# MDP-1: A Novel Eukaryotic Magnesium-Dependent Phosphatase ‡

**ARTICLE** *in* BIOCHEMISTRY · AUGUST 2000

Impact Factor: 3.02 · DOI: 10.1021/bi0005052 · Source: PubMed



**2 AUTHORS:**

Jeremy Selengut
University of Maryland, College Park
**55** PUBLICATIONS **12,291** CITATIONS
SEE PROFILE

Rodney Levine
National Institutes of Health
**186** PUBLICATIONS **16,903** CITATIONS
SEE PROFILE

# MDP-1: A Novel Eukaryotic Magnesium-Dependent Phosphatase‡

Jeremy D. Selengut* and Rodney L. Levine

*Laboratory of Biochemistry, National Heart Lung and Blood Institute, National Institutes of Health, Bethesda, Maryland 20892-0320*

*Received March 6, 2000; Revised Manuscript Received May 10, 2000*

ABSTRACT: We report here the purification, cloning, expression, and characterization of a novel phosphatase, MDP-1. In the course of investigating the reported acid phosphatase activity of carbonic anhydrase III preparations, several discrete phosphatases were discerned. One of these, a magnesium-dependent species of 18.6 kDa, was purified to homogeneity and yielded several peptide sequences from which the parent gene was identified by database searching. Although orthologous genes were identified in fungi and plants as well as mammalian species, there was no apparent homology to any known family of phosphatases. The enzyme was expressed in *Escherichia coli* with a fusion tag and purified by affinity methods. The recombinant enzyme showed magnesium-dependent acid phosphatase activity comparable to the originally isolated rabbit protein. The enzyme catalyzes the rapid hydrolysis of *p*-nitrophenyl phosphate, ribose-5-phosphate, and phosphotyrosine. The selectivity for phosphotyrosine over phosphoserine or phosphothreonine is considerable, but the enzyme did not show activity toward five phosphotyrosine-containing peptides. None of the various substrates assayed (including various nucleotide, sugar, amino acid and peptide phosphates, phosphoinositides, and phosphodiesters) exhibited $K_M$ values lower than 1 mM, and many showed negligible rates of hydrolysis. The enzyme is inhibited by vanadate and fluoride but not by azide, cyanide, calcium, lithium, or tartaric acid. Chemical labeling, refolding, dialysis, and mutagenesis experiments suggest that the enzymatic mechanism is not dependent on cysteine, histidine, or nonmagnesium metal ions. In recognition of these observations, the enzyme has been given the name magnesium-dependent phosphatase-1 (MDP-1).

Phosphomonoesterases are enzymes that catalyze the hydrolysis of phosphoester bonds of a wide variety of phosphorylated substrates such as sugars, proteins, lipids, and nucleic acids (*1*). These enzymes play various crucial roles in the biochemistry of all organisms including basic metabolism, signal transduction, and metabolic regulation. Although the chemistry of phosphoester cleavage is conceptually straightforward, a diverse group of enzymes has evolved utilizing numerous mechanistic strategies to carry out this task (*2*). Various criteria have been employed to classify phosphatases including, substrate type, substrate specificity, pH optimum, size, type of catalytic mechanism, the identity of the enzymatic nucleophile (or lack thereof), and, most recently, sequence homology (*1*−*7*).

In the course of studying the reported phosphatase activity of carbonic anhydrase III (CAIII),[1] several discrete acid phosphatases were discerned (*8*). One of these was notable in its low molecular weight and requirement for magnesium ions. In the present study, this phosphatase has been purified to homogeneity and partially sequenced. The sequence appeared to be novel, so a mouse homologue was cloned and overexpressed. Detailed studies of this recombinant's substrate specificity, inhibitor profile, cofactor requirements, and mechanistic characteristics confirm that it is a heretofore unknown species. Accordingly, the enzyme has been given the name magnesium-dependent phosphatase-1 (MDP-1).

## MATERIALS AND METHODS

*Materials.* Materials were purchased from the following sources: young (8−12 weeks old) rabbit muscle tissue from Pel-Freez Biologicals (Rodgers, AK); *Achromobacter lyticus* lysyl endopeptidase from Wako (Richmond, VA); human thrombin and recombinant enterokinase proteases from Novagen (Madison, WI); cloning vectors, calmodulin affinity purification resin, and competent cells from Stratagene (La Jolla, CA); gels and buffers for SDS−PAGE from Novex (San Diego, CA); *taq* polymerase from Roche Molecular Biochemicals (Indianapolis, IN); deoxynucleotide triphosphates from Perkin-Elmer (Norwalk, CT); magnesium chloride and sodium chloride from Quality Biologicals (Gaith-

‡ The nucleotide sequence coding for mouse MDP-1 has been deposited in the GenBank database (accession number AF230273).

* To whom correspondence should be addressed at Laboratory of Biochemistry, NHLBI, NIH, Building 3-118, 3 Center Drive, Bethesda, MD 20892. Phone: (301)-496-8708. Fax: (301)-496−0599. E-mail: selengut@nih.gov.

[1] Abbreviations: CAIII, carbonic anhydrase isozyme III; CBP, calmodulin binding peptide; CM, carboxymethyl; DEPC, diethylpyrocarbonate; DTNB, 5,5′-dithiobis(2-nitrobenzoic acid); DTT, dithiothreitol; EDTA, ethylenediaminetetraacetic acid; EGTA, ethylene glycol-bis-(*β*-aminoethyl ether) *N*,*N*,*N*′,*N*′-tetraacetic acid; EST, expressed sequence tag; HPLC, high-pressure liquid chromatography; IPTG, isopropyl *β*-D-thiogalactopyranoside; LIC, ligation-independent cloning; Lys-C, lysyl endopeptidase; MALDI-TOF-MS, matrix-assisted laser desorption ionization-time-of-flight-mass spectrometry; MDP-1, magnesium-dependent phosphatase-1; MES, 2-(*N*-morpholino)ethanesulfonic acid; ORF, open reading frame; PCR, polymerase chain reaction; *p*NPP, *para*-nitrophenyl phosphate; PTPase, protein-tyrosine phosphatase; PTP-1B, protein-tyrosine phosphatase 1B; SDS−PAGE, sodium dodecyl sulfate−polyacrylamide gel electrophoresis; STS, sequence-tagged site; $TNB^-$, 5-thio-2-nitrobenzoate anion.

10.1021/bi0005052 
Published on Web 06/21/2000

ersburg, MD); EDTA and Tris from Digene Diagnostics (Beltsville, MD); guanidine hydrochloride and ammonium sulfate from Gibco-BRL (Grand Island, NY); acetone, calcium chloride, and lithium chloride from Mallinckrodt, Inc. (Paris, KY); nickel chloride and manganese chloride from Alfa Products (Danvers, MA); cobalt chloride from Fisher Scientific (Fair Lawn, NJ); ammonium molybdate and HPLC-grade acetonitrile from J. T. Baker (Philipsburg, NJ); IPTG from ICN Biomedicals (Aurora, OH); 2′-deoxy-5′-AMP, 5′-ADP, [Tyr($PO_3H_2$)]$^4$-Angiotensin-II, and H-Ala−Glu−Tyr($PO_3H_2$)−Ser−Ala-OH from Calbiochem/Novabiochem (San Diego, CA); Ac-Asp−Tyr($PO_3H_2$)−Val−Pro−Met−Leu-$NH_2$ and Ac-Ile−Tyr($PO_3H_2$)−Gly-−Glu−Phe-$NH_2$ from Bachem (King of Prussia, PA); H-Gly−Gly−Tyr($PO_3H_2$)−Gly−Gly-OH from SynPep Corporation (Dublin, CA); phosphocholine and phosphoethanolamine from the California Foundation for Biochemical Research (Los Angeles, CA); dibutanoylglyceryl-phosphatidylinositol-3-phosphate and dibutanoylglyceryl-phosphatidylinositol-4,5-bisphosphate from Echelon Research Laboratories (Salt Lake City, UT); and Tween-20 from Pierce (Rockford, IL). All other chemicals were purchased from Sigma (St. Louis, MO). Recombinant human PTP-1B was prepared by Z. Y. Zhang (Albert Einstein College of Medicine, Bronx, NY) and kindly provided by Dr. P. Boon Chock (NHLBI, NIH, Bethesda, MD).

*Purification of MDP-1 from Rabbit Muscle Carbonic Anhydrase III.* Rabbit muscle CAIII was purified following the method of Noltmann et al. (*9*) with modifications as follows. Muscle tissue (5 kg) was ground in a mechanical meat grinder and homogenized with 10 mM KCl solution (10 L) in a Waring blender at 4 °C. The homogenate was allowed to stand in the low ionic strength buffer for 30 min and then clarified by filtration through cheesecloth followed by centrifugation at 3500*g* for 30 min at 4 °C. The clarified homogenate was dialyzed (Spectrapor 1, 6000-8000 MWCO, Spectrum Medical Industries, Los Angeles, CA) twice versus 10 vol of buffer A (pH 6.90 phosphate containing 10 mM $Na^+$, 1 mM EDTA, and 1 mM DTT) and centrifuged a second time.

