YZGD from *Paenibacillus thiaminolyticus*, a pyridoxal phosphatase of the HAD (haloacid dehalogenase) superfamily and a versatile member of the Nudix (nucleoside diphosphate x) hydrolase superfamily

Isaac M. TIRRELL*, Jennifer L. WALL†, Christopher J. DALEY*, Sarah J. DENIAL*, Frances G. TENNIS†, Kevin G. GALENS* and Suzanne F. O’HANDLEY*.

*Department of Chemistry, Rochester Institute of Technology, Rochester, NY 14623, U.S.A., and †Department of Chemistry, University of Richmond, Richmond, VA 23173, U.S.A.

INTRODUCTION

We became interested in YZGD from *Paenibacillus thiaminolyticus* when we discovered, through BLAST searches [1], that YZGD belongs to two separate and distinct enzyme families; the first approx. 250 amino acids show homology to the ‘nitrophenyl phosphatase’ family of the HAD (haloacid dehalogenase) superfamily [2], and the last approx. 160 amino acids contain the signature sequence of the Nudix (nucleoside diphosphate x) hydrolase superfamily [3]. YZGD is the first enzyme found to belong to both of these enzyme superfamilies.

The HAD superfamily was discovered by computer analysis of bacterial HADs [2]. Using an iterative approach to sequence database searching, a number of known enzymes and open reading frames were found to have several conserved amino acids within three motifs [2]. These conserved amino acids have been defined most recently as: an aspartate nucleophile in motif I, a serine or threonine for phosphoryl binding in motif II, a lysine or arginine for phosphoryl binding, and carboxylate residues forming a metal-binding pocket in motif III [4]. These motifs compose the core domain. Many family members have an additional cap domain involved in substrate recognition and positioning [5].

The HAD superfamily is a fairly large superfamily, containing over 3000 members [4]. The characterized enzymes in this superfamily catalyse group transfer and fall within one of the following classifications: HADs, phosphonatases, phosphate monoestersases, ATPases and phosphomutases [4]. The phosphate monoestersases can be further divided into a number of distinct phosphatase families including phosphoglycolate phosphatases, phosphoserine phosphatases, histidinol phosphatases, trehalose phosphatases [2], the MDP-1 (magnesium-dependent phosphatase-1) family [6] and the phospho 1 ethanolamine/choline phosphatase family [7].

One of the phosphatase families that was delineated by cluster analysis based on BLAST scores includes ‘the nitrophenyl phosphatases and related putative proteins’ [2]. Although the nitrophenyl phosphatase family of the HAD superfamily contains over 400 open reading frames, the function of most of these putative proteins remains unknown. When this family was discovered [2], only two enzymes of the family had been identified, the p-nitrophenyl phosphate-specific phosphatases from *Saccharomyces cerevisiae* [8–10] and *Schizosaccharomyces pombe* [11]. Though the enzymes were identified as phosphatases, no biologically relevant substrate could be determined, no obvious phenotype was revealed, and both enzymes appeared dispensable for vegetative growth and sporulation [10,11]. A phosphatase with protein phosphatase activity on histone II-A and casein was suggested to be the p-nitrophenyl phosphatase from *S. cerevisiae* [12], though protein phosphatase activity was sought, but could not be detected in the earlier studies [10,11].

The only other enzyme in this family that has been previously characterized is a human PLPase (pyridoxal phosphatase). This enzyme was initially isolated and purified from human erythrocytes [13] and a 20 amino acid peptide sequence was determined.

Abbreviations used: HAD, haloacid dehalogenase; MDP-1, magnesium-dependent phosphatase-1; NMP, nucleoside monophosphate; Nudix, nucleoside diphosphate x; PLP, pyridoxal 5-phosphate; PLPase, pyridoxal phosphatase.

To whom correspondence should be addressed (email sfosch@rit.edu).
[14]. From this peptide sequence, human ESTs (expressed sequence tags) were uncovered and the corresponding cDNA clones were identified from the brain [15]. Jang et al. [15] identified a number of similar protein sequences, but failed to recognize this human PLPase or these other proteins as members of the HAD superfamily.

The Nudix hydrolases are a family of enzymes that hydrolyse a phosphoanhydride bond in substrates consisting of a nucleoside diphosphate linked to some moiety, x, hence the acronym ‘Nudix’, and are defined by the signature sequence GX, EX, REUXEXGU, where U represents the bulky aliphatic amino acids leucine, isoleucine or valine [3]. Members of the Nudix hydrolase family are ubiquitous throughout all three domains of life with over 1800 putative proteins [16]. With each new enzyme uncovered, the substrates predominantly fall into the categories of nucleoside triphosphates, nucleotide sugars, dinoleside polyphosphates and nucleotide coenzymes [3], with few exceptions [18,19]. Though the substrate specificity is varied, it is believed that a general role of the Nudix hydrolases is to regulate the level of metabolites that would be detrimental to the cell at elevated levels [3].

In the present paper, we describe the cloning and expression of the yzgd gene, as well as the purification and characterization of the YZGD enzyme. We characterize YZGD as both a PLPase of the nitrophenolphosphatase family of the HAD superfamily and a versatile member of the Nudix hydrolase superfamily and show that this family is not a PLPase family, but rather a family of enzymes that so far appears to act upon phosphorylated substrates containing an aromatic ring.

