Structure- and Function-based Characterization of a New Phosphoglycolate Phosphatase from *Thermoplasma acidophilum*

Received for publication, June 9, 2003, and in revised form, October 7, 2003
Published, JBC Papers in Press, October 10, 2003, DOI 10.1074/jbc.M306054200

Youngchang Kim‡, Alexander F. Yakunin§, Ekaterina Kuznetsova§, Xiaohui Xu¶, Micha Pennycooke§, Jun Gyu, Fred Cheung¶, Michael Proudfoot§, Cheryl H. Arrowsmith§, Andzej Joachimiak§ §§, Aled M. Edwards§ §§ §§, and Dinesh Christendat***

From the §Argonne National Laboratory, Structural Biology Center, Argonne, Illinois 60439, §§Banting and Best, Department of Medical Research, University of Toronto, Ontario M5G 1L6, Canada, the ¶Clinical Genomics Center, Proteomics, Toronto, Ontario M5G 1L7, Canada, and the ***Department of Botany, University of Toronto, Toronto, Ontario M5S 3B2, Canada

The protein TA0175 has a large number of sequence homologues, most of which are annotated as unknown and a few as belonging to the haloacid dehalogenase superfamily, but has no known biological function. Using a combination of amino acid sequence analysis, three-dimensional crystal structure information, and kinetic analysis, we have characterized TA0175 as phosphoglycolate phosphatase from *Thermoplasma acidophilum*. The crystal structure of TA0175 revealed two distinct domains, a larger core domain and a smaller cap domain. The large domain is composed of a centrally located five-stranded parallel β-sheet with strand order S10, S8, S8, S1, S2 and a small β-hairpin, strands S3 and S4. This central sheet is flanked by a set of three α-helices on one side and two helices on the other. The smaller domain is composed of an open faced β-sandwich represented by three antiparallel β-strands, S5, S6, and S7, flanked by two oppositely oriented α-helices, H3 and H4. The topology of the large domain is conserved; however, structural variation is observed in the smaller domain among the different functional classes of the haloacid dehalogenase superfamily. Enzymatic assays on TA0175 revealed that this enzyme catalyzed the dephosphorylation of phosphoglycolate in vitro with similar kinetic properties seen for eukaryotic phosphoglycolate phosphatase. Activation by divalent cations, especially Mg²⁺, and competitive inhibition behavior with Cl⁻ ions are similar between TA0175 and phosphoglycolate phosphatase. The experimental evidence presented for TA0175 is indicative of phosphoglycolate phosphatase.

Acid hydrolases are ubiquitous enzymes that catalyze a diverse array of reactions such as phosphate hydrolysis and phosphoryl transfer. Acid hydrolases belong to the haloacid dehalogenase (HAD) superfamily, which includes 1,2-haloacid dehalogenase, epoxide hydrolase, phosphoserine phosphatase, phosphomannomutase, phosphoglycolate phosphatase, and P-type ATPase, all of which utilize a nucleophilic aspartate in their phosphoryl transfer reaction.

HAD superfamily proteins with different catalytic activities have low sequence identity, less than 14%; however, they are characterized by the following three conserved sequence motifs: motif I, DX(D/T/Y)(X(T/V))(L/V); motif II, (S/T); and motif III, K(G/S)/D/S/XXX(D/N) (1, 2). Motif I is the most conserved; only subtle variations in this motif are observed for different functional classes of the HAD superfamily. In motif I, the first Asp, which is conserved in all HAD family members, is the functional nucleophile (2). This nucleophile becomes phosphorylated and is directly involved in the phosphoryl transfer reaction. The second Asp is found in all phosphotransferases and phosphatases but is replaced by a Thr in P-type ATPases and by a Tyr in dehalogenases (2). This Asp residue partakes both in the acid-base catalysis reaction and in the phosphorylation of the first Asp of motif I. Motif II contains a conserved serine or threonine that hydrogen bonds to the phosphoryl oxygen of the substrate and helps to orient it for the nucleophilic attack (3). The conserved residues, except lysine, from motif III interact with an active site divalent metal (in most cases Mg²⁺) and are essential for activity (3, 4). The conserved lysine interacts with the phosphosapartate intermediate, which results in the stabilization of its phosphorylated state. This lysine, although intimately involved in the catalytic mechanism, is located far away in primary sequence from the conserved sequence DXXXD in motif III and is located on a different secondary structure element in the protein.

