Investigation of Metal Ion Binding in Phosphonoacetaldehyde Hydrolase Identifies Sequence Markers for Metal-Activated Enzymes of the HAD Enzyme Superfamily[†,‡]

Guofeng Zhang,[§] Marc C. Morais,[||] Jianying Dai,[§] Wenhai Zhang,[§] Debra Dunaway-Mariano,*[,§] and Karen N. Allen*[,||]

Department of Chemistry, University of New Mexico, Albuquerque, New Mexico 87131, and Department of Physiology and Biophysics, Boston University School of Medicine, Boston, Massachusetts 02118-2394

Received December 23, 2003; Revised Manuscript Received February 23, 2004

ABSTRACT: The 2-haloalkanoic acid dehalogenase (HAD) family, which contains both carbon and phosphoryl transferases, is one of the largest known enzyme superfamilies. HAD members conserve an α,β-core domain that frames the four-loop active-site platform. Each loop contributes one or more catalytic groups, which function in mediating the core chemistry (i.e., group transfer). In this paper, we provide evidence that the number of carboxylate residues on loop 4 and their positions (stations) on the loop are determinants, and therefore reliable sequence markers, for metal ion activation among HAD family members. Using this predictor, we conclude that the vast majority of the HAD members utilize a metal cofactor. Analysis of the minimum requirements for metal cofactor binding was carried out using Mg(II)-activated *Bacillus cereus* phosphonoacetaldehyde hydrolase (phosphonatase) as an experimental model for metal-activated HAD members. Mg(II) binding occurs via ligation to the loop 1 Asp12 carboxylate and Thr14 backbone carbonyl and to the loop 4 Asp186 carboxylate. The loop 4 Asp190 forms a hydrogen bond to the Mg(II) water ligand. X-ray structure determination of the D12A mutant in the presence of the substrate phosphonoacetaldehyde showed that replacement of the loop 1 Asp, common to all HAD family members, with Ala shifts the position of Mg(II), thereby allowing innersphere coordination to Asp190 and causing a shift in the position of the substrate. Kinetic analysis of the loop 4 mutants showed that Asp186 is essential to cofactor binding while Asp190 simply enhances it. Within the phosphonatase subfamily, Asp186 is stringently conserved, while either position 185 or position 190 is used to position the second loop 4 Asp residue. Retention of a high level of catalytic activity in the G185D/D190G phosphonatase mutant demonstrated the plasticity of the metal binding loop, reflected in the variety of combinations in positioning of two or three Asp residues along the seven-residue motif of the 2700 potential HAD sequences that were examined.

In this paper, we examine metal ion activation of phosphonoacetaldehyde hydrolase (phosphonatase[1])-catalyzed hydrolysis of phosphonoacetaldehyde (Pald) to orthophosphate and acetaldehyde (Figure 1). Phosphonatase is a member of the 2-haloalkanoate dehalogenase (HAD) enzyme superfamily (*1*, *2*). This is an exceptionally large family of enzymes (~2700 members), which includes carbon (2-haloalkonoate dehalogenases) and phosphorus hydrolases (ATPases, phosphate monoesterases, and hexose phosphate mutases) (*3*). The members of this family catalyze nucleophilic substitution reactions at phosphorus or carbon centers, using a conserved Asp carboxylate in covalent catalysis. Like phosphonatase (*4*), most, but not all, members of the superfamily possess a cap domain, the fold and function of which are variable. The

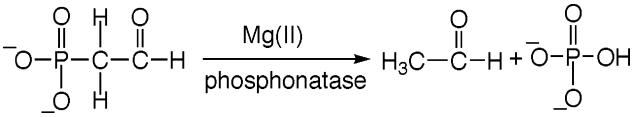

FIGURE 1: Schematic of the reaction catalyzed by phosphonatase.

[†] This work was supported by NIH Grant GM61099 to K.N.A. and D.D.-M.

[‡] Coordinates have been deposited in the Protein Data Bank, as entries 1SWV and 1SWW for the D12A phosphonatase−Mg(II) and D12A phosphonatase−Mg(II)−phosphonoacetaldehyde structures, respectively.

* To whom correspondence should be addressed. K.N.A.: Department of Physiology and Biophysics, Boston University School of Medicine, 715 Albany St., Boston, MA 02118-2394; telephone, (617) 638-4398; fax, (617) 638-4273; e-mail, allen@med-xtal.bu.edu. D.D.-M.: Department of Chemistry, University of New Mexico, Albuquerque, NM 87131; telephone, (505) 277-3383; fax, (505) 277-6202; e-mail, dd39@unm.edu.

[§] University of New Mexico.

[||] Boston University School of Medicine.

[1] Abbreviations: ADH, alcohol dehydrogenase; ATPase, Ca(II)-transporting ATPase; BPGM Ll, *Lactococcus lactis* β-phosphoglucomutase; HAD, 2-haloalkanoic acid dehalogenase; dNT-2 Hs, *Homo sapiens* mitochondrial deoxyribonucleotidase; MLHL, maximum likelihood Hendrickson−Lattman; Pald, phosphonoacetaldehyde; phosphonatase, phosphonoacetaldehyde hydrolase; PSP Mj, *Methanococcus jannaschii* phosphoserine phosphatase; PNK, polynucleotide kinase; rmsd, root-mean-square deviation; sEH, soluble epoxide hydrolase; YrbI Hi, *Haemophilus influenzae* deoxy-D-mannose-octulosonate 8-phosphate phosphatase.