**EXPERIMENTAL**

**Materials**

Oligodeoxynucleotides were from Integrated DNA Technologies, the plasmid pET11b from Novagen, and the plasmid pAN13 [20] containing the yzgd gene was kindly provided by Dr Tadhg Begley (Cornell University, Department of Chemistry and Chemical Biology, Ithaca, NY, U.S.A.). Escherichia coli DH5α competent cells were from Invitrogen and E. coli HMS174(DE3) competent cells were from Novagen. Pfu DNA polymerase was from Stratagen, restriction enzymes NdeI and Blpl, T4 DNA ligase and calf intestinal alkaline phosphatase were from Invitrogen and New England Biolabs and inorganic pyrophosphatase was from Sigma. Media components were from Difco, isopropyl β-D-thiogalactoside was from Research Organics, resins and substrates were from Sigma and general chemicals were from Sigma and J. T. Baker. Thiazole l-phosphate was a gift from Dr Tadhg Begley, The QuikChange® Site-Directed Mutagenesis kit was from Stratagen.

**Cloning**

The yzgd gene was subcloned from the plasmid pAN13 [20]. The yzgd gene was amplified using PCR, where an NdeI restriction site was incorporated at the start of the gene and a Blpl site was incorporated at the end of the gene by using primers containing these sites. The amplified gene was gel-purified, digested with NdeI and Blpl and ligated into plasmid pET11b, which had also been digested with NdeI and Blpl. This put the yzgd gene under the control of a T7 lac promoter for overexpression. The resultant plasmid, pETyzgd, was transformed into E. coli DH5α for storage and E. coli HMS174(DE3) for expression. The sequence of yzgd in pETyzgd was confirmed by the fluorescent dideoxy terminator method on a PerkinElmer ABI 377 automated DNA sequencer at the Johns Hopkins University core facility.

**Expression and purification of the enzyme**

E. coli HMS174(DE3) containing pETyzgd was grown at 37°C in 2 litres of Luria–Bertani medium containing 100 µg/ml ampicillin. When growth reached an A600 (absorbance) of 0.75, the culture was induced with 1 mM isopropyl β-D-thiogalactoside. After an additional 2 h of growth, the cells (4.4 g) were harvested, washed with buffered iso-osmotic saline (50 mM Tris/HCl, pH 7.5, 0.5% KCl and 0.5% NaCl), re-centrifuged and stored at −80°C. A negative control of E. coli HMS174(DE3) containing pET11b was treated in an identical manner.

The cells (4.4 g) were resuspended in 8.8 ml of buffer A (50 mM Tris/HCl, pH 7.5, 1 mM EDTA and 0.1 mM dithiothreitol), sonicated using a Branson sonifier and centrifuged. The supernatant was separated from the pellet and the protein concentration of the supernatant was adjusted to 10 mg/ml with buffer A (fraction I). Fraction I contained 130 mg of total protein with a PLPase specific activity of 1.2 units/mg. Streptomycin sulphate (10%, w/v) was added to fraction I to give a final concentration of 1% streptomycin sulphate and incubated on ice for 15 min, and the precipitate containing nucleic acids was discarded. Most of the enzymatic activity remained in the supernatant (fraction II). Saturated ammonium sulphate was added to fraction II to give a final concentration of 55% ammonium sulphate. The precipitant, containing YZGD, was resuspended in 5 ml of buffer A (fraction III) and loaded on to a 2.5 cm × 60 cm Sephadex G-100 column calibrated with molecular mass protein standards. Protein was eluted from the column with buffer A containing 200 mM NaCl. Fractions containing YZGD were pooled and concentrated with 80% ammonium sulphate, and the precipitant containing YZGD was resuspended in 1 ml of buffer A (fraction IV). Fraction IV was dialysed against 50 mM Tris/HCl (pH 7.6) and 1 mM EDTA (buffer B). The dialysed sample was loaded on to a 1.5 cm × 6 cm DEAE-Sepharose column equilibrated with buffer B and proteins were eluted by rinsing the column with 2 column volumes of buffer B containing 200, 300 and 400 mM NaCl sequentially. YZGD eluted with buffer B containing 300 mM NaCl. Fractions containing YZGD were pooled and concentrated using Centriprep-10 concentrators from Millipore (fraction V). The final purified enzyme contained approx. 15 mg of protein with a PLPase specific activity of 7.0 units/mg.

**Enzyme assays**

For all assays, 1 unit of YZGD hydrolyses 1 µmol of substrate per min.

**Phosphatase assay**

A standard reaction mixture consisted of 50 mM Tris/maleate (pH 5), 5 mM CoCl₂, 1 mM dithiothreitol, 4 mM PLP (pyridoxal 5-phosphate) and 0.5–4.0 m-units of YZGD in 50 µl. Negative controls replaced YZGD with diluent, YZGDD12N, or crude extract of pET11b in HMS174(DE3). The reactions were carried out for 30 min at 37°C, quenched by the addition of 50 µl of 4:1 Norit (20% packed volume)/HClO₄ (7%) and centrifuged, and 50 µl of the supernatant was analysed for inorganic orthophosphate by the colorimetric method of Ames and Dubin [21]. Other phosphorylated substrates were analysed by the
colorimetric method of Ames and Dubin [21] or Fiske and Subbarow [22].

