Supplementary Information:

Novel TPR-containing subunit of TOM complex functions as cytosolic receptor for *Entamoeba* mitosomal transport

Takashi Makiuchi^{a,b}, Fumika Mi-ichi^{a,1}, Kumiko Nakada-Tsukui^a, and Tomoyoshi Nozaki^{a,c,*}

^a Department of Parasitology, National Institute of Infectious Diseases, 1-23-1 Toyama, Shinjuku-ku, Tokyo 162-8640, Japan; ^b Kitasato Institute for Life Sciences and Graduate School of Infection Control Sciences, Kitasato University, Minato-ku, Tokyo 108-8641, Japan; ^c Graduate School of Life and Environmental Sciences, University of Tsukuba, 1-1-1 Tennodai, Tsukuba, Ibaraki 305-8572, Japan

¹ Current address: Department of Biomolecular Sciences, Faculty of Medicine, Saga University, 5-1-1 Nabeshima, Saga 849-8501, Japan * Corresponding author: T. Nozaki

E-mail: nozaki@nih.go.jp; Phone: +81(Japan)-3-4582-2690; Fax: +81-3-5285-1219



Supplementary Figure S1: Fractionation of mitosomes from the *E. histolytica* trophozoites expressing Tom40-HA. A, A scheme of mitosome purification by the Percoll gradient centrifugation, as previously described¹. B, Immunoblot analysis of the fractions after the second Percoll gradient ultracentrifugation. Approximately 9- μ l of each fraction was subjected to SDS-PAGE and immunoblot analyses with anti-Cpn60 and anit-HA antibody.



Supplementary Figure S2: Immunofluorescence analysis of Tom40-HA in *E. histolytica.* The samples were prepared and examined as previously described¹. Arrowheads indicate the mitosomes where Cpn60 and Tom40-HA are co-localized.



Supplementary Figure S3: Immuno-EM analysis of Tom40-HA in *E. histolytica.* The samples were prepared as previously described^{4,5} and examined by electron microscopy at Tokaii Microscopy., Inc (Nagoya, JAPAN). Large (20 nm; blue arrows) and smalle gold particles (10 nm; red arrowheads) indicate Cpn60 and Tom40-HA, respectively.