A total of 50 g of carboxymethyl (CM) Sephadex C-50 resin (Pharmacia, Piscataway, NJ) equilibrated in buffer A was added to the supernatant and allowed to stand at room temperature with occasional stirring for 30 min. Unbound proteins were washed from the beads with a total of 5 L of buffer A. Bound proteins were eluted with a total of 5 L of buffer B (pH 6.90 phosphate containing 40 mM $Na^+$, 1 mM EDTA, and 1 mM DTT). The eluted protein was dialyzed twice versus 3.5 vol of buffer A. One-quarter of this material was filtered through Whatman no. 1 filter paper, applied to a column containing 7.5 g of equilibrated CM-Sephadex C-50, and eluted with a 1600-mL gradient from buffer A to buffer C (pH 6.90 phosphate containing 80 mM $Na^+$, 1 mM EDTA, and 1 mM DTT). After re-equilibration of the column back to buffer A, the remaining dialysate was similarly processed in three additional portions. Eluted fractions were assayed for acid phosphatase assay with *p*NPP as substrate (see below) and CAIII content by SDS−PAGE. Fractions containing CAIII were pooled (CAIII-CM) as were earlier-eluting fractions containing high phosphatase activity (early-CM). Both pooled fractions were concentrated on YM-10 membranes at 40 psi (Amicon, Beverly, MA) to 40 mL. The early-CM pool was further concentrated to 1 mL in a Centriprep-10 concentrator (Amicon, Beverly, MA).

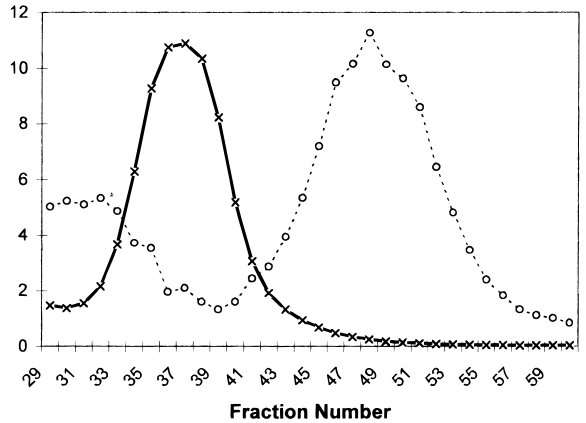


FIGURE 1: Gel filtration separates a magnesium-dependent phosphatase from carbonic anhydrase III. (×), solid line, protein (mg/mL). (○), dashed line, acid phosphatase activity with pNPP substrate (arbitrary units).

The concentrated CAIII-CM pool was gel-filtered in 1.5-mL portions over a Superdex-75 FPLC column (Pharmacia, Piscataway, NJ) equilibrated in buffer D (50 mM Tris-HCl, pH 8.0, 200 mM NaCl, 5 mM DTT, and 1 mM EDTA) at 0.5 mL/min using a model 1050 Hewlett-Packard high-pressure liquid chromatography system. The concentrated early-CM pool was similarly processed in 100-$\mu$L portions. MDP-1 eluted after CAIII in this system (Figure 1). MDP-1-containing fractions were pooled and concentrated using Centricon-10 concentrators to 0.5 mL. This pool was still >90% CAIII, so an additional round of gel filtration was performed and again phosphatase-containing fractions were pooled and concentrated. This final gel filtration pool contained about 20% CAIII along with three other discrete but overlapping low-molecular weight peaks, none of which precisely coeluted with the peak of phosphatase activity.

A portion of the final MDP-1 gel filtration pool was brought to 1 M ammonium sulfate by the addition of 1 vol of 2 M ammonium sulfate and bound to a phenyl Sepharose hydrophobic interaction chromatography column (Phenyl-5PW, HP TosoHaas, Montgomeryville, PA) equilibrated in buffer E (50 mM Tris-HCl, pH 8.0, 1 M ammonium sulfate, 5 mM DTT, and 1 mM EDTA). Proteins were eluted by a descending salt gradient (−125 mM ammonium sulfate/min) at a flow rate of 1 mL/min. Two hundred microliter fractions were collected and assayed for phosphatase activity. A single peak of phosphatase activity eluting at ∼500 mM ammonium sulfate was found, which correlated closely with one of the peaks observed at 276 nm (Figure 2).

A sample of the phenyl Sepharose-purified phosphatase fraction was mixed with 5 vol of 6 M guanidine-HCl containing 500 mM potassium phosphate at pH 2.5, centrifuged at 10000*g* in a tabletop centrifuge, and applied to a model 1100 high-pressure liquid chromatography system equipped with a C18 reverse phase column (218TP54, Vydac, Hesperia, CA). The initial solvent was 0.05% trifluoroacetic acid, and gradient elution was effected with 0.05% trifluoroacetic acid/acetonitrile at 5%/min up to 40% acetonitrile, at 0.5%/min up to 46% acetonitrile, and at 5%/min up to 95% acetonitrile at a flow rate of 1 mL/min. In this system, MDP-1 eluted at 44% acetonitrile.

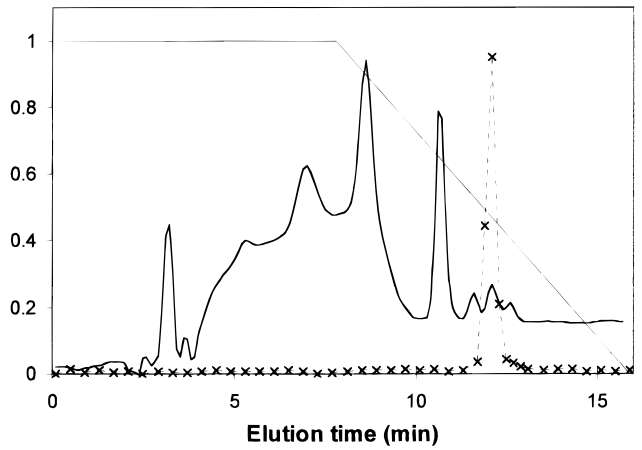


FIGURE 2: Phenyl-Sepharose hydrophobic interaction chromatography resolves the magnesium-dependent phosphatase activity. Thick solid line, protein ($A_{276}$). (×), thin dashed line, acid phosphatase activity with *p*NPP substrate ($A_{405}$). Thin solid line, ammonium sulfate concentration, descending from 1 to 0 M.

*Protease Cleavage and Amino Acid Sequencing*. A portion of the final MDP-1 gel filtration pool was unfolded by the addition of 8 M guanidine-HCl to a final concentration of 5 M guanidine in the presence of 125 mM Tris, pH 8.0, 5 mM DTT, and 1.5 mM EDTA. Protein cysteines were carboxymethylated by the addition of freshly dissolved iodoacetamide to a final concentration of 20 mM followed by incubation at 37 °C for 45 min. Residual unreacted iodoacetamide was quenched by the addition of an excess of DTT. After standing for 5 min at room temperature, the solution was acidified by the addition of a 2 M sodium phosphate solution, pH 2.5, to a final concentration of 500 mM. The carboxymethylated MDP-1 was purified by reverse-phase HPLC as above (note that the protein shifts to an elution at 40.5% acetonitrile after carboxymethylation and was detected by MALDI-TOF of the eluted fractions). The MDP-1-containing fractions were evaporated to dryness in a centrifugal evaporator (Speed-Vac SC110, Savant) and redissolved in 50 $\mu$L of 100 mM Tris, pH 8.5, 1 mM EDTA, and 10% acetonitrile. The protein was cleaved by the addition of 0.6 $\mu$L of 0.5 $\mu$g/$\mu$L *A. lyticus* lysyl endopeptidase and incubation at 37 °C overnight. The cleaved peptide mixture was mixed with 30 $\mu$L of a 4 M guanidine-HCl, 1 M sodium phosphate, pH 2.5 solution, and the peptides were separated by reverse phase HPLC as above using a 1%/min gradient. Masses of the HPLC-purified fragments were determined by MALDI-TOF. The peptide fragments were then subjected to Edman degradation on a Hewlett-Packard model G1000A protein sequencer using the manufacturer's version 3.5 chemistries.

*Database Searching, Translation, and Alignment of Sequences*. The program FASTA3 was used to search the SwissProt database using the computers at the European Bioinformatics Institute (http://www2.ebi.ac.uk). The program TFASTA was used to search the six-frame translation of the GenBank database using the computers at the Institute Génétique Humaine (http://www2.igh.cnrs.fr). The program CLUSTALW was used to align the mouse and human nucleotide sequences and to construct a consensus sequence for each. Some of these ESTs represent immature mRNAs containing apparent introns flanked by GU and AG splice sites; as they were found in minority of the sequences, they were not included in the consensus. The consensus mouse and human nucleotide sequences were conceptually translated in the three forward frames using the program TRANSLATE.