Crystal structures of proteins from the HAD superfamily include 2-haloacid dehalogenase, phosphonoacetalddehyde hydrolase, Ca²⁺-ATPase of sarcoplasm reticulum, and phosphoserine phosphatase (3, 5–8). These proteins all share a conserved α/β-domain classified as a hydrolase fold, which is similar to the Rossman fold. Residues from the conserved sequence motifs are also conserved at the tertiary level as shown from the crystal structures of the proteins for this superfamily.
TA0175 was selected for structural studies as part of our structural proteomic project because it did not have a structural homologue and was annotated as "unknown" in the genome database. Multiple sequence analysis using the non-redundant amino acid sequence data base identified more than 50 similar proteins with sequence identities ranging from 75 to 22%, none of which had a known biological function. However, a subset of these proteins were annotated as "putative hydrolase," and they all contained the three conserved motifs observed for the HAD superfamily, which suggested that TA0175 is a potential acid hydrolase. As stated above, this family comprises a number of different enzymes; it is impossible to determine the catalytic function from sequence analysis because there is low sequence conservation among these different functional classes of proteins. Here we demonstrate how the three-dimensional crystal structure, when combined with sequence and biochemical analysis, led to the functional annotation of TA0175 as phosphoglycolate phosphatase.

### EXPERIMENTAL PROCEDURES

**Cloning, Protein Expression, and Purification**—The TA0175 gene (GenBank™ accession number gi 16081332) was cloned from genomic Thermoplasma acidophilum (TA) DNA, and its gene product was expressed, selenomethionine (SeMet)-labeled, and purified from a bacterial system using the nickel-nitrilotriacetic acid affinity procedure as described elsewhere for Methanobacterium thermoautotrophicum proteins (9). Screening for crystallization conditions was also performed as described elsewhere for M. thermoautotrophicum proteins (9). Crystalization experiments were conducted using the hanging drop vapor diffusion method at room temperature, 296 K. The final crystallization condition consisted of 22% polyethylene glycol 3350 as the precipitant, 0.2M calcium chloride (CaCl₂), and sodium HEPES at pH 7.5. The crystals chosen for x-ray data collection were flash frozen in this buffer containing 12% glycerol, which acted as cryoprotectant. The morphology of the single crystals is rods with maximum dimensions of 0.3 × 0.1 × 0.4 mm³.

**X-ray Diffraction and Structure Determination**—Crystals of the TA0175 belong to the orthorhombic space group I222 with unit cell dimensions a = 88.2 Å, b = 99.2 Å, c = 113.8 Å, α = β = γ = 90°.

Multiwavelength anomalous dispersion data were collected on SBC-2, a

### Table I

**Summary of crystal parameters and multiwavelength anomalous dispersion data collection statistics**

| Unit cell | a = 88.240 Å, b = 99.244 Å, c = 113.781 Å, α = β = γ = 90° |
|-----------|-------------------------------------------------------------|
| Space group | 1222 |
| Molecular mass (Da) (224 residues) | 25,279 |
| Molecules (AU)* | 2 |
| SeMet (AU) | 18 |

| Wavelength (Å) | Edge | Peak | Remote | High resolution |
|----------------|------|------|--------|----------------|
| 0.9795 | 0.9794 | 0.9641 | 0.99187 |

| Resolution range (Å) | 1.8 | 1.8 | 1.8 | 3.8–1.4 (1.45–1.40) |
|----------------------|-----|-----|-----|-------------------|
| No. of unique reflections | 45,433 | 45,536 | 45,364 | 96,123 |
| Completeness (%) | 97.3 | 97.4 | 96.8 | 97.6 (100) |
| R-merge (%) | 9.2 | 8.6 | 8.4 | 5.8 (42.5) |

* AU, asymmetric units.

### Table II

**Summary of phasing and TA0175 model refinement statistics**

| Phasing (resolution range, 3.8–1.8 Å) | Edge | Peak | Remote, Friedel | All |
|---------------------------------------|------|------|-----------------|-----|
| Friedel | Iso | Friedel | Iso | All |
| Phasing power | 2.07 | 1.36 | 2.06 | 1.37 | 1.52 |
| FOM (1.8 Å) | 0.28 | 0.19 | 0.28 | 0.19 | 0.19 | 0.51 |
| Density modification, FOM | 0.88 |

### Table I

**Refinement**

| Resolution range (Å) | 3.8–1.4 |
|----------------------|---------|
| No. of reflections | 92,264 |
| α cutoff | 0.0 |
| R-value (%) | 17.3 |
| Free R-value (%) | 19.5 (9,256)* |
| r.m.s. deviations from ideal geometry | |
| Bond length (1–2) (Å) | 0.007 |
| Angle (°) | 1.30 |
| Dihedral (°) | 23.6 |
| Improper (°) | 0.94 |
| No. of atoms | |
| Protein | 3,576 |
| Formate | 3 |
| Calcium | 5 |
| Water | 715 |
| Mean B-factor (Å²) | |
| All atoms | 18.7 |
| Protein atoms | |
| Protein main chain | 14.7 |
| Protein side chain | 17.5 |
| Formate | 22.7 |
| Calcium | 14.4 |
| Water | 30.2 |
| Ramachandran plot statistics (%) | |
| Residues in most favored regions | 90.8 |
| Residues in additional allowed regions | 9.2 |
| Residues in disallowed region | 0.0 |

