THE JOURNAL OF BIOLOGICAL CHEMISTRY VOL. 285, NO. 7, pp. 4603–4611, February 12, 2010

Crystal Structure and Computational Analyses Provide Insights into the Catalytic Mechanism of 2,4-Diacetylphloroglucinol Hydrolase PhlG from *Pseudomonas fluorescens*[*]

Received for publication, July 15, 2009, and in revised form, November 3, 2009 Published, JBC Papers in Press, December 16, 2009, DOI 10.1074/jbc.M109.044180

Yong-Xing He[‡1], Liang Huang[‡1], Yanyan Xue[‡1], Xue Fei[‡], Yan-Bin Teng[‡], Sheryl B. Rubin-Pitel[§], Huimin Zhao[§¶‖2], and Cong-Zhao Zhou[‡3]

From the [‡]Hefei National Laboratory for Physical Sciences at Microscale and School of Life Sciences, University of Science and Technology of China, Hefei, Anhui 230026, China and the Departments of [§]Chemical and Biomolecular Engineering and [¶]Chemistry and [‖]Institute for Genomic Biology, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801

2,4-Diacetylphloroglucinol hydrolase PhlG from *Pseudomonas fluorescens* catalyzes hydrolytic carbon-carbon (C–C) bond cleavage of the antibiotic 2,4-diacetylphloroglucinol to form monoacetylphloroglucinol, a rare class of reactions in chemistry and biochemistry. To investigate the catalytic mechanism of this enzyme, we determined the three-dimensional structure of PhlG at 2.0 Å resolution using x-ray crystallography and MAD methods. The overall structure includes a small N-terminal domain mainly involved in dimerization and a C-terminal domain of Bet v1-like fold, which distinguishes PhlG from the classical α/β-fold hydrolases. A dumbbell-shaped substrate access tunnel was identified to connect a narrow interior amphiphilic pocket to the exterior solvent. The tunnel is likely to undergo a significant conformational change upon substrate binding to the active site. Structural analysis coupled with computational docking studies, site-directed mutagenesis, and enzyme activity analysis revealed that cleavage of the 2,4-diacetylphloroglucinol C–C bond proceeds via nucleophilic attack by a water molecule, which is coordinated by a zinc ion. In addition, residues Tyr^{121}, Tyr^{229}, and Asn^{132}, which are predicted to be hydrogen-bonded to the hydroxyl groups and unhydrolyzed acetyl group, can finely tune and position the bound substrate in a reactive orientation. Taken together, these results revealed the active sites and zinc-dependent hydrolytic mechanism of PhlG and explained its substrate specificity as well.

The root-colonizing fluorescent *Pseudomonas* species produce a variety of extracellular metabolites with antimicrobial activity and thus are widely utilized as effective biocontrol agents against the plant root diseases caused by soil-borne pathogenic fungi (1, 2). Among these metabolites, the phenolic compound 2,4-diacetylphloroglucinol (DAPG)[4] displays a remarkably broad spectrum of toxic activity toward bacteria, fungi, nematodes, and even plants (3–5). The *Pseudomonas* DAPG biosynthetic cluster includes four genes, termed *phlABCD*, which are transcribed as a single operon (6). Immediately downstream to this operon, the *phlE* gene encodes a putative transmembrane permease that has been implicated in DAPG resistance (7). Adjacent to *phlA*, the genes *phlF* and *phlH* encode two divergently transcribed tetracycline resistance repressor-like regulators involved in DAPG biosynthesis (8–10). In addition to these pathway-specific regulators, a gene located between *phlF* and *phlH*, termed *phlG*, encodes a DAPG hydrolase, which may offer an effective alternative for modulating DAPG levels (11).

The protein encoded by *phlG* was found to catalyze the conversion of DAPG into less toxic monoacetylphloroglucinol (MAPG) by cleaving one of the carbon-carbon (C–C) bonds linking the acetyl groups to the phenolic ring (Fig. 1). This enzyme showed strict substrate specificity as it cannot degrade other compounds of similar structure such as MAPG and triacetylphloroglucinol (11). The amino acid sequence of PhlG displays 25–37% sequence identity to several hypothetical proteins and the phloretin hydrolase (EC 3.7.1.4) from *Eubacterium ramulus*. Phloretin hydrolase is a known C–C bond hydrolase that specifically converts phloretin to phloroglucinol and 3-(4-hydroxyphenyl) propionic acid (12). However, neither PhlG nor phloretin hydrolase possesses sequence homology to other C–C bond-cleaving hydrolases or motifs typical of the α/β hydrolase family (13), suggesting these two enzymes belong to a distinct hydrolase family.

In contrast to common biochemical hydrolysis of amide and ester bonds, hydrolytic cleavage of C–C bonds is relatively rare in nature (14). To date, only 10 enzymes with this primary function have been described in the KEGG data base (EC 3.7.1.1 to EC 3.7.1.10). Nonetheless, these enzymes are generally found to be involved in the degradation pathway of various hydroxylated aromatic compounds. Recently, the hydrolytic enzymes acting on the C–C bonds in β-diketones were characterized in detail

[*] This work was supported by National Natural Science Foundation of China Program 30628011, Ministry of Science and Technology of China Projects 2006CB910202 and 2006CB806501, and the China Postdoctoral Science Foundation.

The atomic coordinates and structure factors (code 3HWP) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (http://www.rcsb.org/).