E. histolytica	1	1522 MAEIVGASIVAGIGFGVFKVICDGVEKMTTPTDQTDKESSENQTYEEINQVKQQEEITKEKQ-DEEDKQADDTKIEEIIKTHYE
E. dispar	1	1561 MAEIVGASIIAGIGFGVFKVICDGVEKITTPNDQRDKEKRENKNEEEINEIVQHSEKEREKQNEEEDNQENDKNIEEIIKKHYE
E. invadens	1	1367 9.2e-04 MGDLIGSALFAGLSFGLFKVACDTLDTFLREKIPEVCDDESKKLETTKNTDFTIENEDVENTTKETHDHKDDVTEIPLEEVKTYVETHYE **** ******* ***
E. histolytica		KAQQYVEMNNIEGAEEEFILIDEINKKIGE-KGMFAAVIKKELASIRFLKEDYCTSKKYLEEVIKIMKDIMKEESNQEVEILYIEVIN
E. dispar		KAQKYVEMNNIEGAEEEFILINEINKKIGE-KGIFAGIIKKEIATIRFIKEDYHTSKKYLEEVIEIMKEIMKEENSKEIEIMYIEVIN
E. invadens		GYKSAINRNEVDVAESELLKVIDLLKNV GQLKGIEG CSLLKEESSLRFMHSDYEASIDFLVMSLELLKDV SSTTEQEQTDVTLLEIESLN ****************
E. histolytica		QMIETRIMNIKYEEPKEMKNEEEMKRIEEEIQENKERIEKMKKEE-KYKEIMKEKE <mark>EENMMGKIQYYIMKEEWKESEKESKALIKKYKEE</mark>
E. dispar		QMIETRIMKEKYEEGKENKEEMKKIYSEIKENKERIEKMKKEE-KYKEIMKEKEEENIMGRIQYYIMKEEWEKSEKESKILIERYKEE
E. invadens		QLVQA <mark>KTLLTKYTNCNNPEKIISEAEMSANEALDKLRKL</mark> PGDYEDLS <mark>FEISLSLIQVKVFQKDWISAEKIGSALLQNHSKI</mark>
E. histolytica		2.0e-02 GKKEEE-KR <u>IIGWLNNIYVIEKNK</u> EGIKENFIRIQELIENEKEEIEVISERAIIEHEIKEYEDSNNSCYECINAIKRKQEK <mark>KEVENKRDI</mark>
E. dispar		5.1e-02 5.2e-04 GKKEEE-KRIIKWLNNIYVIEKNKEGIKENFIRNQEIIENEKEEIEMISERAIIEHEIGEYEDSNNSCYECINKIERKKEKKEEENQRNN
E. invadens		1.3e-03 PETDPKK <mark>VALRTWLSSIYVALNDKDRLEKVLEYSREHSEGV</mark> -EK <mark>IEAYQTSSELLLDCSDNESAIKYGLQTLSFINES</mark> QVELTEDL
E. histolytica		IEIENQMYYTIIENITSSNKEKERKEEIEECMNQIIKLTKEIKETMLISSHFIKTLSIEEINSIDQIYRNLICKILIRNVGKNELPNEYT
E. dispar		IEIENQMYYTIIENIKSSNKEKERKEEIEECMNQIIKLTKEIKETMLISNHFIKTLSIEEINSIDQIYRNLICKILIRNVGKSELPNEYT
E. invadens		ETLMVQTFYTLLDAFENTKNATHYNKYFEMLLQNLPKNKTILVHSHILRTVEIQQIPSAENNIFNIVTKFLVRSVGASALPKTYK
E. histolytica		LQMNIKTVNGTMIKEGEKIKNTGKKQITMIVTNIENMNKDEVYLIEINIQNENKENIFKHLQFFYLN 505
E. dispar		LQMNIKTINGTIIKEGEKIKNKGKKQITMIVTNIENMIKNEVYLIEIIIQNENKENIFKHLQFFYLN 504
E. invadens		ANISYYDTDNTCLCSTIDFTLNQPQTTFTLTSPPLTLTKNKNYYFIVSLK-DNETTFSKHRQFFMIV 497 * * * * * * * ***

Supplementary Figure S4: The multiple alignment and domain prediction of Tom60 from three Entamoeba species. Tom60 homologues from E. dispar and E. invadens (Pathema ID; EDI 218540 and EIN 149090, respectively) acquired were from http://pathema.jcvi.org/cgi-bin/Entamoeba/PathemaHomePage.cgi. The multiple alignment was created by ClustalW2 (http://www.ebi.ac.uk/Tools/msa/clustalw2/). Asterisks and dots indicate identical and similar amino acid residues, respectively. E. histolytica, E. dispar, and E. indavens Tom60s consist of 505, 504, and 497 amino acids, and their predicted isoelectric points are 4.79, 5.02, and 4.61, respectively. EhTom60 shows 84 and 24% identity, and 96 and 76% similarity with Tom60 from E. dispar, and E. indavens, respectively. These proteins are predicted to contain one hydrophobic cluster (depicted by a gray box), and several tetratricopeptide repeats (TPR, depicted by pink and yellow boxes), predicte by TMpred (http://www.ch.embnet.org/software/TMPRED form.html) and TPRpred (http://toolkit.tuebingen.mpg.de/tprpred), respectively. Numerical values indicate the scores by TMpred (in grav boxes, scores above 500 are considered significant) and e-values by TPRpred (in pink boxes). These proteins were predicted to be soluble by the TMHMM (http://www.cbs.dtu.dk/services/TMHMM-2.0/) and the OCTOPUS (http://octopus.cbr.su.se/), and the SOSUI (http://bp.nuap.nagoya-u.ac.jp/sosui/) programs. However, the HMMTOP (http://www.enzim.hu/hmmtop/) program indicated that *E. histolytica* and *E. dispar* Tom60s have one transmembrane helix at the amino terminus. These proteins were also predicted to possess no signal peptide by the SignalP (http://www.cbs.dtu.dk/services/SignalP/) program. The double-underlined region indicates the polypeptide detected by LC-MS/MS analysis in both the 60kD band of the 600 kDa complex immunoprecipitated from the Tom40-HA lystes (Fig. 1c and Supplementary Table S1A) and the band corresponding with the 600 kD complex (Supplementary Table S1B). The single-underlined region indicates the polypeptide detected only in the 600kD complex immunoprecipitated (Supplementary Fig. S5 and Supplementary Table S1B).



