Molecular Cloning and Functional Expression of Mannitol-1-phosphatase from the Apicomplexan Parasite Eimeria tenella*

Paul Liberator‡§, Jennifer Anderson‡, Marc Feiglin‡, Mohinder Sardana‡, Patrick Griffin‡, Dennis Schmatz‡, and Robert W. Myers‡

From the Departments of Parasite Biochemistry and Cell Biology and Molecular Design and Diversity, Merck Research Laboratories, Rahway, New Jersey 07065 and the Department of Biological Chemistry, Merck Research Laboratories, West Point, Pennsylvania 19486

A metabolic pathway responsible for the biosynthesis and utilization of mannitol is present in the seven species of Eimeria that infect chickens, but is not in the avian host. Mannitol-1-phosphatase (M1Pase), a key enzyme for mannitol biosynthesis, is a highly substrate-specific phosphatase and, accordingly, represents an attractive chemotherapeutic target. Amino acid sequence of tryptic peptides obtained from biochemically purified Eimeria tenella M1Pase was used to synthesize degenerate oligonucleotide hybridization probes. Using these reagents, a partial genomic clone and full-length cDNA clones have been isolated and characterized. The deduced amino acid sequence of E. tenella M1Pase shows limited overall homology to members of the phosphohistidine family of phosphatases. This limited homology to other histidine phosphatases does, however, include several conserved residues that have been shown to be essential for their catalytic activity. Kinetic parameters of recombinant M1Pase expressed in bacteria are essentially identical to those of the biochemically purified preparation from E. tenella. Moreover, recombinant M1Pase is subject to active site-directed, hydroxylamine-reversible inhibition by the histidine-selective acylating reagent diethyl pyrocarbonate. These results indicate the presence of an essential histidine residue(s) at the M1Pase active site, as predicted for a histidine phosphatase.

Protozoan parasites of the genus Eimeria are the causative agents of the intestinal disease known as coccidiosis. Coccidiosis occurs in a large number of domesticated and wild animals, but of major economic importance is the impact that Eimeria spp. have on the poultry industry. During acute infections, these parasites cause significant morbidity and mortality, resulting in a reduction in the productivity of broiler chickens. The life cycle of Eimeria is well characterized (1). Mature sporulated oocysts, the vegetative stage of the parasite, are ingested by the chicken. Infective sporozoites are released from oocysts in the intestine and rapidly invade epithelial cells lining the gut with a tropism that is unique for each species. Once intracellular, sporozoites expand in number and develop into merozoites in a process known as schizogony. Merozoites then proceed through a second round of asexual replication termed merogony, which ultimately causes much of the intestinal pathology characteristic of the disease. Upon re-infecting epithelial cells, second generation merozoites develop to form both macrogametes and microgametes. Microgametes travel to and fertilize macrogametes within neighboring cells. The resulting zygote develops intracellularly into an unsporulated oocyst, which is released from the debilitated host cell and is shed in the feces. Unsporulated oocysts must then mature, or sporulate, outside of the animal host to initiate the next round of infection.

Unsporulated oocysts are entirely dependent upon endogenous sources of energy to remain viable and to sporulate efficiently. The source of this energy in Eimeria is the alcohol sugar mannitol (2). Mannitol represents as much as 25% of the total dry weight of an unsporulated Eimeria tenella oocyst. During sporulation, there is greater than a 90% reduction in the level of this polyol. In vivo studies using inhibitors of mannitol biosynthesis have shown that this storage carbohydrate is essential for parasite viability.1

The metabolic pathway that is responsible for the biosynthesis and utilization of mannitol in E. tenella, the mannitol cycle, has been characterized (3, 4). All four mannitol cycle enzymes, namely mannitol-1-phosphate dehydrogenase, mannitol-1-phosphatase (M1Pase),2 mannitol dehydrogenase, and hexokinasase, are present in unsporulated and sporulated oocysts of E. tenella. Aside from protozoan parasites of the subphylum Apicomplexa, the mannitol cycle has only been described in certain fungi (5). Higher eukaryotic hosts do not synthesize or catabolize mannitol. It is this potential for selective intervention that initiated the development of a series of biochemical screens for mannitol cycle inhibitors.

In this regard, E. tenella M1Pase has been purified to homogeneity.3 The sequence of several tryptic peptides from this purified enzyme has allowed us to isolate and characterize genomic and cDNA clones that code for E. tenella M1Pase. This report represents the first molecular description of a mannitol-1-phosphatase from any organism. Each of the tryptic peptide

* The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The nucleotide sequence(s) reported in this paper has been submitted to the GenBank®/EBI Data Bank with accession numbers AF032462.