Protein phosphatase assay

A 10 mg/ml casein solution was prepared and dialysed against Tris/HCl (pH 7.5), 10 % (v/v) glycerol, 1 mM dithiothreitol and 1mM EDTA to remove any free inorganic phosphate. The reaction mixture contained 50 mM Tris/maleate (pH 5) or Tris/HCl (pH 7.2), 5 mM MgCl₂ or MnCl₂, 1 mM dithiothreitol, 50 µg of casein and 0.5 or 30 m-units of YZGD. A positive control contained 0.5 mg of protein phosphatase 1 instead of YZGD. The reactions were incubated at 37 °C for 1 h or overnight and further treated as described under the subsection Phosphatase assay.

Nudix hydrolase assays

A standard reaction mixture consisted of 50 mM Tris/HCl (pH 8.5), 10 mM MnCl₂, 1 mM dithiothreitol, 0.6 unit of calf intestinal alkaline phosphatase, 4 mM substrate and 0.5–4.0 m-units of YZGD in 50 µl. Negative controls replaced YZGD with diluent, YZGD324Q, or crude extract of pET11b in E. coli DH5α and pETYZGD12N were each transformed separately, into a NaCl concentration of 0.3 M. This purification method yielded approx. 15 mg of highly purified YZGD from 2 litres of culture using the model as a search tool. This model was used in ‘hmmssearch’ against known proteins of other families of the HAD superfamily, members of the nitrophenyl phosphatase family as a positive control, and completely unrelated proteins as a negative control.

RESULTS AND DISCUSSION

Subcloning, expression and purification

The sequence of the subcloned yzd from P. thiaminolyticus agrees with the sequence deposited by Costello et al. [28]. Over-expression in E. coli produced a protein product, which appears as a major band at 45 kDa on a denaturing polyacrylamide gel (Figure 1). Additionally, the crude extract of YZGD contained ≥ 1000-fold more PLPase activity than a corresponding extract lacking YZGD. Purification of YZGD resulted in a highly purified protein (Figure 1). Sonication was necessary to extract soluble YZGD from the cells. Streptomycin sulphate was necessary to remove nucleic acids. YZGD eluted from a G100 column as expected for a 90 kDa protein, indicating that YZGD, with a monomeric molecular mass of 45 kDa, is a dimer in solution. YZGD was predicted to have a net charge of −22 and a pI of 4.9 (DNAid). We took advantage of YZGD’s negative charge to purify it over a DEAE anion-exchange column, where YZGD eluted with a NaCl concentration of 0.3 M. This purification method yielded approx. 15 mg of highly purified YZGD from 2 litres of culture with a 6-fold purification and a 67 % yield.

Substrate specificity

Since YZGD was found through BLAST searches to belong to two separate and distinct protein families, the ‘nitrophenyl phosphatase’ family of the HAD superfamily [2] (Figure 2) and the Nudix hydrolase family [3] (Figure 3), we predicted that YZGD would possess both phosphatase activity and Nudix hydrolase activity.

When Koonin and Tatusov [2] delineated the HAD superfamily into subfamilies, the only identified and characterized enzymes of the nitrophenyl phosphatase family of the HAD superfamily

![Figure 1 Expression and purification of the YZGD PLPase](image-url)

A 15 % polyacrylamide gel with 1 % SDS contained the following: lane 1, reference proteins with molecular masses of 66, 45, 29 and 20 kDa; lane 2, crude extract from E. coli cells containing PETygzt; lane 3, fraction I, supernatant obtained from sonication of the cells; lane 4, fraction II, supernatant obtained from fractionation with streptomycin; lane 5, fraction III obtained by fractionation with ammonium sulphate; lane 6, fraction IV obtained by purification over a G100 column; lane 7, fraction V obtained by purification over a DEAE-Sepharose column. Lane 7 contains 5 µg of purified YZGD. Lanes 2–6 contain approximately the same number of PLPase units as present in lane 7.
Figure 2  BLAST search results for the HAD domain of YZGD

A BLAST search of non-redundant databases using the first 250 amino acids of YZGD uncovered 390 hits with E values of < 1. The 20 best matches along with S. pombe and S. cerevisiae p-nitrophenyl phosphatases, hits 23 and 29 respectively, were aligned using CLUSTAL X. Most of the matches were unknown open reading frames. Proteins with characterized phosphatase activity are indicated by a '#', and aligned separately in Figure 6. Absolutely conserved identical amino acids are noted with a '*'. Conserved similar amino acids and highly conserved identical amino acids are identified by the symbols 'a n d' respectively. The amino acids are also highlighted when there is sequence identity or similarity. Similar amino acids are grouped as follows: (1) G; (2) P; (3) I, L, V, A, M, F, W; (4) T, S, N, Q; (5) D, E; R, K; and (6) Y, H. Gaps in the amino acid sequences, inserted by CLUSTAL to optimize alignment, are indicated with a '-'. The height of the peaks below the sequence alignment indicate the degree of overall homology; the higher the peak, the greater the overall homology.
were the two yeast p-nitrophenyl phosphatases. YZGD also hydrolyses p-nitrophenyl phosphate; however, p-nitrophenyl phosphate is not a biologically relevant compound. The only biologically significant phosphorylated substrate identified for YZGD is the essential cofactor PLP, with a specific activity of 7 units/mg (Table 1), a $k_{cat}$ of 19.4 s$^{-1}$, a $K_m$ of 0.33 mM and a $k_{cat}/K_m$ of 5.8 $\times$ $10^4$ M$^{-1}$ s$^{-1}$ (Table 3). Even the related pyridoxamine 5-phosphate is not a substrate for YZGD (Table 1).