Supplementary Figure S5: BN-PAGE and immunoblot analyses of *E. histolytica* trophozoites expressing Tom40-HA. A, Coomassie brilliant blue-stained gels of BN-PAGE analysis. The $100,000 \times g$ organelle fraction (60 mg) was solubilized with 2% digitonin or 2% n-dodecyl-b-D-maltoside (DDM), and subjected to BN-PAGE analysis. After the electrophoresis, the gel was stained with 0.1% Coomassie brilliant blue R-350. **B**, Immunoblot analysis. When the fraction was solubilized by 2% DDM, the specific 600 kD band (arrows) was not detected. The specific 600 kD band detected in the digitonin-solubilized sample was subjected to LC-MS/MS analysis. The specific *E. histolytica* polypeptides detected in only the digitonin sample by LC-MS/MS analysis are described in the Supplementary Table S1B.



Supplementary Figure S6: The secondary structure of *Entamoeba* Tom60s and mitochondrial receptors. The secondary structure of *Entamoeba* Tom60s, rat Tom20 (NCBI ID: Q62760), and yeast Tom70 (NCBI ID: NP_014278) are predicted by PSIPRED. Crystal structure IDs of rat Tom20 and yeast Tom70 are $10M2^2$ and $2GW1^3$ in Protein data bank, respectively. Green and blue boxes indicate putative α -helices and β -strands predicted by PSIPRED, respectively. Gray boxes on rat Tom20 and yeast Tom70 based on crystal structure indicate unknown regions where the crystal structure is not solved. Dashed boxes indicate TPRs predicted by the TPRpred and the crystal structure^{2,3}. Note that the fourth TPR of *E. histolytica* and *E. dispar* Tom60 (also the fourth and fifth TPRs of *E. invadens* Tom60), predicted by TPRpred, contain a β -strand by PSIPRED. In addition, the third TPR of *E. invadens* Tom60 is undetectable in *E. histolytica* and *E. dispar* Tom60. Therefore, the presence and absence of these putative TPRs need to be further verified.



Supplementary Figure S7: *In vitro* binding of His-Tom60 Δ N-HA to the mitosomal matrix protein. The recovery efficiency of His-Tom60 Δ N-HA, AS-FLAG, and CS3-FLAG in the *in vitro* binding assay performed at indicated KCl concentrationsis is shown. Recombinant AS-FLAG and CS3-FLAG was mixed with or without His-Tom60 Δ N-HA in the presence of 0-400mM KCl, and pulled down with Nickel NTA resin. The bound protein was eluted with imidazole. Vertical and horizontal axes indicate the recovery of proteins (%) and the KCl concentrations, respectively.

Supplementary Methods

Immunofluorescence assay and immunoelectron microscopy: Immunofluorescence assay¹ and Immunoelectron microscopy^{4,5} were performed as previously described.

Percoll gradient centrifugation: Mechanical homogenization and fractionation on two series of Percoll gradient ultracentrifugation of 8×10^6 trophozoites were performed as previously described¹. Briefly, cells were washed three times by 2% glucose/PBS. After resuspended in the lysis buffer (10mM MOPS-KOH (pH7.2)/250mM sucrose/protease inhibitors), cells were disrupted mechanically by a Dounce homogenizer. After unbroken cells were removed by centrifugation at 5,000 \times g for 10 min, the supernatant was centrifuged on Percoll gradient.

Immunoblot analysis: Immunoblot analysis was performed with the Immobilon[®]-P membrane, anti-HA mouse monoclonal antibody (16B12; COVANCE, Berkeley, CA), alkaline phosphatase-conjugated anti-mouse IgG antibody (Cell Signaling Technology Beverly, MA), and CSPD (Roche Applied Science).

