

# Discovery of PP<sub>i</sub>-type Phosphoenolpyruvate Carboxykinase Genes in Eukaryotes and Bacteria\*

Received for publication, June 25, 2015, and in revised form, August 3, 2015. Published, JBC Papers in Press, August 12, 2015, DOI 10.1074/jbc.M115.672907

Yoko Chiba (千葉洋子)<sup>‡§</sup>, Ryoma Kamikawa (神川龍馬)<sup>¶1</sup>, Kumiko Nakada-Tsukui (津久井久美子)<sup>§</sup>, Yumiko Saito-Nakano (中野由美子)<sup>§</sup>, and Tomoyoshi Nozaki (野崎智義)<sup>‡§2</sup>

From the <sup>‡</sup>Faculty of Life and Environmental Sciences, University of Tsukuba, 1-1-1 Tennodai, Tsukuba, Ibaraki 305-8572, Japan, the <sup>§</sup>Department of Parasitology, National Institute of Infectious Diseases, 1-23-1 Toyama, Shinjuku-ku, Tokyo 162-8640, Japan, and the <sup>¶</sup>Graduate School of Environmental Studies, Graduate School of Human and Environmental Studies, Kyoto University, Yoshida Nihonmatsu cho, Kyoto, Kyoto 606-8501, Japan

**Background:** Inorganic pyrophosphate-type phosphoenolpyruvate carboxylase (PP<sub>i</sub>-PEPCK) was unidentified.

**Results:** A conserved hypothetical protein was annotated as PP<sub>i</sub>-PEPCK.

**Conclusion:** PP<sub>i</sub>-PEPCK arose independently from the functional homologs ATP/GTP-PEPCKs and PEP carboxylase.

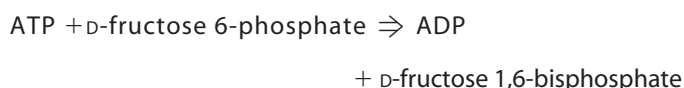
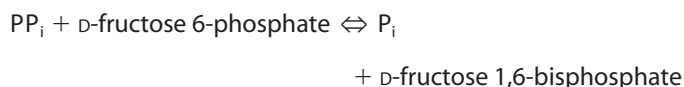
**Significance:** Identification of PP<sub>i</sub>-PEPCK reveals the wide distribution of this enzyme and accelerates understanding the diversity of the central metabolism.

Phosphoenolpyruvate carboxykinase (PEPCK) is one of the pivotal enzymes that regulates the carbon flow of the central metabolism by fixing CO<sub>2</sub> to phosphoenolpyruvate (PEP) to produce oxaloacetate or *vice versa*. Whereas ATP- and GTP-type PEPCKs have been well studied, and their protein identities are established, inorganic pyrophosphate (PP<sub>i</sub>)-type PEPCK (PP<sub>i</sub>-PEPCK) is poorly characterized. Despite extensive enzymological studies, its protein identity and encoding gene remain unknown. In this study, PP<sub>i</sub>-PEPCK has been identified for the first time from a eukaryotic human parasite, *Entamoeba histolytica*, by conventional purification and mass spectrometric identification of the native enzyme, followed by demonstration of its enzymatic activity. A homolog of the amebic PP<sub>i</sub>-PEPCK from an anaerobic bacterium *Propionibacterium freudenreichii* subsp. *shermanii* also exhibited PP<sub>i</sub>-PEPCK activity. The primary structure of PP<sub>i</sub>-PEPCK has no similarity to the functional homologs ATP/GTP-PEPCKs and PEP carboxylase, strongly suggesting that PP<sub>i</sub>-PEPCK arose independently from the other functional homologues and very likely has unique catalytic sites. PP<sub>i</sub>-PEPCK homologs were found in a variety of bacteria and some eukaryotes but not in archaea. The molecular identification of this long forgotten enzyme shows us the diversity and functional redundancy of enzymes involved in the central

metabolism and can help us to understand the central metabolism more deeply.

Inorganic pyrophosphate (PP<sub>i</sub>) is composed of two molecules of phosphate (P<sub>i</sub>) linked by a phosphoanhydride bond. PP<sub>i</sub> has been proposed as an evolutionary precursor of ATP and GTP (1) because this structurally simple compound can be formed spontaneously, and the hydrolysis of PP<sub>i</sub> produces high energy (2, 3). If ancestral organisms could utilize only PP<sub>i</sub> or polyphosphates as an energy and/or phosphate donor for enzymatic reactions, there should have been an event in which new enzymes utilizing ATP/GTP arose. This further raises the question of how organisms change the major substrate: by accumulation of mutation on the enzyme to change the substrate specificity from PP<sub>i</sub> to nucleotide triphosphate or the substitution of PP<sub>i</sub>-utilizing enzymes by an independently emerging functional homolog that utilizes nucleotide triphosphates. To gain insight into the evolutionary transition from PP<sub>i</sub> to nucleotide triphosphate, the evolutionary relationship between the extant PP<sub>i</sub>- and ATP/GTP-utilizing enzymes has been examined (4–11).

In glycolytic/gluconeogenic pathways and the closely related reactions in current organisms, three enzymatic reactions that can utilize PP<sub>i</sub> as the substrate have been reported. The first one is a PP<sub>i</sub>-dependent phosphofructokinase (PFK)<sup>3</sup> (EC 2.7.1.90) reaction. PP<sub>i</sub>-PFK catalyzes a reversible reaction, whereas ATP-dependent PFK (EC 2.7.1.11) catalyzes an irreversible reaction.



REACTIONS 1 AND 2

\* This work was supported in part by Grants-in-Aid for Scientific Research on Innovative Areas 23117001 and 23117005 and Grants-in-Aid for Scientific Research (B) 26293093 from the Ministry of Education, Culture, Sports, Science, and Technology, the Research Program on Emerging and Re-emerging Infectious Diseases of the Japan Agency for Medical Research and Development (AMED), the Science and Technology Research Partnership for Sustainable Development (SATREPS), and the Japan International Cooperation Agency (JICA) and AMED (to T. N.); Grant-in-Aid for Young Scientists (B) 26860275 from the Japan Society for the Promotion of Science (JSPS) (to Y. C.); and Grant-in-Aid for Scientific Research on Innovative Areas 23117006, for Young Scientists (A) 15H05606, and for Challenging Exploratory Research 15K14591 from JSPS (to R. K.). The authors declare that they have no conflicts of interest with the contents of this article.

<sup>1</sup> Supported by the Institute for Fermentation (Osaka, Japan).

<sup>2</sup> To whom correspondence should be addressed: Dept. of Parasitology, National Institute of Infectious Diseases, 1-23-1 Toyama, Shinjuku-ku, Tokyo 162-8640, Japan. Tel.: 81-3-4580-2690; Fax: 81-3-5285-1173; E-mail: nozaki@nih.go.jp.

<sup>3</sup> The abbreviations used are: PFK, phosphofructokinase; PEP, phosphoenolpyruvate; PPK, pyruvate phosphate dikinase; PEPCK, phosphoenolpyruvate carboxylase; Ni-NTA, nickel-nitrilotriacetic acid; WB, Western blot.

## Identification of PP<sub>i</sub>-type Phosphoenolpyruvate Carboxykinase

PP<sub>i</sub>-PFKs from an eukaryotic human parasite, *Entamoeba histolytica* (12), and an anaerobic bacterium, *Propionibacterium freudenreichii* subsp. *shermanii* (13), were proposed to work for fructose 1,6-bisphosphatase synthesis like ATP-PFKs. In higher plants, it has been shown that PP<sub>i</sub>-PFK works in the opposite direction, at least during internode developmental stages (14). Therefore, PP<sub>i</sub>-PFK can catalyze both PP<sub>i</sub>-utilizing and PP<sub>i</sub>-producing reactions not only *in vivo* but also *in vitro*. ATP- and PP<sub>i</sub>-PFK share a common origin, but the evolutionary history is highly complex. Change of the phosphate donors in PFKs occurred more than once, as suggested by phylogenetic analyses of ATP- and PP<sub>i</sub>-PFKs (6, 9, 10, 15). The complex evolution of substrate utility in PFK suggests that, at least in PFK, the transition of PP<sub>i</sub>-utilizing and nucleotide triphosphate-utilizing ability can have occurred relatively easily. Indeed, mutation of just a few amino acids can change the substrate preference from PP<sub>i</sub> to ATP, based on the ratio of  $k_{\text{cat}}/K_m$  (8).

The second reaction utilizing PP<sub>i</sub> is catalyzed by pyruvate phosphate dikinase (PPDK; EC 2.7.9.1).



REACTION 3

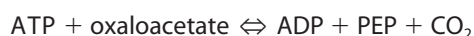
PEP-pyruvate conversion is also catalyzed by PP<sub>i</sub>-independent enzymes, PEP synthase (EC 2.7.9.2; Reaction 4) and pyruvate kinase (EC 2.7.1.40; Reaction 5).