10.1021/bi036309n CCC: $27.50
Published on Web 04/09/2004

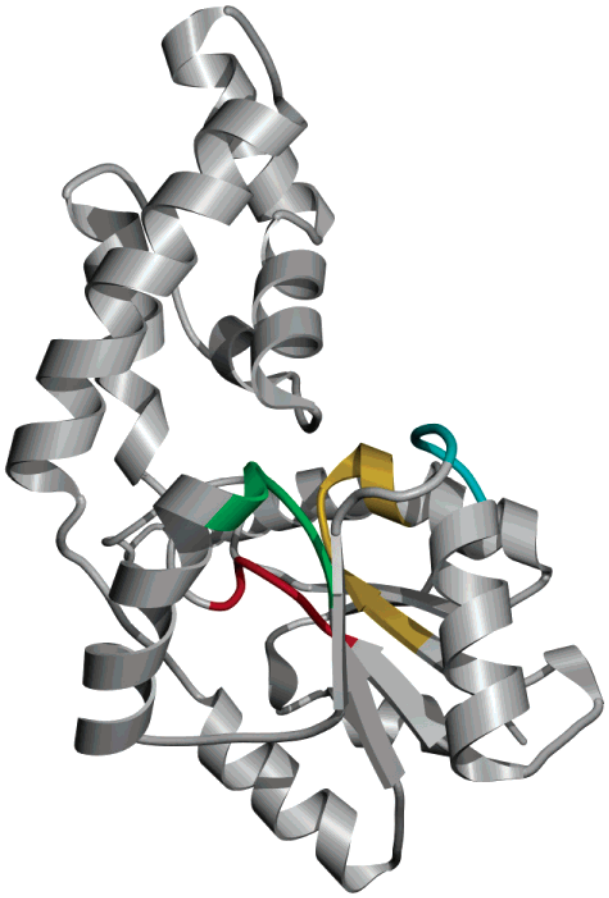

FIGURE 2: Phosphonatase structure (*4*) depicted as a ribbon diagram with catalytic loops colored as follows: red for loop 1, green for loop 2, cyan for loop 3, and gold for loop 4.

α,*β*-core domain, common to all members, is highly conserved in the fold (e.g., see the phosphonatase structure in Figure 2).

The large size of the HAD family attests to its importance to cellular function, yet only a small fraction of the members of the family have known catalytic function. Moreover, the chemical boundaries of catalysis within the family are not yet set. The goal of our work is to identify sequence markers that will assist investigators in function assignment. In the study presented here, we examine the metal-binding motif of the HAD family phosphotransferases as a potential predictor of metal cofactor usage in family members of unknown function. Phosphonatase served as the experimental model.

In all HAD members, the active-site platform is formed by four peptide segments (hereafter termed loops 1−4) donated by the core domain (see the phosphonatase active-site platform highlighted in Figure 2). The loops are comprised of conserved sequences used to identify family members. Each loop contributes one or more catalytic groups. Loop 1 positions the Asp nucleophile, which is used to mediate phosphoryl or carbon group transfer in all HAD members. Loop 2 positions a hydrogen bond donor (Thr or Ser) to assist in substrate binding. Loop 3 contributes a positively charged group (Arg or Lys) to orient the electrophile and/or nucleophile.

The backbone of loop 4 traces a seven-residue hairpin−turn motif in which the carboxylate groups (most commonly Asp), which function in metal binding in the phosphotransferases (*4*−*8*), are positioned (we will refer to these residues as stations 1−7 starting from the N-terminus of the loop). Among the eight known phosphotransferase structures, six position two carboxylate residues and two position three carboxylate residues on loop 4 (see the structure-based sequence alignment in Figure 3). The carboxylate residues are located on station 2, 3, 6, or 7 of loop 4. Position 2 is used in six of the eight structures; position 3 is used in seven of the eight structures, while position 6 or 7 is used in five of the eight structures. This variation in the use of stations suggests plasticity in loop 4 metal binding, a feature favorable to the evolution of new function. On the other hand, it complicates the application of loop 4 stations for distinction between HAD members requiring metal activation and HAD members that do not. Instead, the number of carboxylate residues might be used. Indeed, inspection of the alignment of HAD family sequences shows that the dehalogenases have but a single carboxylate at loop 4 (*9*) while the phosphotransferases have two or three. If it can be demonstrated that metal ion binding requires the participation of at least two carboxylate groups from loop 4, then HAD sequences might be sorted on the basis of the number of carboxylate residues stationed on loop 4.

```
(A)... ...1VG DTVSD7... PHN Bc (pdb: 1FEZ)
(B)... ...1VG DGAND7... PSP Mj (pdb: 1J97)
(C)... ...1LE DSQAG7... BPGM Ll (pdb: 1O03)
(D)... ...1GD DSVDL7... YrbI Hi (pdb: 1K1E)
(E)... ...1ID DRPDI7... dNT-2 Hs (pdb: 1MH9)
(F)... ...1LD DFGSN7... sEH Murine (pdb: 1CQZ)
(G)... ...1ID DRTQV7... PNK T4 phage (pdb: 1LTQ)
(H)... ...1GD GVNDA7... ATPase Rabbit (pdb: 1EUL)
```

FIGURE 3: Positions of the carboxylate metal ligands on loop 4 of (A) phosphonatase, (B) phosphoserine phosphatase, (C) *β*-phosphoglucomutase, (D) deoxy-D-mannose-octulosonate 8-phosphate phosphatase, (E) mitochondrial deoxyribonucleotidase, (F) soluble epoxide hydrolase, (G) polynucleotide kinase, and (H) Ca(II)-ATPase. The carboxylate groups that bind the metal cofactor (directly or through water) are underlined. Assignment of the metal ligands in the apo structures [soluble epoxide hydrolase, polynucleotide kinase, and Ca(II)-ATPase] was carried out by modeling Mg(II) at the binding site by superimposing these three structures with the Mg(II)-bound structures.

In this work, we have used X-ray crystallography in combination with steady-state kinetic analysis of site-directed mutants of the HAD family member phosphonatase to test the hypothesis that both carboxylate residues on loop 4 function in combination with the conserved Asp on loop 1 in Mg(II) binding and substrate positioning. Here we report that replacement of the loop 1 Asp12 with Ala causes Mg(II), and hence substrate, to bind nonproductively. Replacement of the loop 4 Mg(II) ligand Asp186 (station 3) with Ala precludes catalysis, while replacement of loop 4 Asp190 (station 7) with Ala impairs Mg(II) binding and catalysis. In addition, high activity in the phosphonatase double mutant G185D/D190G is reported as evidence of loop 4 plasticity (i.e., stations 2 and 7 are interchangeable).