[1] These authors contributed equally to this work.

[2] To whom correspondence may be addressed: Dept. of Chemical and Biomolecular Engineering, University of Illinois, 600 S. Mathews Ave., Urbana, IL 61801. Tel.: 217-333-2631; Fax: 217-333-5052; E-mail: zhao5@illinois.edu.

[3] To whom correspondence may be addressed. Tel./Fax: 86-551-3600406; E-mail: zcz@ustc.edu.cn.

[4] The abbreviations used are: DAPG, 2,4-diacetylphloroglucinol; MAPG, monoacetylphloroglucinol; C–C, carbon-carbon; SeMet, selenomethionine; HPLC, high performance liquid chromatography; r.m.s.d., root mean square deviation; PDB, Protein Data Bank; MAD, multiwavelength anomalous dispersion.

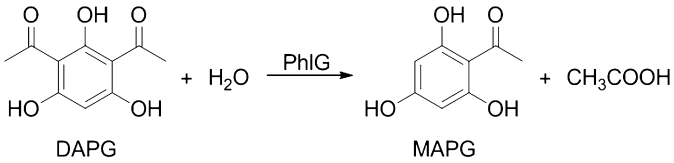

FIGURE 1. **Conversion of DAPG to MAPG by PhlG.**

and showed surprising structural and mechanistic diversity (14, 15). For instance, fumarylacetoacetate hydrolase, an enzyme that cleaves fumarylacetoacetate to fumarate and acetoacetate, is a metalloenzyme characterized by a novel α/β-fold and uses a His/Asp dyad-activated water molecule as the attacking nucleophile. The bound metal ion, identified to be Ca^{2+}, functions both in stabilizing a carbanion leaving group and indirectly positioning the water molecule for nucleophilic attack at a carbonyl group (16). Another β-diketone hydrolase, 2,6-dihydroxy-pseudo-oxynicotine hydrolase, which is involved in the nicotine degradation pathway, belongs to the classical α/β hydrolase family and possesses a catalytic Ser/His/Asp triad for the hydrolytic reaction (17). It was proposed that this diversity possibly reflects the necessity of nature to recruit different protein sequences, structures, and mechanisms to cleave C–C bonds of different susceptibility, for the purpose of detoxification and catabolic processes (15).

In this study, we describe the 2.0-Å resolution crystal structure of PhlG, solved with the MAD method. Moreover, based on structural analysis, computational docking studies, site-directed mutagenesis, and enzyme activity analysis, we propose a catalytic mechanism of the PhlG enzyme and provide a structural interpretation of its strict substrate specificity. To the best of our knowledge, PhlG is the first C–C bond-cleaving hydrolase with a Bet v1-like fold, divergent from other known C–C bond-cleaving enzymes in the classical α/β-fold hydrolase family.

EXPERIMENTAL PROCEDURES

Construction, Expression, and Purification of PhlG and Its Mutants—The *phlG* gene was amplified from purified genomic DNA extracted from *Pseudomonas fluorescens* Pf-5 using a Wizard Genomic DNA purification kit (Promega, Madison, WI). Genes encoding PhlG and its mutants were individually cloned into a pET28a-derived vector and overexpressed in *Escherichia coli* Rosetta (DE3) strain (Novagen, Madison, WI) using 2× YT culture medium (5 g of NaCl, 16 g of Bacto-tryptone, and 10 g of yeast extract/liter). A hexahistidine (His_6) tag was added to the N terminus of each of the recombinant proteins. The PhlG-expressing cells were grown to an $A_{600\ nm}$ of 0.6 at 37 °C. Expression of the recombinant proteins was induced at the exponential phase with 0.2 mM isopropyl β-D-thiogalactoside, and cell growth was continued for another 20 h at 16 °C before harvesting. Cells were collected by centrifugation at 4,000 × *g* for 20 min and resuspended in lysis buffer (20 mM Tris-HCl, 100 mM NaCl, pH 8.0). After 3 min of sonication and centrifugation at 12,000 × *g* for 25 min, the supernatant containing the soluble target protein was collected and loaded onto a nickel-nitrilotriacetic acid column (Qiagen, Valencia, CA) equilibrated with the binding buffer (20 mM Tris-HCl, 100 mM NaCl, pH 8.0). The target protein was eluted with 250 mM imidazole and further loaded onto a Superdex 75 column (GE Healthcare) equilibrated with 20 mM Tris-HCl, 100 mM NaCl, pH 8.0. Fractions containing the target protein were combined and concentrated to 20 mg/ml for crystallization. Samples for enzymatic activity assays were collected at low concentrations (0.5 mg/ml). The purity of protein was evaluated by SDS-PAGE (data not shown), and the protein sample was stored at −80 °C.

The selenium-Met (SeMet)-labeled PhlG protein was expressed in *E. coli* strain B834 (DE3) (Novagen, Madison, WI). A culture of transformed cells was inoculated into LB medium and incubated at 37 °C overnight. The cells were harvested when the $A_{600\ nm}$ reached 0.2 and were then washed twice in the M9 medium. The cells were then cultured in SeMet medium (M9 medium with 25 mg/liter L-SeMet and the other essential amino acids at 50 mg/liter) to an $A_{600\ nm}$ of 0.6–0.8. The remaining steps in protein purification were the same as those for the native protein.