In vitro binding assay: Approximately 100 nmol of AS-FLAG, CS3-FLAG, or CS3-FLAG-EEVD was incubated with or without 50 nmol of His₆-Tom60 Δ N-HA in the assay buffer (10mM HEPES-NaOH, pH7.5) containing HisLinkTH Protein Purification Resin (Promega, Madison, WI), 0-400 mM KCl, and 50 mM imidazole at 35.5°C for 1.5 h. The mixture was transferred into a microspin column, and resins were washed three times with the assay buffer supplemented with KCl and imidazole. The proteins bound to His₆-Tom60 Δ N-HA were co-eluted with the assay buffer containing 500mM imidazole. Approximately 40% of the eluted samples and 1.3 and 2.5 pmol of standard proteins were subjected to SDS-PAGE and immunoblot analyses with anti-HA and anti-FLAG antibodies. The relative amounts of proteins were quantified by the Analysis Toolbox in the ImageQuant TL safeware (GE Healthcare). Briefly, the protein amounts were calculated by subtracting the intensity of the band in the pull-down samples without His-Tom60 Δ N-HA.

Phylogenetic analysis: The Bootstrap probability by the maximum likelihood method was calculated as previously described⁶. Briefly, the TPRs were identified and extracted from representative TPR-containing proteins including 23 yeast proteins (such as Tom20, Tom70, and Sti1p) using the TPRpred program (http://toolkit.tuebingen.mpg.de/tprpred). The TPRs were also extracted from 28 putative TPR-proteins including Tom60 from *E. histolytica, E. dispar*, and *E. invades,* 36 proteins from *Dictyostelium discoideum, Homo sapiens* Tom34, and *Arabidopsis thaliana* Toc64. Data sets to reconstruct individual maximum likelihood trees were prepared as follows: Set 1 (yeast TPRs and TPRs from *Homo sapiens* Tom34, *Arabidopsis thaliana* Tom60s; Supplementary Table S2A); Set 2 (*Dictyostelium discoideum* TPRs including TPRs from *Entamoeba* Tom60s; Supplementary Table S3A); and Set 3 (*E. histolytica* TPRs including TPRs from *Entamoeba* Tom60s; Supplementary Table S4A).

References for Supplementary Methods

1. Mi-ichi, F., Yousuf, M. A., Nakada-Tsukui, K. & Nozaki, T. Mitosomes in *Entamoeba histolytica* contain a sulfate activation pathway. *Proc Natl Acad Sci USA*. **106**, 21731-21736 (2009).

2. Abe, Y. *et al.* Structural basis of presequence recognition by the mitochondrial protein import receptor Tom20. *Cell.* **100**, 551-560 (2000).

3. Wu, Y., & Sha, B. Crystal structure of yeast mitochondrial outer membrane translocon member Tom70p. *Nat Struct Mol Biol.* **13**, 589-593 (2006).

4. Mi-ichi, F. *et al.* Sulfate activation in mitosomes plays an important role in the proliferation of *Entamoeba histolytica*. *PLoS Negl Trop Dis.* **5**, e1263, doi: 10.1371/journal.pntd.0001263 (2011).

5. Baba, M. Electron Microscopy in Yeast . Methods Enzymol. 451, 133-149 (2008).

6. Makiuchi, T. *et al.* Compartmentalization of a glycolytic enzyme in *Diplonema*, a non-kinetoplastid euglenozoan. *Protist.* **162**, 482-489 (2011).

Supplementary Table S1 (1 of 5 pages): Amino acid sequences of polypeptides that were detected exclusively from the 60kD band of immunoprecipitated sample (A), or the 600kD band of organelle fraction solubilized by digitonin (B), but not from corresponding controls, identified by liquid chromatography-tandem mass spectrometry.