*Cloning and Expression of MDP-1*. The mouse clone (g1514302) whose reported sequence spanned the longest stretch of the MDP-1 coding region (as compared to the above consensus, all but the final 18 nucleotides) was chosen as a template. A strain of *Escherichi coli* (DH10B) containing this clone on a pT7T3D plasmid vector was obtained from the IMAGE consortium via ATCC (no. 931495, Manassas, VA) and cultured on LB-ampicillin plates at 37 °C. A ligation-independent cloning strategy was employed to express MDP-1 as an N-terminal fusion with a calmodulin-binding peptide (kit 214405, Stratagene, La Jolla, CA). Primers that are complementary to the kit-supplied linearized cloning vector overhangs as well as the ends (omitting the ATG start codon) of the mouse MDP-1 open reading frame (as determined from the consensus) were obtained from Biosynthesis, Inc. (Lewiston, TX). PCR was carried out as follows. A small number of cells from a single colony of the *E. coli* strain were transferred with a sterile toothpick to 10 $\mu$L of distilled water in a thin-walled PCR tube. To this tube was added 1 $\mu$L of each primer; 10× PCR buffer (H0817, Perkin-Elmer); magnesium chloride to a final concentration of 2.5 mM; dATP, dCTP, dGTP, and dTTP each to a final concentration of 0.1 mM; 0.025 U of Amplitaq polymerase (N808, Perkin-Elmer); and water to give a final volume of 50 $\mu$L. The temperature program consisted of an initial cycle of 97 °C for 2.5 min, 65 °C for 3 min, and 72 °C for 5 min, followed by 10 cycles of 94 °C for 1 min and 72 °C for 1.5 min, followed by 25 cycles of 94 °C for 30 s and 72 °C for 1.5 min, and finally a single step of 72 °C for 10 min. Agarose gel electrophoresis indicated a single product of the expected size (522 base pairs).

The PCR product was purified using the QIAquick PCR purification kit (28104, Qiagen, Valencia, CA) and then digested with *Pfu* DNA polymerase in the presence of dATP (as per the Stratagene kit's instructions) to prepare overhangs complementary to the linearized, similarly digested cloning vector. The digested PCR product and the vector were allowed to anneal and then used to transform Epicurian coli XL1-Blue MRF' competent cells (200249, Stratagene). Colonies that grew in the presence of ampicillin were grown in liquid culture overnight at 37 °C, a portion of which was mixed with sterile glycerol and frozen at −70 °C. The remainder was used to prepare plasmid using a miniprep kit (12123, Qiagen).

This plasmid was used as template for a further round of PCR with the same primers to produce DNA for confirmatory sequencing. PCR products were purified by agarose gel electrophoresis, passively eluted from gel slices, and concentrated by ethanol precipitation. Both strands were sequenced by Bio-Synthesis, Inc. (Lewisville, TX), using a fluorescent dye-based chemistry.

The miniprep purified plasmid was used to transform the expression strain, BL21(DE3)-pLysS. Transformed colonies were grown in liquid culture (LB containing 50 $\mu$g/mL ampicillin and 34 $\mu$g/mL chloramphenicol) to OD = 0.6 and induced with 1 mM IPTG for 3 h. The cells were chilled at 4 °C for 15 min, centrifuged for 10 min at 2500*g*, the supernatant decanted, and the pellet frozen at −70 °C until needed.

*Purification of Recombinant CBP-MDP-1*. Frozen, IPTG-induced BL21(DE3)-pLysS cells expressing CBP-MDP-1 from 1 L of culture were resuspended in 100 mL of buffer F (50 mM Tris-HCl, pH 8.0, 150 mM NaCl, 10 mM $\beta$-mercaptoethanol, 1 mM magnesium acetate, 1 mM imidazole, and 2 mM calcium chloride) additionally containing 5 $\mu$g/mL aprotinin, 1 $\mu$g/mL leupeptin and 200 $\mu$g/mL lysozyme. The cells were rocked for 15 min at room temperature and then sonicated at 40% power for 5 × 30 s with 5 min rests on ice (model VC40 power supply with model ASI probe, Sonics & Materials, Inc, Danbury, CT). Cell debris was removed by centrifugation for 10 min at 7500*g*. The supernatant was decanted, mixed with 40 mL of a 1:1 slurry of calmodulin-affinity resin (214303, Stratagene) in buffer F, and rocked at room temperature for 2 h. The beads were spun down briefly at 100*g*, and the unbound material was decanted. The beads were similarly rinsed with four 80-mL portions of buffer F. Bound proteins were eluted with three portions of buffer G (50 mM Tris-HCl, pH 8.0, 1 M NaCl, 2 mM EGTA, and 10 mM $\beta$-mercaptoethanol) and were then pooled.

The affinity-purified material was chilled on ice, brought to 35% saturation with ammonium sulfate, and centrifuged for 10 min at 17000*g*. The supernatant was decanted, brought to 55% saturation with ammonium sulfate, and centrifuged again. The supernatant was discarded, and the pellet was redissolved in 6 mL of buffer H (50 mM Tris-HCL, pH 8.0, 50 mM NaCl, and 0.1% Tween-20), which was centrifuged again.

*Cleavage of CBP Affinity Tag with Thrombin*. Ammonium sulfate concentrated fusion protein at 0.7 mg/mL was treated with 10 U of thrombin/mg of protein in the presence of 10 mM $CaCl_2$ for 24 h at room temperature. The protein was once again dialyzed versus buffer H supplemented with 1 mM DTT overnight at 4 °C. The resulting solution was centrifuged at 14000*g*, frozen in liquid nitrogen in small aliquots, and stored at −70 °C until use.

*Mass Analysis*. Electrospray mass spectroscopy was performed with a Hewlett-Packard model G1946A instrument interfaced to a model 1100 high-pressure liquid chromatography system equipped with a Vydac narrow bore C18 column (218TP5205, Vydac, Hesperia, CA). The initial solvent was 0.05% trifluoroacetic acid and 10% acetonitrile, and gradient elution was effected with 0.05% trifluoroacetic acid/acetonitrile at 1%/min at a flow rate of 0.2 mL/min. The effluent from the column was mixed in a tee with neat acetic acid delivered by another 1100 series pump (100 $\mu$L/min), and the mixture was introduced into the mass spectrometer (*10*). MALDI-TOF was performed on a Hewlett-Packard model G2025A instrument using a 4-cyano-α-hydroxycinnamic acid matrix (Hewlett-Packard, Palo Alto, CA).

*Acid Phosphatase Assays*. Four different assay methodologies were employed depending on the particular substrates used. Unless otherwise noted, the assay buffer contained 200 mM MES, pH 5.5, 10 mM magnesium chloride, and 2 mM EDTA. When only *p*NPP was used as a substrate (i.e., in the assay of chromatographic fractions), the assay was performed as follows: 5 mg/mL *p*NPP was allowed to react with phosphatase in a total volume of 40 $\mu$L at 37 °C in a 96-well microtiter plate. The reaction was quenched by the addition of 20 $\mu$L of 10M NaOH. $PNP^-$ anion was detected at 405 nm ($\epsilon_M = 1.85 \times 10^4$). For the comparison of various substrates, a modification of the Fiske-SubbaRow (phosphomolybdate) method was used to detect inorganic phosphate in which excess molybdate was complexed by citrate, and the phosphomolybdate was measured without reduction after addition of acetone to the reaction mixture (*11*). Quantitation in this assay is afforded by comparison with a standard curve of inorganic phosphate concentrations run in parallel. In cases in which higher sensitivity was required, a commercial malachite green assay was utilized (17-262, Upstate Biotechnology, Lake Placid, NY). Assay of phosphotyrosine-containing peptides was carried out using a fluorimeter (QuantaMaster Luminescence Spectrometer, Photon Technology International, Monmouth Junction, NJ) to measure the increase in the fluorescence due to tyrosine ($\lambda_{ex} = 280$ nm, $\lambda_{em} = 305$ nm).

*Residue-Specific Reagents*. 5,5′-dithiobis(2-nitrobenzoic acid) (DTNB) is used to specifically label cysteine residues in a spectroscopically quantifiable manner (*12*). Affinity-purified CBP-MDP-1, which contains a single cysteine (the CBP tag does not contain any cysteine residues), was reduced with dithiothreitol (DTT, 10 mM), precipitated three times with 3 M ammonium sulfate, and redissolved in 100 mM Tris, pH 8, 200 mM NaCl, and 2 mM EDTA to remove excess DTT. The enzyme, thus prepared, at a final concentration of 100 $\mu$g/mL, was reacted with 40 $\mu$g/mL DTNB for 30 min at 25 °C. The amount of labeling was determined by monitoring the absorbance of the $TNB^-$ at 412 nm in comparison with the amount of protein as determined by absorbance at 280 nm (relative to controls lacking protein, reagent, or both). Activity after labeling was determined after one additional ammonium sulfate precipitation followed by solubilization in buffer H.