* 9,256 is the number of test reflections.
3 x 3 charged coupled device detector built at Advanced Photon Source, to 1.7 Å resolution from a single crystal containing SeMet-labeled protein at three different wavelengths (peak, inflection, and high remote) near the selenium edge (Table I). The native data were collected to 1.4 Å and used for the later stages of the refinement of the protein structure. All diffraction data were collected at 100 K at the 19ID beamline of the Structural Biology Center at the Advanced Photon Source, Argonne National Laboratory. The multiwavelength anomalous dispersion and native data were processed using the HKL2000 suite of programs (10). Data collection statistics are presented in Table I. Both multiwavelength anomalous dispersion phasing and phase improvement by density modification were done using CNS version 1.0 software (11). The initial protein model was autobuilt using ARP/wARP at 1.7 Å (12). The initial refinement was made against the data collected at the peak wavelength of selenium, up to 1.7 Å, by several rounds of CNS refinements consisting of B-group, minimize (or anneal), and water pick steps. After every round of refinement, the model was manually adjusted using both 2F – F and F – Fo difference electron density maps followed by individual B-factor refinement. The annealing step used in the refinement includes the simulated annealing with torsion angle dynamics and starting temperature of 5000 K. The water picking was done using the following criteria: a peak of at least 3.0 eÅ⁻³ in the F – Fo difference map with acceptable bonding distance in the range of 2.0–5.0 Å with other atoms. In the later stages of the refinement, individual B-factor refinement steps were also included which gave R of 0.224 and free R of 0.246. The final refinements were done with the native data collected to 1.4 Å using the same refinement procedure as above. The final R is 0.171 with the free R of 0.195. All refinements were done using CNS version 1.0, and electron density visualization and model building were done with the software package O (13). Refinement statistics are included in Table II. The programe MOLSCRIPT (14), RASTER 3D (15), and SPOCK (16) were used for preparation of the figures.

Enzymatic Assays—Phosphatase activity with 5 mM p-nitrophenyl phosphate (pNPP) was determined in 50 mM HEPES buffer (pH 7.5) as described previously (17). Phosphoglycolate phosphatase activity was determined as described by Rose et al. (18) except that bovine serum albumin was omitted. Phosphatase activity with phosphorylated substrates was determined spectrophotometrically using reaction mixtures (0.8-ml final volume) containing 50 mM HEPES buffer (pH 7.5), 0.1 M NaCl, 5 mM MgCl2, 50 μM substrate, and 0.15–1.0 μg of TA0175. After a 10-min incubation at 70 °C, the reaction was stopped by the addition of 0.2 ml of Malachite Green reagent (19), and after 10 min at room temperature the absorbance at 630 nm was measured. Kinetic parameters were determined by non-linear curve fitting using the saturation kinetic equation v = (Vmax [S])/(Km + [S]). Saturation plots and kinetic parameters were obtained with the software package GraphPad Prism®.

RESULTS AND DISCUSSION

Amino Acid Sequence Analysis—A position-specific iterative BLAST (PSI-BLAST) sequence analysis of TA0175 with other non-redundant amino acid sequences in the protein data base identified more than 50 bacterial proteins that are functionally annotated as hypothetical, conserved, or unknown. A small subset, fewer than 10 proteins, were tentatively classified as belonging to the HAD superfamily of hydrolases. All of these proteins share between 76 and 22% of sequence identity with TA0175. The three conserved motifs, highlighted in black, are observed in proteins belonging to the HAD superfamily. TA, gi 16081332; TV, Thermoplasma volcanium gi 13541111; MT, M. thermoautotrophicum gi 7429215; CA, Clostridium acetobutylicum gi 15683986; MF, Mycoplasma pneumoniae gi 13508166; MC, Myco-
plosma capricolum gi 609250.

Structure Overview—TA0175 is composed of two distinct do-
mains, a larger “core” domain, comprising the αβ hydrolase fold, and a smaller “cap” domain. The core domain is composed of a centrally located five-stranded parallel β-sheet with strand order S10, S9, S8, S1, S2 and a small β-hairpin, strands S3 and S4. The β-sheet is flanked by three α-helices on the convex side and by two α-helices on the concave side of the barrel. On the concave side, the loop L5 connecting β-strands S8 and S9 contains a 9-residue stretch (residues 177–185) assuming a pseudo-α-helix (closely mimicking an α-helix). This arrangement of secondary structure in the core domain, involving the six parallel β-strands (S10, S9, S8, S1, S2, and S3) and the six helices including the pseudo-α-helix in the concave side, is similar to that observed for other proteins belonging to the HAD superfamily, phosphoserine phosphatase (PSP) and YrbI from Hae-

![Fig. 1. ClustalW alignment of proteins identified by a BLAST search of the National Center for Biotechnology Information (NCBI) sequence data base. The three conserved motifs, highlighted in black, are observed in proteins belonging to the HAD superfamily. TA, gi 16081332; TV, Thermoplasma volcanium gi 13541111; MT, M. thermoautotrophicum gi 7429215; CA, Clostridium acetobutylicum gi 15683986; MF, Mycoplasma pneumoniae gi 13508166; MC, Myc-
plosma capricolum gi 602030.](image-url)
mophilus influenzae (3, 20) (Fig. 2). Such secondary organization is also observed for the core domain of CheB, a bacterial response regulator (21).