Sample	Annotation	Protein ID	M.W.	Detected polypeptide	<i>Mr</i> (expt)	<i>Mr</i> (calc)
А	Tom60	XP_657124	60658	IIGWLNNIYVIEK	1573.92	1573.89
В	Tom60	XP_657124	60658	IIGWLNNIYVIEK	1573.89	1573.89
				AEIVGASIVAGIGFGVFK	1775.98	1775.98
	Tom40	XP_655014	32441	GSFLSSFNLLSK	1298.69	1298.69
				AIYGLLAFEHSSQDSSFK	1998.97	1998.97
	Glycogen debranching enzyme	XP_653608	167577	TFEGQLR	849.43	849.43
				GRNPEELK	941.49	941.49
				HKEEYLK	945.49	945.49
				EGSPEELAK	958.46	958.46
				YLVGPLGMK	992.54	992.54
				GNYSSNPSTK	1053.47	1053.47
				IMNELHER	1056.5	1056.5
				VLHWINTSK	1096.6	1096.6
				GYDELVPYR	1110.54	1110.53
				LDSSHEGITR	1113.54	1113.54
				MAPQGIEILK	1114.61	1114.61
				KYLVGPLGMK	1120.63	1120.63
				GKEGSPEELAK	1143.58	1143.58
				IEESIYQLR	1149.6	1149.6
				EAVVWEDSLK	1174.59	1174.59
				KIMNELHER	1184.6	1184.6
				YVMVDGSYEK	1205.53	1205.53
				GIYKDVYGSTK	1229.63	1229.63
				AIDVLNVDNYR	1290.66	1290.66

Sample	Annotation	Protein ID	M.W.	Detected polypeptide	<i>Mr</i> (expt)	<i>Mr</i> (calc)
В	Glycogen debranching enzyme	XP_653608	167577	ELQEIKEYPR	1303.68	1303.68
				IIVMYVHTSFK	1352.72	1352.72
				DGGDVEIEGLLMK	1390.66	1390.66
				EGLIPEHIHIEK	1413.76	1413.76
				GFEEVPNLEGFIK	1477.74	1477.75
				HGLIPNLLDGGDHSR	1599.81	1599.81
				YIIPGEFWLYYVIDVNK	2131.11	2131.1
				VSKEEHILPYYFEVIDGK	2165.1	2165.1
				FLFFGALTTGMLSSMENNEK	2268.04	2268.04
				IIVPYGTYPYEIEVETNYSNEHNFQGSK	3290.54	3290.54
	Multidrug resistance protein	XP_651702	182959	AMLIAIIPR	1012.61	1012.61
				LLTEVLNAIK	1112.68	1112.68
				KHDNDLYEK	1160.55	1160.55
				TGFDIEEGNVEK	1336.61	1336.61
				LLNDTMQGVTSMR	1496.7	1496.7
				GMTMDVSDIDQLDIK	1711.76	1711.76
				VLNSPMSFFQATPIGR	1779.9	1779.9
				IIIDGVDTSTLALETLR	1829.01	1829.01
				HTDEELIQALDDVNIR	1879.93	1879.93
				SPLYNTFQETLLGLDTIR	2080.08	2080.08
	Cysteine protease binding family protein 1	XP_655218	102930	VYYILLSMK	1144.62	1144.62
				GGWYAFLSVTEK	1356.67	1356.67
				GYIDGDSYSNPYK	1477.64	1477.64
				VSFYYDEPPINSK	1557.73	1557.74
				YVGMWYTIVGDGGR	1588.73	1588.73
	Hypothetical protein	XP_652282	37468	TNNVILK	800.48	800.48
				AEVAIPLIYNAK	1300.74	1300.74
				GEDDYNKFDELTK	1572.69	1572.69
				AYDLTFALYVNDVFK	1777.89	1777.89

Supplementary Table S1, continued (2 of 5 pages)