1 J. Allocco and H. Profous-Juchelka, unpublished data.
2 The abbreviations used are: M1Pase, mannitol-1-phosphatase; kb, kilobase(s); nt, nucleotide(s); USO, unsporulated oocysts; ORF, open reading frame; PCR, polymerase chain reaction; LC-ESI-MS, liquid chromatography-electrospray ionization-mass spectrometry; PAGE, polyacrylamide gel electrophoresis; PaseNR, recombinant expressed nonfusion M1Pase; PaseNRH, recombinant expressed COOH-terminal hexahistidine-tagged M1Pase; PNPP, p-nitrophenyl phosphate; DEPC, diethyl pyrocarbonate.
3 R. W. Myers, M. N. Feiglin, J. J. Allocco, and D. M. Schmatz, manuscript in preparation.
sequences can be found in the amino acid sequence deduced from the cDNA clone. We have functionally expressed recombinant M1Pase in bacteria both as a nonfusion protein and with six histidine residues at the COOH-terminal end. The recombinant enzymes are enzymatically indistinguishable from each other and, in turn, are very similar to native M1Pase purified from *E. tenella* unpurified oocytes. Using functional recombinant enzyme, we provide data which indicate that *E. tenella* M1Pase is a new member of the phosphohistidine phosphotransferase family of phosphatases.

**EXPERIMENTAL PROCEDURES**

**Materials**—Oligonucleotides were synthesized using an Applied Biosystems model 384 DNA synthesizer and purified according to the manufacturer's recommendations. Oligonucleotides were radioactively end-labeled with [α-32P]dATP (5000 Ci/mmol, Amersham Corp.) using polynucleotide kinase from Boehringer Mannheim. Enzymes used for the polymerase chain reaction (PCR) were purchased from Perkin-Elmer and Life Technologies, Inc. PCR products were gel-purified using Qiagen II (Qiagen, Inc.) and subcloned directly into the TA-clone vector (Invitrogen). Gel purified DNA fragments to be used as hybridization probes were random-primer labeled with [α-32P]dCTP (3000 Ci/mmol, Amersham Corp.). Unlabeled DNA was restriction enzyme digested and modifying enzymes were purchased from Life Technologies, Inc. and New England Biolabs. Electrocompetent bacterial cells were from Life Technologies, Inc. Precast polyacrylamide gels and protein molecular weight markers were purchased from Novex. The generation and characterization of polyclonal antisera prepared in rabbits against native *E. tenella* M1Pase is described elsewhere. Other standard DNA manipulations were carried out as described (9).

**Construction and Screening of *E. tenella* Subgenomic Library**—The following three degenerate oligonucleotides were synthesized based upon M1Pase tryptic peptide sequences: from peptide 21, 5'-TG GGT GCC GTC GAC YTC YTC GCG GAA GAA-3'; from peptide 36, 5'-GTT GTC CCA YTC GCC GAA GCC CAT YTC-3'; and from peptide 38, 5'-GTA GTC GCC GTC GAA GCC NAG GAA-3'. Only the oligonucleotide derived from peptide 36 (36-a) specifically hybridized to restriction enzyme digested *E. tenella* genomic DNA. Genomic DNA was preparatively digested with ApaI and resolved by gel electrophoresis, and the gel slice corresponding to the 3.5-kb hybridization signal was selected. Genomic DNA in this gel slice was extracted, ligated into ApaI-digested pBluescript KS (Stratagene), and used to transform electrocompetent DH10B bacteria. This subgenomic plasmid library was screened by colony lift hybridization using end-labeled oligonucleotide 36-a.

**Construction and Screening of cDNA Libraries**—Total RNA was prepared from unpurified oocytes (USO) of *E. tenella* (10). Poly(A) RNA was twice selected using oligo(dT)-cellulose chromatography and used as a template for first strand cDNA synthesis with SuperScript RT (Life Technologies, Inc.) and oligo(dT) or random hexamer primers. Following RT-cDNA synthesis, DNA synthesis and EcoRI digestion of double-stranded cDNA was size fractionated by column chromatography using Sephacryl S-500 HR and then ligated into the phage vector λ ZAP II (Stratagene). Ligation reaction products were then packaged using Gigapack III Gold packaging extract (Stratagene) according to the manufacturer's recommendations.

A total of 1.25 × 109 recombinant phage were screened using the 190-nat Pase 1–2 PCR product prepared using genomic clone 6L-11 as template. Several plaque-pure positive clones were isolated from each library and subcloned into pBluescript KS (Stratagene) by in vivo excision. Phagemid clones were characterized by restriction enzyme mapping and partial nucleotide sequence analysis.

Automated DNA sequencing was performed using an Applied Biosystems model 373 instrument with the Prism FS cycle sequencing kit. Nucleotide sequence analysis was carried out with the University of Wisconsin Genetics Computer Group Software Package (12) and the TFASTA protein homology program (13).

**Expression of Recombinant M1Pase**—Constructs were made for expression in bacteria of both a nonfusion M1Pase (PaseNR) and a M1Pase fusion protein with six histidine residues at the COOH-terminal end (PaseNRH). Inserts were prepared by PCR using cDNA clone T2a15 as template. Oligonucleotide primer Pase-N2F (5'-GGC GCC GAG ACT GAG TGG-3'), which includes the initiator methionine, was used for the 5' end of both constructs. The 3'-PCR primers were 5'-GGC CAA GCT TAG GGT TTA GCG TGT GGT-3' and 5'-GGC CAA GCT TAG TGG TGA TGG TGA TGG GTC GGT TTA GCG TGT GGT-3', for PaseNR and PaseNRH, respectively. Both of these span the translocation termination codon depicted in bold type. PCR reaction products were gel-purified, sequentially digested with Nool and HindIII, directionally cloned as 5'-Nool/HindIII fragments into the bacterial expression vector pQE60 (Qiagen, Inc.), and used to transform electrocompetent bacteria. Nucleotide sequence of the respective constructs was determined to verify the fidelity of the amplification reaction. Plasmids were then transformed into the bacterial host M15pREP4 for expression.