REACTIONS 4 AND 5

PPDK and PEP synthase share conserved domains, whereas overall amino acid sequence similarity is not high (16). On the other hand, neither PPDK nor PEP synthase shows substantial similarity to pyruvate kinase at the primary structure level. This case is apparently distinguishable from that of PFK because transition of substrate utility for the PEP-pyruvate conversion has not occurred by simple amino acid substitutions, as seen in PFK.

The third mechanism is a PEP-oxaloacetate interconverting reaction catalyzed by phosphoenolpyruvate carboxykinase (PEPCK). Because PEP is a key intermediate in a variety of metabolic processes in all living organisms (17, 18), PEPCK works as a major crossroad that connects glycolysis/gluconeogenesis and organic acids metabolisms like the tricarboxylic acid cycle and fumarate fermentation. According to the phosphate donor to oxaloacetate, PEPCK can be divided into three types: GTP-PEPCK (EC 4.1.1.32), ATP-PEPCK (EC 4.1.1.49), and PP<sub>i</sub>-PEPCK (EC 4.1.1.38). PEPCK reactions are fundamentally reversible; however, ATP- and GTP-PEPCKs prefer a PEP-producing reaction, whereas PP<sub>i</sub>-PEPCK prefers an oxaloacetate-producing reaction, at least *in vitro* (19). PEP carboxylase (EC 4.1.1.31) also catalyzes PEP-oxaloacetate interconversion, but this reaction is irreversible and requires HCO<sub>3</sub><sup>-</sup> instead of CO<sub>2</sub> (18, 20).



REACTIONS 6–9

ATP- and GTP-PEPCKs and PEP carboxylase have been well characterized. ATP-PEPCKs are mainly present in bacteria, yeast, and plants, whereas GTP-PEPCKs are mostly present in higher eukaryotes, archaea, and some bacteria (21). PEP carboxylase exists in various bacteria and in limited archaea and eukaryotes. Although there is no discernible similarity in the overall structure of these proteins, the residues implicated for the binding to metal, oxaloacetate/PEP, and nucleotides are conserved between ATP- and GTP-PEPCKs (22, 23), whereas conservation of these residues has not been reported in the case of PEP carboxylase. In contrast, the information on PP<sub>i</sub>-PEPCK, which is also called PEP carboxytransphosphorylase, is limited. Although PP<sub>i</sub>-PEPCK was previously purified from *P. freudenreichii* subsp. *shermanii* (19, 24–26) and *E. histolytica* (27) and extensively characterized enzymologically, the gene encoding PP<sub>i</sub>-PEPCK has remained unknown since the first description in 1961 (24). Both the distribution of PP<sub>i</sub>-PEPCK among the three domains of life and the evolutionary relationship between other PEP-oxaloacetate interconverting enzymes were totally obscure.

In the present study, we identified the protein for PP<sub>i</sub>-PEPCK activity and its encoding gene from *E. histolytica* and *P. freudenreichii*. The identification of PP<sub>i</sub>-PEPCK finally enabled us to compare the amino acid sequences of all three types of PEPCKs as well as PEP carboxylase. It also allowed us to estimate the distribution of PP<sub>i</sub>-PEPCK in the tree of life. Last, we described the complex evolutionary history of proteins utilizing different substrates for the same catalytic reaction.

### Experimental Procedures

**Microorganisms and Cultivation**—Trophozoites of *E. histolytica* clonal strain HM-1:IMSS Cl6 were maintained axenically in Diamond's BI-S-33 medium at 35.5 °C as described previously (28). Cells were grown to the late logarithmic phase (2–3 days after inoculation), harvested by the addition of ice-cold PBS buffer to culture flasks, after discarding the medium, and followed by centrifugation at 300 × *g* for 5 min at 4 °C. For protein purification, the harvested cell pellets were stored at –80 °C until use.

**Enzyme Assays**—PP<sub>i</sub> and oxaloacetate-producing direct PEPCK activity was assayed by measuring the oxidation of NADH by detecting the decrease of absorbance at 340 nm as described previously (27). The reaction mixture contained 10 mM potassium phosphate buffer (pH 6.5), 0.4 mM PEP, 0.1 mM CoCl<sub>2</sub>, 20 mM KHCO<sub>3</sub>, 10 mM MgCl<sub>2</sub>, 2.0 units/ml malate dehydrogenase (from rabbit muscle; Sigma), 0.25 mM NADH, and an enzyme solution. Reaction was started by adding PEP, and the NADH oxidizing activity without the addition of PEP was subtracted as the base line. To obtain the absolute active value, 400 μl of the reaction mixture was added in 1.0-cm path length cuvettes, and absorbance was monitored by a spectrometer. To obtain the relative activity during the purification, 100 μl of the mixture was put in a 96-well plate and incubated at 37 °C, and

## Identification of PP<sub>i</sub>-type Phosphoenolpyruvate Carboxykinase

340 nm was monitored by a microplate reader. One unit of activity was defined as the amount of enzyme oxidizing 1  $\mu$ mol of NADH/min.

**Purification of Native *E. histolytica* PEPCK—**EhPEPCK was purified from 5 g of wet cells as follows. The cells were suspended in 15 ml of 20 mM Tris-HCl (pH 8.0) containing 0.1 mM CoCl<sub>2</sub> (buffer A) and 0.5 mg/ml E-64 and disrupted by freezing and thawing, and cell debris was then removed by centrifugation at 100,000  $\times$  g for 1 h. After the supernatant was diluted to 30 ml with buffer A, ammonium sulfate was added to give 30% saturation. The samples were then applied to a Butyl-Toyoppearl column (22  $\times$  15 cm; Tosoh, Tokyo, Japan) equilibrated with buffer A supplemented with ammonium sulfate at 30% saturation. All chromatography steps were performed using an ÄKTA purifier system (GE Healthcare) at room temperature. Proteins were eluted with a gradient of ammonium sulfate from 30 to 0% at a flow rate of 4 ml/min. The active fractions were desalted using a PD10 column (GE Healthcare) and then applied to a MonoQ HR 5/5 column (bed volume, 1 ml; GE Healthcare) equilibrated with buffer A. Proteins were eluted with a gradient of NaCl from 0 to 500 mM for 20 column volumes at a flow rate of 1.0 ml/min. The active fractions were pooled and loaded onto a Superdex 200 (10/300) column (GE Healthcare). Proteins were eluted with buffer A supplemented with 150 mM NaCl.

**Identification of EhPEPCK by LC-MS/MS Analysis—**~130 kDa bands were excised from silver-stained gel and subjected to LC-MS/MS analysis at the W. M. Keck Biomedical Mass Spectrometry Laboratory (University of Virginia, Charlottesville, VA). The analysis was performed in almost the same manner as before (29); the only difference was that 5  $\mu$ l of the extract was injected to a reversed-phase capillary column, and peptides were eluted over 0.3 h.

To estimate the relative ratio of EhPEPCK1, -2, and -3 in the analyzed sample, the peak area for the selected ion chromatogram of the monoisotopic mass of the most abundant charge state ( $\pm$ 0.02 Da) for each peptide was quantified. Three peptides conserved in all three proteins, specific to EhPEPCK1, and conserved only in EhPEPCK2 and -3, respectively, and one peptide unique to EhPEPCK2 or -3, respectively, were selected for the analysis because these peptides were detected in all three samples: immunoprecipitated samples (see below) using EhPEPCK1-HA, EhPEPCK2-HA, and purified PP<sub>i</sub>-PEPCK from the wild-type amoeba lysate. The sum of the peak area value of the three peptides conserved in all was used for normalization.

**Construction of Plasmids—**For expression in *Escherichia coli*, the protein coding sequences of EhPEPCK1 (XP\_654765.1) and EhPEPCK2 (XP\_650862.1) genes were PCR-amplified from *E. histolytica* cDNA using the following pair of forward and reverse primers: 5'-TCGAAGGTAGGCATATGTTTAATCAAGAAAAAGGTACC-3' (PEPCK\_1\_pCold\_F) and 5'-ATTCGGATCCCTCGATTAATGTTTCATGCATTTGTATG-3' (PEPCK\_1\_2\_pCold\_R) for PEPCK1 and 5'-TCGAAGGTAGGCATATGTTTAATCAAGAAACAAGGTA-3' (PEPCK\_2\_pCold\_F) and PEPCK\_1\_2\_pCold\_R for PEPCK2. For expression in *E. histolytica* trophozoites with a hemagglutinin (HA) tag at the C termini, the protein coding sequences of EhPEPCK1 and -2 genes were PCR-amplified

using the following primers: 5'-ACACATTAACAGATCATGTTTAATCAAGAAAAAGGTAC-3' (PEPCK\_1\_pEhExHA\_F) and 5'-ATGGATACATAGAATGTTTCATGCATTTGTATG-3' (PEPCK\_1\_2\_pEhExHA\_R) and 5'-ACACATTAACAGATCATGTTTAATCAAGAAACAAGGTA-3' (PEPCK\_2\_pEhExHA\_F) and PEPCK\_1\_2\_pEhExHA\_R, respectively. pColdI (Takara, Tokyo, Japan) and pEhExHA (30) plasmids were used for expression in *E. coli* and *E. histolytica*, respectively. The amplified fragments were inserted into the plasmids cut with NdeI and XhoI (pColdI) or BglII (pEhExHA) using the In-Fusion HD cloning system (Takara).