Sample	Annotation	Protein ID	M.W.	Detected polypeptide	<i>Mr</i> (expt)	<i>Mr</i> (calc)
В	Hypothetical protein	XP_651464	140450	LVEYDSALR	1064.55	1064.55
				ALVEFIFWK	1151.64	1151.64
				LPQEFIEFLEER	1548.78	1548.78
				YVEPESIVEDDIIK	1647.82	1647.82
	Hypothetical protein	XP_655913	109507	SIVITGIFNLK	1203.72	1203.72
-				ISTATSQFSTGVQFLQR	1869.96	1869.96
				LGLVLQMQDEPYALELIK	2088.12	2088.12
	Hypothetical protein	XP_651234	26833	EIAQNAISK	972.52	972.52
				LDGYLLEYAK	1183.61	1183.61
				LSNIPASVDETVIK	1484.81	1484.81
				AADPNIYVPVHMR	1497.74	1497.74
	Gpi16 subunit, GPI transamidase component	XP_653165	47693	SNLPTILLVLR	1237.78	1237.78
				QSPDEFIADILLVR	1597.83	1597.84
-				YYYVSKPSPDSISDGTLR	2046.99	2046.99
	Hypothetical protein	XP_650881	23244	DIIQVSTADFDK	1350.67	1350.67
				SGNLWSSSIYIK	1353.69	1353.69
				GIAVLPLYSIFHK	1456.84	1456.84
	Vacular protein sorting 33A	BAE94829	76232	LGRPIIIGSTHIIR	1544.95	1544.95
		XP_649746		DVGILNEIYSSLNTFQK	1939.99	1939.99
	Steroid 5-alpha reductase	XP_654922	34218	GNTYTVETELTK	1354.66	1354.66
				TLQESGVTEDTPIEMK	1792.84	1792.84
	Gal/GalNAc lectin heavy subunit	XP_650534	141442	AIETVYNQMSSK	1385.65	1385.65
				NTVDNTVNNPQFTAK	1661.8	1661.8
				YAAKPPLTAAYFLEK	1681.91	1681.91
	Cysteine protease binding family protein 8	XP_652899	101577	YTIFAMNAQR	1229.59	1229.59
				ALPIQELPFSR	1269.71	1269.71
				VQIPDVAYTVTVR	1459.8	1459.8
	Plasma membrane calcium- transporting ATPase	XP_653525	120546	AADVGLAMGIR	1088.56	1088.56
		XP_651287				

Supplementary Table S1, continued (3 of 5 pages)

Sample	Annotation	Protein ID	M.W.	Detected polypeptide	<i>Mr</i> (expt)	<i>Mr</i> (calc)
В	Hypothetical protein	XP_649559	35397	YQQKDEFALLMK	1528.76	1528.76
	Cation-dependent mannose-6-phosphate	XP_656907	24994	IVLGSTSTESLELINDGK	1874.98	1874.98
	Histidine acid phosphatase	XP_651963	49636	HFDEIYR	978.46	978.46
				TLQSAQAFLQTFYPLER	2012.04	2012.04
	Cysteine protease binding family protein 6	XP_653036	100103	GLWYTFVGDGR	1269.61	1269.61
				TEAVWYTFSVEK	1458.7	1458.7
	Gaa1-like, GPI transamidase component	XP_649469	68794	LSLQEFLGR	1061.59	1061.59
				IISGFEFDIQPMVAGPTK	1964.99	1964.99
	Hypothetical protein	XP_654389	37476	KHFGDYGNLN	1163.54	1163.54
	Proteasome alpha subunit	XP_653086	28386	HVSDNAQVNTQK	1339.65	1339.65
				YGLRPYGVGLLVIGYDTKPR	2236.24	2236.24
	Histidine acid phosphatase	XP_649718	50017	LDAGWFLR	976.51	976.51
				DILSFQQMK	1124.55	1124.55
				KGLNQLATLGDK	1256.71	1256.71
	Small GTPase Rab11B	XP_652776	24056	GALGALVVYDITK	1318.75	1318.75
	COPII-coated vesicle membrane protein P24	XP_656857	22597	LQQVSTYLNDAK	1378.71	1378.71
				LQLEQIQPATEQEVK	1752.92	1752.93
	Hypothetical protein	XP_656399	68314	LYDGYGFIEDLEK	1560.73	1560.73
	Glycerophosphoryl diester phosphodiesterase	XP_654143	45642	VFVALGLR	873.54	873.54
				GITMTTLNELLDEFYGK	1959.95	1959.95
	Malic enzyme	XP_648590	54139	EAADVAMQAIK	1161.57	1161.57
		XP_001913406		GNFVGVVSDSTR	1236.61	1236.61
	60S ribosomal protein L18a	XP_648709	20364	VFAPNYVVAK	1106.61	1106.61
				IQLISITTVGDK	1286.74	1286.74
	Hypothetical protein	XP_648553	65525	WTQGNEANVER	1302.6	1302.6
				GSDGYAIFHLLR	1347.69	1347.69
				ADTFHVLNFLQIDQLK	1901.01	1901
	Signal peptidase	XP_654079	38369	EKLTESMDLK	1208.6	1208.6
				FFITIPFINSK	1325.74	1325.74