Once we discovered this PLPase activity, we identified a human pyridoxal phosphatase. This is not surprising, as the HAD superfamily consists of a wide variety of phosphatases [4].

**Table 1** Substrate specificity of YZGD for phosphorylated substrates

<table>
<thead>
<tr>
<th>Organism (Enzyme)</th>
<th>Accession</th>
<th>Nudix Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>P. thiaminolytica (YZGD)</td>
<td>P46351</td>
<td>0.1</td>
</tr>
<tr>
<td>Bacillus thuringiensis</td>
<td>YP_036157</td>
<td>0.4</td>
</tr>
<tr>
<td>Bacillus cereus ZK</td>
<td>YP_038340</td>
<td>0.1</td>
</tr>
<tr>
<td>Bacillus anthracis AeA</td>
<td>NP_840402</td>
<td>0.1</td>
</tr>
<tr>
<td>Bacillus cereus ATCC14978</td>
<td>NP_978387</td>
<td>0.1</td>
</tr>
<tr>
<td>Streptomyces coelicolor</td>
<td>BAD21724</td>
<td>0.1</td>
</tr>
<tr>
<td>Lactobacillus plantarum</td>
<td>NP_629476</td>
<td>0.1</td>
</tr>
<tr>
<td>Bacillus cereus 0524</td>
<td>NP_0040525</td>
<td>0.2</td>
</tr>
<tr>
<td>Bacillus cereus ATCC14579</td>
<td>NP_831762</td>
<td>0.2</td>
</tr>
<tr>
<td>Lactobacillus johnsonii</td>
<td>NP_956074</td>
<td>0.2</td>
</tr>
<tr>
<td>Bacillus thuringiensis</td>
<td>YP_036373</td>
<td>0.2</td>
</tr>
<tr>
<td>Bacillus cereus ZK</td>
<td>YP_038364</td>
<td>0.2</td>
</tr>
<tr>
<td>Bacillus anthracis AeA</td>
<td>NP_846650</td>
<td>0.2</td>
</tr>
<tr>
<td>Bacillus cereus ATCC14978</td>
<td>NP_978600</td>
<td>0.2</td>
</tr>
<tr>
<td>Chlorellaeus aurantius</td>
<td>FP_0002067</td>
<td>0.2</td>
</tr>
<tr>
<td>Escherichia coli</td>
<td>SF_0018434</td>
<td>0.2</td>
</tr>
<tr>
<td>Methanobacterium sp.</td>
<td>LF_064311</td>
<td>0.2</td>
</tr>
<tr>
<td>S. coelicolor</td>
<td>NF_625039</td>
<td>0.2</td>
</tr>
<tr>
<td>Bacillus thuringiensis</td>
<td>YP_036851</td>
<td>0.2</td>
</tr>
<tr>
<td>Bacillus anthracis AT2012</td>
<td>NP_656641</td>
<td>0.2</td>
</tr>
</tbody>
</table>

**Table 2** Substrate specificity of YZGD for Nudix substrates

<table>
<thead>
<tr>
<th>Substrate</th>
<th>Specific activity (units/mg)</th>
<th>Relative activity (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CDP-glycerol</td>
<td>11.4</td>
<td>100</td>
</tr>
<tr>
<td>CDP-ethanolamine</td>
<td>10.5</td>
<td>92</td>
</tr>
<tr>
<td>CDP-choline</td>
<td>7.0</td>
<td>61</td>
</tr>
<tr>
<td>TDP-glucose</td>
<td>5.5</td>
<td>48</td>
</tr>
<tr>
<td>FAD</td>
<td>9.1</td>
<td>80</td>
</tr>
<tr>
<td>ADP-ribose</td>
<td>9.1</td>
<td>80</td>
</tr>
<tr>
<td>NAD</td>
<td>6.7</td>
<td>59</td>
</tr>
<tr>
<td>NADH</td>
<td>6.6</td>
<td>58</td>
</tr>
<tr>
<td>GDP-mannose</td>
<td>1.6</td>
<td>14</td>
</tr>
<tr>
<td>ADP-glucose</td>
<td>0.8</td>
<td>7</td>
</tr>
<tr>
<td>Ap5A</td>
<td>0.7</td>
<td>6</td>
</tr>
<tr>
<td>UDP-glucose</td>
<td>4.2</td>
<td>37</td>
</tr>
<tr>
<td>UDP-mannose</td>
<td>1.5</td>
<td>13</td>
</tr>
<tr>
<td>UDP-A-acetylglucosamine</td>
<td>0.8</td>
<td>7</td>
</tr>
<tr>
<td>UDP-galactose</td>
<td>0.3</td>
<td>3</td>
</tr>
<tr>
<td>GDP-mannose</td>
<td>0.2</td>
<td>2</td>
</tr>
<tr>
<td>GDP-glucose</td>
<td>&lt;1</td>
<td>&lt;1</td>
</tr>
<tr>
<td>CoA</td>
<td>&lt;1</td>
<td>&lt;1</td>
</tr>
</tbody>
</table>

* One unit of enzyme hydrolyses 1 µmol of substrate per min.