The PP<sub>i</sub>-PEPCK gene from *P. freudenreichii* subsp. *freudenreichii* (PfPEPCK; LC062511) was PCR-amplified from the genomic DNA (DSM 20271) purchased from the German Collection of Microorganisms and Cell Cultures using the primers 5'-TCGAAGGTAGGCATAATGTCCGTAGTCGAACGC-3' (Pf\_PEPCK\_pCold\_F) and 5'-ATTCGGATCCCTCGATCAGACGAACCTGGGCTG-3' (Pf\_PEPCK\_pCold\_R) and cloned into pColdI as described above.

**Overexpression and Purification of Recombinant PP<sub>i</sub>-PEPCKs—**The plasmids for bacterial expression were introduced into *E. coli* BL21 CodonPlus (DE3)-RIL (EhPEPCK1 and -2; Agilent) and *E. coli* BL21 Star (DE3) (PfPEPCK; Life Technologies, Inc.). The hosts transformed with the expression plasmids were inoculated into 400 ml or Luria-Bertani medium in a 1-liter conical flask containing 50  $\mu$ g/ml ampicillin and 34  $\mu$ g/ml chloramphenicol if necessary. After cultivating the cells aerobically at 37  $^{\circ}$ C until the A<sub>600</sub> reached ~0.5, protein expressions were induced by cooling the culture on ice for 30 min and adding 0.1 mM isopropyl thio- $\beta$ -D-galactopyranoside to the medium, followed by overnight cultivation at 15  $^{\circ}$ C. Harvested cells (~8 g of wet cells from 1.2 liters of culture) were disrupted by adding BugBuster (~5 ml/g of wet cells; Merck), and cell debris was removed by centrifugation. Imidazole and NaCl concentrations in the supernatant were adjusted to 10 and 300 mM, respectively. The supernatant was then applied to an open column packed with Ni-NTA-agarose (1.5-ml bed volume; Qiagen). After washing the column with 16 bed volumes of 50 mM Tris-HCl, 300 mM NaCl, and 20 mM imidazole-HCl (pH 8.0), the His-tagged protein was eluted with 4 bed volumes of the same buffer containing 250 mM imidazole-HCl. The eluted proteins were further purified using a MonoQ column as described above.

**Antibodies—**To make anti-EhPEPCK antisera, full-length EhPEPCK1 with a His tag at the N terminus was expressed in *E. coli*, purified using Ni-NTA and MonoQ columns as described above, and used as the antigen. The anti-EhPEPCK1 antisera from rabbit were commercially raised by Operon Biotechnology (Tokyo, Japan). The specificity of these antibodies was confirmed using *E. histolytica* lysates by Western blotting (WB) analysis (data not shown).

**Amoeba Transformation—**The plasmids generated as described above were introduced into *E. histolytica* trophozoites by lipofection as described previously (31) with minor modifications. Approximately 40  $\mu$ l of the transfection medium containing 3–5  $\mu$ g of one of the plasmids was mixed with 10  $\mu$ l of PLUS<sup>TM</sup> reagent (Life Technologies) and kept at room tem-



perature for 15 min. This mixture was combined with 20 μl of Lipofectamine<sup>TM</sup> transfection reagent mixed with 30 μl of transfection medium, kept at room temperature for 15 min, diluted with 400 μl of transfection medium, and added to the seeded trophozoites attached to a 12-well plate after removing the medium. The plate was then incubated at 35.5 °C for 5 h, and cells were transferred to a tube containing 5.5 ml of BI-S-33 medium. About 24 h after transfection, BI-S-33 medium was changed into fresh medium with 1 μg/ml G418 and gradually increased for ~2 weeks until the G418 concentration reached 10 μg/ml.

**Protein Assay**—Protein concentrations were measured using a Bio-Rad protein assay DC dye. Bovine γ-globulin was used as a standard.

**WB**—Protein samples were subjected to SDS-PAGE after boiling at 95 °C for 3 min with SDS-PAGE loading buffer. Proteins in the gel were transferred to nitrocellulose membrane and then blocked with 5% (w/v) skim milk in Tris-buffered saline with Tween 20 (TBST) overnight at 4 °C. Proteins on the membrane were reacted with antibodies in TBST for 1 h at room temperature. Primary antibodies were used at a 1:5000 dilution for anti-EhPEPCK rabbit antibody and at a 1:1000 dilution for anti-cysteine synthase 1 (32) and anti-Cpn60 (33) rabbit antibodies and anti-HA mouse monoclonal antibody (clone 11MO, Covance (Princeton, NJ)). The blots were visualized using alkaline phosphatase-conjugated anti-rabbit or mouse IgG antibody (1:2000 dilution; Cell Signaling Technology, Danvers, MA) with AP Color Reagent (Bio-Rad) according to the manufacturer's protocol.

**Immunoprecipitation**—To examine the interaction between EhPEPCKs, 0.08 g of wet cells expressing PEPCK1-HA, PEPCK2-HA, or only HA were lysed with 1 ml of lysis buffer (50 mM Tris-HCl, pH 8.0, 150 mM NaCl, 0.1 mM CoCl<sub>2</sub>, 0.2% Triton X-100, 0.5 mg/ml E-64) for 10 min on ice. After centrifugation at 20,000 × g for 10 min, the supernatants were precleared with 150 μl of Protein G-Sepharose (GE Healthcare) for 1 h. After removing the beads by centrifugation, the precleared supernatant samples were mixed and incubated with 50 μl of anti-HA monoclonal antibody-conjugated agarose (Sigma-Aldrich) for 4 h. The agarose beads were washed with 400 μl of lysis buffer without E-64 five times (9300 × g for 1 min each) to remove the unbound proteins, and the bound proteins were eluted by incubating with 200 μl of lysis buffer containing 0.2 mg/ml HA peptide (Sigma-Aldrich) overnight. The eluted samples were collected by centrifugation at 800 × g for 3 min. All of the centrifugations and incubations were done at 4 °C otherwise mentioned.

**Phylogenetic Analysis**—Homologues of PP<sub>i</sub>-PEPCK were searched by BLASTP analysis against the non-redundant protein sequence database using EhPEPCK1 as a query (May 5, 2015). In addition, we also retrieved the homologues of *Mastigamoeba balamuthi* and *Spironucleus barkhanus* by tBLASTN analysis against the whole genome shotgun reads and the expressed sequence tag database, respectively. Other eukaryotic counterparts were retrieved by tBLASTN analysis against the Marine Microbial Eukaryote Transcriptome Sequencing Project database (34). No more than two sequences with 97–100% amino acid sequence similarity were included, except

**TABLE 1**  
Purification of PP<sub>i</sub>-PEPCK from *E. histolytica*

1 unit represents the amount of enzyme required to oxidize 1 μmol of NADH/min.

	Activity	Specific activity	Purification	Yield
	units	units/mg	-fold	%
100,000 × g supernatant	75.4	0.165	1.0	100
Butyl-Toyopearl	32.1	4.86	29.4	42.6
MonoQ	8.06	34.0	206	10.7

for *Entamoeba* spp. and *Giardia intestinalis* (for details, see “Discussion”). Sequences were aligned by MAFFT (35), and ambiguously aligned sites were excluded, resulting in a data set comprising 158 taxa 749 positions. The data set was subjected to the maximum likelihood method with the LG + Γ + F model. The maximum likelihood tree was heuristically searched from 10 distinct maximum parsimony trees. In maximum likelihood bootstrap analyses (100 replicates), a heuristic tree search was performed from a single maximum parsimony tree per replicate. The maximum likelihood phylogenetic analyses described above were conducted by RAxML version 7.2.8 (36).