EXPERIMENTAL PROCEDURES

Preparation of Wild-Type and Mutant Phosphonatase Enzymes. Pald was prepared according to the published procedure (*10*). The wild-type and D12A mutant *Bacillus cereus* enzymes were prepared from the *Escherichia coli* clone as previously described (*2*). The D12E, D186A, D186E, D190A, D186A/D190A, and G185D/D190G phos-

phonatase mutant genes were generated by PCR using the pKK223-3 plasmid, containing the wild-type phosphonatase gene as a template, and verified by DNA sequencing. The D186E mutant phosphonatase was insoluble and not characterized further. The other mutants were purified using the same procedure that was used to purify the wild-type enzyme in yields of 10−20 mg/g of wet cell paste. The chromatographic behavior, solubility, and stability to storage of the mutants were similar to those of the wild-type enzyme.

Steady-State Kinetic Constants. The K_m and V_{max} values for wild-type and mutant phosphonatases were determined from the initial velocity data measured as a function of Pald concentration. The 1 mL reaction solutions contained Pald ($0.5-10K_m$), 10 mM $MgCl_2$, 0.225 mM NADH, and 5 units of alcohol dehydrogenase (ADH) dissolved in 50 mM K^+-HEPES (pH 7.5 and 25 °C). The concentration of phosphonatase was in the range of 0.02−2 μM, depending on the mutant that was being studied. Reactions were monitored at 340 nm ($\Delta\epsilon = 6200\ M^{-1}\ cm^{-1}$) for the conversion of acetaldehyde and NADH to ethanol and NAD^+. The initial velocity data were analyzed using eq 1.

$$V_0 = V_{max}[S]/(K_m + [S]) \quad (1)$$

where [S] is the substrate concentration, V_0 is the initial velocity, V_{max} is the maximum velocity, and K_m is the Michaelis constant. The k_{cat} was calculated from V_{max} and the enzyme concentration using the equation $k_{cat} = V_{max}/[E]$. The enzyme concentration was determined using the Bradford method (*11*).

Mg(II) Activation. Metal-free wild-type, D190A, and G185D/D190G phosphonatase were prepared by dialysis (1−2 mL of protein) against four (wild-type phosphonatase) or two (mutant phosphonatase) changes (3 h each) of 0.5 L of buffer A [50 mM K^+HEPES containing 30 mM EDTA and 1 mM DTT (pH 7.5 and 4 °C)] and then against 1 L of buffer B [50 mM K^+HEPES containing 1 mM DTT (pH 7.5 and 4 °C)] to remove the EDTA. When assayed in the absence of added (10 mM) $MgCl_2$, the mutant enzymes had no activity. To test for loosely bound Mg(II), the wild-type and mutant enzymes were dialyzed, in the absence of EDTA, against four changes (3 h each) of 0.5 L of buffer B (pH 7.5 and 4 °C), and enzyme activity was assayed. Whereas the dialyzed wild-type enzyme retained its full activity when assayed in the absence of added $MgCl_2$, the dialyzed G185D/D190G and D190A mutants were not active in the absence of added $MgCl_2$, but regained full activity when 10 mM $MgCl_2$ was included.

The original activities of the dialyzed enzymes were restored when they were assayed in the presence of 10 mM $MgCl_2$. The k_{cat} and K_m values for Mg(II) activation of the metal-free enzymes were measured using 1 mL reaction solutions containing 15 units of ADH, 225 μM NADH, 1 mM Pald, and varying concentrations of Mg(II) ($0.5-10K_m$) in 50 mM K^+HEPES (pH 7.5 and 25 °C).

Crystallization and Data Collection. Wild-type and D12A phosphonatase were prepared at a concentration of 10 mg/mL dissolved in 1 mM K^+HEPES, 10 mM $MgCl_2$, and 0.1 mM DTT (pH 7.5 and 4 °C). Protein crystals were obtained by using the vapor diffusion method (hanging-drop geometry at 18 °C) as previously described for crystallization of the wild-type enzyme (*12*) [viz., 10 μL each of the protein solution and the well solution consisting of 30% PEG 4000, 100 mM Tris-HCl (pH 7.4), and 100 mM $MgCl_2$]. Large (0.4 mm per side) crystals grew within 1 week. For the substrate complex, crystals were soaked overnight in mother liquor containing 500 μM Pald. Before data collection, crystals were soaked (1−12 h) in well solution with 20% glycerol and frozen in a stream of nitrogen gas cooled by liquid nitrogen at −180 °C. The 2.3 and 2.5 Å data sets were collected for the D12A phosphonatase−Mg(II) complex and D12A−Mg(II)−Pald complex, respectively, on beamline X12B at Brookhaven National Laboratory's National Synchrotron Light Source (Upton, NY) using a 60 mm MAR detector. A 2.5 Å Pt derivative data set used only for empirical phasing was collected on an R-AXIS II area detector using monochromatic Cu Kα radiation generated by a Rigaku RU300 rotating-anode X-ray generator.

DENZO and *SCALEPACK* (*13*) were used for data indexing, reduction, and scaling. Crystals of the D12A mutant are monoclinic, belonging to space group $P2_1$. The unliganded D12A mutant has the following unit cell dimensions: $a = 64.50$ Å, $b = 62.50$ Å, $c = 73.76$ Å, and $\beta = 108.8°$. The substrate-bound D12A mutant was isomorphous, with the following unit cell dimensions: $a = 63.82$ Å, $b = 62.52$ Å, $c = 72.62$ Å, and $\beta = 108.0°$. If a Matthews coefficient of 2.3 is assumed, the unit cell dimensions are consistent with a dimer in the asymmetric unit for both the D12A mutant enzyme crystals (*14*). Data collection statistics are summarized in Table 1.