Crystallization, Data Collection, and Processing—Crystals of native PhlG were grown at 16 °C using the hanging drop vapor-diffusion technique, with the initial condition of mixing 1 μl of 20 mg/ml protein sample with an equal volume of the reservoir solution (20% polyethylene glycol 4,000, 17% isopropyl alcohol, 0.1 M sodium citrate, pH 5.6). Typically, crystals appeared in 1 or 2 days and reached the maximum size in 1 week. The SeMet derivative crystals were grown under the same conditions. The crystals were transferred to cryoprotectant (reservoir solution supplemented with 25% glycerol) and flash-frozen with liquid nitrogen. Multiwavelength anomalous dispersion data for a SeMet derivative crystal was collected at radiation wavelengths of 0.9794, 0.9799, and 0.9252 Å at Beijing Synchrotron Radiation Facility, Institute of High Energy Physics, Chinese Academy of Sciences, using beamline 3W1A at 100 K with a MAR 165-mm CCD (MARresearch, Germany). All diffraction data were indexed, integrated, and scaled with HKL2000 (18).

Structure Determination and Refinement—The crystal structure of PhlG was determined using the MAD phasing method from a SeMet-substituted protein crystal to a maximum resolution of 2.0 Å. The SHELX program suite (19) was used to locate the heavy atoms and 22 selenium atoms were identified. The phase was calculated and further improved with the program SOLVE/RESOLVE (20). Electron density maps calculated from solvent-flattened experimental phases showed clear features of secondary structural elements, allowing automatic model building of most residues with the ARP/wARP program (21). The initial model was refined by using the maximum likelihood method implemented in REFMAC5 as part of the CCP4i program suite (22) and rebuilt interactively by using the σ_A-weighted electron density maps with coefficients $2F_o - F_c$ and $F_o - F_c$ in the program COOT (23). Refinement converged to an *R*-factor of 19.1% and *R*-free of 21.9% at the resolution of 2.0 Å. The final model was evaluated with the programs MOLPROBITY (24) and PROCHECK (25). The final coordinates and structure factors were deposited in the Protein Data Bank under the accession code of 3HWP. The data collection and structure refinement statistics are listed in Table 1. Sequence alignment was performed using the programs ClustalW (26) and ESPript (27). Structure comparison was

TABLE 1
Data collection, phasing, and refinement

	MAD peak	MAD edge	MAD remote
Data collection statistics			
Wavelength	0.9794	0.9799	0.9252
Space group	$C222_1$	$C222_1$	$C222_1$
Unit cell (Å)	a = 127.71 Å, b = 132.50 Å, c = 95.02 Å, $\alpha = \beta = \gamma$ = 90 Å	a = 124.75 Å, b = 132.58 Å, c = 95.03 Å, $\alpha = \beta = \gamma$ = 90 Å	a = 124.75 Å, b = 132.58 Å, c = 95.03 Å, $\alpha = \beta = \gamma$ = 90 Å
Resolution limit	50.00 to 2.00 (2.07 to 2.00)	50.00 to 2.00 (2.07 to 2.00)	30.00 to 2.00 (2.07 to 2.00)
Unique reflections	96,750 (6,164)	95,198 (5,516)	101,065 (9,051)
Completeness	95.3[a] (66.0)	94.3 (60.7)	99.1 (93.4)
R_{merge}[b]	12.7 (39.9)	11.5 (39.0)	10.3 (44.7)
$I/\sigma I$	14.5 (2.6)	16.0 (2.6)	18.1 (2.3)
Redundancy	4.5	4.2	5
MAD phasing			
Heavy atom sites	22	22	22
Mean figure of merit after solve			0.3
Mean figure of merit after resolve			0.61
Refinement statistics			
Resolution limit			26.00 to 2.00
R-factor[c]			19.1%
R-free[d]			21.9%
r.m.s.d. bond length[e]			0.008 Å
r.m.s.d. bond angles			1.051°
Average of B-factors			22.94 Å^2
Ramachandran plot[f]			
Most favored			97.6%
Additional allowed			2.4%
PDB entry			3HWP

[a] The values in parentheses refer to statistics in the highest bin.
[b] $R_{merge} = \Sigma_{hkl}\Sigma_i|I_i(hkl) - \langle I(hkl)\rangle|/\Sigma_{hkl}\Sigma_i I_i(hkl)$, where $I_i(hkl)$ is the intensity of an observation and $\langle I(hkl)\rangle$ is the mean value for its unique reflection; Summations are over all reflections.
[c] R-factor $= \Sigma_h|F_o(h) - F_c(h)|/\Sigma_h F_o(h)$, where F_o and F_c are the observed and calculated structure-factor amplitudes, respectively.
[d] R-free was calculated with 2% of the data excluded from the refinement.
[e] Root mean square deviation was from ideal values.
[f] Categories were defined by Molprobity.

carried out using the DALI server. All structure figures were created using the program PyMol (28).