YZGD HAD pyridoxal phosphatase and Nudix hydrolase from S. cerevisiae [12]. Therefore we analysed YZGD for protein phosphatase activity. YZGD is not active on the naturally phosphorylated milk protein casein (Table 1). A positive control of protein phosphatase 1 acting on casein demonstrated that enough free phosphate could be released to detect readily in our assays. Likewise, YZGD is not active on serine phosphate, threonine phosphate, tyrosine phosphate, histidine phosphate or arginine phosphate (Table 1). Thus we have no evidence that YZGD is a protein phosphatase. This is not surprising, as the HAD superfamily consists of a wide variety of phosphatases [4].

Substrates of other members of the HAD superfamily were analysed including glycolate 2-phosphate, trehalose 6-phosphate, serine phosphate, phosphoethanolamine and phosphocholine. No activity was observed for YZGD with these substrates (Table 1).

We have determined that NAGD from E. coli is active on several substrates besides UMP, including N-acetylgulcymamine 6-phosphate, glucosamine 6-phosphate and fructose 6-phosphate (I. M. Tirrell, A. T. Nguyen, R. J. Mentz, E. J. Slivka, and S. F. O’Handley, unpublished work). Conversely, YZGD is not active on these substrates (Table 1). Because $yzgd$ is just upstream of the thiaminase I gene $yzgc$ [28], thiamine phosphate, thiamine pyrophosphate and a thiamine precursor, thiazole phosphate, were analysed for activity. However, none of these are substrates for YZGD (Table 1), and the relationship of $yzgd$ and $yzgc$, if any, remains unknown (T. P. Begley, personal communication). Coenzymes, sugar phosphates, NMPs, inositol phosphates and other biologically significant phosphorylated compounds were analysed (Table 1) and we especially took note of the phosphorylated substrates that would be products of YZGD’s Nudix hydrolase activity, but none of these compounds were substrates for YZGD’s phosphatase activity.

Whereas YZGD’s phosphatase activity is highly specific, its Nudix hydrolase activity is extremely broad (Table 2). YZGD displays comparable activity on CDP-alcohols, ADP-coenzymes, ADP-ribose and TDP-glucose and some activity on UDP-sugars.
Table 3 Kinetic parameters for YZGD

For pyridoxal phosphate, the standard phosphatase assay was used with 0.1–40 mM substrate. For the other eight substrates, the standard Nudix hydrolase assay was used with substrate concentrations ranging from 0.5–20 mM for TDP-glucose to 2–80 mM for CDP-choline and CDP-ethanolamine; the other Nudix substrate concentrations were within these two ranges. $V_{max}$ and $K_m$ were determined from a nonlinear regression analysis and $k_{cat}$ was calculated from $V_{max}$ assuming that YZGD is a dimer with a molecular mass of 90.8 kDa. One unit of enzyme hydrolyses 1 µmol of substrate per min.

<table>
<thead>
<tr>
<th>Substrate</th>
<th>$V_{max}$ (units/mg)</th>
<th>$k_{cat}$ (s$^{-1}$)</th>
<th>$K_m$ (mM)</th>
<th>$k_{cat}/K_m$ ($\times 10^3$ M$^{-1}$·s$^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pyridoxal phosphate</td>
<td>12 ± 1</td>
<td>19 ± 1</td>
<td>0.33 ± 0.07</td>
<td>57 ± 16</td>
</tr>
<tr>
<td>FAD</td>
<td>15 ± 1</td>
<td>22 ± 1</td>
<td>3.7 ± 0.6</td>
<td>6.0 ± 2.3</td>
</tr>
<tr>
<td>ADP-ribose</td>
<td>11 ± 1</td>
<td>17 ± 2</td>
<td>6.9 ± 2.5</td>
<td>2.4 ± 0.8</td>
</tr>
<tr>
<td>NADH</td>
<td>5.3 ± 0.3</td>
<td>7.8 ± 0.5</td>
<td>4.7 ± 1.0</td>
<td>1.7 ± 0.5</td>
</tr>
<tr>
<td>NAD</td>
<td>5.3 ± 0.5</td>
<td>8.0 ± 0.8</td>
<td>8.5 ± 1.3</td>
<td>0.9 ± 0.6</td>
</tr>
<tr>
<td>CDP-glycerol</td>
<td>19 ± 1</td>
<td>29 ± 2</td>
<td>7.1 ± 1.0</td>
<td>4.1 ± 1.5</td>
</tr>
<tr>
<td>CDP-ethanolamine</td>
<td>33 ± 2</td>
<td>50 ± 3</td>
<td>22 ± 3</td>
<td>2.3 ± 0.9</td>
</tr>
<tr>
<td>CDP-choline</td>
<td>14 ± 1</td>
<td>22 ± 1</td>
<td>22 ± 3</td>
<td>1.0 ± 0.4</td>
</tr>
<tr>
<td>TDP-glucose</td>
<td>4.1 ± 0.2</td>
<td>6.3 ± 0.2</td>
<td>1.3 ± 0.2</td>
<td>4.9 ± 1.1</td>
</tr>
</tbody>
</table>

and GDP-sugars, YZGD is the first Nudix hydrolase reported to have comparable ADP-x, CDP-x and TDP-x hydrolase activities, the first Nudix hydrolase reported to have a TDP-sugar hydrolase activity and the second Nudix hydrolase reported to have CDP-alcohol hydrolase activity (a CDP-choline hydrolase from Bacillus cereus has been described [16]). There is a rat liver enzyme with substrate specificity similar to the Nudix hydrolase activity of YZGD [29], but its sequence has not been determined, and thus it has not yet been classified as a Nudix hydrolase.