Diethylpyrocarbonate (DEPC) was used to specifically label histidine residues in a spectroscopically quantifiable manner (*13*). Affinity-purified CBP-MDP-1 that contains six histidines (the CBP tag does not contain any histidine residues) was dialyzed versus 50 mM MES, pH 5.5, 50 mM NaCl, 2 mM $MgCl_2$, and 0.1% Tween-20 and then centrifuged at 12000*g*. The enzyme, at 400 $\mu$g/mL, was reacted sequentially with three aliquots of DEPC, each equal to 2 molar equivalents. Light scattering prevented accurate monitoring of the reaction by spectroscopy at 240 nm. After each aliquot, the reaction was allowed to proceed for 30 min at 25 °C, at which time a sample was removed for phosphatase activity assays and LC-MS determination of labeling.

*Unfolding/Refolding Experiments*. Affinity-purified CBP-MDP-1 was unfolded by the addition of guanidine hydrochloride to a final concentration of 6 M. After 5 min at 25 °C, the enzyme was refolded by a 50-fold dilution in buffer H additionally including 1 mM DTT. The presence of Tween-20 in the dilution buffer is critical; no activity is recovered in its absence. The residual guanidine (120 mM) caused a 40% decrease in activity as determined in separate experiments. The addition of EDTA in the refolding solution improves the recovered activity, perhaps by chelating inhibitory metal ions.

*Site-Specific Mutants*. Single amino acid mutations in MDP-1 were introduced using the QuickChange site-directed mutagenesis kit (200518, Stratagene). In brief, miniprep-purified plasmid from the wild-type MDP-1 expression strain (see above) was used as a template for PCR using a pair of self-complementary primers containing the desired mutation.

Wild-type template DNA was selectively digested with *Dpn*I, a restriction enzyme recognizing methylated DNA. The resulting PCR product was a nicked circular double-stranded plasmid which, when transformed into Epicurian Coli (XL1-Blue) supercompetent cells (200236, Stratagene), was repaired and amplified. Subsequent manipulations were identical to the wild-type expression methodology (see above).

## RESULTS

*Purification and Characterization of Rabbit Muscle MDP-1*. CAIII purified by the method of Noltmann et al. (*9*) had a specific acid phosphatase activity of 0.36 nmol $min^{-1}$ $mg^{-1}$ in the absence of magnesium and 0.68 nmol $min^{-1}$ $mg^{-1}$ in the presence of 8 mM magnesium. Assay of the individual fractions eluting from a Superdex-75 gel filtration column during the purification shows clearly that the magnesium-dependent activity was due to a separable lower molecular weight species (Figure 1). Pooling of these low molecular weight gel filtration fractions and rechromatography by gel filtration enriched this low molecular weight phosphatase with respect to CAIII and other components but did not completely resolve the phosphatase. Phenyl-Sepharose hydrophobic interaction chromatography separates this mixture into several components, one of which appeared to coelute with the phosphatase activity (Figure 2). When subjected to reverse-phase chromatography, this fraction was dominated by a single peak (>90% pure as judged by the area of the eluted peaks). This reverse-phase eluted protein, when allowed to refold by dilution in aqueous buffer, recovered 16% of the original phosphatase activity, while no other fraction showed significant activity. The phenyl-Sepharose-purified phosphatase shows a molecular mass of 18.6 kDa by MALDI-TOF-MS, a specific activity of 20 $\mu$mol *p*NPP $min^{-1}$ (mg of protein)$^{-1}$ (6 $s^{-1}$) at pH 5.5 in the presence of 8 mM $MgCl_2$ at 37 °C, and a pH optimum of 5.1 and is inhibited by vanadate and fluoride ion but not by calcium or lithium (each at 1 mM). The enzyme is not inactivated by treatment with the cysteine-reactive reagents, iodoacetamide or DTNB. Activity was not detectable in the absence of added magnesium, and magnesium exhibited an apparent dissociation constant of 1 mM. There are no significant differences between these properties and those of the mouse recombinant enzyme, which is discussed below.

*Sequence Determination*. A Lys-C protease peptide map was prepared from carboxymethylated, reverse-phase purified MDP-1, and three fragments were sequenced by Edman degradation (Table 1). The SwissProt protein database contains a single protein with significant homology to these peptides, the "hypothetical" yeast protein, YEW4_YEAST (Yeast Protein Database ID: YER134c). Alignment of the three peptide fragments with this sequence is shown in Figure 3. It became apparent from this alignment that the lysine terminating peptide 2 is probably the one preceding peptide 3.

Joining these two peptides to give a longer search query, and using a TFASTA search (which translates a nucleotide database in all six reading frames) versus GenBank found a large number of EST and cDNA sequences from mouse and human in addition to three rat sequences from various tissues (Table 2). One human STS (sequence-tagged site) has been published (GenBank Accession No. 1343614) localizing the gene to chromosome 14. The mouse and human nucleotide

Table 1: Edman Sequencing of Reverse Phase-Separated Lys-C Fragments[a]

| pept | elution (% acetonitrile) | Edman sequence[b] |
|---|---|---|
| 1 | 35 | [K]TGLPFSQMIFFDDEK |
| 2 | 37 | [K]LAVFDLDYTLWPFWVDTHVDPPFHK |
| 3 | 42 | [K]GSDGTVRD(RG)R(VLG)Q(NH)... |

[a] A total of 675 pmol of IAA-treated, reverse phase-purified phosphatase was treated with 0.3 $\mu$g of Lys-C protease at 37 °C for 16 h and quenched with 8 M guanidine, and the resulting peptides were separated on a C-18 reverse phase column in a 1% acetonitrile/min gradient. [b] Amino acids in brackets are implied by cleavage with the Lys-C enzyme and the fact that the uncleaved phosphatase is N-terminally blocked as isolated. Amino acids in parentheses are ambiguous calls at a single sequencing cycle.

```
             |-------Peptide #2--------||--Peptide #3-->

RABBIT     KLAVFDLDYTLWPFWVDTHVDPPFH..KGSDG...TVRDRRGQN
            :|:||||||:|| : |||:  ||:  |:|:|   |:  | | :
YEW4_YEAST DVAAFDLDYTIWPCYCDTHLHGPFKPVKSSNGEVLTIICRDGYE
               |         |         |         |
              10        20        30        40

             |-Peptide #1--|

RABBIT     KTGLPFSQMIFFDDEK
            : |   ::::||||
YEW4_YEAST TSDLKSKKICLFDDES
             |         |
            120       130
```

FIGURE 3: Alignment of rabbit MDP-1 peptides with corresponding regions of the *S. cerevisiae* homologue, YEW4_YEAST (SwissProt #P40081). Vertical bars indicate identical amino acids, and colons indicate homologous residues.

sequences were grouped, and consensus sequences were determined. Translation of these consensus sequences each yielded a single open reading frame containing all three of the regions homologous to the rabbit peptides. These open reading frames were of identical length (163 amino acids, not including the initiating methionine) and correspond to proteins whose expected molecular masses (mouse, 18.45 kDa; human, 18.63 kDa) were similar to that observed for the purified rabbit protein (18.6 kDa, by MALDI-TOF). These results lend confidence that the consensus nucleotide sequence contains no frame-shift errors.

The full-length YER134c and mouse (hypothetical) protein sequences, used as queries for further searches, yield sequences from cognate MDP-1's in *Arabidopsis thaliana* (mouse-ear cress), *Lycopersicon esculentum* (tomato), *Sorghum bicolor* (sorghum wheat), *Zea mays* (corn), and *Schizosaccharomyces pombe* (Table 3). Alignment of all of these MDP-1 protein sequences produces a consensus that indicates several regions of particularly high conservation (Figure 4).