Computational Docking—Because an enzyme is usually oriented to recognize the reaction intermediate form of a substrate rather than the ground-state form of a substrate, the reaction intermediate form of DAPG was used as the small molecule for the docking studies. Almost all structurally characterized enzymes catalyzing the carbon-heteroatom bond hydrolytic reactions involve a catalytic water or hydroxyl acting as a nucleophile, forming a tetrahedral intermediate (29). The high energy tetrahedral intermediate three-dimensional structures of DAPG of opposite chirality were transformed with the program phniex.elbow (30) from the simplified molecular input line entry specification (SMILES) level representations (c1(O)c ([C@](C)(O)[O-])c(O)cc(O)c1C(C) = [O]) and c1(O)c([C@@](C)(O)[O-])c(O)cc(O)c1C(C) = [O]). These high energy intermediates reflect the electronics of the transition state as described by Hermann *et al.* (31), *i.e.* the oxygen atom of the added hydroxyl was made negative by moving its proton to the carbon atom that was double-bonded to the reactive center in the ground-state structure. Enzyme structures for docking were prepared with the program CHIMERA (32), and the bound metal ions were assigned formal +2 charge. Docking grids were calculated with the program GRID (33), and docking runs were performed using DOCK6.1 (33). Up to 10,000 initial poses per molecule were generated, for each of which multiple conformations were scored. The best scoring pose was rigid-body minimized and scored for electrostatic and van der Waals interactions.

Site-directed Mutagenesis—Site-directed mutagenesis was performed by using the QuikChange site-directed mutagenesis kit (Stratagene, La Jolla, CA) with the plasmid encoding the wild-type PhlG as the template. The mutant proteins were expressed, purified, and stored in the same manner as the wild-type protein.

Enzymatic Activity Assay—The enzyme kinetic parameters of native PhlG and its mutants were measured as described in Ref. 11 with minor changes. All the assays were performed at 25 °C in 100 mM NaH_2PO4/Na_2HPO4 buffer, pH 7.1. The substrate DAPG (Toronto Research Chemicals Inc., North York, Canada) was diluted to a series of concentrations from 100 mM stock solution in methanol. The reaction was triggered by adding the purified protein solution, and the reaction was terminated by mixing with an equal volume of methanol/H_2O/acetic acid (50:45:5 volume ratio). The mixture was centrifuged at 12,000 × *g* for 10 min, and the supernatant was subjected to high performance liquid chromatography (HPLC) analysis. DAPG and MAPG were quantified by HPLC analysis. Acetonitrile and water (both containing 0.1% trifluoroacetic acid) served as the mobile phase. Five percent acetonitrile was used to calibrate the Eclipse XDS-C18 column, 4.6 × 150 mm (Agilent). The gradient was formed as follows: from 0 to 40% acetonitrile in 3 min and then from 40 to 100% acetonitrile in 3 min, at a flow rate of 1.5 ml/min. The initial velocities and substrate concentrations were used to linearly fit the Lineweaver-Burk plot to calculate the K_m and k_{cat} values. Three independent kinetic determinations were made to calculate means ± S.D. for the reported K_m and k_{cat} values.

Liquid Chromatography-Tandem Mass Spectrometry Analysis—Tryptic peptides were analyzed by liquid chromatography-tandem mass spectrometry with a Thermo LTQ linear ion trap mass spectrometer (Thermo Electron, San Jose, CA) equipped with a Thermo Surveyor HPLC system and an autosampler. Peptides were separated on line by a fused silica capillary column packed with C18 resin (5 μm, 300 Å, Jupiter, Phenomenex, Torrance, CA) using a linear gradient of 0.1% (v/v) formic acid in water (solvent A) and acetonitrile (solvent B) (2–25% acetonitrile for 35 min and 25–90% acetonitrile for 15 min). MS-MS spectra were acquired by a multiple reaction monitoring (MRM) method, in which the m/z corresponding to the cysteine-containing tryptic peptides was targeted for MS-MS. A window of ± 1.0 m/z was used for precursor isolation.

Metal Content Determination of PhlG—To determine the metal content, 400 μl of PhlG wild-type enzyme was transferred to a 50-ml beaker, digested at 90 °C for 2 h with 1 ml of concentrated HNO_3 and 0.1 ml of 30% H_2O_2, and diluted to final volume of 10 ml with deionized water. 400 μl of the buffer (20 mM Tris-HCl, pH 8.0, 200 mM NaCl) was also treated with the same procedure and used as the control. The metal contents of the samples were analyzed with inductively coupled plasma-mass spectrometry (Thermo Electron Corp.) at the University of Science and Technology of China.

Analysis of Substrate Entrance—The on-line version of CAVER was used to explore the putative substrate access tunnel of PhlG. The position of the bound metal ion in the interior pocket was specified to identify tunnels directly connecting the active site to the surface. The tunnel profile, which is the average tunnel cross-section radius along the length, was calculated from the detected accessible path.

RESULTS

Overall Structure—The PhlG crystal contains two identical subunits, termed A and B, in an asymmetric unit. Subunits A and B are structurally similar, with an overall root mean square deviation (r.m.s.d.) of ~0.5 Å over the 287 Cα atoms. The two subunits associate to form a tight dimer related by a noncrystallographic 2-fold axis (Fig. 2*A*), which was confirmed in solution by the gel filtration method (data not shown). The dimerization is mainly formed by the small N-terminal domain (Arg^4–Gly^{82}), which is composed of a two-stranded anti-parallel β-sheet, three α-helices (α1–α3), and three 3_{10} helices (η1–η3). Nineteen pairs of hydrogen bonds were involved in the dimerization, most of which come from residues of the N-terminal domain such as Thr^{10}, Tyr^{11}, Phe^{12}, Lys^{18}, Glu^{39}, Arg^{48}, Ile^{50}, Gln^{51}, Gly^{53}, and Tyr^{83}. In contrast, only three residues from the C-terminal helix α10, Glu^{261}, Tyr^{265}, and Glu^{266}, contribute to hydrogen bonding of the dimer interface. The total buried area of the dimer interface is up to ~2500 Å^2, indicating a stable dimeric PhlG structure (Fig. 2*A*).