Structure Solution and Model Refinement. The molecular replacement method was used to phase both data sets (*15*, *16*). The 3.0 Å structure of the tungstate-bound, wild-type phosphonatase (PDB entry 1FEZ), with ligand and water molecules removed, was used as the search model in *AMoRE*

Table 1: Data Collection, Refinement Statistics, and Geometry of the Final Models for D12A Phosphonatase Bound to Mg(II) Alone and to Mg(II) and Pald

	D12A−Mg(II)	D12A−Mg(II)−Pald
Data Collection Statistics		
resolution range (Å)	100−2.3	100−2.55
no. of unique reflections	27388	22144
completeness (%)[a]	91.5 (86.7)	98.3 (97.3)
$I/\sigma(I)$	12.5 (2.9)	12.3 (3.0)
R_{merge}[b] (%)	6.4 (44.2)	8.8 36.0)
Refinement and Model Statistics		
no. of reflections used in refinement	24386	20143
final R_{work}[c]	0.249	0.252
final R_{free}[d]	0.265	0.274
average B value (Å^2)	30.9	36.9
main chain	28.8	34.3
side chain	33.0	39.4
rmsd from ideality		
bond lengths (Å)	0.009	0.009
bond angles (deg)	1.39	1.42
dihedral angles (deg)	21.0	22.2
B rmsd for bonded atoms (Å^2)	4.2	6.6
main chain	6.4	9.3
side chain		
estimated coordinate error (Å)	0.37	0.43

[a] Statistics for the outermost shell (2.38−2.30 and 2.65−2.55 Å for the unliganded and liganded Asp12Ala phosphonatase, respectively) are given in parentheses. [b] $R_{merge} = \sum_{hkl}\sum_i|I_{hkl,i} - \langle I_{hkl}\rangle I|/\sum_{hkl}\sum_i|I_{hkl,i}|$, where $\langle I_{hkl}\rangle$ is the mean intensity of the multiple $I_{hkl,i}$ observations for symmetry-related reflections. [c] $R_{work} = \sum_{hkl}|F_{obs} - F_{calc}|/\sum_{hkl}|F_{obs}|$. [d] $R_{free} = \sum_{hkl\sum T}|F_{obs} - F_{calc}|/\sum_{hkl}|F_{obs}|$, where test set T includes 7% of the data.

Table 2: Steady-State Kinetic Constants for Wild-Type and Mutant Phosphonatase Measured in Solutions Containing Varying Concentrations of Phosphonoacetaldehyde, 10 mM $MgCl_2$, 0.23 mM NADH, and 15 Units/mL Alcohol Dehydrogenase in 50 mM K^+HEPES (pH 7.5 and 25 °C)[a]

enzyme	k_{cat} (s^{-1})	K_m (μM)	k_{cat}/K_m (s^{-1} M^{-1})
wild type	15 ± 1	33 ± 2	5.0×10^5
D12E	$(1.2 \pm 0.3) \times 10^{-3}$	72 ± 4	1.6×10
D186A	$(8.4 \pm 0.2) \times 10^{-5}$	–	–
D190A	$(2.20 \pm 0.04) \times 10^{-2}$	520 ± 30	4.2×10
D186A/D190A	not active		
G185D/D190G	1.7 ± 0.3	54 ± 3	3.1×10^4

[a] Note that in the case of D186A, the activity of the enzyme was too low to accurately measure K_m.

(*16*) using data between 10.0 and 4.2 Å. The wild-type dimer is composed of one "open" monomer and one "closed" monomer (*4*). For the D12A−Mg(II) and the Pald-bound D12A data sets, the search model corresponding to two closed monomers resulted in the best initial *AMoRE* solution. The resulting *AMoRE* solution was subjected to rigid body refinement in *CNS* (*17*).

To facilitate refinement and reduce model bias for the D12A phosphonatase, the phases from molecular replacement were supplemented with experimental phases. A 2.5 Å data set was collected on a Pt isomorphous derivative obtained by soaking crystals for 1.5 h (in the dark) in a freshly prepared artificial mother liquor solution containing 1 mM K_2PtCl_4. Seven Pt sites were located by examination of a difference Fourier map using phases obtained from the initial molecular replacement solution. Hendrickson−Lattman coefficients from the Pt solution were used in the initial stages of refinement as part of a maximum likelihood Hendrickson−Lattman (MLHL) minimization target (*18*). Iterative cycles of minimization and simulated annealing using slow-cool torsional molecular dynamics and individual *B*-factor refinement were implemented in CNS [excluding 7% of the data for the calculation of R_{free} (*19*)] followed by manual rebuilding using $2F_o - F_c$ and $F_o - F_c$ maps in the graphics program O (*20*). NCS restraints (300 kcal mol^{-1} $Å^{-2}$ in initial rounds and 50 kcal mol^{-1} $Å^{-2}$ in the final round) between the two monomers in the dimer were used in all stages of refinement. Waters were added using a 3.0σ cutoff in $F_o - F_c$ maps (162 total and 120 total for the unliganded and Pald-bound D12A structures, respectively). For the liganded structure, a model of Pald was built in QUANTA (Molecular Simulations Inc.) and topology and parameter files for refinement were generated using XPLO2D (*21*, *22*). Residues 1−4 and 261−267 (of 267 residues) were not visible in the electron density map, and were omitted from the final model. Analysis of the Ramachandran plot of the final model showed good geometry as defined by PROCHECK (*23*) for both structures. Refinement statistics and model geometry are summarized in Table 1.

RESULTS AND DISCUSSION

Catalytic Activities of Site-Directed Mutants. The steady-state kinetic properties of the phosphonatase mutants are listed together with those of wild-type phosphonatase in Table 2. In anticipation that metal binding might be weakened in these mutants, 10 mM $MgCl_2$ was included in the assay mixtures. The loop 1 Asp12 carboxylate group coordinates Mg(II) and serves in nucleophilic catalysis (Figure 4A). Replacement of the carboxylate side chain with a methyl group, as in the D12A mutant, eliminates both functions. The D12A mutant was found to be devoid of catalytic activity, but otherwise thermodynamically stable. This mutant was crystallized and then soaked with the substrate Pald to determine the structure. The crystal structure is described in the following section.

Substitution of Asp12 with Glu conserved the carboxylate group for metal binding and nucleophilic catalysis in the D12E mutant, but the extended side chain of the Glu residue was expected to alter the orientation of the carboxylate group relative to the metal ion and nucleophile and, consequently, impair catalysis. Indeed, the k_{cat} is reduced 10000-fold and the k_{cat}/K_m 30000-fold in the D12E mutant.

The loop 4 Asp186 carboxylate group coordinates to the Mg(II) cofactor (Figure 4A). Substitution of Asp186 with Glu resulted in an insoluble enzyme. Substitution with Ala produced a soluble but "essentially inactive" enzyme. Clearly, a carboxylate at position 186 (station 2) is essential for catalytic function.