*Cloning, Expression, and Affinity Purification of Mouse CBP-MDP-1 Fusion*. A ligation independent cloning (LIC) strategy was employed, as detailed in Materials and Methods, to create an N-terminal fusion containing a calmodulin-binding peptide affinity purification tag. This fusion contains a thrombin protease cleavage site 10 amino acids from the fusion joint and an enterokinase cleavage site flush to the fusion joint. The recombinant protein was expressed in *E. coli* [BL21(DE3)-pLysS] induced with IPTG. The expressed protein was purified by affinity chromatography and concentrated by ammonium sulfate precipitation. SDS−PAGE shows the resulting protein to be >95% pure (Figure 5). The plasmid insert was proofed by nucleotide sequencing (Ac-

Table 2: Distribution of Mammalian MDP-1 ESTs by cDNA Library Tissue Source[a]

| tissue source | GenBank accession nos. | | | | | | |
|---|---|---|---|---|---|---|---|
| blastocyst | *2670347* | | | | | | |
| B-cell | 2788466[d] | | | | | | |
| brain | *4967588* | 574605[c,d] | 765910[c,d,e] | 888647 | 888648[d] | 1979057[e] | |
| colon | 2025062[b] | 4081177[b,d] | | | | | |
| diaphragm | *2346217* | | | | | | |
| germ cell | 6475479[b] | | | | | | |
| heart | 2835038[c,d] | **3397125** | 3676042[c] | | | | |
| kidney | 3058680[d] | 3094719 | *3260823*[d] | *3299558*[d] | 3872885 | 3870058 | 4002814[b] |
| | 4005179[b,d] | 4071566[b] | 4371586[d] | 4664652[d] | 5234166 | 5424528[b] | 5525993 |
| | 5740785[b] | 5741472[b] | 6140436 | 6504822[b] | 6569775[e] | 6576458 | 6658680 |
| | 6709842 | 6946375[b] | | | | | |
| liver/spleen | 3250254[c] | | | | | | |
| lung | 4988444[b] | 4991234[b] | 6143425[b,e] | 6838768[b] | | | |
| lymph node | *1777079*[d] | | | | | | |
| mammary | *2102380* | *2164928* | *2813699* | *3158643* | *3685110* | *4768008* | |
| ovary | **3830070** | **4252222** | | | | | |
| prostate | 5234166 | 5769299 | | | | | |
| stem cell | *5241027*[c] | *5244887*[c] | | | | | |
| thymus | *1681498*[d] | *3175333* | | | | | |
| T-cell (Jurkat) | 1963141[d] | | | | | | |
| uterus | 1645284 | 1645396 | 3645635 | | | | |
| mixed tissues | *1282277*[c] | *1290625*[c] | *1325095*[c] | *1333151*[c] | *1381995*[c] | *1429147*[c] | *1500704*[c] |
| | *1504764*[c] | *1514302*[c] | *1542021*[c] | *1542662*[c] | 1894518 | 1892155 | 3871649 |
| | 3837694 | 4087931[f] | 4189552 | 5369752[c] | 6139292[d] | | |

[a] Italicized entries are from mouse, bold entries are from rat, and other entries are from human. [b] Cancerous tissue source. [c] Embryonic, fetal, or infant tissue source. [d] Contains an apparent intron, which, if translated, would result in transcript truncation. [e] Contains a deletion of 46 amino acids and a frame-shift resulting in an alternative C-terminus.

Table 3: Nonmammalian MDP-1 Genes

| species | GenBank accession no. | DNA source | chromosome |
|---|---|---|---|
| *Saccharomyces cerevisiae* | 1384128 | genomic (ORF predicted) | V |
| *Schizosaccharomyces pombe* | 3810855 | genomic (ORF predicted) | II |
| *Arabidopsis thaliana* | 4587681 | genomic (ORF predicted)[a] | II |
| *Lycopersicon esculentum* | 4379953 4382873<br>5276153 5278720<br>5279112 5279113<br>5896195 5899583<br>6533614 6534029 | cDNA library | |
| *Sorghum bicolor* | 5410347 | genomic (ORF predicted) | J |
| *Zea mays* | 5915346 6056748<br>6194890 6828196 | cDNA library | |

[a] The predicted coding sequence in the GenBank database, when compared with the consensus sequence assembled herein (Figure 4), has not identified one exon corresponding to (mouse) amino acids 94−133.

cession No. AF230273), and the mass of the expressed fusion was confirmed by LC−MS (expected, including the initial methionine: 22 928.1, observed: 22 927.9).

*Cleavage of Affinity Tag.* Cleavage of the fusion construct with enterokinase, which would afford a protein without any overhanging amino acids at the N-terminus, was attempted but was unsuccessful as monitored by SDS−PAGE and LC−MS. When MDP-1 was recloned with a glycine spacer between the natural N-terminal threonine and the enterokinase cleavage site, treatment with 4 U of enterokinase/$\mu$g fusion[2] at 37 °C overnight resulted in >90% cleavage as observed by SDS−PAGE. Increase of the spacer to two glycines allowed complete cleavage at 25 °C but did not alter the requirement for large amounts of protease. The CBP-MDP-1 fusion and the enterokinase-cleaved *N*-glycyl-MDP-1 were nearly indistinguishable in their ability to hydrolyze *p*NPP at a concentration of 5 mM. A more detailed analysis of kinetic parameters, however, revealed a much tighter binding of this substrate after removal of the CBP fusion tag (Figure 6). Thus, removal of the calmodulin fusion was deemed necessary prior to further characterization. The fusion tag also includes a thrombin site, which produces a 10-amino acid N-terminal overhang. Cleavage with thrombin is facile, proceeding with low amounts of protease at 25 °C and goes to completion as confirmed by SDS−PAGE and LC−MS (mass = 19 384.2, expected = 19 384.8). MDP-1 carrying this short N-terminal extension exhibits kinetic parameters ($K_M$ and $V_{max}$, as listed in Table 5) indistinguishable from the *N*-glycyl-MDP-1 with *p*NPP as a substrate. Subsequent characterization was performed on this thrombin-cleaved species.

*Requirement for Divalent Cations*. In the absence of divalent cations such as Ca, Mg, Mn, Co, Ni, or Zn, MDP-1 had no observable activity with either *p*NPP or ribose-5-phosphate as a substrate at pH 5.5. Monovalent cations such

[2] This is 400 times the amount of protease recommended by the supplier (Stratagene) for cleavage of this fusion.

```
HUMAN          ----------------MARLPKLAVFDLDYTLWPFWVDTHVDPPFH--KSSDG---TVRDRRGQDVRLYPEVPEV
MOUSE          ----------------MTRLPKLAVFDLDYTLWPFWVDTHVDPPFH--KSSDG---TVRDRRGQNIQLYPEVPEV
TOMATO         MGDEKVKNEAMEIMGLFQVLPRLVVFDLDYTLWPFYCECRSK----------------NEKPS----MYPQAKGI
ARABIDOPSIS    MAEEKVKDEAMQIMGMFQILPRLVVFDLDYTLWPFYCECRSK----------------REMPS----MYPQAKGI
SORGHUM        MGDERVKAEALQILGLFQVLPRLVVFDLDYTLWPFYCECRSK----------------RDSPS----LFRHARGI
S. POMBE       -------------MVKNIEFPKCVVFDLDYTLWPLWIDTHVTAPFKPSKNDPG---VLIDKYGTEICFYSDITGI
S. CEREVISIAE  ----------------MTGYPDVAAFDLDYTIWFCYCDTHLHGPFKPVKSSNGEVLTIICRDGYELTIYKDIPRI

consensus      ..................hPchssFDLDYTlWPhah-s+.p.................p..u....ha.ph..l

HUMAN          LKRLQSLGVPGAAASRTSEIEGANQLLELFDLFR---------YFVHREIYP--GSKITH----FERLQQKTGIP
MOUSE          LGRLQSLGVPVAAASRTSEIQGANQLLELFDLGK---------YFIQREIYP--GSKVTH----FERLHHKTGVP
TOMATO         LHAFKDKGVNVAIASRSPTPDIANAFLQKLEIKS---------MFVAQEIFSSWTHKTEH----FQKINRKTTVP
ARABIDOPSIS    LSALKEKGIEMAIASRSPTSDIANTFLDKLNIKP---------MFVAQEIYSSWSHKTEH----FQKIHTRTGVP
SORGHUM        IYALKEKGIDMAIASRSPTPDKAKVFLDKLELQS---------MFVAQEIFSSWTHKTEH----FQKIQRKTGIP
S. POMBE       LQELRNQKVTLCVASRTCAPKYAKQALNLMKVPIDGSLKPAIEFFTYVKAWP--GSKMDH----FKEIHNESGID
S. CEREVISIAE  LGDLKDNGVKLMTASRTWAPEIAQEILKIFKVKYAGVVTPLANLFDEFQWGE--RSKIGHLRDGLKDLYNTSDLK

consensus      l..hpp.hl.hhhASRo...phAp.hLphhpl..........hF...phh....pKh.H....hp+l.pposl.

HUMAN          FSQMIFFDDERRNIVDVSKLGVTCIHIQNGMNQQTLSQGLETFAKAKAGL----------------------
MOUSE          FSQMVFFDDENRNIIDVGRLGVTCIHIRDGMSLQTLTQGLETFAKAQAGL----------------------
TOMATO         YNEMLFFDDENRNIQAVSKMGVTSILVGKGVNLGALRQGLSEFVKNSASVEKNKEQWRKFSKRTGSSKSSDAQ
ARABIDOPSIS    FTAMLFFDDEDRNIKQVSKMGVTSILVGDGVTLGAFRQGLTEFTQNHNSIEKNKQVWRDKYSGKPTSSETDKG
SORGHUM        YKSMLFFDDEDRNIDPVSNLGVTSVLVENGVNLDMFKLGLSNFATKYAASSRKEDE----------------
S. POMBE       YREMVFFDDESRN-REVERLGVTFCLVPDGLNRASFEQGINKWRKLKKTH----------------------
S. CEREVISIAE  SKKICLFDDESRN-KEVEKYGVKFVYVRDPENGPSWKLYQDYLSGKV-------------------------

consensus      .p.hhhFDDEpRN...V.phGVp.hhl.ps.s...hp.h.p.h.............................
```

FIGURE 4: Alignment of MDP-1 deduced protein sequences from various species. The deduced protein sequences were aligned using the program CLUSTALW (http://molbio.info.nih.gov/molbio/gcglite/clustal.html). Invariant residues are highlighted in black, and highly conserved residues are highlighted in gray. The consensus sequence was calculated with the program CONSENSUS (http://www.bork.embl-heidelberg.de/cgi/consensus) using the default amino acid groupings: a = aromatic (FHYW), c = charged (DEHKR), h = hydrophobic (ACFGHIKLMRTVWY), l = aliphatic (ILV), o = alcohol (ST), p = polar (CDEHKNQRST), s = small (ACDGNPSTV), u = tiny (AGS), (+) = basic (HKR), and (−) = acidic (DE).