Supplementary Table S1, continued (4 of 5 pages)

Sample	Annotation	Protein ID	M.W.	Detected polypeptide	<i>Mr</i> (expt)	<i>Mr</i> (calc)
В	60S ribosomal protein L7	XP_649035	26131	LVDPYVTYGYPTLETVQK	2085.07	2085.07
		XP_650340				
		XP_653065				
		XP_656887				
	Hypothetical protein	XP_655238	112544	ILDLFSYIR	1138.64	1138.64
				EIPADGLALFTK	1273.69	1273.69
	Hypothetical protein	XP_001913453	45179	IAAPFFDYSHLDAAK	1664.82	1664.82
				KAEEGHETKPGQSHFK	1808.88	1808.88
	beta-keto acyl reductase	XP_649970	36899	EGFNLIVIAR	1130.65	1130.64
	ARP2/3 complex 16 kDa subunit	XP_652420	15050	GEEAPAVDYSAVVEK	1604.76	1604.76
	Nucleotide-sugar transporter	XP_001913453	45836	QPSVSNEEK	1016.48	1016.48
	Hypothetical protein	XP_653391	19648	VLDLGLEEILK	1240.73	1240.73
		XP_656635				
	Hypothetical protein	XP_653235	25847	LSDQITLFTSPFR	1523.8	1523.8
	Hypothetical protein	XP_655479	73327	SPISYGLATVQK	1262.69	1262.69
	60S ribosomal protein L13	XP_649667	33280	AFDAVAHLR	998.53	998.53
		XP_656471				
	Signal peptidase	XP_653142	21107	GDNNPVDDR	1000.42	1000.42
	Hypothetical protein	XP_650421	28815	LLSDETYETIAASPNR	1778.87	1778.87
				VSVDIHPSSPLK	1277.7	1277.7
				IYAGFIVPSVNDR	1449.76	1449.76
	Hypothetical protein	XP_650004	21783	GEAIPAQYQVK	1202.63	1202.63
	Hypothetical protein	XP_657536	30964	AQQEATILQK	1128.61	1128.61
	Hypothetical protein	XP_654842	23448	SIPIIYLISLYNPSK	1719.98	1719.98
	Galactose-inhibitable lectin small subunit	XP_001913520	31550	VVVQEIIQR	1082.64	1082.64
	Hypothetical protein	XP_654584	22249	FIVSTYGQSGK	1185.6	1185.6

Supplementary Table S1, continued (5 of 5 pages)

M.W. indicates the molecular wieght (dalton). Mr (expt) and Mr (calc) denote the experimentally determined and calculated molecular masses, respectively. Proteins in bold were previously detected in the mitosomal proteome¹. Delta values (experimental minus calculated) of less than ±0.03 are considered specific.