The loop 4 Asp190 carboxylate group forms a hydrogen bond with the Mg(II) water ligand (Figure 4A). The catalytic efficiency of the D190A mutant was first evaluated by measuring the activity in reaction solutions containing (near) saturating Mg(II) and varying Pald concentrations. Under these conditions, the k_{cat} was reduced 700-fold and (Pald) k_{cat}/K_m 10000-fold (Table 2). To determine the impact of the mutation on Mg(II) binding, the enzyme activity was measured with reaction solutions containing saturating Pald and varying Mg(II) concentrations. The K_m of Mg(II), which is the thermodynamic equivalent of the K_d for dissociation of Mg(II) from the mutant phosphonatase−Pald−Mg(II) complex, was determined to be 3.5 mM (Table 3). The corresponding K_d for the complex formed from the wild-type phosphonatase is not known precisely. However, it may be assumed to be less than 1 μM, because the Mg(II) is not removed by dialysis against a buffer. Therefore, it can be concluded that the loss of Asp190 function has a dramatic effect on Mg(II) binding and catalysis.

The impact of Asp190 substitution on the catalytic efficiency of the *B. cereus* phosphonatase might not have been anticipated given the fact that in five of the seven reported phosphonatase sequences the residue at the equivalent position is not Asp but rather Gly. However, each of those five phosphonatase sequences has an Asp residue stationed at position 185. A simple modeling exercise, in which Gly185 in the *B. cereus* enzyme is replaced with Asp, demonstrates that if the side chain position at 185 is retained, the carboxylate group of Asp185 would superimpose with, and therefore replace, the water liganded to Mg(II) in the *B. cereus* phosphonatase structure. To describe the plasticity of site usage on loop 4, we will refer to the seven positions in loop 4 as stations 1−7 from the N- to C-terminus. Thus, within the phosphonatase subfamily, two different loop 4 metal binding configurations are seen: Asp at stations 3 and 7 in one case and Asp at stations 2 and 3 in the other. The phosphonatase sequences with Asp at loop 4 stations 2 and 3 are from Gram-negative bacteria, while the phosphonatases with Asp at loop 4 stations 3 and 7 are from Gram-positive bacteria. The divergence of the loop 4 metal binding residues

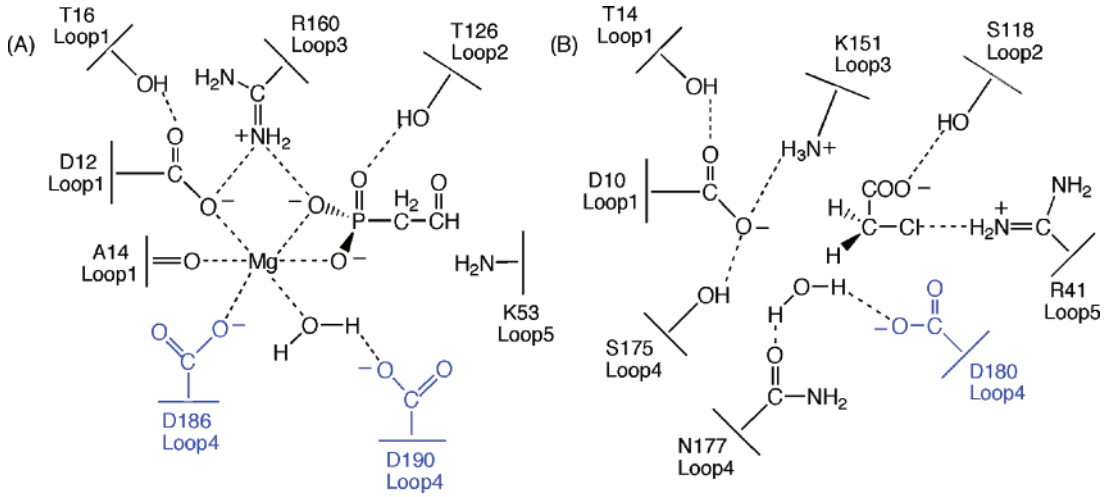

FIGURE 4: Schematic of the active-site residues utilized in substrate and metal binding in the HAD family members (A) phosphonatase and (B) 2-haloalkanoate dehalogenase (*9*, *32*, *33*). Loop 4 carboxylates are shown in blue.

Table 3: Steady-State Kinetic Constants for (dialyzed) Mutant Phosphonatase Measured in Reaction Solutions Containing Varying Concentrations of $MgCl_2$ (0.5-10K_m), 1 mM Phosphonoacetaldehyde, 0.23 mM NADH, and 15 Units/mL Alcohol Dehydrogenase in 50 mM K^+HEPES (pH 7.5 and 25 °C)

enzyme	k_{cat} (s^{-1})	K_m (μM)	k_{cat}/K_m (s^{-1} M^{-1})
D190A	0.0472 ± 0.0008	3500 ± 200	1.4×10
G185D/D190G	0.67 ± 0.02	700 ± 50	1.0×10^3

apparently followed the divergence of Gram-negative and Gram-positive bacteria.

The G185D/D190G double mutant was prepared to test the hypothesis that loop 4 stations 2 and 7 are interchangeable. The catalytic efficiency of the double mutant was first evaluated by measuring the activity in reaction solutions containing saturating Mg(II) and varying Pald concentrations. Under these conditions, the k_{cat} and (Pald) k_{cat}/K_m were reduced by ~10-fold (Table 2). To determine the impact of the mutation on Mg(II) binding, the enzyme activity was measured with reaction solutions containing saturating Pald and varying Mg(II) concentrations. The K_m value of 0.7 mM for Mg(II) suggested significantly weaker binding to the mutant enzyme than to the wild-type enzyme, yet significantly stronger binding (and catalytic activity) than that found for the D190A mutant (Table 3). From these results, it is evident that station 2 of loop 4 can be used for placement of an Asp that will assist Asp186 at station 3 in metal cofactor binding. It is also clear that the "Asp185 for Asp190" swap is not an equal one, as significant Mg(II) binding affinity is lost. This can be attributed to the fact that the structural context of loop 4 in the phosphonatase from Gram-positive bacteria is certainly different from that of loop 4 in Gram-negative bacteria, with a direct effect on the exact positioning of the Mg(II) ligands.