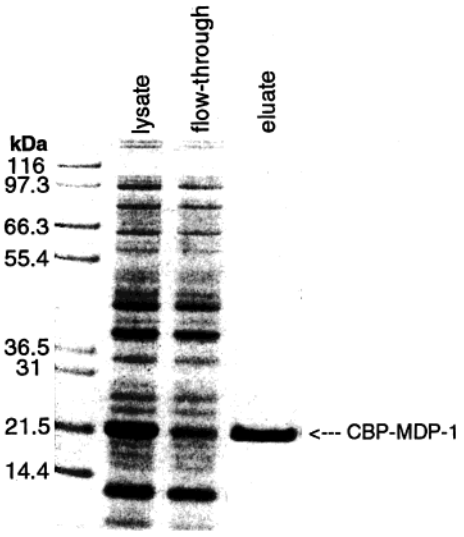


FIGURE 5: Affinity purification of the calmodulin-binding peptide−MDP-1 fusion. Lane 1, molecular weight standards. Lane 2, *E. coli* lysate. Lane 3, flow through (unbound fraction) from calmodulin affinity resin. Lane 4, CBP−MDP-1 fusion eluted from resin with buffer containing EGTA and 1 M NaCl. 4−12% SDS−PAGE gel stained with Coomassie blue.

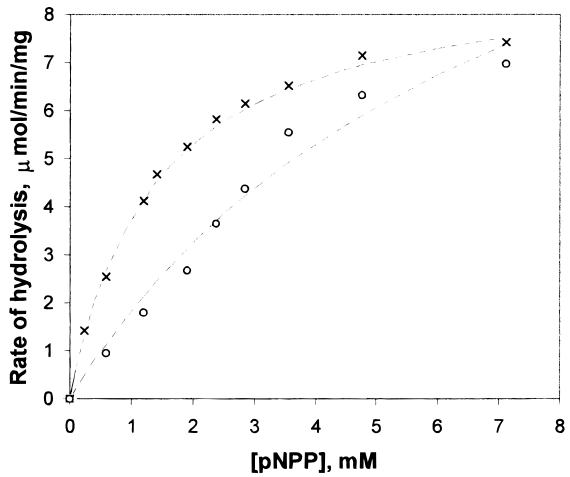


FIGURE 6: Presence of the fusion tag perturbs the kinetics of MDP-1 resulting in a higher $K_M$ for *p*NPP. Assay conditions are described under Materials and Methods. N-glycyl-MDP-1 was prepared as indicated in the text. (×), N-glycyl-MDP-1; (○), CBP−MDP-1.

as Na or K did not support activity. Apparent dissociation constants (the concentration of cation resulting in half-maximal activity) and maximal rates were determined by fitting activity versus cation concentration plots to the Michaelis−Menten equation. With 5 mM *p*NPP as substrate, magnesium supported the highest activity at 5.5 $s^{-1}$ and had an apparent dissociation constant of 1.1 mM (Table 4), comparable to normal physiological magnesium concentrations. Physiological manganese concentrations are also comparable to the measured $K_D$ of 200 $\mu$M and may function as a physiological cofactor to some degree. The dissociation constant for magnesium was unaffected by changing the substrate from *p*NPP to substrates with higher $K_M$ values such as ribose-5-phosphate or fructose-6-phosphate ($K_M$ = 1.7, 9.5, and 21 mM, respectively; see Table 5). The apparent dissociation constant of magnesium was uniformly reduced

Table 4: Divalent Cation Binding and Kinetics of Mouse MDP-1[a]

| cation | $V_{max}$ ($s^{-1}$) | $K_{D,\ apparent}$ (mM) |
|---|---|---|
| Mg | 5.5 | 1.1 |
| Co | 2.9 | 0.1 |
| Mn | 1.2 | 0.2 |
| Zn | 1.1 | 0.1 |
| Ni | 1.2 | 0.1 |
| Ca | 0.7 | 3.0 |

[a] Assay buffers in these experiments did not include EDTA.

Table 5: Substrate Kinetics of Mouse MDP-1

| substrate | rate ($s^{-1}$)[a,b] | $K_M$ (mM)[c] | $V_{max}$ ($s^{-1}$)[c] |
|---|---|---|---|
| ribose-5-phosphate | 10.0 | 9.5 | 26.0 |
| 2-deoxy-ribose-5-phosphate | 12.4 | 5.6 | 23.5 |
| phosphotyrosine | 5.9 | 15 | 20.0 |
| arabinose-5-phosphate | 13.8 | 1.1 ± 0.3 | 16.5 |
| fructose-6-phosphate | 2.3 | 21 ± 5 | 14.0 |
| 5′-CMP | 2.0 | 12.0 | 5.6 |
| *p*-nitrophenyl phosphate (*p*NPP) | 4.0 | 1.7 | 4.8 |
| 5′-AMP | 0.6 | 26 | 4.8 |
| glucose-6-phosphate | 0.22 | 31 | 1.6 |
| 2′-deoxy-5′-AMP | 1.0 | ND | |
| α-glycerophosphate | 0.5 | >20 | |

[a] Substrates were assayed at 5 mM in the presence of 8 mM magnesium at pH 5.5 in MES buffer. [b] The following substrates showed activity of less than 0.05 $s^{-1}$ (maximum concentration assayed is indicated in parentheses): bis(*p*-nitrophenyl) phosphate (5 mM), phosphoserine (5 mM), phosphothreonine (25 mM), glucose-1,6-diphosphate (12.5 mM), 2′-AMP (5 mM), 5′-ATP (5 mM), 3′(+2′)-CMP (60 mM), β-glycerophosphate (5 mM), phosphocholine (5 mM), phosphoethanolamine (5 mM), NADP (5 mM), dibutanoylglycerylphosphatidylinositol-3-phosphate (80 μM), and dibutanoylglycerylphosphatidylinositol-4,5-bisphosphate (50 μM). [c] Michaelis−Menten constants were calculated using the program EnzFitter from rate data collected at at least five different substrate concentrations. Error was less than 10% unless otherwise noted.

to 0.6 mM by a 4-fold increase of the substrate concentrations from 5 to 20 mM irrespective of which of these substrates was used. The $K_M$ of *p*NPP was the same whether 0.5 or 8 mM magnesium was present. In all subsequent assays, 8 mM magnesium was included.

*Dependence of Activity on pH*. With *p*NPP, ribose-5-phosphate, and phosphotyrosine as substrates (all at 5 mM), the pH optimum of MDP-1 was near 5.3 (Figure 7). The activity of MDP-1 toward *p*NPP decreased sharply at pH values higher than 5.3, but this did not occur with other substrates. It is unclear from these data whether this is an effect on $K_M$, $V_{max}$, or both. Ribose-5-phosphate and phosphotyrosine both showed a broader peak with considerable activity still in evidence at physiological pH. All subsequent assays were performed at pH 5.5 in MES buffer.