Bacteria harboring recombinant plasmid were grown in 2 × YT medium under carbenicillin (100 μg/ml) and kanamycin (25 μg/ml) pressure at 37 °C to an absorbance of 0.5 at 600 nm. Protein expression was induced with the addition of isopropyl-β-D-thiogalactopyranoside to a final concentration of 1 mM. After 3 h, the bacteria were collected by centrifugation and washed once with phosphate-buffered saline. For most applications, the cells were sonicated on ice (5 × 60 s) in buffer A (50 mM Tris-HCl, pH 7.5, 500 mM NaCl). Protein concentration was determined (Bio-Rad protein assay, microassay protocol with bovine serum albumin as standard) and samples were resolved by denaturing SDS-PAGE on a linear 4–20% (w/v) gradient gel. Fractionated denatured proteins were either stained or electroblotted to nitrocellulose for Western blot analysis. In some cases, nondenaturing PAGE was performed as described previously and gels were processed for *in situ* M1Pase activity. Native PAGE gels were also electroblotted to nitrocellulose for Western blot analysis, but in transfer buffer without methanol.

**Purification of Recombinant Expressed M1Pase**—Following sonication of the bacterial pellet containing histidine-tagged M1Pase, the suspension was centrifuged (15 min, 40,000 × g). The vast majority of immunopositive M1Pase remained in the supernatant, and this was injected onto a Hitrap chelating Sepharose column (1 ml, nickel form). The column was washed with 20 ml of buffer A at a flow rate of 1.0 ml/min, and fractions (1.0 ml) were collected immediately. The column was eluted using a 0–200 mM gradient of imidazole-HCl, pH 8.0, increasing at 5 mM/ml. Fractions 33–40 (nominal imidazole concentration of 65–110 mM) containing the major peak of purified M1Pase, as determined both by enzyme activity and Western blot immunoreactivity, were pooled, yielding 15.1 mg of protein from 0.56 g wet weight bacterial pellet.

**Bacterial Cells Expressing the Nonfusion M1Pase Construct** (1.17 g wet weight) were collected as described above and were sonicated (6 × 60 s) in 4 ml of buffer B (10 mM Hpes sodium, pH 7.4). The suspension was centrifuged (15 min, 40,000 × g), and the supernatant volume was increased with buffer B to 10.1 ml. The supernatant (85.9 mg of protein) was injected onto a Hitrap Q anion exchange column (5 ml, chloride form) equilibrated with buffer B. The flow rate was 0.2 ml/min, and fractions (1.0 ml) were collected immediately. The column was washed with buffer B (15 ml) and then eluted using a 0–150 mM gradient of NaCl in buffer B increasing at 5 mM/ml. Fractions 29–35 (nominal NaCl concentration of 65–110 mM) containing the major peak of purified M1Pase were pooled, yielding 49.4 mg of protein.

**Kinetic Analysis of M1Pase Species**—Methods for the continuous assay of M1Pase as well as the determination of *K*₅₀, *K*₉₀, and IC₅₀ values for inhibitors are described elsewhere.

**Diethyl Pyrocarbonate (DEPC) Inactivation Studies**—Unless otherwise noted, reactions (500 μl) were performed at 25 °C in 10 mM HEPES sodium buffer, pH 7.5, containing p-nitrophenyl phosphate (PNPP, 0.5 mM, 0.22 × *K*₅₀, adjusted to pH 7.5) and DEPC (0–1330 μM). DEPC was quantitated spectrophotometrically by reaction with iodoacetate (6) and delivered in acetoneitrile (final concentration, 1% v/v). Reactions were initiated with the addition of purified PaseNR (36 μg). Absorbance at 400 nm was recorded as a function of time for 20 min using a Beckman DU70 spectrophotometer. M1Pase activity in the absence of DEPC remained linear under these conditions. Background rates of nonenzymatic PNPP hydrolysis under identical conditions were subtracted. Corrected *A*₄₀₀ versus time data were analyzed by computer fitting to either linear (no DEPC present) or integrated first order equations (DEPC present) using methods previously described (14).