The only common Nudix substrates for which YZGD shows no activity are the (d)NTPs, the diadenosine polyphosphates and CoA. The (d)NTPs are relatively small, while the diadenosine polyphosphates and CoA are considerably larger compared with YZGD’s substrates; thus (d)NTPs, diadenosine polyphosphates and CoA may not fit YZGD’s active site properly.

When kinetic assays were carried out, the $k_{cat}/K_m$ values were within a factor of six of one another for the eight best Nudix substrates (Table 3). The specific activities for YZGD’s Nudix activities are comparable with those of other Nudix hydrolases seen today. The specific activities for YZGD’s Nudix substrates were within a factor of six of one another and well comparable with one another and well comparable with those of other Nudix hydrolases [30–32]. Considering YZGD’s broad substrate specificity, the higher $K_m$ values are not unexpected. Proof that all of these Nudix hydrolase activities can be attributed to YZGD include the observations that the specific activities are comparable with one another and well above the activities of an E. coli extract lacking YZGD and are affected similarly by bivalent metal cations, pH and site-directed mutagenesis.

With its broad substrate specificity, YZGD is the most versatile Nudix hydrolase to date, capable of hydrolysing derivatives of all of the nucleotides. Thus it is interesting to speculate that YZGD is representative of an evolutionarily ancient Nudix hydrolase with rather non-specific activity from which evolved the more specific Nudix hydrolases seen today.

Products of the reaction

YZGD hydrolyses phosphate from pyridoxal phosphate and the reaction can be written as:

$$\text{PLP + H}_2\text{O} \rightarrow \text{pyridoxal + P}_i$$  

(YZGD also catalyses the hydrolysis of Nudix substrates as follows: ADP-ribose and NADH yield AMP and ribose 5-phosphate or nicotinamide mononucleotide respectively, CDP-ethanolamine, CDP-choline and CDP-glycerol yield CMP, and phosphoethanolamine, phosphocholine or phosphoglycerol respectively, UDP-mannose and GDP-mannose yield UMP or GMP respectively and mannose 1-phosphate. TDP-glucose yields TMP and glucose 1-phosphate. Regardless of the substrate, one of the products is always an NMP. Thus a general reaction can be written as:

$$\text{N-P-P-X + H}_2\text{O} \rightarrow \text{NMP} + \text{X-P}$$  

(reaction 2)

where N is any of the canonical nucleosides, P is phosphate and X is any moiety. Most other Nudix hydrolases are pyrophosphohydrolases [3,33], catalysing nucleophilic attack of phosphorous, and YZGD follows suit.

Properties of the enzyme

YZGD’s PLPase activity is characterized as an acid phosphatase with optimal activity at pH 5 (Figure 4A). The same pH optimum was obtained for YZGD with p-nitrophenyl phosphate. The pH optimum for YZGD’s Nudix hydrolase activity is pH 8.5 (Figure 4B). This held true for the three substrates tested: CDP-ethanolamine, ADP-ribose and TDP-glucose. This alkaline pH optimum is characteristic of the Nudix hydrolase family with pH optima typically ranging from pH 8.5 to 9.5 [23].

YZGD absolutely requires a bivalent metal ion for both its phosphatase activity and its Nudix hydrolase activity; however, the metal ion requirements differ for the two activities. For YZGD’s phosphatase activity, Co$^{2+}$ is most effective. Mg$^{2+}$ is almost as effective as Co$^{2+}$, while Zn$^{2+}$ and to a lesser extent Mn$^{2+}$ can partially substitute for Co$^{2+}$ (Figure 5A). Ni$^{2+}$ and Fe$^{2+}$ support catalytic activity with activities reaching 40 and 20% maximal activity respectively, while Ca$^{2+}$ and Cu$^{2+}$ do not support activity.

The Nudix hydrolases absolutely require a bivalent metal cation for activity and YZGD is no exception. However, while most Nudix hydrolases are maximally activated by Mg$^{2+}$ [23,30,32–35], YZGD is absolutely dependent on Mn$^{2+}$ for its Nudix hydrolase activity (Figure 5B). No other bivalent metal ion tested (Mg$^{2+}$, Co$^{2+}$, Zn$^{2+}$ or Ca$^{2+}$) supported hydrolysis of Nudix substrates by YZGD. Several Nudix hydrolases from B. cereus and the rat liver enzyme with similar Nudix substrate specificity to YZGD are also Mn$^{2+}$-dependent enzymes [16,29].

We also investigated the effect of two metals at equal concentrations. There was no enhancement of either activity with two metals. Co$^{2+}$ and Zn$^{2+}$ nearly completely inhibited the Mn$^{2+}$-supported Nudix hydrolase activity, whereas Mg$^{2+}$ and Ca$^{2+}$ had little effect. Mn$^{2+}$ lowered the activity of the Co$^{2+}$-supported phosphatase activity, but did not eliminate the activity, since Mn$^{2+}$ supports phosphatase activity itself, albeit at a lower level than does Co$^{2+}$.