Comparison of the Structures of the Wild-Type and D12A Phosphonatase−Mg(II) Complexes. The structures of the D12A phosphonatase−Mg(II) and D12A phosphonatase−Mg(II)−Pald complexes were determined at 2.3 and 2.5 Å resolution, respectively. Both structures are homodimers of 30 kDa subunits with the cap in the closed conformation relative to the core. The α-carbons of the D12A phosphonatase−Mg(II) structure can be superimposed on those of the D12A phosphonatase−Mg(II)−Pald structure with an rmsd of 0.48 Å. Thus, substrate binding does not appear to induce a significant global conformation change in the mutant enzyme. The α-carbons of D12A phosphonatase monomers in the Pald-bound and unbound complexes superimpose with those of the cap-closed monomer of the wild-type phosphonatase−Mg(II) complex (*4*) with rmsds of 0.61 and 0.58 Å, respectively, which indicates that the replacement of the active-site Asp12 with Ala destroys the catalytic turnover of the enzyme but does not significantly change the backbone conformation.

The electron density map of the active-site region of the D12A phosphonatase−Mg(II)−Pald complex is shown in stereo in Figure 5. This active site is compared with that of the wild-type phosphonatase−Mg(II)−tungstate complex in Figure 6 (*4*). Three ordered, active-site water molecules (Wat111, Wat120, and Wat168) are clearly seen in the mutant enzyme structure that were not previously observed in the structure of the wild-type enzyme (probably due to the resolution of the wild-type structure which was determined at 3.0 Å resolution). The most striking difference between the wild-type and mutant structures is the position of the bound Mg(II). Whereas in the wild-type phosphonatase−Mg(II)−tungstate complex, Mg(II) is coordinated by the carboxylates of Asp12 of loop 1 and Asp186 of loop 4, in the D12A phosphonatase−Mg(II)−Pald complex, the Mg(II) has shifted to additionally allow innersphere coordination of the carboxylate of Asp190 of loop 4 while maintaining coordination with Asp186 of loop 4. The change in the relative position of Mg(II) was most evident from the overlay of the two active sites, shown in Figure 6C. It can be seen that there would be a direct clash between the position of Mg(II) in the D12A mutant and that of the Asp12 side chain in the wild type.

The shift in the Mg(II) position is accompanied by a shift in the position of the "phospho" group of the ligand [assessed by comparison of the phosphate analogue tungstate in the wild type−Mg(II)−tungstate complex vs the Pald phosphonyl group in the D12A−Mg(II)−Pald complex]. As shown in Figure 6C, the phosphonyl group of Pald coordinates Mg(II), even though the Mg(II) has shifted in position toward Asp190. The result is a change in the relative position of the Pald substrate phosphonyl group when compared to that of the tungstate.

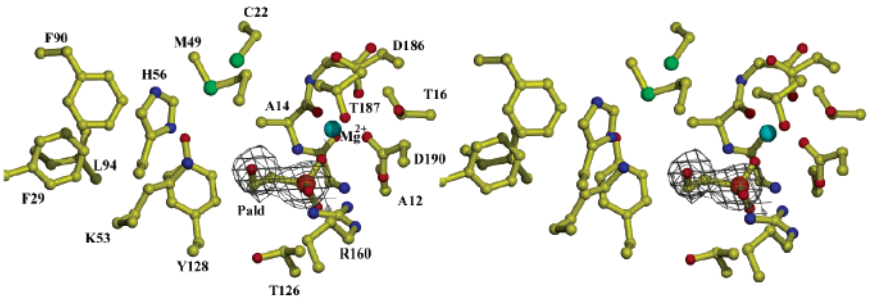

FIGURE 5: Stereodiagram of the electron density map of Pald bound at the active site of D12A phosphonatase. The amino acid residues are depicted as ball-and-stick models, and Mg(II) is depicted as a cyan sphere.

