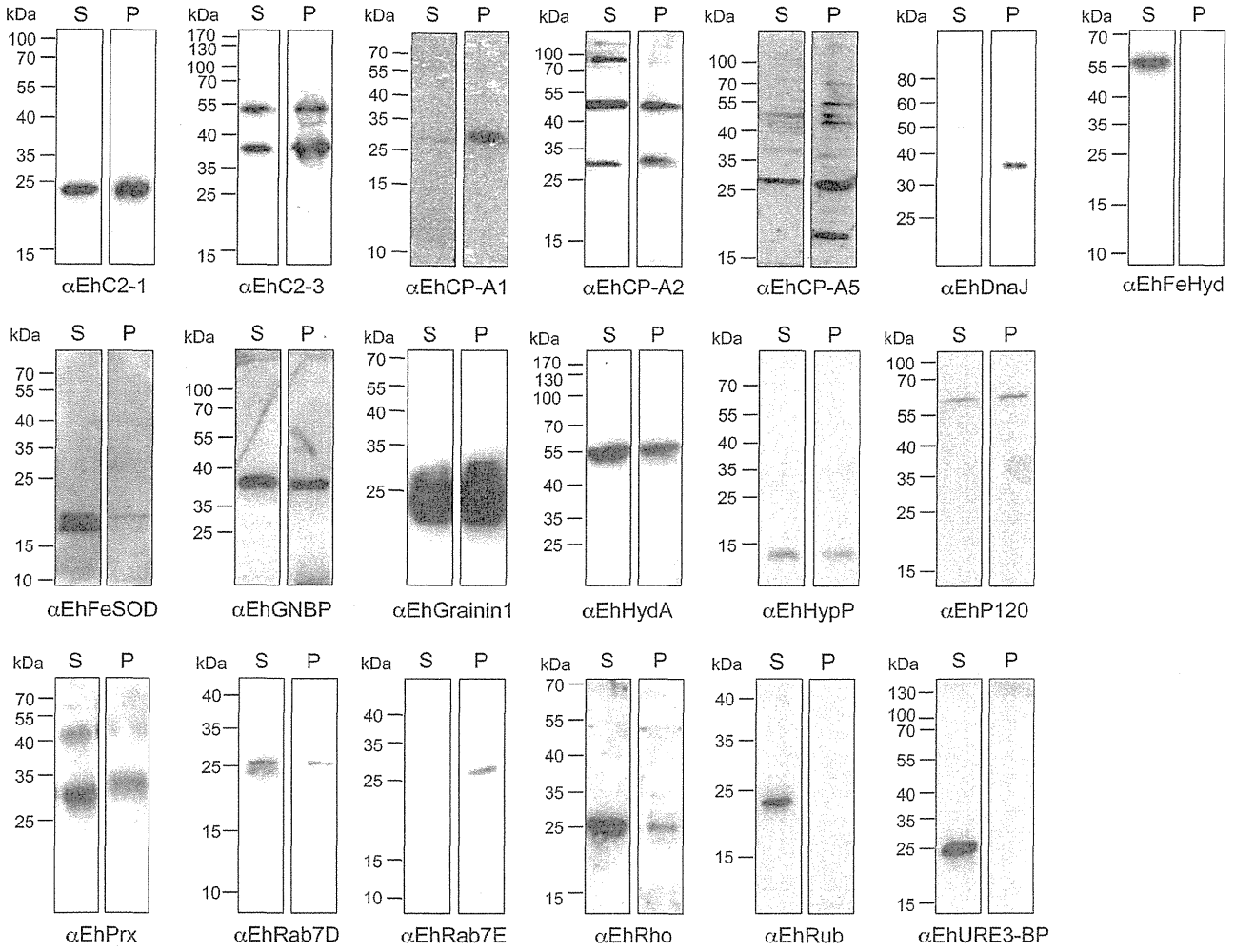

FIG. 5. **Detection of various proteins in the soluble and membrane fractions of *E. histolytica* extracts.** After lysis of the amoebae, the extracts were separated by centrifugation. The supernatant (S) contains the PBS-soluble proteins, whereas the pellet (P) contains the PBS-insoluble proteins of *E. histolytica.* The extracts were separated on 12% SDS-PAGE gels and transferred to nitrocellulose membranes. These Western blots were incubated with antibodies against various *E. histolytica* proteins and specific secondary antibodies, and were developed using ECL.