The core domain is disrupted by the insertion of a small α/β-domain (cap domain). The disruption of the core domain extends from residue 50 to residue 168 between strand S4 and helix H5. The insertion is an open faced β-sandwich represented by three antiparallel β-strands, S5, S6, and S7, and two oppositely oriented α-helices, H3 and H4 (Fig. 2). The first helix, H3, precedes strand S5, and the second helix, H4, is followed by strands S6 and S7, which form a β-hairpin.

The electrostatic surface analysis of a subunit of this protein revealed a relatively acidic surface (Fig. 2). Interestingly the active site of this protein is lined with acidic residues and is continuous as a channel through the monomer. This high localization of acidic residues in the active site is consistent with those observed for other acid phosphatases.

Structure Comparison—Structural analysis of TA0175 was done by comparing its coordinates with those of other proteins in the Protein Data Bank using the DALI algorithm. The closest structural matches to TA0175 are PSP, YrbI, phosphonoacetaldehyde hydrolase, calcium-transporting ATPase from sarcoplasmic reticulum, L-2-haloacid dehalogenase, and epoxide hydrolase (3, 5–8) (Table III). All of these proteins belong
to the HAD superfamily; their structures consist of an $\alpha/\beta$-hydrolase (core) domain with an insertion (cap) domain. The topology of the core domain within proteins in this superfamily is well conserved. The superposition of the structure of TA0175 with that of PSP (Fig. 3) revealed that the core $\alpha/\beta$-domain aligns very well with a root mean square deviation value of 3.4 Å from the DALI analysis. In addition, residues from the three sequence motifs superpose particularly well (Fig. 3).

Although the core domain shows structural similarity with the $\alpha/\beta$-hydrolases, the cap domain shows no structural similarity with any of the different functional classes of HAD enzymes identified by DALI. In TA0175, the cap domain is an open $\beta$-sandwich consisting of three antiparallel $\beta$-sheets flanked by two helices. In PSP, the cap domain is a four-helix bundle; in phosphonoacetaldehyde hydrolase it is a five-helix bundle. These structural differences suggest that the biological function of TA0175 differs from that of the above proteins.

Three known bacterial response regulators identified by the DALI analysis share structural similarities to TA0175. Two of them, CheY and CheB, catalyze Mgs$^{2+}$-dependent phosphoryl transfer to an active site aspartate residue. CheY and CheB are members of bacterial two-component signaling cascades; each interacts with CheA as part of a phosphorelay reaction. CheY has a single domain with an $\alpha/\beta/\alpha$-fold similar to the HAD superfamily. CheY and CheB also share active site residues similar to the active site aspartate of motif I from the HAD superfamily (21, 22). However, the cap domain is not seen in the above proteins, therefore it is unlikely that TA0175 belongs to this class of bacterial response regulators.

**Toward Functional Characterization of TA0175**—Using both sequence analysis and structure information of TA0175, we have refined the possible functions of TA0175 and of its sequence homologues to a subset of four proteins in the HAD superfamily. The sequence analysis of this class of proteins was performed by searching for the N-terminal motif containing DGT found in TA0175 and its close sequence homologues. In the HAD superfamily, the first aspartate (Asp-8), which is the characteristic nucleophile for this class of hydrolases, is strictly conserved. The second conserved aspartate (Asp-10) in this motif, which is present in TA0175, is found only in phosphotransferases and phosphatases and is replaced by a threonine in ATPases and by a tyrosine in dehalogenases, suggesting that TA0175 does not have ATPase or dehalogenase activity.

The threonine found at position 5 in the DGT motif is conserved in all enzymes except $\beta$-phosphoglucomutases and $p$-nitrophenylphosphatases. The glycine residue found at position 4 in the motif is replaced by a serine in PSP whose amide interacts with the phosphate oxygen of the substrate, thus increasing the electrophilicity of the phosphorus atom to facilitate attack by the nucleophilic aspartate (23, 24). Accordingly, if these substitutions are coordinately changed with function, TA0175 is unlikely to be any of the following acid hydrolases: a $\beta$-phosphoglucomutase, a phosphoserine phosphatase, a histidinol phosphatase, a $p$-nitrophenylphosphatase, a P-type ATPase, or a halocid dehalogenase.

To summarize, sequence and structural analysis suggested that the biological function of TA0175 is different from a number of phosphatases from the list of potential functional candidates, including PSP, phosphonoacetaldehyde hydrolase, calcium-transporting ATPase, $\beta$-halocid dehalogenase, YrbI, epoxide hydrolase, CheY, and CheB.