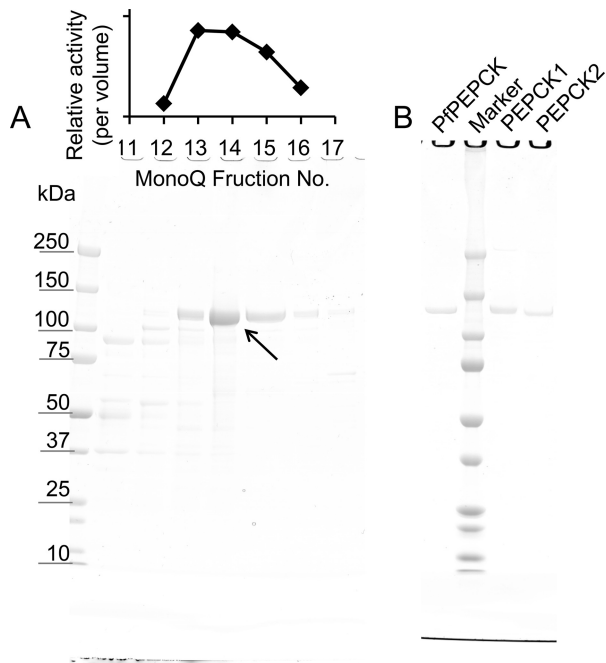
## Results

**Identification of PP<sub>i</sub>-PEPCK from *E. histolytica***—An *in silico* search of potential PEPCK homologs in the *E. histolytica* genome (AmoebaDB) using canonical ATP- and GTP-PEPCK from *Dictyostelium discoideum* (XP\_645490.1 and XP\_645396.1, respectively) failed to detect any candidate of PEPCK (data not shown), although PP<sub>i</sub>-PEPCK activity was detected from *E. histolytica* trophozoite lysates (0.165 units/mg of protein) as previously reported by Reeves (27). PP<sub>i</sub>-PEPCK activity was sequentially separated using hydrophobic (Butyl-Toyopearl) and anion exchange (MonoQ) columns (Table 1), and a MonoQ fraction containing peak PP<sub>i</sub>-PEPCK activity gave one major protein band at ~130 kDa on SDS-PAGE analysis (Fig. 1A). PP<sub>i</sub>-PEPCK activity detected in this MonoQ fraction indicated 206-fold purification with a specific activity of 34.0 units/mg, which agrees well with the previous results of PP<sub>i</sub>-PEPCK from *P. freudenreichii* that showed 12.8 units/mg of activity after 99-fold purification (25). The activity level was also similar to that of ATP-PEPCK from *E. coli* (24 units/mg) (37). Phosphate, PEP, and MDH dependence of the NADH oxidizing activity in the MonoQ fraction was also confirmed. NADH oxidizing activity was not detected when phosphate was substituted to ADP, indicating that the fraction did not contain ATP-PEPCK activity.

The fraction containing PEPCK activity was further subjected to size exclusion chromatography using Superdex 200. The PP<sub>i</sub>-PEPCK activity was detected in the fractions corresponding to two major peaks around ~143 and ~440 kDa (Fig. 2). The integrated areas of the absorbance at 280 nm corresponding to the 440- and 143-kDa proteins (the first and second peaks, respectively) and their enzymatic activity levels are comparable, suggesting that active PP<sub>i</sub>-PEPCK is present as both monomer and trimer in *E. histolytica* (Fig. 2A). PP<sub>i</sub>-PEPCK from *P. freudenreichii* has been previously reported to exist as a monomer, dimer, and tetramer with activity (26). Formation of several quaternary structures may be a characteristic of PP<sub>i</sub>-PEPCK.

## Identification of PP<sub>i</sub>-type Phosphoenolpyruvate Carboxykinase

Because both of the fractions seemed to contain an identical ~130-kDa protein band on SDS-PAGE, the protein band from the fractions corresponding to the first peak (~440 kDa) was

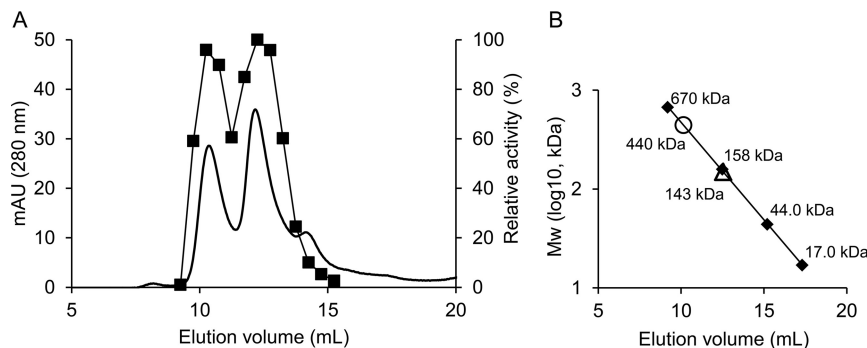


**FIGURE 1. SDS-PAGE analysis of purified native EhPEPCKs (A) and *E. coli*-expressed recombinant EhPEPCK1, EhPEPCK2, and PfPEPCK (B).** A, relative PP<sub>i</sub>-PEPCK activities in the selected MonoQ fractions (per volume) are shown at the top. Approximately 15  $\mu$ l of the MonoQ fractions were applied to SDS-PAGE and stained with Coomassie Brilliant Blue (bottom). The arrow indicates the band corresponding to EhPEPCK1. B, homogeneity of purified *E. coli*-expressed recombinant EhPEPCK1 (PEPCK1), EhPEPCK2 (PEPCK2), and PfPEPCK. One  $\mu$ g of recombinant proteins after the purification by MonoQ column was subjected to SDS-PAGE and Coomassie Brilliant Blue staining.

excised and subjected to TOF MS/MS analysis. Peptides detected from the band correspond to a “cluster of hypothetical proteins,” which consist of three homologous proteins: EHI\_166920 (XP\_654765.1), EHI\_030750 (XP\_650862.1), and EHI\_198620 (XP\_655201.1) (Table 2). The calculated molecular mass of these proteins (131 kDa) agreed well with that expected from the SDS-PAGE result, and the detected peptides covered 63–67% of the three proteins. Hereafter these proteins are referred to as EhPEPCK1, -2, and -3, respectively. EhPEPCK1 and EhPEPCK2 or -3 showed 91% amino acid identity, whereas EhPEPCK2 and -3 showed 98% identity. Amino acid variations between EhPEPCK1 and EhPEPCK2/3 were mainly found in the first 60 residues at the N-terminal region, showing only 45% identity between EhPEPCK1 and EhPEPCK2/3. Neither of these proteins show any similarity to known proteins or domains as predicted by Pfam, indicating that the purified proteins were novel proteins.

**Confirmation of PP<sub>i</sub>-PEPCK Activity of the Amebic Enzymes Expressed in *E. coli***—EhPEPCK1, EhPEPCK2, and PfPEPCK were expressed as soluble recombinant proteins with the histidine tag at the N terminus in *E. coli*. A Coomassie Brilliant Blue-stained SDS-polyacrylamide gel of the proteins purified using Ni-NTA and MonoQ chromatography showed that the full-length proteins of the expected size were purified to homogeneity (Fig. 1B). PP<sub>i</sub>-PEPCK activity of the purified EhPEPCK1, EhPEPCK2, and PfPEPCK was 8.9, 3.9, and 23.0 units/mg, respectively, confirming that the identified genes indeed encode proteins with PP<sub>i</sub>-PEPCK activity.

**Expression of EhPEPCK1 and -2 with HA Tag in *E. histolytica***—Amebic transformants expressing either EhPEPCK1 or EhPEPCK2 with the HA tag at the carboxyl terminus (EhPEPCK1-HA or EhPEPCK2-HA) using the episomal plasmid were established. Expression of the epitope-



**FIGURE 2. Molecular weight of purified native EhPEPCKs.** A, chromatograph of native EhPEPCKs on Superdex200. The optical absorbance at 280 nm is shown with a thick line, and the relative activity per volume of each fraction is shown with squares and a thin line. The highest activity per volume in all fractions was defined as 100%. B, estimation of molecular weight of the proteins corresponding to the two peaks in A (circle and triangle). Standard proteins are shown with diamonds.

**TABLE 2**  
Proteins detected by MS/MS analysis

Peptide thresholds were standard, and protein thresholds were 90.0% minimum and 1 peptide minimum.

Identified proteins	Accession numbers	Molecular mass	Quantitative value
		kDa	
Hypothetical protein (EhPEPCK1)	EHI_166920; XP_654765	131	214
Hypothetical protein (EhPEPCK2)	EHI_030750; XP_650862	131	209
Hypothetical protein (EhPEPCK3)	EHI_198620; XP_655201	131	198
Hypothetical protein	EHI_192430; XP_656829	29	1
Hypothetical protein	EHI_189920; XP_652326	62	1

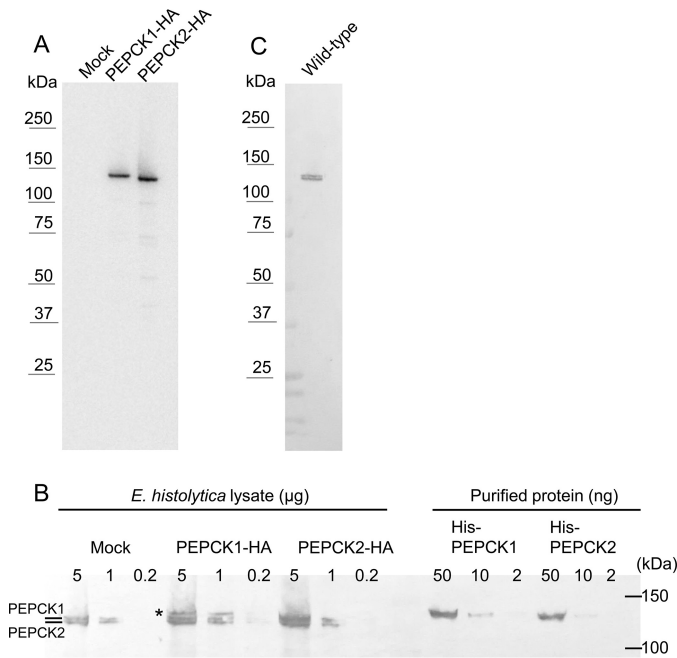
tagged proteins were confirmed by WB using anti-HA antibody (Fig. 3A). The molecular mass estimated from the band corresponding to EhPEPCK1-HA was slightly larger than that of EhPEPCK2-HA. We hypothesize that this subtle difference in the mobility between EhPEPCK1 and -2 is attributable to the

difference in the isoelectric point between the two proteins. The isoelectric point values of EhPEPCK1 and -2 are calculated as 6.15 and 6.06, respectively. At this stage, we do not know whether the slight difference in the isoelectric point and the molecular size between EhPEPCK1 and -2 is physiologically important.