Besides the small N-terminal domain involved in dimerization, PhlG possesses a large C-terminal catalytic domain (Tyr^{83}–Thr^{291}) showing the so-called helix-grip fold (34), which is characterized by a partly opened β-barrel (β3–β11) wrapped around a long C-terminal helix α10 (Fig. 3*A*). This partly opened β-barrel is sealed by helices α4 and α5 on one end and helix α9 on the other end, leading to the formation of a deep narrow pocket in the center of the protein. Structure-based alignment using the DALI server revealed that the C-terminal catalytic domain closely resembles the tetracenomycin aromatase/cyclase (PDB code 2REZ, r.m.s.d. 2.7 Å), cytokinin-specific binding protein (PDB code 3C0V, r.m.s.d. 2.9 Å), and cholesterol-regulated start protein 4 (PDB code 1JSS, r.m.s.d. 3.4 Å), despite their low sequence homology (less than 9% of sequence identity). According to the structural classification of proteins (SCOP) data base, all these structures belong to the Bet v1-like superfamily (SCOP:d.129.3), all members of which share the prototypic structure of a seven-stranded β-barrel (topology 1765432) wrapped around a long C-terminal α-helix with the secondary structural arrangement of β-α2-β6-α (35). Conforming to the prototype of Bet v1-like fold, the PhlG C-terminal catalytic domain contains an additional β-strand (β3) adding to the β-barrel on the N-terminal β-strand (topology 12876543) and several internal insertions of α-helices (helices α6 and α7 between strand β5 and β6, α8 between β7 and β8, and α9 between β9 and α10) (Fig. 3*B*). Because no significant sequence homology is detected between PhlG and known members of Bet v1-like superfamily, PhlG defines a new sequence family of the Bet v1-like superfamily.

Additionally, based on structural and sequence similarities, the Bet v1-like superfamily is further classified into 13 families, including three families of enzymes (polyketide cyclases, ring hydroxylase α-chain, and homotrimeric ring hydroxylase), three families of lipid transfer/ligand-binding proteins (plant pathogen-related proteins Bet v1, phosphatidylinositol transfer proteins, and steroidogenic acute regulatory protein-related lipid transfer domains), and seven families of unclear biochemical profile (35). Because no members of these families are reported to exhibit hydrolytic activity, PhlG is therefore the first discovered hydrolase whose catalytic domain belongs to the Bet v1-like fold.

Active Site, Substrate Docking Simulation, and Mutational Analysis—The most important characteristic of Bet v1-like proteins is the presence of an interior hydrophobic/amphiphilic pocket accessible to the exterior. As mentioned above, this pocket is also present in the C-terminal domain of PhlG, and is ~20 Å deep and 480 Å^3 large. The fine distribution of hydrophobic and polar residues in the pocket creates an ideal environment for accommodating both hydrophobic and polar surfaces of the DAPG molecule. Interestingly, a metal ion was identified inside the pocket, coordinated by His^{129}(β5), Glu^{160}(β6), His^{270}(α10), Glu^{274}(α10), and a water molecule (Wat1) with pentagonal bipyramidal coordination geometry. It may play an important structural role in stabilizing the PhlG catalytic domain by holding the strands β5 and β6 and the helix α10 together. By having a high occupancy as indicated by the strong electron density peak, this metal ion is present in the structure despite the presence of metal ion chelators (0.1 M sodium citrate) in the crystallization buffer. It interacts with the Nϵ atoms of His^{129} (2.15 Å) and His^{270} (2.16 Å), the Oδ atoms of Glu^{160} (2.18 Å) and Glu^{274} (2.13 Å), and the water molecule Wat1 (2.31 Å) with pentagonal bipyramidal configuration, where the two Glu residues occupy the axial positions (Fig. 4*A*). Because the electron density in this metal-binding site is so strong that even Ca^{2+} cannot be well fitted into electron density