	0650 - 10				
Described section of				Mentio	no
Method	Name	Sequence	Derectin	Underline	double-underline
RNAs and cDNAs	Primer-1	5'- AAAATGAACAATTTTATTCA -3'	antisense		
Cloning of <i>Tom40</i> and <i>Tom60</i> genes for the expression in <i>E. histolytica</i>	Primer-2	5'- GCA <u>AGATCT</u> ATGGAGGGGATTGTTGGATCC -3'	sense	Bg/ II restriction site	
	Primer-3	"5- GGA <u>AGATCT</u> ACGAAAGATTTGAATAGAAG CACC-3'	antisense	Bg/ II restriction site	
	Primer-4	5- GTTG <u>AGATCT</u> ATGGCAGAGATAGTAGGTG CATCAATTG-3'	sense	Bg/ II restriction site	
	Primer-5	5- GTTG <u>AGATCT</u> ATTAAGATAAAAAAATTGAA GATGTTTAAAGATG-3'	antisense	Bg/ II restriction site	
	Primer-6	5- <u>GCGGCCGCTCTAGAA</u> AGATAATTAC- 3'	sense	The sequencw for In-Fusion [®] system	
	Primer-7	5- AAGCTGGGGGATCCAACCGTATTGTCTTC -3'	antisense	 The sequencw for In-Fusion [®] system	
Gene silencing	Primer-8	5'- ATCC <u>AGGCCT</u> ATGGAGGGATTGTTGGATC C-3'	sense	Stu I restriction site	
	Primer-9	5- AGGC <u>GAGCTC</u> CATATCTCATTCCAAAACA C-3'	antisense	Sac I restriction site	
	Primer-10	5- ATCC <u>AGGCCT</u> ATGGCAGAGATAGTAGGT GCATCAATTG-3'	sense	Stu I restriction site	
	Primer-11	5- AGGC <u>GAGCTC</u> CTTCTTTGAGAAATCTAAT TGATGCTAG-3'	antisense	Sac I restriction site	
qRT-PCR	Primer-12	5'-CATCTCAATACCTTACTTATGGCTGTC-3'	sense		
	Primer-13	5-GAAAGATTTGAATAGAAGCACCCCAAG- 3'	antisense		
	Primer-14	5-GAAGAAGAAACGAATAATTGGTTGG- 3'	sense		
	Primer-15	5-CTATAATGTCTCTTTTATTCTCTACTTC- 3'	antisense		
	Primer-16	5'- CGAATAATTGCAAGAAATGCAGGAATTG-3'	sense		

Supplementary Table. S5 (1 of 2 pages): List of primers.

ners.
of prii
List o
ages):
of 2 p
S5 (2
Table.
nentary
upplen

ouppiellielitary lable. 33 (4	z ur z paye:	ej. List di priners.			
Described section of				Men	tion
Method	Name	Sequence	Derectin	Underline	double-underline
qRT-PCR	Primer-17	5'-GGTTCATCAGTTATTAATGCTTCTGAC- 3'	antisense		
	Primer-18	5- CCTATGAAAATCGATTGGACATTCTATTGC C-3'	sense		
	Primer-19	5- CATCACCAGTAGCAAACTTTTGTAACTTG- 3'	antisense		
	Primer-20	5- GCCCCAATTGCACCATATCGTGAAATTAG- 3'	sense		
	Primer-21	5- GCACATTGATCAACAGACTTACCAGCAG- 3'	antisense		
	Primer-22	"5- GATCCTCTTGCTCAAACCATTACATCTG-3'	sense		
	Primer-23	5- GTCTAACGCCAATTTTGATAACTTCTTTTG AG-3'	antisense		
	Primer-24	5'-GATCCAACATATCCTAAAACAACA-3'	sense		
	Primer-25	5'-TCAATTATTTTCTGACCCGTCTTC-3'	antisense		
Production of recombinant proteins	Primer-26	5- AA <u>GGATCC</u> CCAACCGATCAAACAGACAA AG-3'	sense	BamH I restriction site	
	Primer-27	5- AA <u>CTGCAG</u> GAGTTCAACTCGAGTTACCC G-3'	antisense	Pst I restriction site	
	Primer-28	5- AAGGATCCGAAAACCTGTATTTTCAGGGA ATGAGCATTCAAGAAAAC-3'	sense	BamH I restriction site	for addition of the Tobacco Etch Virus (TEV) protease recognition site
	Primer-29	5- AACTGCAGTTACTTATCGTCGTCATCCTTA <u>IAGTC</u> TTTCATGGCATCACCAGF-3'	antisense	Pst I restriction site	for addition of the FLAG epitope tag
	Primer-30	5- AAGGATCCGAAAACCTGTATTTTCAGGGA ATGCAAAATATTACTATCAAC-3'	sense	BamH I restriction site	for addition of the TEV protease recognition site
	Primer-31	5- AACCTGCAGGTTACTTATCGTCGTCATCC TTATAGTCAAGTAATAATGAGTCTAATAC-3'	antisense	Sbf I restriction site	for addition of the FLAG epitope tag