**Mass Spectrometry**—Liquid chromatography-electrospray ionization-mass spectrometry (LC-ESI-MS) analysis was performed using a Finnigan TSQ-7000 (San Jose, CA), as described previously (15). Protein was incubated in the presence of 100 mM dithiothreitol at 52 °C for 15 min, prior to analysis to LC-MS. Following treatment with dithiothreitol, samples were loaded on a C₁₃ reverse phase column (1 × 100 mm) at a flow rate of 100 μl/min with 0.075% aqueous trifluoroacetic acid and eluted with a gradient of 2–60% acetonitrile over 40 min. The effluent was fed directly to the electrospray interface of the mass spectrometer. Ions were detected throughout the entire LC gradient over a *m/z* (mass to charge ratio) range of 500–2000.
TABLE I

| Peptide | Amino acid sequence | Corresponding residues in the deduced amino acid sequence from the cDNA clones |
|---------|---------------------|--------------------------------------------------------------------------------|
| 21      | VFGQQVQDYYANNQLTQQQQQQQAAA | Val32–Ala117                                                                 |
| 22      | SSCPQQLEAYA          | Ser20–Ala10                                                                  |
| 24      | VAELKDKDPAH          | Val177–His187                                                                 |
| 25      | AVHTAQFVPDVNHPKFL    | Ala142–Phe156                                                                 |
| 28      | ICCQLWQSE           | Ile285–Glu294                                                                 |
| 34      | EFKPLELPDVICYIR      | Glu85–Arg83                                                                  |
| 36      | VLPILAEAPGDNWNA      | Val160–Arg175                                                                 |
| 38      | DQLGFLGFEDGY        | Asp202–Tyr203                                                                 |
| 39      | EFIPD                | Gly138–Asp132                                                                 |

RESULTS

Protein Sequence Analysis—Repeated attempts to sequence biochemically purified native M1Pase were unsuccessful, implying that the amino terminus was blocked. Accordingly, the protein was subjected to trypsic digestion and peptide fragments were resolved by reverse phase high performance liquid chromatography. Edman degradation analysis yielded the nine peptide sequences listed in Table I. The corresponding positions of the respective peptides within the protein sequence deduced from the nucleotide sequence of the cDNA clones encoding M1Pase are also illustrated beneath the schematic in Fig. 1B.

Cloning of the cDNA Encoding E. tenella M1Pase—Three degenerate antiensi oligonucleotides were synthesized based upon sequence contained within peptides 21, 36, and 38. These oligonucleotides were used as hybridization probes on Southern blots of E. tenella genomic DNA. Using several experimental conditions, only oligonucleotide 36-a (from peptide 36) gave discrete hybridization signals (Fig. 2A). Oligonucleotide 36-a was then used to screen an E. tenella cDNA library prepared from USO mRNA. Using identical hybridization conditions, no positives were detected in the 1.25 × 10⁶ recombinants that were screened.

Acknowledging the possibility that mRNA coding for M1Pase might not be present in this developmental stage of the parasite’s life cycle, E. tenella genomic DNA was preparatively digested with Apal and size-fractionated DNA was extracted from agarose gel slices. A subgenomic DNA library, characterized by restriction enzyme mapping, and four of the clones were fully sequenced.

The nucleotide and deduced amino acid sequence of the longest cDNA clone (T2a15, 1160 nt) are shown in Fig. 1B. Each of the nine trypptic peptide sequences generated from native M1Pase purified from E. tenella USOs can be located within this ORF (Table I and schematically illustrated in Fig. 1B). On the assumption that the first methionine within this ORF represents the translation start site, the cDNA codes for a protein of 309 amino acid residues with a calculated molecular mass of 34,737 Da. This is in good agreement with the biochemically purified protein. LC-ESI-MS analysis of native M1Pase gave a molecular mass value of 34,648 Da. From the deduced amino acid sequence, the theoretical mass for M1Pase that has lost its NH₂-terminal methionine residue and has been subsequently acetylated at the amino terminus is 34,648 Da, identical to the experimental value found for the native enzyme. Support for post-translational modification at the amino terminus of M1Pase comes from our earlier observation that the native enzyme is refractory to NH₂-terminal Edman sequencing. These data provide compelling evidence to suggest that the deduced amino acid sequence of M1Pase is correct and that the methionine residue corresponding to nucleotide positions 40–42 represents the initiation codon.

E. tenella M1Pase Shares Sequence Similarity with Phosphohistidine Phosphotransferases—The deduced amino acid sequence predicted by clone T2a15 did not exhibit significant homology to any data base entry. Enzymatic characterization of native M1Pase suggested that it might be related to the phosphohistidine phosphotransferase family of enzymes (17, 18). The deduced amino acid sequence of E. tenella M1Pase shows limited overall homology to members of this group of enzymes. Accordingly, manual alignments were done with some representative members (Table II). A highly conserved diagnostic sequence motif among this family is the so-called “RHG” domain (18). Phosphoenzyme trapping experiments have shown that the histidine residue in this motif is phosphorylated as a reaction intermediate (19–21). Site-directed mutagenesis has confirmed the critical role that both the arginine and the histidine residues play in the catalytic reaction (22, 23). E. tenella M1Pase shares this sequence motif at residues 81–83. Furthermore, a second arginine residue at a relatively conserved distance from the RHG domain has been functionally identified as a contributor to substrate phosphate binding (22). The arginine residue at position 141 of E. tenella M1Pase is in a compatible location and could prove to serve a similar role (Table II).