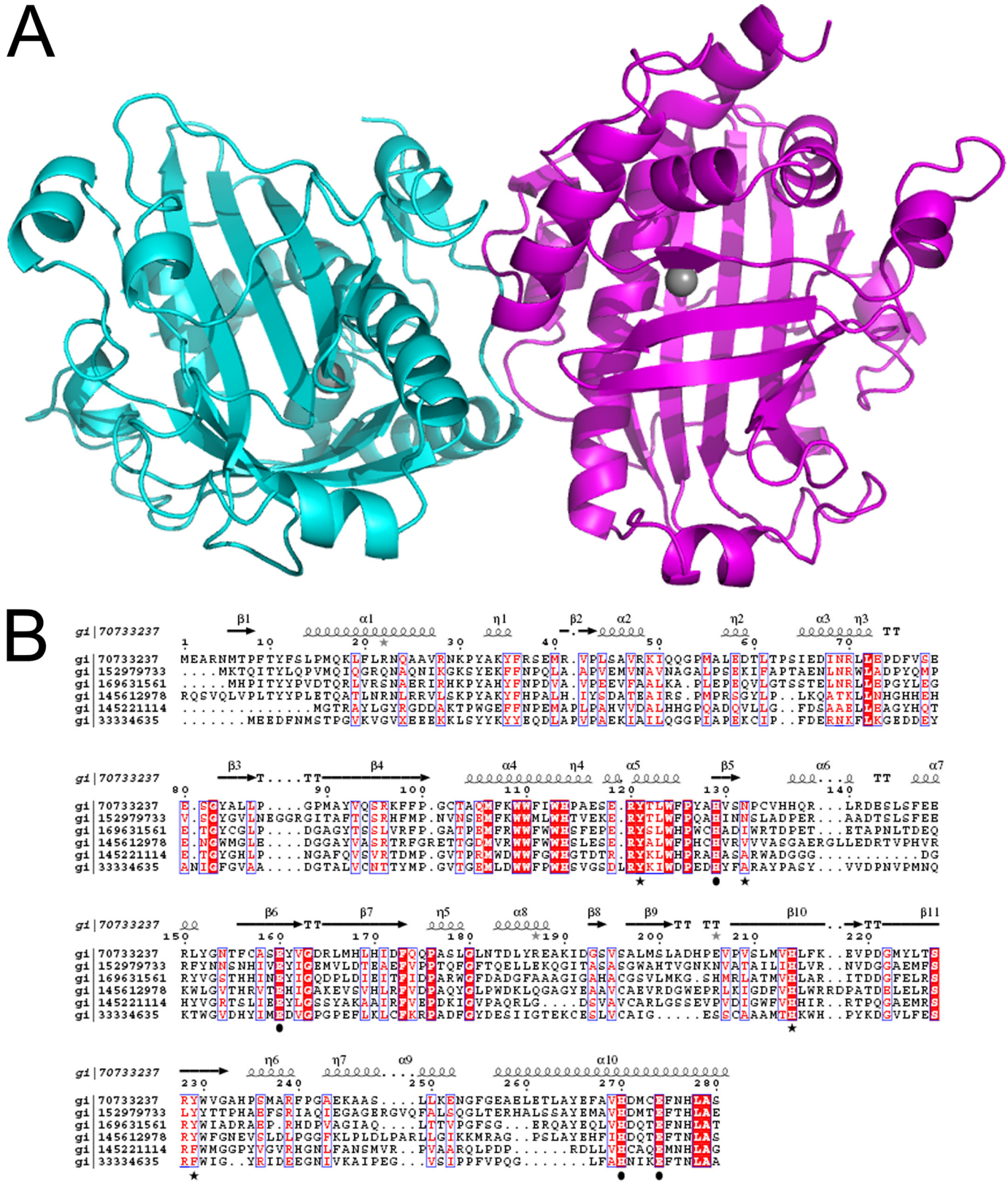

FIGURE 2. **Overall structure of PhlG dimer.** *A,* schematic representation of the PhlG dimer. The metal ions are denoted as *gray spheres*. *B,* multisequence alignment of PhlG homologs. The alignment was performed using ClustalW and ESPript. The residues involved in coordinating zinc ions are marked by *black filled circles,* and residues involved in catalysis are marked by *black stars*.

assuming the occupancy equals 1, we speculate the bound metal ions should have more electrons than Ca^{2+} and are most likely to be transition metal ions, as is also supported by the anomalous difference signal at the wavelength of 0.9250 Å (Fig. 4*A*). By using the inductively coupled plasma-mass spectrometry method, we further determined that the molar ratio of Zn^{2+}