Our results from the above structural and sequence analysis approach are consistent with TA0175 belonging to one of the remaining four class of hydrolases (a trehalose-6-phosphate phosphatase, a deoxyglucose-6-phosphate phosphatase, a phosphomannomutase, or a phosphoglycolate phosphatase) or that TA0175 may belong to a novel class of acid hydrolases altogether.

TA0175 was screened for biological activity of the four possible enzymes by monitoring for the release of free phosphate in the presence of their appropriate substrates. We found that TA0175 catalyzed only the dephosphorylation of phosphoglycolate and displayed kinetic characteristics similar to the enzyme phosphoglycolate phosphatase (18, 25). The pH dependence, divalent metal ion requirement, chloride ion dependence, and kinetic parameters of TA0175 were then compared with those of phosphoglycolate phosphatase obtained from spinach and red blood cells. Results from these studies are consistent with TA0175 functioning as a phosphoglycolate phosphatase (Table IV and Fig. 4.).

Although enzymatic activity of TA0175 with phosphoglycolate was observed, further kinetic characterization was undertaken. We also investigated the ability of TA0175 to catalyze the release of phosphate from the two nonspecific phosphatase substrates $p$NPP and inorganic pyrophosphate (PP$_i$). As expected, enzyme TA0175 catalyzed the release of phosphate from both $p$NPP and PP$_i$; however, the catalytic behavior of this enzyme is markedly different. With substrate analogue $p$NPP, saturation kinetics are observed but with reduced binding affinity, reduced catalytic activity, and markedly reduced catalytic efficiency with the $k_{cat}/K_m$ value 1000-fold lower than with phosphoglycolate as substrate. With PP$_i$, the enzyme catalytic properties are slightly altered with its catalytic efficiency reduced by less than 10-fold. However, we observed a sigmoidal curve, instead of a classic saturation curve, with a Hill’s coefficient $n_H = 1.7 \pm 0.2$ indicating positive cooperativity in PP$_i$ binding. This observation is consistent with the crystal structure of the dimer, which revealed different distinct conformational states of the individual subunits within the protein.

Phosphoglycolate phosphatase from both spinach and red blood cells is activated by Mg$^{2+}$ ions (18, 26). The phosphoglycolate phosphatase activity of TA0175 was activated by Co$^{2+}$, Mn$^{2+}$, and Mg$^{2+}$ with Mg$^{2+}$ yielding the maximal turnover.

### Table III

| Protein Data Bank code | Z (Å) | r.m.s.d. | Identity % | Protein |
|-----------------------|-------|---------|------------|---------|
| 1j8d                  | 16.8  | 1.9     | 24         | YrbD (hypothetical protein hi1679) |
| 1fsa                  | 12.6  | 2.5     | 22         | PSP     |
| 1fez                  | 11.6  | 2.9     | 12         | Phosphonoacetaldehyde hydrolase |
| 1eul                  | 10.3  | 2.4     | 21         | Calcium-transporting ATPase P-type |
| 1q6q                  | 9.9   | 2.6     | 19         | l-2-Haloacid dehalogenase |
| 1cr6                  | 7.3   | 2.7     | 18         | Epoxide hydrolase |
| 1chd                  | 6.8   | 3.1     | 8          | CheB methyltransferase |

**Summary of seven closest structural homologues to protein TA0175 obtained with DALI analysis**

Although the Z score and root mean square deviation (r.m.s.d.) are indicative that these proteins are structurally related, structural conservation is only observed in the core domain. The structure of the smaller cap functional domain is different amongst these proteins indicating different biological function.
rate (Fig. 5). Ca$^{2+}$ is known to inhibit phosphoglycolate phosphatase (27). The addition of 10 mM Ca$^{2+}$ resulted in greater than 90% inhibition of the phosphoglycolate phosphatase activity of TA0175 (Fig. 6). Rose et al. (18) have shown that Cl$^{-}$ ions are activating at low concentration (up to 25 mM KCl) and are competitive inhibitors with respect to phosphoglycolate at higher concentrations. Our analysis with TA0175 is also consistent with the above observation; at low chloride ion concentration (up to 50 mM) the enzyme was activated 2-fold, but with increasing chloride concentration, its activity was inhibited, and at 600 mM KCl, total phosphoglycolate phosphatase inhibition occurs (Fig. 7). In summary, the results from our kinetic characterization of TA0175 support the proposal that this protein is a phosphoglycolate phosphatase.