Site-directed mutagenesis studies

YZGD’s PLPase activity exhibited by YZGD is consistent with YZGD being a member of the HAD superfamily as there are a number of phosphatases within this superfamily [4]. Likewise, the hydrolysis of CDP-alcohols, ADP-coenzymes and nucleotide sugars is consistent with YZGD being a member of the Nudix hydrolase superfamily [3]. Conversely, the Nudix hydrolases are not expected to be phosphatases nor are HAD superfamily members expected to be pyrophosphohydrolases [3,4]. To establish unequivocally that for YZGD, it is the HAD superfamily motif that is responsible for PLPase activity, and the Nudix
signature sequence that is responsible for the Nudix hydrolase activity, site-directed mutagenesis studies were carried out.

In the first mutant, YZGDD12N, Asp12 was changed to an asparagine residue (Figure 6). The corresponding aspartate residue in other HAD superfamily members has been shown to be within the active site, to co-ordinate the bivalent metal cation necessary for catalysis, and to be essential for catalysis [4]. As expected, there was no change in the expression, solubility or Nudix hydrolase activity of YZGDD12N when compared with wild-type YZGD, but there was a complete loss of PLPase activity.

For the second mutant, YZGDE324Q, Glu324 was changed to a glutamine residue (Figure 6). This amino acid is one of the four glutamates conserved among Nudix hydrolases [3] and essential for Nudix hydrolase activity [36,37]. When this corresponding glutamate was mutated in MutT, the activity of MutT decreased more than 100 000-fold [36]. Upon NMR analysis, Lin et al. [36] discovered that this glutamate is necessary for co-ordinating the bivalent metal cation required for catalysis. When YZGDE324Q was analysed, there was no difference in the expression, solubility or PLPase activity between the mutant or the wild-type enzyme; however, there was complete loss of all Nudix hydrolase activity for all Nudix substrates.

Potential role of YZGD
The only other purified and characterized PLP-specific phosphatase is the human PLP-specific phosphatase [13,14,38,39]. In many cases, alkaline phosphatase is responsible for PLPase activity [40,41]. Although the importance of PLPase activity has been established [42], the exact nature of its role in the cell has not. Fonda and Zhang [39] described how the human PLP phosphatase might be important in vitamin B6 metabolism. Erythrocytes can
I. M. Tirrell and others

Figure 6 Alignment of YZGD with other members of the nitrophosphatase family

The amino acid sequences of YZGD, NAGD from E. coli, the p-nitrophosphatase (pNpase) from S. pombe and S. cerevisiae and the human PLPase were aligned using CLUSTAL X [26]. Similar and identical amino acids are in boldface. Gaps in the amino acid sequences inserted to optimize alignment are indicated by a ‘-’. Motifs I, II and III of the HAD superfamily are underlined and labelled. The 23 amino acid Nudix signature sequence in YZGD is in boldface. The amino acids Asp12 (within the HAD domain of YZGD) and Glu324 (within the Nudix signature sequence of YZGD), which were modified by site-directed mutagenesis, are highlighted in grey. The [S/T]NN and [F/W/Y][I/L/V] motifs, which appear to be predictive markers for the nitrophosphatase family are denoted by a ‘*’, as are the conserved serine between motifs I and II and the conserved GKP sequence of motif III.

Because YZGD hydrolyses CDP-alcohols, ADP-coenzymes, ADP, TDP, UDP and GDP sugars and one of the products of the reaction is always CMP, AMP, TMP and GMP from a wide array of nucleotide derivatives. YZGD’s ability to generate all of the ribonucleoside monophosphates from a number of sources suggests a role for the enzyme in nucleotide salvage, Designating a new function for a Nudix hydrolase.

Bioinformatics and comparison of YZGD to other family members

We initially uncovered YZGD from P. thiaminolyticus through BLAST searches of Nudix hydrolases, since YZGD contains the Nudix signature sequence GX5EX7REUXEEXGU (where U = I, L or V). A subsequent BLAST search of YZGD uncovered the two yeast p-nitrophosphatases and a number of open reading frames that compose the nitrophosphatase family of the HAD superfamily, in addition to members of the Nudix hydrolase family. Therefore we performed two separate BLAST searches, one on the first 250 amino acids of YZGD with sequence similarity to nitrophosphatase family members, and another on the last 163 amino acids containing the Nudix signature sequence. Results from these two BLAST searches can be seen in Figures 2 and 3 respectively.

Although the number of identified and characterized enzymes of the nitrophosphatase family is currently limited (Figure 6), there are several hundred open reading frames that show sequence similarity to the identified enzymes (Figure 2). Thus enzymes of this family appear to be ubiquitous, existing in organisms ranging from bacteria to humans. There is a potentially wide variety of functions that the phosphatases of this family may possess, as the substrates so far include pyridoxal phosphate, UMP and p-nitrophosphatase. All of these substrates are structurally similar in that they all contain an aromatic ring.