Recombinant Expression of E. tenella M1Pase in Bacteria—Two plasmid constructs were assembled for expression purposes in the bacterial host M15pREP4. Both a non-fusion M1Pase recombinant (PaseNR) and a fusion protein tagged at the COOH-terminal end with six histidine residues (PaseNRH) have been expressed. Upon induction of expression from these plasmids, a tremendous amount of protein with an approximate molecular mass of 35 kDa can be visualized by Coomassie Blue staining (Fig. 3A, lanes 5 and 7). These bands co-migrate with native M1Pase (Fig. 3A, lane 8) and are not apparent in uninduced bacteria (lanes 4 and 6), or in bacteria that carry only the nonrecombinant expression plasmid pQE60 (lanes 1 and 2). Western blot analysis of this gel (Fig. 3B) with polyclonal antisera raised against native M1Pase supports the conclusion that these predominant protein products in induced bacteria are indeed recombinant expressed enzyme. A single Western blot-positive species co-migrates in this gel system with the native enzyme. The Western blot also demonstrates that there is some limited expression from the plasmid in the uninduced controls (Fig. 3B, lanes 4 and 6). However, bacteria
carrying the nonrecombinant vector (Fig. 3B, lanes 1 and 2) show no immunoreactivity.

Many recombinant proteins have their NH₂-terminal methionine removed when expressed in a bacterial host (24). This appears to be the case with E. tenella M1Pase. Following purification, LC-ESI-MS analysis of non-histidine-tagged recombinant M1Pase yielded an experimental molecular mass of 34,603 Da. This is in excellent agreement with the theoretical molecular mass of 34,606 Da calculated for M1Pase that has lost its NH₂-terminal methionine residue. These results further serve to indicate that M1Pase is accurately synthesized in bacteria and that the recombinant protein is essentially identical in primary structure to the native enzyme.

Recombinant expressed M1Pase has also been fractionated by native PAGE and then assayed for enzymatic activity in situ, within the polyacrylamide matrix. The in-gel activity assay in Fig. 4A demonstrates that, following induction, bacteria harboring the E. tenella M1Pase ORF are characterized by robust enzyme activity (lanes 5 and 7). Furthermore, this activity comigrates with the activity stain generated with purified M1Pase.

Fig. 1. Genomic and cDNA clones encoding E. tenella M1Pase. Schematic representation of the alignment between the 3.42-kb genomic clone 6L-11 (A) and the 1160-nt cDNA clone T2a15 (B). Noncoding regions are depicted as narrow lines, and the ORF corresponds to the thick black lengths. The relative position of the hybridization probes used to isolate the respective clones, degenerate oligonucleotide 36-a and Pase 1–2, are indicated above clone 6L-11. The location of these probes is also underscored in the nucleotide sequence and deduced amino acid sequence of clone T2a15. Oligonucleotide 36-a corresponds to nucleotide positions 535–561, while Pase 1 (nt 94–117) and Pase 2 (nt 283–259). The locations of the tryptic peptide sequences within the ORF (see Table I) are illustrated below the schematic of cDNA clone T2a15.
fied native M1Pase (lane 8). Much like M1Pase purified from *E. tenella*, native PAGE is able to resolve two distinct bands of activity for the recombinant expressed enzymes. We have not been able to identify the biochemical feature responsible for this fractionation of activity. However, it is reproducible and it is observed for both native and recombinant enzyme preparations, but only in nondenaturing electrophoretic conditions.

A Western blot of the native activity gel illustrates that the regions of M1Pase activity are also immuno-reactive with anti-sera raised against the native enzyme (Fig. 4B). Bacteria carrying the nonrecombinant pQE60 expression vector have no detectable M1Pase enzyme activity (Fig. 4A, lanes 1 and 2), nor are they Western blot-positive (Figs. 3B and 4B, lanes 1 and 2).

**Enzymatic Properties of Recombinant Expressed M1Pase**—Selected enzymatic properties of purified recombinant expressed M1Pase (both the nonfusion and the COOH-terminally tagged versions) were measured and compared with those obtained for the native enzyme. As shown in Table III, the kinetic constants for the two recombinant enzyme preparations are virtually indistinguishable, suggesting that the COOH-terminal six histidine residues do not compromise phosphatase activity. Furthermore, the kinetic parameters are in good agreement with those reported for native M1Pase prepared from *E. tenella* USO. Sodium fluoride, an inhibitor of phosphatase activity, inhibits native M1Pase with an *I*_{50} value of 28 μM. The recombinant expressed enzymes are also sensitive to inhibition by sodium fluoride, with *IC*_{50} values of 54 μM and 44 μM for the nonfusion and His-tagged proteins, respectively. Taken together, these experiments convincingly demonstrate that *E. tenella* M1Pase is accurately expressed in bacteria and that it adopts a higher order structure that is similar to the native enzyme. This results in a catalytically active recombinant enzyme that is essentially indistinguishable in its properties from native M1Pase.

**Evidence for Essential Active Site Histidine Residue(s)**—Both the biochemical properties of M1Pase and its limited deduced amino acid sequence homology suggested that the enzyme may be a new member of the phosphohistidine phosphotransferase family. To further support this conclusion, inhibition of M1Pase by DEPC, a highly selective acylating reagent for histidine at pH 6–8 (25), was examined. These studies were facilitated by the availability of substantial quantities of purified PaseNR and by the observation that PNPP is an alternative, albeit poor M1Pase substrate suitable for continuously monitoring the time course of DEPC inhibition.