*Substrate Specificity*. The activity of MDP-1 toward various nonpeptide substrates is summarized in Table 5. MDP-1 exhibited a greater than 100-fold rate difference between phosphotyrosine and either phosphoserine or phosphothreonine. Primary pentose sugar phosphates were good substrates, and MDP-1 showed a preference for these over hexoses. Nucleotide phosphates, however, were poorer substrates than ribose-5-phosphate. There was no evidence of any activity toward the phosphodiester substrate bis(*p*NPP) nor toward secondary phosphates such as β-glycerophosphate, 2′- or 3′-ribonucleotides, and phosphoinositides. Table 6 summarizes the relative activity of MDP-1 and protein-

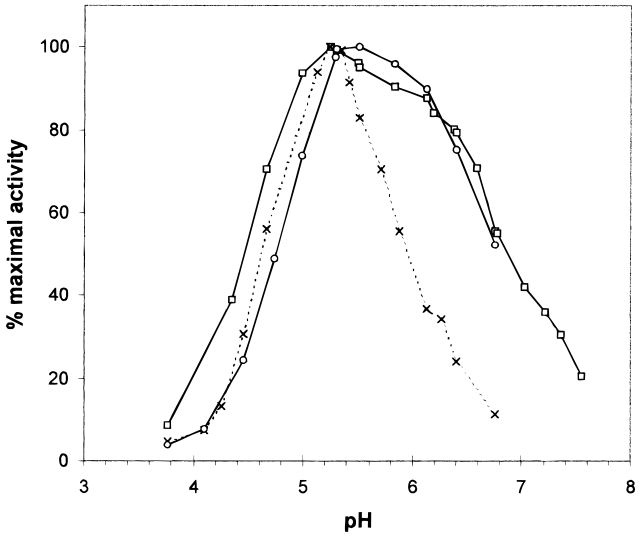


FIGURE 7: Effect of pH on the activity of MDP-1 with various substrates. Assays below pH 5.3 were run with acetate buffer, between pH 5.3 and pH 6.5 with MES buffer, and above pH 6.5 with PIPES buffer. Substrate concentration was 5 mM, and free magnesium concentration was 8 mM. Otherwise, the assay system was as described in Materials and Methods. No buffer effects on phosphatase activity were observed in regions of overlapping buffer capacity. (×), pNPP; (○), phosphotyrosine; and (□), ribose-5-phosphate.

tyrosine phosphatase 1B (PTP-1B) toward a panel of short phosphotyrosine-containing peptides and phosphotyrosine. The activity of MDP-1 toward phosphotyrosine at 200 μM was similar to that of PTP-1B. Unlike PTP-1B, however, MDP-1 hydrolyzed the tested substrates much more slowly than phosphotyrosine.

*Inhibitors*. Inhibitor susceptibility was assayed with 5 mM *p*NPP substrate at pH 5.5 in MES buffer (Table 7). *o*-Vanadate inhibited the enzyme but only at much higher concentrations than is typical of PTPase enzymes such as PTP-1B, in which a covalent cysteine−sulfur−vanadate adduct acts as a transition state analogue (*14*). The inhibition by vanadate can be overcome by EDTA, which is capable of chelating vanadium (*15*). The weak inhibition by calcium and strong inhibition by zinc are consistent with competition with magnesium (see Table 4), and phosphate is most likely acting by product inhibition. The lack of inhibition by tyrosine up to its 2 mM solubility limit is consistent with the $K_M$ of phosphotyrosine (Table 5). L-(+)-Tartaric acid, an inhibitor of certain acid phosphatases, has no effect.

*Exclusion of a Cysteine-Based Catalytic Mechanism*. MDP-1 does not contain the signature C−$X_5$−R motif of the protein-tyrosine phosphatases (PTPases), and alignment of the presumptive MDP-1 sequences (Figure 4) did not reveal any conserved cysteine. The mouse MDP-1 contains only a single cysteine (C137). Treatment of this enzyme with 40 μg/mL DTNB quantitatively labeled the protein with 1 equiv of TNB without significantly altering the specific activity. Similarly, large excesses of iodoacetamide had no effect on the activity. Finally, the single cysteine was mutated to serine, alanine, or glycine without significant effect on the activity (Table 8).

*Evidence against a Histidine-Based Catalytic Mechanism*. Although the mouse MDP-1 contains six histidines, only one of these (H102) is completely conserved among all the MDP-1 sequences, while a second (H23) is replaced by argi-

Table 6: Comparison of MDP-1 and PTP-1B Activity toward Phosphotyrosine-Containing Peptide Substrates[a]

| substrate | relative rate[b] | | ratio |
|---|---|---|---|
| | MDP-1 | PTP-1B | $k_{PTP-1B}/k_{MDP-1}$ |
| H-Tyr($PO_3H_2$)-OH | 1.0 | 2.3 | 2.3 |
| H-Gly−Gly−Tyr($PO_3H_2$)−Gly−Gly-OH | 0.10 | 4.2 | 42 |
| H-Asp−Arg−Tyr($PO_3H_2$)−Ile−His−Pro−Phe-OH | 0.007 | 43 | 6100 |
| H-Ala−Glu−Tyr($PO_3H_2$)−Ser−Ala-OH | 0.042 | 34 | 810 |
| Ac-Ile−Tyr($PO_3H_2$)−Gly−Glu−Phe-$NH_2$ | 0.066 | 100 | 1500 |
| Ac-Asp−Tyr($PO_3H_2$)−Val−Pro−Met−Leu-$NH_2$ | 0.035 | 51 | 1500 |

[a] Substrates (200 $\mu$M) were incubated with the indicated enzyme at 37 °C in 100 mM PIPES, pH 6.5, 100 mM NaCl, 20 mM $MgCl_2$, 10 mM EDTA, 1 mM DTT, and 0.01% Tween-20. Hydrolysis rate was monitored by fluorimetry ($\lambda_{ex} = 280$, $\lambda_{em} = 305$) over 10 min. [b] Rates normalized to that of phosphotyrosine with MDP-1.

Table 7: Inhibition of Mouse MDP-1[a]

| inhibitor | $IC_{50}$ (mM) | no inhibition up to (mM) |
|---|---|---|
| *o*-vanadate | 0.05 | |
| fluoride | 0.7 | |
| calcium | 3.2[b] | |
| phosphate | 10 | |
| 5′-AMP | 50[c] | |
| azide | | 25 |
| cyanide | | 25 |
| lithium | | 25 |
| tyrosine | | 2 |
| tartaric acid | | 125 |

[a] Assay in pH 5.5 MES buffer at 37 °C with 5 mM *p*NPP substrate in the presence of 8 mM $MgCl_2$ (no EDTA). [b] Consistent with competition versus magnesium; see table of $K_D$ values (Table 4). [c] Consistent with competition versus *p*NPP; see table of $K_M$ values (Table 5).

Table 8: Mouse MDP-1 Is Resistant to Conditions That Inactivate Cysteine-, Histidine- and (Non-Magnesium) Metal-Dependent Enzymes

| treatment | % residual activity |
|---|---|
| none | 100 |
| iodoacetamide, 34 mM, pH 6, 1 h, 37 °C | 90 |
| 34 mM, pH 7, 1 h, 37 °C | 94 |
| dithio-bis(nitrobenzoic acid), 40 $\mu$g/mL, 30 min[a] | 77 |
| mutant[b]: C137S | 87 |
| C137A | 110 |
| H102K | 56 |
| H102A | 53 |
| diethylpyrocarbonate, 2 equiv, 30 min[c] | 92 |
| 6 equiv, 60 min | 83 |
| unfolded in 6 M GuHCl, refolded | 91 |
| unfolded in 6 M GuHCl, refolded with Mg-EDTA | 98 |

[a] Enzyme was shown spectroscopically to be labeled at a 1:1 ratio. [b] Mass of mutants was checked by electrospray mass spectrometry and found to be correct to less than one mass unit. [c] Enzyme was shown by electrospray mass spectrometry to be labeled at a ratio of 0.5:1 at 2 equiv and to be multiply labeled at 6 equiv.

nine in the plant sequences (Figure 4). The conserved histidine (H102) has been mutated to alanine and lysine, resulting in the loss of only about one-half of the enzymatic activity toward *p*NPP (Table 8). DEPC is a histidine-specific reagent, which has been used to inactivate histidine-dependent enzymes such as thermolysin and ribonuclease (*13*). Up to 6 equiv of DEPC can be reacted with mouse MDP-1 without significant loss of catalytic activity (Table 8). Under these conditions, LC−MS confirms the labeling of the protein.

*Evidence against the Participation of Metals Other than Magnesium.* MDP-1 showed no activity in the absence of magnesium in the assay buffer. To determine whether any other tightly bound metals were present, MDP-1 was unfolded in the presence of 6 M guanidine hydrochloride both with and without the addition of EDTA to bind any released metals. Subsequent refolding resulted in complete recovery of activity when assayed in the presence of magnesium. This was also the case when the enzyme was refolded and assayed in the presence of magnesium titriplex, an equimolar mixture of magnesium and EDTA (Table 8).

## DISCUSSION

We have identified MDP-1 genes in four mammalian species (rabbit, mouse, rat, and human), four plant species (corn, sorghum, tomato, and *Arabidopsis*) and two yeast species (*S. cerevisiae* and *S. pombe*) through database searching. These EST and genomic data do not, however, contain any information about possible function, expression pattern, intracellular localization, or motifs. Protein-level similarity among these MDP-1 enzymes is high, although the two yeast enzymes are as distantly related from one another as they are from the plant and mammalian species (data not shown). Although the complete genome of several bacterial species are now known, no prokaryotic homologue of this gene is apparent, nor is a homologue evident in the genomes of the eukaryotes, *Drosophila*, or *Caenorhabditis elegans*.