The original name of the family was the nitrophosphatase family, because initially only the yeast nitrophosphatases were known. Jang et al. [15] assumed that this was a PLPase family. Many open reading frames of the family are designated as sugar phosphatases, probably because NAGD is within the N-acetylglucosamine operon, yet none of the characterized enzymes including NAGD are sugar phosphatases. Misannotation is common when the number of known enzymes is limited.

The nitrophosphatase family is a distinct family of the HAD superfamily [1]; BLAST searches of the known enzymes of
only the UDP-sugar hydrolases contain the ‘NGD’ UDP-sugar signature sequence. Note, however, that both YZGD and many of the Nudix hydrolases from B. cereus require Mn²⁺ instead of the more typical Mg²⁺. Vital to our characterization of YZGD as YZGD also absolutely

the family retrieve one another and putative family members well before they retrieve other HAD superfamily members. Enzymes of the HAD superfamily are distinguished by three conserved motifs: motif I (DXDX[T/V][L/V]), motif II ([S/T][N/D]) and motif III (K-[G/S][D/S][I/X][D/N]) [45], or more generally, an aspartate in motif I, a serine or threonine in motif II, and a lysine or arginine and one to three aspartates or glutamates in motif III [4]. While members of the nitrophenyl phosphatase family contain these motifs (Figure 6), there is considerably more sequence similarity amongst these family members (Figure 2). Much of this sequence similarity lies between motifs I and II. This region corresponds to the substrate specificity loop described by Lahiri et al. [5] and thus these additional conserved residues may dictate this family’s substrate specificity. This family appears to recognize aromatic phosphates, and consistent with this substrate specificity, there is a conserved aromatic amino acid within an [F/W/Y][I/L/V] motif. There are also a number of conserved glycines and bulky aliphatics. More interestingly, there is a conserved [S/T][N/N] motif and a conserved serine residue within this region. Additionally, OKP is conserved within motif III instead of just a conserved lysine.

The [S/T][N/N] and [F/W/Y][I/L/V] motifs appear to be unique to the nitrophenyl phosphatase family. The following families of the HAD superfamily were analysed for these motifs: HADs, epoxide hydrolases, MDP-1 family members, Phospho-1 ethanolamine/choline phosphatases, P-type ATPases, histidinol phosphatases, phosphoglycolates and trehalose phosphatases. A BLAST search was performed on a known enzyme of each family, and five unique sequences were aligned for each. Upon visual observation, the [S/T][N/N] and [F/W/Y][I/L/V] motifs are absent in these families, thus verifying that these motifs can be used as predictive markers of the family.

When a hidden Markov model was built from the nitrophenyl phosphatase family, a mouse nitrophenyl phosphatase [15], not used to build the model, gave an E (expectation) value of 6 × 10⁻¹⁹, while enzymes from the other HAD families gave E values of approx. 1, comparable with the results for a completely unrelated DNA ligase. This hidden Markov model search verifies that the nitrophenyl phosphatases are considerably more conserved with one another than they are with other HAD superfamily members.

A standard BLAST search of non-redundant databases, using the last 163 amino acids of YZGD, was performed to uncover the related Nudix hydrolases. Of the top 20 hits, 12 are from B. anthracis, B. cereus or B. thuringiensis (Figure 3), presumably due to close evolutionary relationships of the organisms. Fortuitously, the Nudix hydrolases of B. cereus have been characterized [16], and we anticipated that this study would help in determining the activity of YZGD. We did a BLAST search of the last 163 amino acids of YZGD, searching only the B. cereus ATCC14579 database, to identify as many unique paralogues of known function as possible. The first five hits are UDP-sugar hydrolases (Figure 7). Though YZGD does hydrolyse UDP-sugars, they are not the best Nudix substrates for YZGD (Table 2). The bioinformatics data corroborate this finding as the UDP-sugar hydrolases contain an ‘NGD’ triad [16] that is missing in YZGD (Figure 7). However, the bioinformatics data were useful in that many of the B. cereus enzymes use Mn²⁺ instead of Mg²⁺ for their activation [16]. This observation was vital to our characterization of YZGD as YZGD also absolutely requires Mn²⁺. These bioinformatics results are consistent with YZGD being a unique Nudix hydrolase.

This work was supported by a grant (J-497) from the Thomas and Kate Jeffress Trust Foundation, a Cottrell College Science Award (CC5180) from Research Corporation (Tucson, AZ, U.S.A.), and an AREA grant (GM066715-01) from the National Institutes of Health. We also acknowledge the expertise of Dr Shuba Gopal (Department of Biological Sciences, Rochester Institute of Technology, Rochester, NY, U.S.A.) in the use of the hidden Markov model analysis.

REFERENCES

16 O’Handley, S. F., Dunn, C. A. and Bessman, M. J. (2001) Orf135 from Escherichia coli is a Nudix hydrolase specific for CTP, dCTP and 5-methyl-dCTP. J. Biol. Chem. 276, 24861–24865
17 Reference deleted
23 O’Handley, S. F., Dunn, C. A. and Bessman, M. J. (2001) Orf135 from Escherichia coli is a Nudix hydrolase specific for CTP, dCTP and 5-methyl-dCTP. J. Biol. Chem. 276, 5421–5426

Received 19 July 2005; accepted 8 December 2005
Published as BJ Immediate Publication 8 December 2005, doi:10.1042/BJ20051172

© 2006 Biochemical Society