munofluorescence analysis because no signal was present in cells that were not treated with saponin (Fig. 7). Antibody against EhFeHyd was used as a control because, to date, the protein has only been detected in the PBS-soluble extract of *E. histolytica* (Fig. 5; (30)). This localization was confirmed as the protein was exclusively found in the cytoplasm of *E. histolytica* (Fig. 7). Interestingly, the bead-like protrusion observed on the surface of biotin-labeled amoebae, could also be observed for the majority of the live trophozoites labeled with the various antibodies (Fig. 7).

DISCUSSION

Surface-associated molecules are important for interactions of parasites with their hosts. To date, ~20 surface-exposed proteins or protein families of *E. histolytica* have been described (1–4, 6–15, 17–23, 36) (5, 16). To identify the entire set of plasma membrane-associated proteins, the surface molecules of *E. histolytica* were labeled with biotin, purified, and analyzed by mass spectrometry. Use of this non-membrane soluble biotin reagent preferentially labeled plasma membrane and plasma membrane-associated proteins. Of the total of 693 putative surface-exposed proteins identified in the *E. histolytica* strain HM-1:IMSS, 94 were found to possess a predicted transmembrane domain and/or signal peptide, whereas 86 were associated with the plasma membrane by palmitoylation, prenylation, or myristoylation, or were GPI anchored. Nevertheless, *in silico* analyses indicated that 175 of these 693 proteins were nonclassically secreted, whereas no prognosis for a plasma membrane association could be made for the remaining 339 proteins.