Incubation of recombinant M1Pase with DEPC resulted in rapid, time-dependent loss of catalytic activity exhibiting pseudo first order kinetics. The calculated rate constants (k_{obs}) increased linearly with increasing DEPC concentrations (Fig. 5), corresponding to a bimolecular rate constant for PaseNR modification by DEPC of 9.02 ± 0.41 M^{-1} s^{-1} in the presence of subsaturating levels of PNPP (0.5 mM, 0.22 * K*_{m}); near-saturating levels of PNPP (25 mM; 10.8 * K*_{m}; 91.5% occupancy) significantly reduced the rate of M1Pase inhibition by 440 μM DEPC (*k*_{obs} = 5.72 × 10^{-3} min^{-1}, 98.4% protection). The excellent agreement between the predicted and observed protection afforded by substrate (91.5% versus 98.4%) strongly suggests that M1Pase inhibition by DEPC is due to ethoxyformylation of active site residue(s).

Under mild conditions, hydroxylamine treatment deacylates only N-ethoxyformylhistidine in DEPC-modified proteins (25). Reactivation of DEPC-inhibited enzyme activity under these conditions is therefore considered compelling evidence that the DEPC-mediated inhibition is due to histidine modification. PaseNR incubated with DEPC and hydroxylamine was not reactivated. FHC inhibition by DEPC was reactivated in the presence of hydroxylamine (68% of control activity was obtained using 20 mM hydroxylamine). Facile reactivation of DEPC-inhibited PaseNR under these conditions is most consistent with ethoxyformylation of active site histidine residue(s) as the cause of enzyme inhibition.

**DISCUSSION**

We have isolated and characterized a partial genomic clone and several full-length cDNA clones from *E. tenella* that code for M1Pase, an enzyme critical for the physiological regulation of mannitol metabolism in this protozoan parasite. This report represents the first molecular description of a mannitol-1-phosphatase from any organism. The alignment between the genomic and cDNA clones serves to map the location of two intervening sequences within the genomic clone. Both of these introns (335 and 910 nt in length) possess characteristic sequence features that are critical for efficient pre-mRNA splicing in higher eukaryotes (reviewed in Ref. 26). These include canonical splice donor and splice acceptor sequences at the intron-exon boundaries, as well as a target sequence for 2'-5' lariat formation located upstream of the right splicing junctions. The genomic clone contains 1.56 kb of 5'-flanking sequence. Unfortunately, a functional assay to scan this length for promoter elements is not available for *Eimeria*.

Several pieces of data lead us to conclude that the ORF within clone T2a15 represents the primary structure of *E. tenella* M1Pase. A combined total of 22 clones isolated from two *Eimeria* cDNA libraries have in large part been sequenced. Eighteen of these have 5' ends, which map between −14 nt and −40 nt with respect to the proposed initiation codon. The 5' ends of the remaining four cDNA clones lie within the protein coding region. Each of the cDNAs has identical nucleotide sequence in the areas that overlap with clone T2a15. Inspection of the presumptive 5'-flanking sequence from genomic clone 61-11 uncovers termination codons in each reading frame prior to encountering the next upstream methionine residue. Unless there is an additional intron in this region of the genomic clone, it appears that the ATG triplet beginning at nucleotide position 40 is indeed the point of translation initiation. In support of this conclusion, this ATG codon is in a desirable nucleotide sequence context for eukaryotic translation initiation (27, 28).

The sequences of nine tryptic peptide fragments have been determined from native *E. tenella* M1Pase, and each of these can be found within the single extended ORF of clone T2a15.
The deduced amino acid sequence from clone T2a15 predicts a molecular weight value that is in good agreement with the LC-ESI-MS determination for the native enzyme. \(^3\) The recombinant protein expressed in \(E.\ coli\) from clone T2a15 is immunologically detected using polyclonal antisera raised against native M1Pase and has a molecular weight determined by LC-ESI-MS that is consistent with the deduced amino acid sequence. Taken together, these data strongly support the contention that cDNA clone T2a15 codes for a full-length M1Pase protein product. Most importantly, recombinant expressed M1Pase is enzymatically active with kinetic properties that are virtually indistinguishable from those of the native enzyme.