**Functional Role of Active Site Residues**—Superposition of the three-dimensional structure of PSP with that of this newly identified *T. acidophilum* phosphoglycolate phosphatase revealed strong similarities in their active sites. The residues identified in motifs I, II, and III of the HAD superfamily were all located in the active site of this phosphoglycolate phosphatase. These active site residues superpose with analogous active site residues of PSP (Fig. 3). The functional roles of a

---

**Table IV**

Comparison of steady-state kinetic parameters for TA0175 with those of eukaryotic phosphoglycolate phosphatases

| Variable substrate                  | $K_m$    | $k_{cat}$ | $k_{cat}/K_m$ |
|------------------------------------|----------|-----------|---------------|
| Phosphoglycolate                   | 0.037 ± 0.003 | 8.26 ± 1.54 | 2.2 × 10^3   |
| pNPP                               | 0.41 ± 0.5 | 0.58 ± 0.04 | 0.19 × 10^3  |
| Pyrophosphate                      | 0.17 ± 0.02 | 4.8 ± 0.3  | 0.3 × 10^3   |
| Spinach phosphoglycolate phosphatase | 0.029 | 292       | 97 × 10^3    |
| Phosphoglycolate                   | 0.061 | 19        | 3.0 × 10^5   |

| Phosphoglycolate                   | 0.037 ± 0.003 | 8.26 ± 1.54 | 2.2 × 10^3   |
| pNPP                               | 0.41 ± 0.5 | 0.58 ± 0.04 | 0.19 × 10^3  |
| Pyrophosphate                      | 0.17 ± 0.02 | 4.8 ± 0.3  | 0.3 × 10^3   |
| Spinach phosphoglycolate phosphatase | 0.029 | 292       | 97 × 10^3    |
| Phosphoglycolate                   | 0.061 | 19        | 3.0 × 10^5   |

* From Rose et al. (18).
Fig. 4. Kinetic characterization of TA0175 with phosphoglycolate, pNPP and PP\textsubscript{i}. A, saturation curve showing that classical Michaelis-Menten kinetics are observed with phosphoglycolate and that kinetic parameters obtained are consistent with values obtained for the enzyme phosphoglycolate phosphatase from red blood cells. The inset is a reciprocal plot using the Lineweaver-Burke method for approximation of kinetic parameters that was compared with results obtained with the non-linear regression method. B, saturation kinetic profile with pNPP as substrate for TA0175. As expected, a marked reduction in catalytic efficiency with this substrate was seen as a result of its lower binding affinity to the enzyme. C, kinetic profile of TA0175 with substrate PP\textsubscript{i} showing a sigmoidal curve with positive cooperativity for substrate binding. This observation is consistent with the crystal structure of the dimer. We observed different conformational states of the individual subunits within the dimer of TA0175. U, units.
number of these conserved residues have been confirmed for PSP by chemical modification/proteolysis and mass spectrometry (4). In addition, crystal structures of complexes with $\text{BeF}_3\cdot\text{H}_2\text{O}_4\cdot\text{Mg}_2\cdot\text{H}_2\text{O}_1$ and now with $\text{Ca}_2\cdot\text{H}_2\text{O}_1$ have all contributed to the characterization of binding sites within the HAD superfamily of proteins (3, 7, 24).

Identification of Active Site and Active Site Residues—

The crystal structure of TA0175 revealed five $\text{Ca}_2\cdot\text{H}_2\text{O}_1$ ions per dimer. The protein was crystallized in the presence of 1 mM $\text{CaCl}_2$, therefore its substitution for $\text{Mg}_2\cdot\text{H}_2\text{O}_1$, the likely physiological divalent metal as seen from our kinetic analysis, is not surprising. The active site of TA0175, which was identified by the bound $\text{Ca}_2\cdot\text{H}_2\text{O}_1$ ions, was located between loops 1, 3, 5, 12, and 14 in the three-dimensional structure. The active sites are located on opposite faces of the dimer, and residues from each subunit contribute to form a complete active site. In one subunit, we identified two well ordered $\text{Ca}_2\cdot\text{H}_2\text{O}_1$ ions that interact with conserved residues seen in motifs I, II, and III. The first of these $\text{Ca}_2\cdot\text{H}_2\text{O}_1$ ions interacts with Asp-8, Asp-174, Asp-176, Ser-175, and two water molecules (Fig. 3). The position and interactions of this $\text{Ca}_2\cdot\text{H}_2\text{O}_1$ is analogous to the location of the functionally im-

FIG. 5. Divalent metal ion requirement for TA0175. Saturation curves showing the activation of TA0175 with different divalent metal ions. $\text{Mg}^{2+}$ is the likely physiological metal for this enzyme since the highest level of enzyme activity was observed with it. Indeed this catalytic behavior is also observed with phosphoglycolate phosphatase from red blood cells and spinach. U, units.

FIG. 6. Comparison of $\text{Mg}^{2+}$ and $\text{Ca}^{2+}$ as cofactor for TA0175 activity. Top panel, addition of $\text{Mg}^{2+}$ to the reaction chamber results in activation of TA0175. However, $\text{Ca}^{2+}$ is shown to have the reverse effect (bottom panel). More than 90% inhibition was observed with $\text{Ca}^{2+}$ ions in the reaction chamber. Similar behavior was also observed with phosphoglycolate phosphatase from red blood cells. U, units.