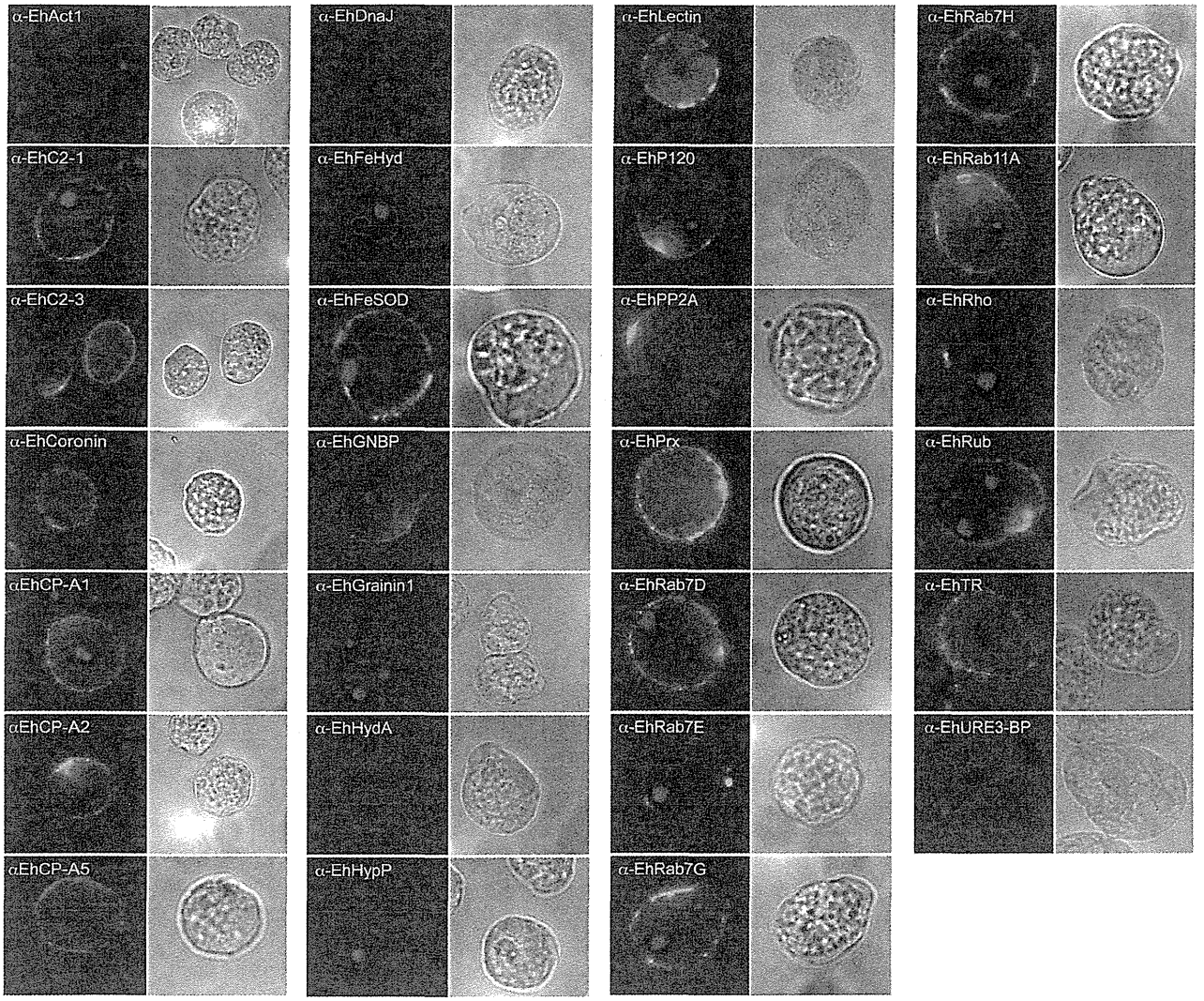

FIG. 6. **Surface association of various *E. histolytica* proteins by immunofluorescent staining of live cells.** Trophozoites were harvested at 4 °C and blocked with FCS prior to incubation with antibodies against various *E. histolytica* proteins. After washing, the amoebae were incubated with Alexa Fluor®488 goat anti-mouse or anti-rabbit antibody (green) or Alexa Fluor®598 goat anti-chicken antibody (red), fixed with paraformaldehyde, and visualized by fluorescence microscopy (Leica, DM BR, Wetzlar, Germany). Hoechst staining was used to visualize nuclei.

Twelve of the ~20 surface-exposed molecules/protein families previously identified in *E. histolytica* were identified in this study. These include the Gal/GalNAc lectin (1, 2), calreticulin (6), transmembrane kinases (8, 9), a Bsp-A like molecule (10), some Rab7 molecules (12), an actinin-like protein, KERP-1 (13), EhCP-A5 (14), EhCP-A2 (15), peroxiredoxin (37, 38), EhMSP-1 (21), and ADH3 (22).

We could not identify EhCP-B9, which previously was shown to be exposed on the *E. histolytica* cell surface (3, 4). However, the gene encoding EhCP-B9 is expressed at very low levels in *E. histolytica* strain HM-1:IMSS (39), which may explain our inability to detect the protein on the surface of these cells. In addition, we were unable to identify the serine-rich *E. histolytica* protein (SREHP) (40), rhomboid protease (11), ARIEL antigen (18, 19), and a LIM protein of *E. histolytica* (EhLimA) (20). Furthermore, we found that the protein URE3-BP, which was previously shown to be localized to the cytoplasm and inner surface of the plasma membrane (17), was present on the outer surface. By contrast, we did not detect syntaxin 1 and SNAP-25, both of which were shown to be surface-associated (23). Nevertheless, other proteins thought to be part of the SNARE machinery, including the v-SNARE protein VampF and the alpha-soluble NSF attachment protein, were detected.

In addition to transmembrane proteins localized to the plasma membrane, many transmembrane proteins were pre-