The large number of phosphatase sequences presently represented in databases should make it a straightforward matter to relate most new phosphatase sequences to well-characterized species. However, since the sequence motifs that define the various classes of phosphatase tend to be short stretches of amino acids, sequence homology alone is not sufficient to categorize a novel phosphatase such as MDP-1. Thus, the question of classification hinges on two primary concerns: what is the nature of the substrate and what is the nature of the catalytic mechanism?

The substrate specificity of MDP-1, like its sequence, does not correlate with any previously described phosphatases. The preference shown by the enzyme for phosphotyrosine over phosphoserine or phosphothreonine initially suggested to us that it might function as a protein-tyrosine phosphatase. The five peptides assayed here were not substrates, suggesting that MDP-1 is not competent to accept peptide substrates into its active site. It is possible that MDP-1 is either highly selective for its specific substrate, or requires the whole native protein substrate or an activator to "open up" the active site to accept a peptide in a way that PTP-1B does not.

The close homology among the MDP-1's from diverse eukaryotic species suggests that the true substrate for this enzyme is ancient and the lack of a prokaryotic cognate suggests that the substrate may be distinctive of eukaryotes. The enzyme's activity appears to be limited to primary phosphates and the lack of activity toward bis(*p*NPP) argues

against a phosphodiesterase function. The substrates having the lowest $K_M$, *p*NPP and arabinose-5-phosphate, do not immediately suggest a biological role. The apparent preference for pentose sugar phosphates over hexoses might suggest a role in nucleotide biochemistry, but the weak activity toward AMP, dAMP, and CMP seems to rule this out. MDP also shows little or no activity toward phosphoinositides, glycerophosphates, phosphocholine, or phosphoethanolamine (Table 5).

On the basis of mechanisms (type of nucleophile) and sequence homologies, acid phosphatases can be divided into four groups: cysteine-nucleophile, histidine-nucleophile, aspartate-nucleophile, and water (hydroxide ion) nucleophile. Enzymes utilizing cysteine as the active site nucleophile encompass a large group of protein phosphotyrosine phosphatases (PTPases) and are distinguished by the motif Cys−$X_5$−Arg (*16*). Enzymes utilizing histidine as the active site nucleophile fall into two subgroups: some, such as prostatic acid phosphatase and fructose-2,6-bisphosphatase, are distinguished by the motif Arg−His−Gly (*4*, *17*), while others, such as glucose-6-phosphatase and phosphatidic acid phosphatase 2A are distinguished by a rather more complex motif (*6*). Neither of these classes of enzyme generally requires magnesium for activity, and MDP-1 does not exhibit any of the conserved motifs characteristic of these groups. In combination with the cysteine and histidine labeling results and the cysteine and histidine mutagenesis described here (Table 8), we conclude that MDP-1 must function by some other mechanism.

All acid phosphatases utilizing a water molecule as the nucleophile are metalloenzymes and can be subdivided into four groups. The purple acid phosphatases utilize either iron alone or iron and zinc (*18*). Similarly, the PPP family of serine-threonine protein phosphatases including PP1, PP2A, and PP2B utilize iron and manganese (*19*). Considering the refolding experiments reported here which show that MDP-1 requires only magnesium as a cofactor (Table 8), MDP-1 is not a member of either of these two families. A family of magnesium-requiring phosphatases including inositol mono- and polyphosphate phosphatases, nucleotidases, and fructose-1,6-bisphosphatase are defined by the motif Asp−Pro−(Ile or Leu)−Asp−(Gly or Ser)−(Ser or Thr) and are inhibited by lithium ion (*20*), neither being a characteristic of MDP-1. The PPM (PP2C) family of serine-threonine phosphatases also can function with only magnesium as a cofactor (*21*). However, there is no apparent sequence homology between PP2C and MDP-1.

The fourth well-defined group of phosphatases uses aspartate residues as the active site nucleophile. These enzymes, which include phosphoserine and phosphoglycolate phosphatases and certain nonspecific acid phosphatases and phosphosugar mutases and are a subset of the larger haloacid dehalogenase superfamily (*22*), contain a characteristic N-terminal Asp−$X_1$−Asp−$X_2$−(Thr or Val) motif (*7*, *23*). MDP-1 does contain this motif (human and mouse residues 10−14). However, although the $X_1$ residue, consistent with the other identified enzymes with this motif, is hydrophobic (leucine) in MDP-1, the $X_2$ residue is tyrosine, whereas, with rare exception, only glycine is found in the other known phosphatases of this group (*7*). Most members of this superfamily, and all of the phosphatases within it, contain a C-terminal Gly−Asp−X(3)−Asp motif (*7*, *22*, *24*), but this is absent in MDP-1. Epoxide hydrolase, a nonphosphatase member of the superfamily, is an exception in this respect and contains the sequence Asp−Asp−X(3)−Asn, which is indeed found in a conserved C-terminal region of MDP-1 (human and mouse residues 121−125). Thus, if MDP-1 is a member of this class of phosphatases, it represents a different branch of the family than is presently known. This distant relationship is further demonstrated by the lack of significant homology between members of this group and MDP-1 when compared on a one-to-one basis. Although this assignment appears to be the most likely, until the mechanistic details of MDP-1 can be elucidated, the possibility that it operates with a novel mechanism must be held open.

## ACKNOWLEDGMENT

We thank Dr. Terry Stadtman and Dr. P. Boon Chock for the donation of various reagents.

## REFERENCES

1. Boyer, P. D., Lardy, H., and Mayback, K., Eds. (1961) *The Enzymes*, Vol. V, Academic Press, New York.
2. Walsh, C. (1979) *Enzymatic Reaction Mechanisms*, pp 179−209, W. H. Freeman.
3. Vincent, J. B., Crowder, M. W., and Averill, B. A. (1992) *Trends Biochem. Sci. 17*, 105−110.
4. Van Etten, R. L., Davidson, R., Stevis, P. E., MacArthur, H., and Moore, D. L. (1991) *J. Biol. Chem. 266*, 2313−2319.
5. Stone, R. L., and Dixon, J. E. (1994) *J. Biol. Chem. 269*, 31323−31326.
6. Stukey, J., and Carman, G. M. (1997) *Protein Sci. 6*, 469−472.
7. Thaller, M. C., Schippa, S., Rossolini (1998) *Protein Sci. 7*, 1647−1652.
8. Kim, G., Selengut, J. S., and Levine, R. L. (2000) *Arch. Biochem. Biophys. 377*, 334−340.
9. Blackburn, M. N., Chirgwin, J. M., James, G. T., Kempe, T. D., Parsons, T. F., Register, A. M., Schnackerz, K. D., and Noltmann, E. A. (1972) *J. Biol. Chem. 247*, 1170−1179.
10. Apfel, A., Fischer, S., Goldberg, G., Goodley, P. C., and Kuhlmann, F. E. (1995) *J. Chromatog. A 712*, 177−190.
11. Heinonen, J. K., and Lahti, R. J. (1981) *Anal. Biochem. 113*, 313−317.
12. Sliwkowski, M. X., and Levine, R. L. (1985) *Anal. Biochem. 147*, 369−373.
13. Miles, E. W. (1977) in *Methods in Enzymology Vol. XVLII, Enzyme Structure, Part E* (Hirs, C. H. W., Timasheff, S. N., Eds.) pp 431−442, Academic Press, New York.
14. Zhang, M., Zhou, M., Van Etten, R. L., and Stauffacher, C. V. (1997) *Biochemistry 36*, 15−23.
15. Skorey, K. I., Johnson, N. A., Huyer, G., and Gresser, M. J. (1999) *Protein Expression Purif. 15*, 178−187.
16. Zhang, Z.-Y. (1998) *Crit. Rev. Biochem. Mol. Biol. 33*, 1−52.
17. Bazan, J. F., Fletterick, R. J., and Simon, J. P. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 9642−9646.
18. Ek-Rylander, B., Barkhem, T., Ljusberg, J., Ohman, L., Andersson, K. K., and Andersson, G. (1997) *Biochem. J. 321*, 305−311.
19. Barford, D., Das, A. K., and Egloff, M. P. (1998) *Annu. Rev. Biophys. Biomol. Struct. 27*, 133−164.
20. York, J. D., Ponder, J. W., and Majerus, P. W. (1995) *Proc. Natl. Acad. Sci. U.S.A. 92*, 5149−5153.
21. Das, A. K., Helps, N. R., Cohen, P. T., and Barford, D. (1996) *EMBO J. 15*, 6798−6809.
22. Ridder, I. S., and Dijkstra, B. W. (1999) *Biochem. J. 339*, 223−226.
23. Collet, J.-F., Stroobant, V., Pirard, M., Delpierre, G., and Van Schaftingen, E. (1998) *J. Biol. Chem. 274*, 33985−33990.
24. Aravind, L., Galperin, M. Y., and Koonin, E. V. (1998) *Trends Biochem. Sci. 23*, 127−129.

BI0005052