FIG. 7. Biological Mechanism from Crystal Structure—The crystal structure of TA0175 reveals residues from the three highly conserved sequence motifs interacting with a $\text{Ca}^{2+}$ ion, which most likely replaced the physiological $\text{Mg}^{2+}$ ion in the protein. Interestingly a similar interaction with $\text{Mg}^{2+}$ has been observed for other acid hydrolases, and the role of these residues in the catalytic mechanism has been described elsewhere (24). We propose that the enzymatic mechanism observed for other acid hydrolases, such as PSP and P-type ATPase, is also conserved for phosphoglycolate phosphatase based on conserved structural features, conserved key active site residues, conserved $\text{Mg}^{2+}$ binding site, and conserved sequence motifs in this HAD superfamily (Fig. 3) (23). In brief, $\text{Mg}^{2+}$ or other divalent ions help to orient the nucleophilic aspartate, Asp-8, which attacks the phosphorus atom of phosphoglycolate. The susceptibility of the P–glycolate bond is increased by two prom-
molecules interacting with the third Ca\(^{2+}\) ion and propose that they occupy the position of the biological substrate. In the other subunit, the third Ca\(^{2+}\) is absent, and loop 1 closes off the active site. This closed conformation is stabilized by ionic interaction between Arg-18 and Asp-10. It is likely that the conformation of loop 1 most likely regulates the passage of substrate and product to and from the active site, and the interaction between Arg-18 and Asp-10 modulates this behavior.

**Biological Significance and Role of Phosphoglycolate Phosphatase in Microorganisms**—The enzyme phosphoglycolate phosphatase has been isolated from plants, from animal tissues, and from red blood cells. In plants, it is involved in the dephosphorylation of phosphoglycolate, which is produced when ribulose-bisphosphate carboxylase reacts with oxygen during photorespiration. Phosphoglycolate phosphatase is localized in the chloroplast where its substrate, phosphoglycolate, is a potent inhibitor of phosphofructokinase and triose-phosphate isomerase (27). In red blood cells, phosphoglycolate is synthesized by pyruvate kinase and has been shown to activate the breakdown of 2,3-bisphosphoglycerate, which is a regulator of the oxygen affinity of hemoglobin (18). In autotrophic organisms, the role of phosphoglycolate phosphatase is to prevent accumulation of phosphoglycolate and permit it to be recycled by the Calvin cycle. However, in heterotrophs, the role of phosphoglycolate phosphatase is still not understood. An *Escherichia coli* mutant lacking a functional phosphoglycolate phosphatase did not show an altered phenotype (28). Interestingly the gene encoding this enzyme is found in the same operon as the *dam* gene whose gene product is involved in the methylation of adenine in the GATC sequence. Double-stranded DNA breaks produced by ionizing radiation create terminal 3’-phosphate or 3’-phosphoglycolate. These 3’-phosphoglycolate overhangs are inhibitors to the DNA repair process and must be removed for DNA polymerase to function (29). The enzyme involved in removal of the 3’-phosphoglycolate has yet to be identified. Therefore, further studies need to be conducted to determine whether the identified phosphoglycolate phosphatase could play a role in DNA repair in prokaryotes.

**Acknowledgments**—We thank all members of the Structural Biology Center at Argonne National Laboratory and the Ontario Center for Structural Proteomics for help in conducting experiments.

**REFERENCES**

1. Aravind, L., Galperin, M. Y., and Koonin, E. V. (1998) Trends Biochem. Sci. 23, 127–129
2. Collet, J. F., Stroobant, V., Pirard, M., Delpierre, G., and Van Schaftingen, E. (1998) J. Biol. Chem. 273, 14107–14112
3. Wang, W., Kim, R., Jancarik, J., Yokota, H., and Kim, S. H. (2001) Structure (Camb.) 9, 65–71
4. Collet, J. F., Stroobant, V., and Van Schaftingen, E. (1999) J. Biol. Chem. 274, 33985–33990
5. Regn, C., Tipton, P. A., and Beamer, L. J. (2002) Structure (Camb.) 10, 269–279
6. Morais, M. C., Zhang, W., Baker, A. S., Zhang, G., Dunaway-Mariano, D., and Allen, K. N. (2000) Biochemistry 39, 10385–10396
7. Toyoshima, C., Nakasako, M., Nomura, H., and Ogawa, H. (2000) Nature 405, 647–655
8. Hisano, T., Hata, Y., Fujii, T., Liu, J. Q., Kurihara, T., Esaki, N., and Sada, K. (1996) J. Biol. Chem. 271, 20322–20330
9. Christendat, D., Sardakis, V., Dharamsi, A., Bochkarev, A., Pai, E. F., Arrowsmith, C. H., and Edwards, A. M. (2000) J. Biol. Chem. 275, 24608–24612
10. Otwinski, Z., and Minor, W. (1997) Methods Enzymol. 276, 307–326