**Heteromeric Configuration of EhPEPCKs**—To investigate whether EhPEPCKs form homo- or heterotrimer, EhPEPCK1 and -2 were immunoprecipitated using lysates from transformants expressing corresponding EhPEPCK with the HA tag using anti-HA antibody. Both of the immunoprecipitated samples contained peptides from all of the three EhPEPCKs (Table 3). When proteins were immunoprecipitated using EhPEPCK1-HA, EhPEPCK2/3-specific peptides were also detected, whereas the relative ratio decreased about 1.7-fold compared with that of trimetric EhPEPCK purified from the wild-type *E. histolytica*. Similarly, the immunoprecipitated sample using EhPEPCK2-HA also contained peptides unique to EhPEPCK1, whereas the relative ratio decreased 2.0-fold. These data suggest that some proportion of the trimer consists of a mixture of EhPEPCK1, -2, and -3. The detailed composition of the trimeric complex should be analyzed in the future.

**Detection of PP<sub>i</sub>-PEPCKs by WB Using Anti-EhPEPCK Antiserum**—The anti-EhPEPCK antiserum reacted with recombinant PEPCK1 and -2 with comparable efficiency (Fig. 3B). Major bands of the predicted molecular mass (~130 kDa) were observed in the lysate of *E. histolytica* (Fig. 3C), confirming the specificity of this anti-EhPEPCK antiserum. We estimated that EhPEPCKs account for ~1% of the total protein of the amebic lysate by comparing the band intensity of serially diluted recombinant PEPCKs and amebic lysates in WB analysis (Fig. 3B). This estimation also agrees well with the finding that about 200-fold purification of *E. histolytica* lysate yields highly purified PP<sub>i</sub>-PEPCK (Table 1 and Fig. 1A).

**Localization of PP<sub>i</sub>-PEPCKs in *E. histolytica* Trophozoites**—*E. histolytica* trophozoites were mechanically disrupted in sucrose-MOPS buffer and centrifuged to separate the cytosol and organelle/membrane fractions. The cytosolic marker, cys-



**FIGURE 3. WB analysis of EhPEPCKs with anti-HA (A) and anti-EhPEPCK (B and C) antibodies.** A, expression of EhPEPCK1-HA and EhPEPCK2-HA was confirmed. Note that proteins with the expected size (~130 kDa) were expressed. Amebic lysates containing 5  $\mu$ g of protein were applied per lane on SDS-PAGE. B, expression of EhPEPCKs in EhPEPCK1-HA- or EhPEPCK2-HA-expressing and mock control *E. histolytica* transformants. Total lysates of the above mentioned transformants (0.2–5  $\mu$ g/lane) and purified histidine-tagged PEPCK1 and -2 (2–50 ng/lane) were electrophoresed on SDS-PAGE and subjected to WB analysis with anti-EhPEPCK antiserum. Note that lysates from mock transformant showed two bands, whereas EhPEPCK1-HA-expressing transformant showed one additional band (shown with an asterisk) corresponding to PEPCK1 with the HA tag. Recombinant PEPCKs possess the His tag at the N terminus and were purified to homogeneity using Ni-NTA and MonoQ columns. C, specificity of anti-EhPEPCK antibody. Approximately 1  $\mu$ g of lysate from the wild-type *E. histolytica* was reacted with anti-EhPEPCK antibody.

**TABLE 3**  
Relative quantification of EhPEPCK isotypes

Peptide	Peak area of monoisotopic ion			Normalized relative ratio against common peptides		
	EhPEPCK1-HA IP	EhPEPCK2-HA IP	Purification from cells	EhPEPCK1-HA IP	EhPEPCK2-HA IP	Purification from cells
<b>Unique to EhPEPCK1</b>						
GTDYPILNIQEALADLK	4.54E+07	3.96E+06	4.61E+06			
SDAIDVVAPLVDIIAEGDQESTAPIDAR	4.09E+07	6.45E+06	2.17E+06			
NISDAIFEGK	2.41E+09	3.91E+08	6.71E+08			
Sum	2.50E+09	4.01E+08	6.78E+08	3.0E-01	5.5E-02	1.1E-01
<b>Specific to EhPEPCK2 and -3</b>						
ITTAFFANHFLR	9.22E+07	6.07E+08	1.68E+08			
KLESLANLK	3.81E+08	1.76E+09	8.54E+08			
LAGPLLEEVEESEVNHTTAPIDAR	2.62E+08	1.03E+09	2.54E+08			
Sum	7.35E+08	3.40E+09	1.28E+09	8.8E-02	4.6E-01	2.1E-01
<b>Unique to EhPEPCK2</b>						
AVQEIFDHDFSKR	5.79E+06	2.38E+07	2.05E+07	6.9E-04	3.3E-03	3.3E-03
<b>Unique to EhPEPCK3</b>						
GANLSSQYLR	2.14E+08	9.04E+06	3.37E+08	2.6E-02	1.2E-03	5.5E-02
<b>Common to all</b>						
YLIEHGYLEPCPDVITYNGK	2.15E+08	1.87E+08	3.52E+08			
LFQRPDDAVFR	4.41E+09	4.39E+09	2.91E+09			
DFFPAAK	3.71E+09	2.74E+09	2.92E+09			
Sum	8.34E+09	7.32E+09	6.18E+09	1.0	1.0	1.0



## Identification of PP<sub>i</sub>-type Phosphoenolpyruvate Carboxykinase

teine synthase 1 and the organelle marker, Cpn60, were demonstrated in the corresponding fractions (Fig. 4). When these fractions were reacted with anti-EhPEPCK antiserum, clear bands were detected from the supernatant fraction. Therefore, major parts of both EhPEPCK1 and EhPEPCK2/3 localize in the cytosol (Fig. 4), whereas two isoforms of GTP-PEPCK exist in the cytosol and mitochondria in the case of animals (38). *In silico* analysis to find organelle targeting signals (TargetP version 1.1 and SignalP 4.1) or transmembrane regions (TMHMM Server version 2.0) did not detect any targeting signals and thus was consistent with the result from the wet experiment described above.

**Distribution of PP<sub>i</sub>-PEPCK in the Tree of Life**—Homologs of EhPEPCKs were found in limited lineages of eukaryotes and bacteria but not in archaea (Tables 4 and 5). PP<sub>i</sub>-PEPCK genes were found in all of the three genomes of *Entamoeba* species that were currently present in GenBank<sup>TM</sup>, although the copy numbers are different among the three. The PP<sub>i</sub>-PEPCK gene was also found in the genome data of a free living amoebozoan, *M. balamuthi*, whereas the other amoebozoans *Acanthamoeba* and *Dictyostelium* lacked the gene. In addition to the two amoebozoan genera bearing PP<sub>i</sub>-PEPCK, PP<sub>i</sub>-PEPCK homologs were also present in some eukaryotes, such as excavates, dinoflagel-

lates, diatoms, and cryptophytes, which are distantly related to each other (39). In bacteria, a limited number of species in Actinobacteria, including *P. freudenreichii* used in this study, Verrucomicrobia, Bacteroidetes, Planctomycetes, Proteobacteria, Firmicutes, and Spirochaetes, possessed PP<sub>i</sub>-PEPCK homologs (Table 5); many species were found to lack PP<sub>i</sub>-PEPCK even if many of them were closely related to PP<sub>i</sub>-PEPCK-bearing species.