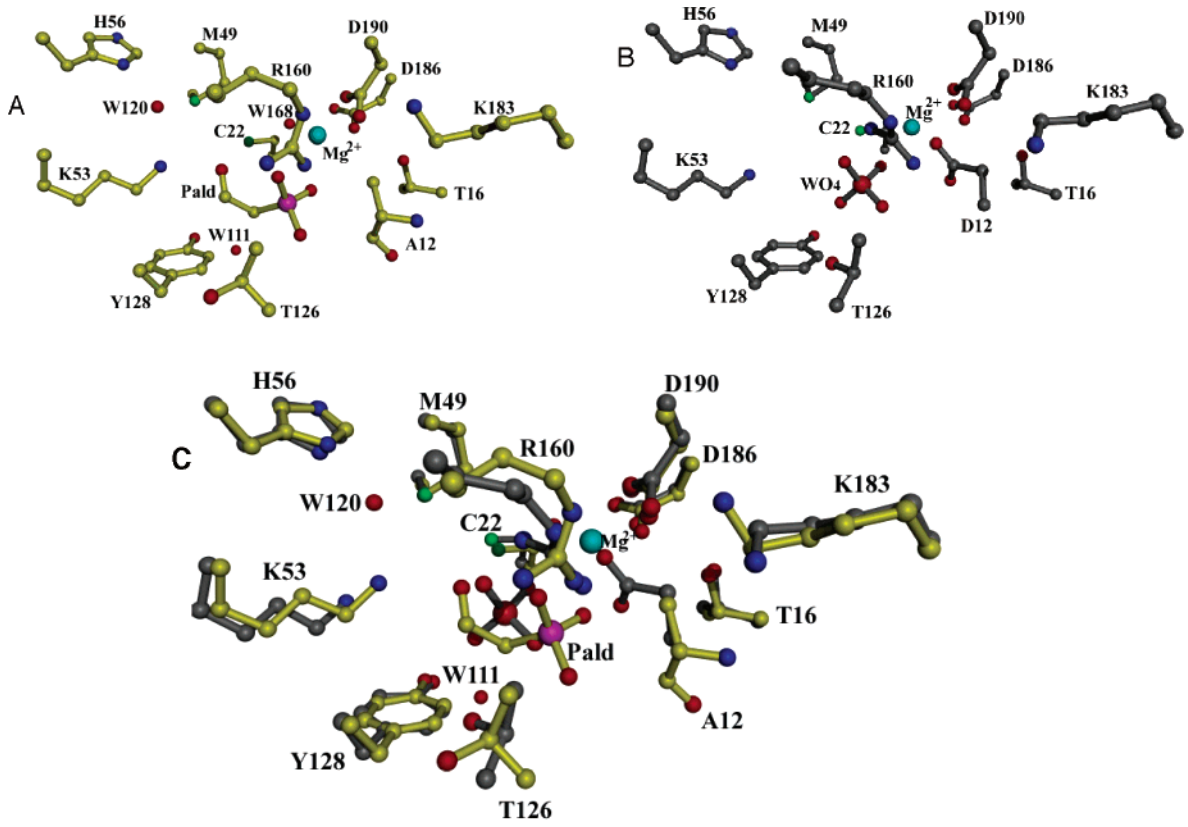

FIGURE 6: Active-site view of (A) D12A phosphonatase complexed with Pald and Mg(II) and (B) wild-type phosphonatase complexed with tungstate and Mg(II). (C) Overlay of the active site of D12A−Mg(II)−Pald (yellow) and wild type−Mg(II)−tungstate (dark gray) complexes (*4*). The Mg(II) from only the D12A−Mg(II)−Pald structure is depicted.

The observation that in the D12A structure both the Mg(II) and substrate ligand shift their positions in the active site to fill the space formerly occupied by the carboxylate ligand of Asp12 has important implications. First, it demonstrates the significant role that Mg(II) plays in anchoring and orienting the substrate ligand for catalysis. The substrate observed in the D12A mutant complex is bound in a nonproductive mode. Use of the D12A mutant with radiolabeled Pald in a single-turnover experiment analogous to that described previously for the wild-type enzyme (*24*) did not trap the Pald-derived Schiff base intermediate (data not shown). For the wild-type phosphonatase, the formation and consumption of a kinetically competent Schiff base intermediate had been clearly demonstrated.

Second, the altered binding mode demonstrates that the orientation of Mg(II) binding is dictated by the available enzyme ligands and not by the substrate. It can be anticipated that the Mg(II) will assume the most stable coordination structure and that this structure can be manipulated by the number of loop 4 carboxylates and their positions along the seven stations. The observation that, within the phosphonatase subfamily, Mg(II) coordination by Asp185 is interchanged with coordination by the outer sphere Asp190 and water is indicative of redundancy in sites and/or the plasticity of the coordination cage.

Occurrence of Multiple Carboxylate Groups in Loop 4 Sequence Motifs in HAD Superfamily Members. Group usage on loop 4 may be used to distinguish HAD members that bind a metal ion cofactor from those that do not. Loop 4 of the phosphotransferases (see, for example, the phosphonatase metal-binding site, in Figure 4A) is equipped with two or three carboxylate groups, as are the metal binding sites (pictured in Figure 1 of the Supporting Information) of phosphoserine phosphatase (*5*), *β*-phosphoglucomutase (*25*), deoxyribonucleotidase (*7*), sugar phosphate hydrolase (*8*), the lipid phosphatase domain of bifunctional phosphatase/epoxide hydrolase (unpublished results and ref *26*), the phosphatase domain of bifunctional T4 polynucleotide ki-

nase/phosphatase (*27*), and the P domain of membrane-bound Ca(II)-ATPase (*28*). Two loop 4 carboxyl groups function in combination with the loop 1 Asp nucleophile to bind the metal cofactor. In the 2-haloalkanoic acid hydrolases (Figure 4B), a metal cofactor is not required, and loop 4 is used to position groups that orient the Asp and water nucleophiles for attack at the respective carbon centers.

Inspection of the HAD family sequences grouped under pfam 00702 (*29*) [NCBI database, http://www.ncbi.nlm.nih.gov/Structure/cdd/cddsrv.cgi?uid=pfam00702 (July 14, 2003, release), and Sanger Institute protein family database, http://www.sanger.ac.uk/cgi-bin/Pfam/getacc?PF00702] and those specified in the list of related COGs (*30*, *31*) reveals that the vast majority contain two or more carboxylate residues on loop 4 (stations 1−7) and many conform to the sequence motif (GDXXXD). The haloalkanoate dehalogenases predominate among the sequences that contain fewer than two carboxylate residues. The nonconformists, confirmed by 3D-PSSM alignment (*29*), may have novel functions that do not employ metal ion catalysis.

If our prediction that metal binding requires two or more loop 4 carboxylate residues is correct, then it follows that most HAD members are metal-activated. The question then becomes, "are all metal-activated HAD members phosphotransferases?" The attack of the loop 1 Asp carboxylate at phosphorus requires charge shielding by the metal ion that bridges the nucleophile and electrophile, whereas the attack of the carboxylate nucleophile at carbon does not. Although we will continue our search for acyl transferases [coordination of Mg(II) to the carbonyl would serve to bind and orient the substrate as well as polarize it for nucleophilic attack] among its members, we view the HAD family as a diverse collection of phosphoryl transferases, dominated by phosphohydrolases. Ironically, the 2-haloalkonate dehalogenase group, from which the superfamily name was derived, is but a minor branch off the family tree.