**Table V**

| Incubation conditions | P<sub>i</sub> release (nmol P<sub>i</sub>/nmol TA0175 monomer) |
|-----------------------|----------------------------------------------------------|
| 2 N HCl              | 0.17 ± 0.01                                             |
| 2 x NaOH             | 0.49 ± 0.03                                             |
| Calf intestine phoshatase | 0.57 ± 0.02                                           |

**FIG. 7** The dual roles of Cl\(^{-}\) ions in TA0175 activity. Top panel, a low concentration of Cl\(^{-}\) ions, up to 25 mM NaCl, is shown to activate TA0175, the activity of the enzyme almost doubles. However, at higher concentrations Cl\(^{-}\) ions competitively inhibit TA0175. The double reciprocal plot (lower panel) exemplifies the competitive inhibition nature of high concentration of Cl\(^{-}\) ions. These plots intersect at the origin, which is characteristic of competitive inhibition behavior. Cl\(^{-}\) ions were shown to have a similar inhibitory effect on phosphoglycolate phosphatase from red blood cells. U, units.

**TABLE V** Phosphate (P<sub>i</sub>) release from purified TA0175 under different conditions

| Incubation conditions | P<sub>i</sub> release (nmol P<sub>i</sub>/nmol TA0175 monomer) |
|-----------------------|----------------------------------------------------------|
| 2 N HCl              | 0.17 ± 0.01                                             |
| 2 x NaOH             | 0.49 ± 0.03                                             |
| Calf intestine phoshatase | 0.57 ± 0.02                                           |

**TABLE V** Phosphate (P<sub>i</sub>) release from purified TA0175 under different conditions

| Incubation conditions | P<sub>i</sub> release (nmol P<sub>i</sub>/nmol TA0175 monomer) |
|-----------------------|----------------------------------------------------------|
| 2 N HCl              | 0.17 ± 0.01                                             |
| 2 x NaOH             | 0.49 ± 0.03                                             |
| Calf intestine phoshatase | 0.57 ± 0.02                                           |
11. Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grasser-Kanstieleve, R. W., Jiang, J. S., Kuszewski, J., Nave, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) *Acta Crystallogr. Sect. D Biol. Crystallogr.* 54, 905–921
12. Perrakis, A., Morris, R., and Lamzin, V. S. (1999) *Nat. Struct. Biol.* 6, 458–463
13. Jones, T. A., Zou, J. Y., Cowan, S. W., and Kjeldgaard, M. (1991) *Acta Crystallogr. Sect. A* 47, 110–119
14. Kraulis, P. J. (1991) *J. Appl. Crystallogr.* 24, 946–950
15. Merritt, E. A., and Murphy, M. E. P. (1991) *Acta Crystallogr. Sect. D Biol. Crystallogr.* 50, 869–873
16. Christopher, J. A. (1998) SPOCK, The Center for Macromolecular Design, Texas A&M University, College Station, TX
17. Kuo, C. H., Gilon, H., Blumenthal, A. B., and Sedat, J. W. (1982) *Exp. Cell Res.* 142, 141–154
18. Rose, Z. B. (1981) *Arch. Biochem. Biophys.* 208, 602–609
19. Baykov, A. A., Kryuchenko, O. A., and Avaeva, S. M. (1988) *Anal. Biochem.* 171, 266–270
20. Parsons, J. F., Lim, K., Tempczyk, A., Krajewski, W., Eisenstein, E., and Herzberg, O. (2002) *Proteins* 46, 393–404
21. West, A. H., Martinez-Hackert, E., and Stock, A. M. (1995) *J. Mol. Biol.* 250, 276–290
22. Cho, H. S., Lee, S. Y., Yan, D., Pan, X., Parkinson, J. S., Kastu, S., Wemme, D. E., and Pelton, J. G. (2000) *J. Mol. Biol.* 297, 543–551
23. Thompson, P. R., and Cole, P. A. (2001) *Proc. Natl. Acad. Sci. U. S. A.* 98, 8170–8171
24. Wang, W., Cho, H. S., Kim, R., Jancarik, J., Yokota, H., Nguyen, H. H., Grigoriev, I. V., Wemme, D. E., and Kim, S. H. (2002) *J. Mol. Biol.* 319, 421–431
25. Seal, S. N., and Rose, Z. B. (1987) *J. Biol. Chem.* 262, 13496–13500
26. Hunic, H. D., and Tolbert, N. E. (1984) *Arch. Biochem. Biophys.* 229, 64–72
27. Mamedov, T. G., Suzuki, K., Miura, K., Kuch, K.-i., and Fukuzawa, H. (2001) *J. Biol. Chem.* 276, 45573–45579
28. Lyngstadaas, A., Lohner-Olesen, A., Grelland, E., and Boye, R. (1999) *Biochim. Biophys. Acta* 1472, 376–384
29. Inamdar, K. V., Pouliot, J. J., Zhou, T., Lees-Miller, S. P., Rasouli-Nia, A., and Povir, L. F. (2002) *J. Biol. Chem.* 277, 27162–27168