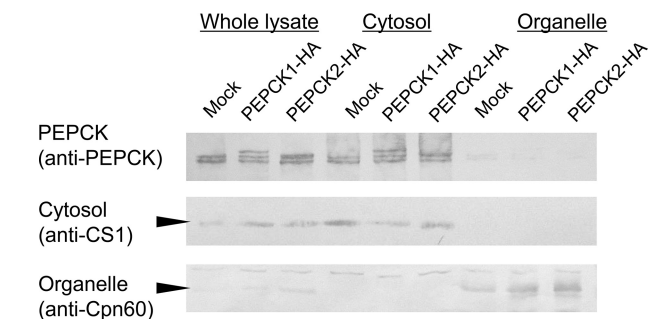
## Discussion

In the present study, we successfully identified the gene encoding PP<sub>i</sub>-PEPCK, which is one of the key enzymes in the central metabolism and connects sugar and organic acids metabolisms, from *E. histolytica* and *P. freudenreichii*. We also found that the homologs of EhPEPCK were distributed in lim-

**TABLE 5**  
Distribution of PP<sub>i</sub>-PEPCK in bacteria

Phylum	No. of genomes <sup>a</sup>	No. of PP <sub>i</sub> -PEPCK-containing genomes <sup>a</sup>
Actinobacteria	910	104
Aquificae	16	0
Armatimonadetes	3	0
Bacteroidetes/Chlorobi group	470	19
Caldiserica	2	0
Chlamydiae/Verrucomicrobia group	59	15
Chloroflexi	25	0
Chrysiogenetes	2	0
Cyanobacteria	104	0
Deferribacteres	6	0
Deinococcus-Thermus	43	0
Dictyoglomi	2	0
Elusimicrobia	2	0
Fibrobacteres/Acidobacteria group	46	1
Firmicutes	1129	1
Fusobacteria	25	0
Gemmatimonadetes	5	0
Nitrospinae	2	0
Nitrospirae	10	0
Planctomycetes	22	18
Proteobacteria	2200	33
Spirochaetes	80	2
Synergistetes	18	0
Tenericutes	140	0
Thermodesulfobacteria	7	0
Thermotogae	24	0
Unclassified bacterium	210	1

<sup>a</sup> In GenBank<sup>TM</sup>, accessed May 5, 2015.



**FIGURE 4. Distribution of *E. histolytica* PEPCKs by fractionation and WB analysis.** Trophozoites of mock, EhPEPCK1-HA-expressing, or EhPEPCK2-HA-expressing transformants were mechanically disrupted and centrifuged to separate into cytosolic and organelle fractions. Those fractions were reacted with anti-EhPEPCK, cysteine synthase 1 (CS1, a cytosolic marker), or Cpn60 (a mitochondrial matrix protein, used as an organelle marker) antibodies. Whereas ~5 μg of lysates were used for anti-PEPCK antibody, 20 μg of lysates were reacted with anti-CS1 and anti-Cpn60 antibodies.

**TABLE 4**  
Distribution of PP<sub>i</sub>-PEPCK and functional homologues in 10 represented genomes

One organism from every phylum that has a PP<sub>i</sub>-PEPCK homolog with the highest identity to EhPEPCK in a BLASTP search was listed. For bacteria, only phyla in which more than two organisms possess a PP<sub>i</sub>-PEPCK homolog were shown. +, ATP, and GTP, the organism has a protein that shows 35% or higher identity to biochemically analyzed PEP carboxylase (from *E. coli*, UniProtKB; P00864), ATP-PEPCK (from *E. coli*, UniProtKB; P22259), or GTP-PEPCK (from *G. intestinalis*, XP\_001709869.1), respectively.

Organism	Phylum	Identity <sup>a</sup>	PEPCK ATP/GTP	PEP carboxylase
%				
<b>Eukaryote</b>				
<i>E. histolytica</i> HM-1:IMSS	Amoebozoa	100		
<i>Naegleria gruberi</i> strain ATCC 30224	Heterolobosea	45	ATP	
<i>Guillardia theta</i> CCMP2712	Cryptophyta	44	ATP	+
<i>Thalassiosira pseudonana</i> CCMP1335	Heterokontophyta	44	ATP	+
<i>G. intestinalis</i> strain ATCC 50581	Metamonada	39	GTP	
<b>Bacteria</b>				
<i>Pedospaera parvula</i> Ellin514	Verrucomicrobia	48	GTP	
<i>Planctomyces maris</i> DSM 8797	Planctomycetes	46	ATP	
<i>Thioalkalivibrio nitratreducens</i> DSM 14787	Proteobacteria	46	GTP	+
<i>Actinomyces graevenitzi</i> C83	Actinobacteria	45	GTP <sup>b</sup>	
<i>Formosa agariphila</i> KMM 3901	Bacteroidetes	42	ATP	+ <sup>b</sup>

<sup>a</sup> Identity against EhPEPCK1.

<sup>b</sup> A protein with significant homology to the query, although the identity is lower than 35%.

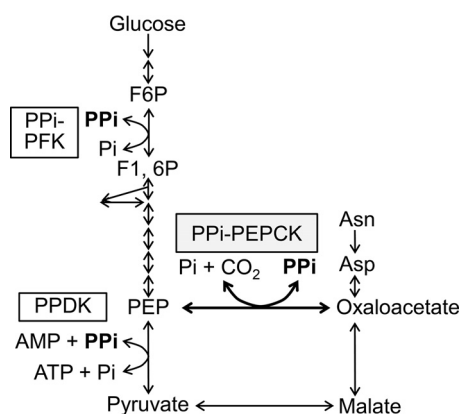


FIGURE 5. **Glycolytic pathway and oxaloacetate-malate node of *E. histolytica*.** Arrows, reaction steps and their directions conducted by the enzymes in *E. histolytica* (46, 47). F6P, fructose 6-phosphate; F1,6P, fructose 1,6-bisphosphate.

ited but diverse lineages of unicellular eukaryotes and bacteria but not in archaea.

In *E. histolytica*, which lacks ATP/GTP-PEPCK and PEP carboxylase,  $PP_i$ -PEPCK has been believed to work in the  $PP_i$ - and oxaloacetate-producing direction (27, 40). This is because a supply of  $PP_i$  is required to replenish  $PP_i$  consumed by the  $PP_i$ -utilizing enzymes in glycolysis,  $PP_i$ -PFK and PPDK (4, 12, 41). However, it is uncertain whether  $PP_i$  produced by  $PP_i$ -PEPCK is truly essential because  $PP_i$  is supplied as a by-product of many reactions *in vivo* (3, 42). Alternatively, it would be possible that  $PP_i$ -PEPCK works for the  $PP_i$ -utilizing direction when there is enough supply of  $PP_i$  because of the following reason. *E. histolytica* is able to consume asparagine and aspartate from media (43), and these amino acids can be converted to oxaloacetate by aminotransferase (Fig. 5) (44). Although the oxaloacetate catabolic pathway has not been experimentally confirmed in this organism, which lacks the tricarboxylic acid cycle, the amoeba may be capable of producing ATP from these amino acids by using the energy of  $PP_i$  if oxaloacetate is further converted to PEP by  $PP_i$ -PEPCK and then pyruvate by PPDK. Future experiments to investigate which direction is critical (or whether both directions are important) for viability in *E. histolytica* would give us deeper insight into the central metabolic pathway in this human parasite.

*E. histolytica* has three isoforms of  $PP_i$ -PEPCK, which are strongly suggested to form heterotrimers among the isoforms. This may be worthy of mention because the two other species of *Entamoeba* have two isoforms, and the four strains of the other eukaryotic human gut parasite *Giardia intestinalis*, for example, possess the single copy of this gene in the genomes (Fig. 6). Thus, it is still a key question whether each isoform of  $PP_i$ -PEPCK from *E. histolytica* has different roles in the heterotrimer and whether their roles are, if different, also shared among *Entamoeba* species. These issues should be examined in the future to consider the roles of  $PP_i$ -PEPCK in the human parasite *E. histolytica*.

Most of the organisms possessing  $PP_i$ -PEPCK homologs have canonical ATP/GTP-PEPCK genes, and some of them have PEP carboxylase as well (Table 4), although *E. histolytica* possesses the three cytosolic  $PP_i$ -PEPCKs but lacks the others. How these functional homologs are utilized/regulated in each

organism is an open question. They might work in different directions from each other, or they might work in different environmental conditions. Especially, in some eukaryotes, these functional homologs might be utilized in different cellular compartments. For instance, a potential  $PP_i$ -PEPCK from the unicellular alga *Thalassiosira pseudonana* (XP\_002285950.1) is predicted to be cytosolic according to the *in silico* predictions (data not shown) as well as those of *E. histolytica*, whereas its ATP-PEPCK was confirmed to exist in mitochondria, and two types of PEP carboxylase localize in mitochondria and the chloroplast, respectively (45). Although the localization of ATP-PEPCK and PEP carboxylase indicates that PEP-oxaloacetate conversion occurs only in organelles of the diatom, our findings suggest that this reaction important for the central carbon metabolism can also occur in the cytosol of *T. pseudonana*. Our findings in this study could further open a new era for deeper understanding the central metabolism of various organisms.