Enzymatic characterization of native M1Pase \(^3\) suggested that it might be a member of the phosphohistidine phosphotransferase family, enzymes whose reaction mechanism is characterized by the formation of a phosphohistidinyl-enzyme intermediate (17, 18). Detailed mechanistic information is available for several members of this group of enzymes, including phosphoglycerate mutase, acid phosphatase, and the bifunctional enzyme 6-phosphofructo-2-kinase/fructose-2,6-bisphosphatase. Each of these enzymes has a common active site array of amino acid residues (18). Covalent modification

Table II

| Enzyme species | Sequence surrounding Arg\(^{81}\) of M1Pase | Spacing between conserved arginyl residues | Sequence surrounding Arg\(^{141}\) of M1Pase |
|----------------|---------------------------------------------|-------------------------------------------|---------------------------------------------|
| \(E.\ tenella\) M1Pase | CYIRHGK | 59 | SSPLLRAVH |
| Rat liver Fru-2,6-P2 | YLCHRGES | 49 | TSHMKRTIQ |
| Human liver Fru-2,6-P2 | YLCHRGES | 49 | TSHMKRTIQ |
| Yeast PGM | VLRHIGQS | 51 | TSKLSRAIQ |
| Human muscle PGM | VMVRHGET | 51 | TSVLRKRA |
| \(E.\ coli\) acid Pase | IVSRHGV | 75 | ADVDERTK |
| Human prostatic acid Pase | LVPRHGD | 67 | STDVDRTL |

\(^a\) Accession numbers: rat liver Fru-2,6-P2, J04197; human liver Fru-2,6-P2, M19938; yeast PGM, X58789; human muscle PGM, M18172; \(E.\ coli\) acid Pase, X05471; human prostatic acid Pase, M34840.

**FIG. 3.** Expression of recombinant M1Pase in bacteria. Recombinant M1Pase expressed both as a nonfusion protein (PaseNR) and with a COOH-terminal tag of six histidine residues (PaseNRH) were induced for 3 h at 37 °C with isopropyl-\(\beta\)-D-galactopyranoside in the bacterial host M15pREP4. Following sonication, proteins were resolved by electrophoresis. A Coomassie stain of a 4–20% SDS-PAGE gel with the molecular size standards (kDa) noted in the right margin is shown in A. Western blot analysis of a similar SDS-PAGE gel probed with polyclonal antisera raised against native M1Pase is shown in B. Samples loaded are as follows: uninduced bacteria carrying the nonrecombinant pQE60 vector (1), induced bacteria carrying nonrecombinant pQE60 (2), native MCI extract (3), uninduced PaseNR (4), induced PaseNR (5), uninduced PaseNRH (6), induced PaseNRH (7), purified native M1Pase (8), and prestained protein molecular weight standards (9).

**FIG. 4.** Functional expression of recombinant M1Pase in bacteria. M1Pase expressed in bacteria has also been resolved in nondenaturing PAGE gels. In A, fractionated proteins were assayed for M1Pase activity within the acrylamide matrix. A Western blot of a native PAGE gel screened with anti-M1Pase sera is illustrated in B. Samples loaded are as follows: uninduced bacteria carrying the nonrecombinant pQE60 vector (1), induced bacteria carrying nonrecombinant pQE60 (2), native MCI extract (3), uninduced PaseNR (4), induced PaseNR (5), uninduced PaseNRH (6), induced PaseNRH (7), and purified native M1Pase (8).
TABLE III
Comparison of the enzymatic properties of recombinant M1Pase with native enzyme prepared from E. tenella unsporulated oocysts

| Property | Native M1Pase | Nonfusion M1Pase | His-tagged M1Pase |
|----------|---------------|------------------|------------------|
| $K_m$ ($\mu$m) | 69.0 ± 16.4 | 57.3 ± 17.3 | 67.4 ± 6.2 |
| $V_{max}$ ($\mu$mol mg$^{-1}$ min$^{-1}$) | 953 ± 57 | 723 ± 54 | 738 ± 14 |
| $k_{cat}$ (s$^{-1}$) | 350 ± 30 | 420 ± 30 | 430 ± 10 |
| pH optimum | 6 | 6 | 6 |

FIG. 5. Inhibition of M1Pase by DEPC. Pseudo first order rate constants ($k_{obs}$) describing the time dependent inhibition of PaseNR (36 µg) by various concentrations of DEPC at 25 °C in 10 mM HEPES-sodium buffer, pH 7.5, containing 0.5 mM PNPP as substrate. Circles represent experimental data points. The slope of the linear fit corresponds to a second order rate constant of 9.02 ± 0.41 M$^{-1}$ s$^{-1}$ ($r^2 = 0.992$).

and phosphoenzyme trapping experiments have demonstrated the presence of one or more histidine and arginine residues at the active site (19, 29–33). Sequence alignments have allowed for the identification of an absolutely conserved sequence motif of RHG at the active site histidine. The adjacent arginine residue is believed to function in substrate phosphate binding. Site-directed mutagenesis experiments have confirmed the critical roles that these residues perform (22, 23, 34, 35). More recently, the catalytic mechanism suggested by chemical and mutagenic studies has been reinforced by solving crystal structures (36–38). While the overall sequence similarity is not striking, E. tenella M1Pase does share this signature sequence at Arg$^{140}$-His$^{141}$-Gly$^{142}$. A second arginine, located 40–50 residues toward the COOH-terminal end in this family of enzymes, is critical for stabilization of the substrate phosphate group during its transfer to water. Arg$^{141}$ within E. tenella M1Pase might act in this capacity.