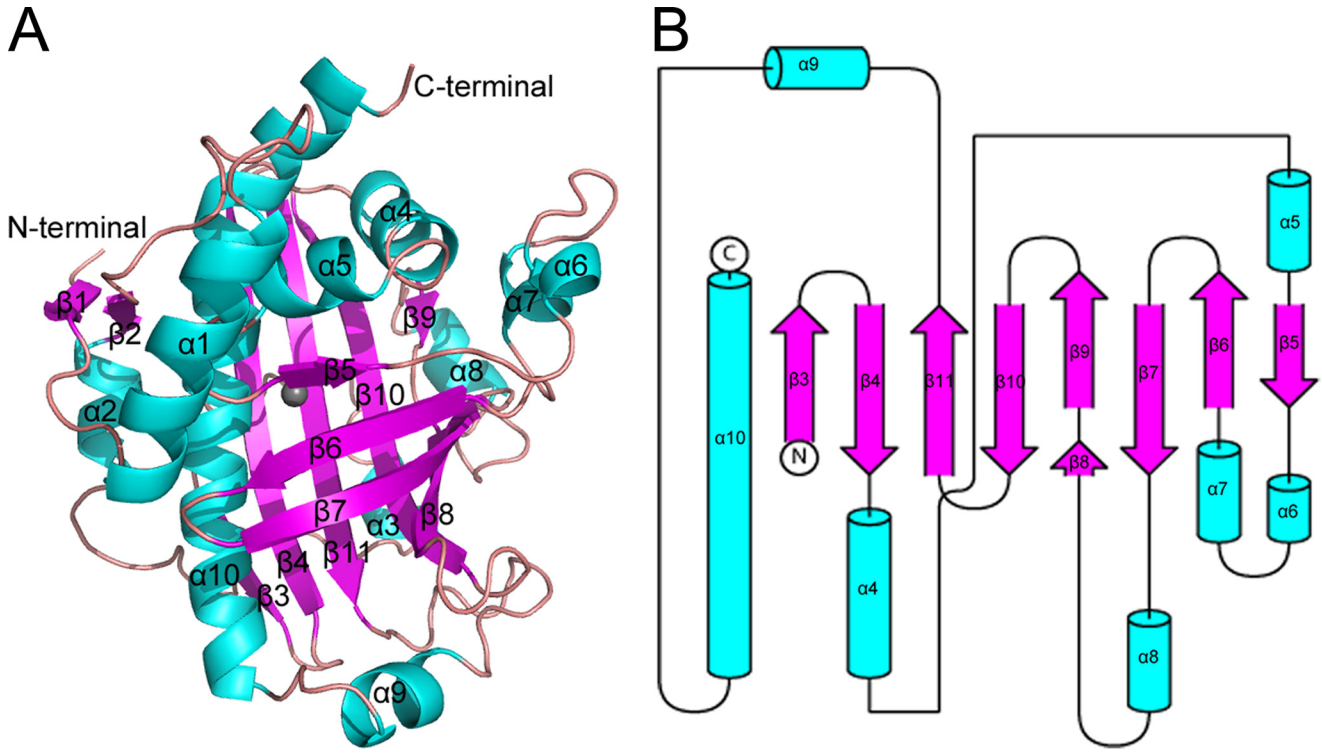

FIGURE 3. **Structure and topology of PhlG monomer.** *A,* schematic representation of PhlG monomer, colored and labeled according to secondary structural elements. The metal ion is denoted as a *gray sphere*. *B,* topology diagram of the C-terminal Bet v1-like domain of PhlG. Compared with the prototypic Bet v1-like fold, the PhlG C-terminal domain has an additional β-strand (β3) in the β-barrel and several insertions of helices (helices α6, α7, α8, and α9).

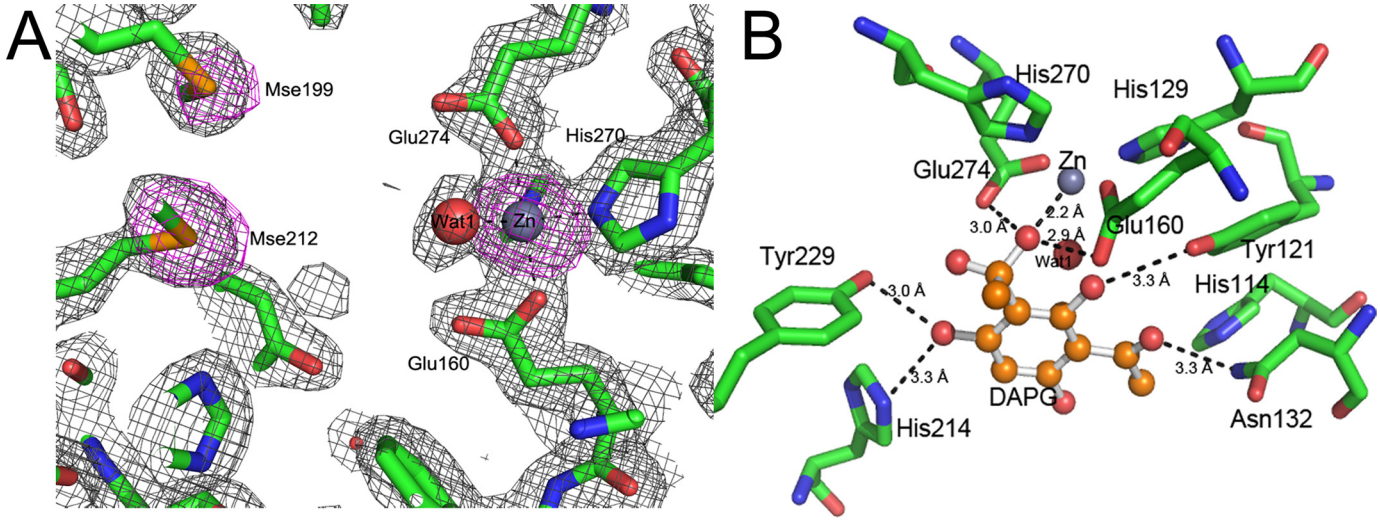

FIGURE 4. **Active site, substrate docking model, and mutational analysis of PhlG.** *A,* anomalous difference Fourier map confirms the zinc ion. The anomalous difference map density of the metal is comparable with that of selenium atoms of SeMet residues (*Mse,* SeMet). The anomalous difference Fourier map calculated with data collected at the wavelength of 0.9250 Å is shown in *pink* contoured at 5σ. At this wavelength, zinc has a modest anomalous signal ($f'' = 2.24$), whereas magnesium ($f'' = 0.06$) and calcium ($f'' = 0.52$) do not. The $2F_o - F_c\,\sigma_A$-weighted map is shown in *gray* and contoured at 1.5σ. *B,* model of DAPG intermediate bound to the active site of PhlG. The docked DAPG intermediate is shown in *ball* and *stick* format, and the residues within 4 Å of DAPG are shown in *stick* format. The distance between the hydroxyl group of the acetyl group and the zinc ion is 2.2 Å. The water molecule coordinated to the zinc ion may serve as the attacking nucleophile because its distance to the carbon atom of the carbonyl group is 2.3 Å. Hydrogen bonds between the active site and docked DAPG are also shown in *black dashes* and the distances are labeled.

to the protein is 1.15:1, which strongly suggests that the Zn^{2+} ion is the cognate metal ion of PhlG.