Identification of  $PP_i$ -PEPCK also allows us to compare the amino acid sequences between  $PP_i$ -PEPCK and the functional homologs, ATP/GTP-PEPCK and PEP carboxylase.  $PP_i$ -PEPCK bears none of the known catalytic domains, such as oxaloacetate binding sites, kinase 1a/P-loop, kinase-2 region, and nucleotide binding sites conserved in ATP- and GTP-PEPCK (22). In addition, no similarity in primary structure was observed between EhPEPCK and the functional homologs. Therefore, it is unlikely that  $PP_i$ -PEPCK shares a common origin with ATP- and GTP-PEPCK/PEP carboxylase. This is similar to the case between pyruvate kinase and PPDK/PEP synthase but different from the cases of ATP-PFK/ $PP_i$ -PFK and PPDK/PEP synthase that have most probably emerged from single origins. Our findings further illuminate the complexity of evolution among the proteins, which are functionally redundant but utilize different phosphate donors (*e.g.*  $PP_i$  and nucleotide triphosphates).

The origins and evolutionary histories of  $PP_i$ -PEPCKs were one of the most interesting aspects for us. However, it was very difficult to reconstruct a rooted tree of  $PP_i$ -PEPCK because we could not find any proteins with sister relationships to  $PP_i$ -PEPCK. In addition, whether the current distribution of  $PP_i$ -PEPCK in the tree of life has been shaped by lateral gene transfer or by vertical inheritance followed by independent gene loss basically remains unclear because of the low resolution of the  $PP_i$ -PEPCK tree (Fig. 6) with one exception for  $PP_i$ -PEPCK in *Dehalobacter* sp.; the *Dehalobacter* homologue was nested in the clade of Actinobacteria, although *Dehalobacter* belongs to Firmicutes, strongly suggesting actinobacterium-to-*Dehalobacter* lateral gene transfer.

In summary,  $PP_i$ -PEPCK was identified at the molecular level for the first time since the activity was first demonstrated more than 50 years ago (24). It was also demonstrated that  $PP_i$ -PEPCK was distributed in phylogenetically diverse but limited unicellular eukaryotes and bacteria, which would remodel the central metabolic pathways in various organisms.  $PP_i$ -PEPCK has no clear evolutionary relationship with ATP- or GTP-PEPCK, in sharp contrast to the case of  $PP_i$ - and ATP-PFK and that of PPDK and PEP synthase, which have evolved from single origins. These data suggest that, in order to substitute the



# Identification of PP<sub>i</sub>-type Phosphoenolpyruvate Carboxykinase



FIGURE 6. **Phylogenetic tree of PP<sub>i</sub>-PEPKs and their homologs.** Sequences were aligned by MAFFT, and the data set was subjected to maximum likelihood analysis with RAxML version 7.2.8. Only bootstrap supports equal to or more than 85 are shown on each node. *White letters in black boxes and black letters* indicate eukaryotic and bacterial phyla, respectively.

energy and phosphate donors from PP<sub>i</sub> to nucleotide triphosphates, proteins have employed various strategies.

**Author Contributions**—Y. C. and T. N. designed this study. Y. C. performed biochemical analysis. K. N.-T. and Y. S.-N. designed molecular experiments. R. K. and Y. C. performed the phylogenetic analysis. Y. C., R. K., and T. N. wrote the manuscript. All of the authors reviewed the results and approved the final version of the manuscript.

**Acknowledgment**—We thank Dr. Nicholas E. Sherman (W. M. Keck Biomedical Mass Spectrometry Laboratory, University of Virginia, Charlottesville, VA) for mass spectrometric analysis.

## References

- Liu, C. L., Hart, N., and Peck, H. D. (1982) Inorganic pyrophosphate: energy source for sulfate-reducing bacteria of the genus *Desulfotomaculum*. *Science* **217**, 363–364
- Gest, H. (1972) Energy conversion and generation of reducing power in bacterial photosynthesis. *Adv. Microb. Physiol.* **7**, 243–282
- Baykov, A. A., Malinen, A. M., Luoto, H. H., and Lahti, R. (2013) Pyrophosphate-fueled Na<sup>+</sup> and H<sup>+</sup> transport in prokaryotes. *Microbiol. Mol. Biol. Rev.* **77**, 267–276
- Mertens, E. (1991) Pyrophosphate-dependent phosphofructokinase, an anaerobic glycolytic enzyme? *FEBS Lett.* **285**, 1–5
- Alves, A. M., Meijer, W. G., Vrijbloed, J. W., and Dijkhuizen, L. (1996) Characterization and phylogeny of the *pfp* gene of *Amycolatopsis methanolica* encoding PP<sub>i</sub>-dependent phosphofructokinase. *J. Bacteriol.* **178**, 149–155
- Siebers, B., Klenk, H. P., and Hensel, R. (1998) PP<sub>i</sub>-dependent phosphofructokinase from *Thermoproteus tenax*, an archaeal descendant of an ancient line in phosphofructokinase evolution. *J. Bacteriol.* **180**, 2137–2143
- Mertens, E., Lador, U. S., Lee, J. A., Miretsky, A., Morris, A., Rozario, C., Kemp, R. G., and Müller, M. (1998) The pyrophosphate-dependent phosphofructokinase of the protist, *Trichomonas vaginalis*, and the evolutionary relationships of protist phosphofructokinases. *J. Mol. Evol.* **47**, 739–750
- Chi, A., and Kemp, R. G. (2000) The primordial high energy compound: ATP or inorganic pyrophosphate? *J. Biol. Chem.* **275**, 35677–35679
- Müller, M., Lee, J. A., Gordon, P., Gaasterland, T., and Sensen, C. W. (2001) Presence of prokaryotic and eukaryotic species in all subgroups of the PP<sub>i</sub>-dependent group II phosphofructokinase protein family. *J. Bacteriol.* **183**, 6714–6716
- Baptiste, E., Moreira, D., and Philippe, H. (2003) Rampant horizontal gene transfer and phospho-donor change in the evolution of the phosphofructokinase. *Gene* **318**, 185–191
- Pocalyko, D. J., Carroll, L. J., Martin, B. M., Babbitt, P. C., and Dunaway-Mariano, D. (1990) Analysis of sequence homologies in plant and bacterial pyruvate phosphate dikinase, enzyme I of the bacterial phosphoenolpyruvate: sugar phosphotransferase system and other PEP-utilizing enzymes: identification of potential catalytic and regulatory motifs. *Biochemistry* **29**, 10757–10765
- Reeves, R. E., South, D. J., Blytt, H. J., and Warren, L. G. (1974) Pyrophosphate:D-fructose 6-phosphate 1-phosphotransferase: a new enzyme with the glycolytic function of 6-phosphofructokinase. *J. Biol. Chem.* **249**, 7737–7741
- O'Brien, W. E., Bowien, S., and Wood, H. G. (1975) Isolation and characterization of a pyrophosphate-dependent phosphofructokinase from *Protonibacterium shermanii*. *J. Biol. Chem.* **250**, 8690–8695
- van der Merwe, M. J., Groenewald, J.-H., Stitt, M., Kossmann, J., and Botha, F. C. (2010) Downregulation of pyrophosphate:D-fructose-6-phosphate 1-phosphotransferase activity in sugarcane culms enhances sucrose accumulation due to elevated hexose-phosphate levels. *Planta* **231**, 595–608
- Michels, P. A., Chevalier, N., Opperdoes, F. R., Rider, M. H., and Rigden, D. J. (1997) The glycosomal ATP-dependent phosphofructokinase of *Trypanosoma brucei* must have evolved from an ancestral pyrophosphate-dependent enzyme. *Eur. J. Biochem.* **250**, 698–704
- Niersbach, M., Kreuzaler, F., Geerse, R. H., Postma, P. W., and Hirsch, H. J. (1992) Cloning and nucleotide sequence of the *Escherichia coli* K-12 *ppsA* gene, encoding PEP synthase. *Mol. Gen. Genet.* **231**, 332–336
- Davies, D. (1979) The central role of phosphoenolpyruvate in plant metabolism. *Annu. Rev. Plant Physiol.* **30**, 131–158
- Sauer, U., and Eikmanns, B. J. (2005) The PEP-pyruvate-oxaloacetate node as the switch point for carbon flux distribution in bacteria. *FEMS Microbiol. Rev.* **29**, 765–794
- Siu, P. M., and Wood, H. G. (1962) Phosphoenolpyruvic carboxytransphosphorylase, a CO<sub>2</sub> fixation enzyme from propionic acid bacteria. *J. Biol. Chem.* **237**, 3044–3051
- Bandurski, R. S., and Greiner, C. M. (1953) The enzymatic synthesis of oxaloacetate from phosphoryl-enolpyruvate and dioxide. *J. Biol. Chem.* **204**, 781–786
- Fukuda, W., Fukui, T., Atomi, H., and Imanaka, T. (2004) First characterization of an archaeal GTP-dependent phosphoenolpyruvate carboxykinase from the hyperthermophilic archaeon *Thermococcus kodakaraensis* KOD1. *J. Bacteriol.* **186**, 4620–4627
- Aich, S., and Delbaere, L. T. (2007) Phylogenetic study of the evolution of PEP-carboxykinase. *Evol. Bioinform. Online* **3**, 333–340
- Matte, A., Tari, L. W., Goldie, H., and Delbaere, L. T. (1997) Structure and mechanism of phosphoenolpyruvate carboxykinase. *J. Biol. Chem.* **272**, 8105–8108
- Siu, P. M., Wood, H. G., and Stjernholm, R. L. (1961) Fixation of CO<sub>2</sub> by phosphoenolpyruvic carboxytransphosphorylase. *J. Biol. Chem.* **236**, PC21–PC22
- Lochmüller, H., Wood, H. G., and Davis, J. J. (1966) Phosphoenolpyruvate carboxytransphosphorylase II: crystallization and properties. *J. Biol. Chem.* **241**, 5678–5691
- Haberland, M. E., Willard, J. M., and Wood, H. G. (1972) Phosphoenolpyruvate carboxytransphosphorylase: study of the catalytic and physical structures. *Biochemistry* **11**, 712–722
- Reeves, R. E. (1970) Phosphopyruvate carboxylase from *Entamoeba histolytica*. *Biochim. Biophys. Acta* **220**, 346–349
- Diamond, L. S., Harlow, D. R., and Cunnick, C. C. (1978) A new medium for the axenic cultivation of *Entamoeba histolytica* and other *Entamoeba*. *Trans. R. Soc. Trop. Med. Hyg.* **72**, 431–432
- Marumo, K., Nakada-Tsukui, K., Tomii, K., and Nozaki, T. (2014) Ligand heterogeneity of the cysteine protease binding protein family in the parasitic protist *Entamoeba histolytica*. *Int. J. Parasitol.* **44**, 625–635
- Nakada-Tsukui, K., Okada, H., Mitra, B. N., and Nozaki, T. (2009) Phosphatidylinositol-phosphates mediate cytoskeletal reorganization during phagocytosis via a unique modular protein consisting of RhoGEF/DH and FYVE domains in the parasitic protozoan *Entamoeba histolytica*. *Cell. Microbiol.* **11**, 1471–1491
- Nozaki, T., Asai, T., Sanchez, L. B., Kobayashi, S., Nakazawa, M., and Takeuchi, T. (1999) Characterization of the gene encoding serine acetyltransferase, a regulated enzyme of cysteine biosynthesis from the protist parasites *Entamoeba histolytica* and *Entamoeba dispar*: regulation and possible function of the cysteine biosynthetic pathway in *Entamoeba*. *J. Biol. Chem.* **274**, 32445–32452
- Nozaki, T., Asai, T., Kobayashi, S., Ikegami, F., Noji, M., Saito, K., and Takeuchi, T. (1998) Molecular cloning and characterization of the genes encoding two isoforms of cysteine synthase in the enteric protozoan parasite *Entamoeba histolytica*. *Mol. Biochem. Parasitol.* **97**, 33–44
- Mi-ichi, F., Abu Yousuf, M., Nakada-Tsukui, K., and Nozaki, T. (2009) Mitosomes in *Entamoeba histolytica* contain a sulfate activation pathway. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 21731–21736
- Keeling, P. J., Burki, F., Wilcox, H. M., Allam, B., Allen, E. E., Amaral-Zettler, L. A., Armbrust, E. V., Archibald, J. M., Bharti, A. K., Bell, C. J., Beszteri, B., Bidle, K. D., Cameron, C. T., Campbell, L., Caron, D. A., Catolico, R. A., Collier, J. L., Coyne, K., Davy, S. K., Deschamps, P., Dyhrman, S. T., Edvardsen, B., Gates, R. D., Gobler, C. J., Greenwood, S. J., Guida, S. M., Jacobi, J. L., Jakobsen, K. S., James, E. R., Jenkins, B., John, U., John-