Direct evidence for the existence of the predicted, essential active site histidine residue(s) in M1Pase was obtained by chemical modification studies using DEPC, a selective acylating reagent for histidine (25). Both prostatic acid phosphatase (29) and fructose-2,6-bisphosphatase (19) are subject to active site-directed inhibition via DEPC modification of histidine. In the latter case, DEPC was shown to prevent phosphohistidine intermediate formation. M1Pase was similarly found to be effectively inhibited by DEPC in a time- and concentration-dependent manner exhibiting pseudo first order kinetics. Lack of saturation kinetics established that a kinetically significant M1Pase-DEPC complex did not form at DEPC concentrations sufficient to cause rapid inhibition. This is not surprising, as M1Pase is expected to have low affinity for DEPC. However, substrate (PNPP) competitively protected M1Pase from DEPC inhibition, indicating that ethoxyformylation of active site residue(s) is responsible for the inhibition. When corrected for the presence of substrate, the calculated bimolecular rate constant for M1Pase inhibition by DEPC was 10.97 ± 0.50 M$^{-1}$ s$^{-1}$. This is in excellent agreement with the bimolecular rate constant for acylation of imidazole by DEPC, 54 ± 2 M$^{-1}$ s$^{-1}$ (6). Reversal of DEPC inhibition by mild hydroxylamine treatment is considered diagnostic for histidine modification as the cause of the observed inhibition (25). Accordingly, the essentially instantaneous partial reactivation of DEPC inhibited M1Pase by 10 mM hydroxylamine provides strong additional evidence that inhibition is due to $N$-ethoxyformylation of histidine. Thus M1Pase appears to possess an essential active site histidine residue(s), as predicted for a histidine phosphatase. Additional support for this conclusion is based on the observation that site-directed mutagenesis of M1Pase His$^{82}$, the predicted phosphate acceptor, ablates enzyme activity (7). Furthermore, homology-based modeling predicts several other potential active site residues in M1Pase that correspond to conserved catalytic residues in this family of enzymes (8). Using this information to guide us, we are currently evaluating the enzymatic consequences of mutating these residues. Taken together, these results suggest that M1Pase is a new member of the phosphohistidine phosphotransferase family.

Free-living unsporulated Eimeria oocysts cannot exchange organic nutrients with their environment. Mannitol, present in enormously high concentrations in these oocysts, serves as the primary metabolic fuel to support the developmental process of sporulation. The biosynthesis and consumption of mannitol is accomplished in a metabolic pathway that is quite unique evolutionarily. Due to the fact that higher eukaryotic hosts do not synthesize or catabolize mannitol, the mannitol cycle represents an attractive drug target in coccidian parasites. In vivo studies in E. tenella using an inhibitor of mannitol biosynthesis have demonstrated that this storage carbohydrate is essential for parasite viability and that selective intervention is indeed possible. Accordingly, screens for inhibitors of the enzymes involved in Eimeria mannitol metabolism using biochemically prepared native enzymes have been established. Several first generation inhibitors of M1Pase have been identified in one such screen. These compounds are also efficacious with comparable IC$_{50}$ values when the recombinant enzyme is assayed (6). Since the recombinant enzyme is much easier to prepare and has properties that are very similar to those of the native enzyme, the recombinant is now the source of enzyme used in our drug discovery effort.

M1Pase has been biochemically characterized in each of the major species of Eimeria that infect chickens. The enzyme is also present in Toxoplasma gondii and Cryptosporidium parvum oocysts, both apicomplexan parasites that infect multiple mammalian hosts including man. Ultimately, the information that is gained from mechanistic and structural studies using recombinant expressed E. tenella M1Pase should prove useful in the rational design of inhibitors for M1Pases from apicomplexan protozoa. The ability to inhibit M1Pase activity will have significant benefits in the efforts to control these human and animal pathogens.

REFERENCES
1. Levine, N. D. (1982) in The Biology of Coccidia (Long, P. L., ed) pp. 1–30, University Park Press, Baltimore, MD
2. Schmatz, D. M., Arison, B. H., Dashkevicz, M. P., and Turner, M. J. (1988) Mol. Biochem. Parasitol. 29, 29–36
3. Schmatz, D., Baginsky, W. F., and Turner, M. J. (1989) Mol. Biochem. Parasitol. 32, 263–270
4. Schmatz, D. M. (1989) Parasitol. Today 5, 205–208
5. Hult, K., and Gatenbeck, S. (1978) Eur. J. Biochem. 88, 697–612
6. Melchior, W. B., and Fahrney, D. (1970) Biochemistry 9, 251–258
7. J. Allesco, M. A. Powles, and R. Myers, unpublished data.
8. M. Feiglin and S. Flattery, unpublished data.
Cloning and Expression of Parasite Mannitol-1-phosphatase
Molecular Cloning and Functional Expression of Mannitol-1-phosphatase from the Apicomplexan Parasite *Eimeria tenella*

Paul Liberator, Jennifer Anderson, Marc Feiglin, Mohinder Sardana, Patrick Griffin, Dennis Schmatz and Robert W. Myers

*J. Biol. Chem.* 1998, 273:4237-4244.
doi: 10.1074/jbc.273.7.4237

Access the most updated version of this article at http://www.jbc.org/content/273/7/4237

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

This article cites 36 references, 13 of which can be accessed free at http://www.jbc.org/content/273/7/4237.full.html#ref-list-1