SUPPORTING INFORMATION AVAILABLE

Stereoviews of the metal-binding sites of metal-dependent HAD superfamily members. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Koonin, E. V., and Tatusov, R. L. (1994) Computer analysis of bacterial haloacid dehalogenases defines a large superfamily of hydrolases with diverse specificity. Application of an iterative approach to database search, *J. Mol. Biol. 244*, 125−132.
2. Baker, A. S., Ciocci, M. J., Metcalf, W. W., Kim, J., Babbitt, P. C., Wanner, B. L., Martin, B. M., and Dunaway-Mariano, D. (1998) Insights into the mechanism of catalysis by the P−C bond-cleaving enzyme phosphonoacetaldehyde hydrolase derived from gene sequence analysis and mutagenesis, *Biochemistry 37*, 9305−9315.
3. Aravind, L., Galperin, M. Y., and Koonin, E. V. (1998) The catalytic domain of the P-type ATPase has the haloacid dehalogenase fold, *Trends Biochem. Sci. 23*, 127−129.
4. Morais, M. C., Zhang, W., Baker, A. S., Zhang, G., Dunaway-Mariano, D., and Allen, K. N. (2000) The crystal structure of *Bacillus cereus* phosphonoacetaldehyde hydrolase: insight into catalysis of phosphorus bond cleavage and catalytic diversification within the HAD enzyme superfamily, *Biochemistry 39*, 10385−10396.
5. Cho, H., Wang, W., Kim, R., Yokota, H., Damo, S., Kim, S. H., Wemmer, D., Kustu, S., and Yan, D. (2001) BeF_3^- acts as a phosphate analog in proteins phosphorylated on aspartate: structure of a BeF_3^- complex with phosphoserine phosphatase, *Proc. Natl. Acad. Sci. U.S.A. 98*, 8525−8530.
6. Lahiri, S. D., Zhang, G., Dunaway-Mariano, D., and Allen, K. N. (2002) Caught in the act: the structure of phosphorylated β-phosphoglucomutase from *Lactococcus lactis*, *Biochemistry 41*, 8351−8359.
7. Rinaldo-Matthis, A., Rampazzo, C., Reichard, P., Bianchi, V., and Nordlund, P. (2002) Crystal structure of a human mitochondrial deoxyribonucleotidase, *Nat. Struct. Biol. 10*, 779−787.
8. Parsons, J. F., Lim, K., Tempczyk, A., Krajewski, W., Eisenstein, E., and Herzberg, O. (2002) From structure to function: YrbI from *Haemophilus influenzae* (HI1679) is a phosphatase, *Proteins 46*, 393−404.
9. Ridder, I. S., Rozeboom, H. J., Kalk, K. H., and Dijkstra, B. W. (1999) Crystal structures of intermediates in the dehalogenation of haloalkanoates by L-2-haloacid dehalogenase, *J. Biol. Chem. 274*, 30672−30678.
10. Isbell, A. F., Englert, L. F., and Rosenberg, H. (1969) Phosphonoacetaldehyde, *J. Org. Chem. 34*, 755−756.
11. Bradford, M. M. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding, *Anal. Biochem. 72*, 248−254.
12. Morais, M. C., Baker, A. S., Dunaway-Mariano, D., and Allen, K. N. (2000) Crystallization and preliminary crystallographic analysis of phosphonoacetaldehyde hydrolase, *Acta Crystallogr. D56* (Part 2), 206−209.
13. Otwinowski, Z., and Minor, W. (1997) Processing of X-ray Diffraction Data Collected in Oscillation Mode, *Methods Enzymol. 276*, 307−326.
14. Matthews, B. W. (1985) Determination of protein molecular weight, hydration, and packing from crystal density, *Methods Enzymol. 114*, 176−187.
15. Rossmann, M. G. (1990) The molecular replacement method, *Acta Crystallogr. A46* (Part 2), 73−82.
16. Navaza, J. (2001) Implementation of molecular replacement in AMoRe, *Acta Crystallogr. D57*, 1367−1372.
17. Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) Crystallography & NMR system: A new software suite for macromolecular structure determination, *Acta Crystallogr. D54* (Part 5), 905−921.
18. Pannu, N. S., and Read, R. J. (1996) Improved structure refinement through maximum likelihood, *Acta Crystallogr. A52*, 659−668.
19. Brünger, A. T. (1992) The Free R Value: a Novel Statistical Quantity for Assessing the Accuracy of Crystal Structures, *Nature 355*, 472−474.
20. Jones, T. A., Zou, J. Y., Cowan, S. W., and Kjeldgaard, G. J. (1991) Improved methods for building protein models in electron density maps and the location of errors in these models, *Acta Crystallogr. A47* (Part 2), 110−119.
21. Kleywegt, G. J. (1995) Dictionaries for Heteros, *CCP4/ESF-EACBM Newsletter on Protein Crystallography 31*, 45−50.
22. Collaborative Computational Project Number 4 (1994) *Acta Crystallogr. D50*, 760−768.
23. Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) PROCHECK: a program to check the stereochemical quality of protein structures, *J. Appl. Crystallogr. 26*, 283−291.
24. Zhang, G., Allen, K. N., and Dunaway-Mariano, D. (2003) Enzymatic synthesis of radiolabeled phosphonoacetaldehyde, *Anal. Biochem. 322*, 233−237.
25. Lahiri, S. D., Zhang, G., Dunaway-Mariano, D., and Allen, K. N. (2003) The Pentacovalent Phosphorus Intermediate of a Phosphoryl Transfer Reaction, *Science 299*, 2067−2071.
26. Argiriadi, M. A., Morisseau, C., Hammock, B. D., and Christianson, D. W. (1999) Detoxification of environmental mutagens and carcinogens: structure, mechanism, and evolution of liver epoxide hydrolase, *Proc. Natl. Acad. Sci. U.S.A. 96*, 10637−10642.
27. Galburt, E. A., Pelletier, J., Wilson, G., and Stoddard, B. L. (2002) Structure of a tRNA repair enzyme and molecular biology workhorse: T4 polynucleotide kinase, *Structure 10*, 1249−1260.
28. Toyoshima, C., Nakasako, M., Nomura, H., and Ogawa, H. (2000) Crystal structure of the calcium pump of sarcoplasmic reticulum at 2.6 Å resolution, *Nature 405*, 647−655.
29. Kelley, L. A., MacCallum, R. M., and Sternberg, M. J. (2000) Enhanced genome annotation using structural profiles in the program 3D-PSSM, *J. Mol. Biol. 299*, 499−520.

30. Tatusov, R. L., Galperin, M. Y., Natale, D. A., and Koonin, E. V. (2000) The COG database: a tool for genome-scale analysis of protein functions and evolution, *Nucleic Acids Res. 28*, 33−36.
31. Tatusov, R. L., Koonin, E. V., and Lipman, D. J. (1997) A genomic perspective on protein families, *Science 278*, 631−637.
32. Ridder, I. S., Rozeboom, H. J., Kalk, K. H., Janssen, D. B., and Dijkstra, B. W. (1997) Three-dimensional structure of L-2-haloacid dehalogenase from *Xanthobacter autotrophicus* GJ10 complexed with the substrate-analogue formate, *J. Biol. Chem. 272*, 33015−33022.
33. Li, Y. F., Hata, Y., Fujii, T., Hisano, T., Nishihara, M., Kurihara, T., and Esaki, N. (1998) Crystal structures of reaction intermediates of L-2-haloacid dehalogenase and implications for the reaction mechanism, *J. Biol. Chem. 273*, 15035−15044.

BI036309N