## Identification of PP<sub>i</sub>-type Phosphoenolpyruvate Carboxykinase

- son, M. D., Juhl, A. R., Kamp, A., Katz, L. A., Kiene, R., Kudryavtsev, A., Leander, B. S., Lin, S., Lovejoy, C., Lynn, D., Marchetti, A., McManus, G., Nedelcu, A. M., Menden-Deuer, S., Miceli, C., Mock, T., Montresor, M., Moran, M. A., Murray, S., Nadathur, G., Nagai, S., Ngam, P. B., Palenik, B., Pawlowski, J., Petroni, G., Piganeau, G., Posewitz, M. C., Rengefors, K., Romano, G., Rumpho, M. E., Rynearson, T., Schilling, K. B., Schroeder, D. C., Simpson, A. G., Slamovits, C. H., Smith, D. R., Smith, G. J., Smith, S. R., Sosik, H. M., Stief, P., Theriot, E., Twary, S. N., Umale, P. E., Vault, D., Wawrik, B., Wheeler, G. L., Wilson, W. H., Xu, Y., Zingone, A., and Worden, A. Z. (2014) The marine microbial eukaryote transcriptome sequencing project (MMETSP): illuminating the functional diversity of eukaryotic life in the oceans through transcriptome sequencing. *PLoS Biol.* **12**, e1001889
35. Katoh, K., Misawa, K., Kuma, K., and Miyata, T. (2002) MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res.* **30**, 3059–3066
36. Stamatakis, A. (2006) RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* **22**, 2688–2690
37. Medina, V., Pontarollo, R., Glaeske, D., Tabel, H., and Goldie, H. (1990) Sequence of the *pckA* gene of *Escherichia coli* K-12: relevance to genetic and allosteric regulation and homology of *E. coli* phosphoenolpyruvate carboxykinase with the enzymes from *Trypanosoma brucei* and *Saccharomyces cerevisiae*. *J. Bacteriol.* **172**, 7151–7156
38. Croniger, C. M., Olswang, Y., Reshef, L., Kalhan, S. C., Tilghman, S. M., and Hanson, R. W. (2002) Phosphoenolpyruvate carboxykinase revisited: insights into its metabolic role. *Biochem. Mol. Biol. Educ.* **30**, 14–20
39. Adl, S. M., Simpson, A. G., Lane, C. E., Lukeš, J., Bass, D., Bowser, S. S., Brown, M. W., Burki, F., Dunthorn, M., Hampl, V., Heiss, A., Hoppenrath, M., Lara, E., Le Gall, L., Lynn, D. H., McManus, H., Mitchell, E. A., Mozley-Stanridge, S. E., Parfrey, L. W., Pawlowski, J., Rueckert, S., Shadwick, R. S., Shadwick, L., Schoch, C. L., Smirnov, A., and Spiegel, F. W. (2012) The revised classification of eukaryotes. *J. Eukaryot. Microbiol.* **59**, 429–493
40. Müller, M., Mentel, M., van Hellemond, J. J., Henze, K., Woehle, C., Gould, S. B., Yu, R.-Y., van der Giezen, M., Tielens, A. G., and Martin, W. F. (2012) Biochemistry and evolution of anaerobic energy metabolism in eukaryotes. *Microbiol. Mol. Biol. Rev.* **76**, 444–495
41. Reeves, R. E. (1968) A new enzyme with the glycolytic function of pyruvate kinase. *J. Biol. Chem.* **243**, 3202–3204
42. Kulaev, I. S., and Vagabov, V. M. (1983) Polyphosphate metabolism in microorganisms. *Adv. Microb. Physiol.* **24**, 83–171
43. Zuo, X., and Coombs, G. H. (1995) Amino acid consumption by the parasitic, amoeboid protists *Entamoeba histolytica* and *E. invadens*. *FEMS Microbiol. Lett.* **130**, 253–258
44. Clark, C. G., Alsmark, U. C., Tazreiter, M., Saito-Nakano, Y., Ali, V., Marion, S., Weber, C., Mukherjee, C., Bruchhaus, I., Tannich, E., Leippe, M., Sicheritz-Ponten, T., Foster, P. G., Samuelson, J., Noël, C. J., Hirt, R. P., Embley, T. M., Gilchrist, C. A., Mann, B. J., Singh, U., Ackers, J. P., Bhattacharya, S., Bhattacharya, A., Lohia, A., Guillén, N., Duchêne, M., Nozaki, T., and Hall, N. (2007) Structure and content of the *Entamoeba histolytica* genome. *Adv. Parasitol.* **65**, 51–190
45. Tanaka, R., Kikutani, S., Mahardika, A., and Matsuda, Y. (2014) Localization of enzymes relating to C4 organic acid metabolisms in the marine diatom, *Thalassiosira pseudonana*. *Photosynth. Res.* **121**, 251–263
46. Pineda, E., Encalada, R., Vázquez, C., González, Z., Moreno-Sánchez, R., and Saavedra, E. (2015) in *Amebiasis: Biology and Pathogenesis of Entamoeba* (Nozaki, T., and Bhattacharya, A., eds) pp. 351–372, Springer, Berlin
47. Jeelani, G., and Nozaki, T. (2014) Metabolomic analysis of *Entamoeba*: applications and implications. *Curr. Opin. Microbiol.* **20**, 118–124