Besides a structural role, this metal ion may also be essential for catalytic activity because most hydrolytic enzymes such as metalloprotease and phosphatase utilize zinc ions to either stabilize the transition intermediate or activate a water molecule as a nucleophile for catalysis (36–38). To confirm the catalytic role of the zinc ion, three metal ligand residues, namely H270A, E274A, and E160A, were constructed, and enzyme kinetics were measured using an HPLC-based assay (Table 2). Because H270A has no detectable activities, further kinetic characterization of this enzyme was not attempted. All these mutants showed at least by 10^{-5}-fold reduced k_{cat} values compared with the k_{cat} value of the wild-type enzyme, indicating that the

TABLE 2

Enzyme kinetics of PhlG wild type and mutants

The activity assays were performed at 25 °C in 100 mM NaH_2PO_4/Na_2HPO_4 buffer, pH 7.1. The values were from at least three independent assays. The K_m and k_{cat} values were calculated using linear regression of a Lineweaver-Burk plot.

Enzyme	K_m	k_{cat}	k_{cat}/K_m
	μM	s^{-1}	$\mu M^{-1}s^{-1}$
Wild type	272 ± 42	4.7 ± 0.1	17,000
E274A	42 ± 7	0.000026 ± 0.000001	0.62
E160A	41 ± 7	0.000081 ± 0.00001	2
H270A	ND[a]	ND	ND
Y121A	242 ± 60	0.050 ± 0.010	210
N132A	610 ± 50	0.74 ± 0.05	1200
Y229A	873 ± 97	0.21 ± 0.03	1100
Y229F	268 ± 63	1.4 ± 0.1	5200
H214A	320 ± 54	1.4 ± 0.1	4300
H214Q	180 ± 24	0.58 ± 0.01	3300

[a] The activity of H270A cannot be detected.

bound metal ion was essential for the conversion of DAPG into MAPG. The K_m values for the mutants E270A and E160A were also decreased, implying that the metal ion does not play a key role in immobilizing the substrate.

To further explore substrate binding in the interior pocket of PhlG, we docked the two reaction tetrahedral intermediates of opposite chirality, because it was unknown in which direction the nucleophilic hydroxyl attacks the electrophilic carbonyl. The energy of the best docked poses of both DAPG intermediates was evaluated, and one of the intermediates obviously surpassed the other, strongly suggesting that our docking results approximately represent the productive binding pose of the reaction intermediate. The docked DAPG intermediate structure shows that the oxygen atom of the attacking hydroxyl is coordinated by the bound zinc ion at a distance of 2.2 Å. The water molecule (Wat1) in the PhlG structure is very close to the docked attacking hydroxyl, suggesting that it could act as the nucleophile in the hydrolysis reaction (Fig. 4*B*). Several aromatic residues such as Phe^{111}, Trp^{124}, Phe^{173}, and Phe^{267} stabilize the phenyl group of DAPG via aromatic π-π stacking interactions and hydrophobic interactions. In addition, several polar residues interact with the intermediate through hydrogen bonds as follows: (i) Tyr^{121} forms a hydrogen bond with one of the three DAPG hydroxyls near the reacting acetyl group; (ii) His^{214} and Tyr^{229} position the substrate by forming hydrogen bonds with another DAPG hydroxyl group near the reaction center; and (iii) Asn^{132} forms a hydrogen bond with the oxygen atom of the unhydrolyzed acetyl group.

Based on the above docking results, we constructed six single mutants of PhlG, including Y229A, Y229F, H214A, H214Q, Y121A, and N132A, the altered residues of which are polar residues that were predicted to form hydrogen bonds with the polar groups of DAPG. The enzyme kinetic parameters of those mutants are listed in Table 2. Apparently, all of these mutants have decreased k_{cat} values, confirming that all these predicted residues are relevant to the catalytic activity. Prominently, the k_{cat} value of Y121A was decreased by 2 orders of magnitude, whereas its K_m value was barely affected, implying an important role for Tyr^{121} in the catalytic reaction. It is tempting to propose that Asn^{132} participates in positioning and orienting the substrate because the K_m value of the mutant N132A was increased by ~2-fold. To investigate the role of Tyr^{229}, we constructed the Y229F and Y229A mutant and measured their catalytic kinetic parameters. The k_{cat} of Y229F decreased by ~4-fold, although the K_m value is barely affected, suggesting Tyr^{229} is involved in facilitating the catalytic conversion of DAPG into MAPG. However, Tyr^{229} is not essential for the catalysis, for the Y229F retains ~30% activity compared with the wild-type enzyme. Furthermore, from the decreased K_m value of the mutant Y229A, we can infer that Tyr^{229} is most likely to be involved in immobilizing and positioning the substrate via π-π stacking interactions or hydrophobic interactions, for the loss of aromaticity at Tyr^{229} causes loosened binding of the substrate. For mutants H214A and H214Q, the K_m values were not significantly different from that of the wild-type enzyme, indicating that His^{214} does not play a key role in the catalytic reaction.

Putative Substrate Access Tunnel and Its Implication in Glutathionylation-induced Inhibition of PhlG Activity—Because the catalytic zinc ion is buried more than 15 Å from the protein surface, there should be a direct path connecting the catalytic site with the exterior solvent. A putative substrate tunnel starting from the bound zinc ion to the exposed protein surface was identified by using the on-line version of the CAVER program. The tunnel is dumbbell-shaped, with a small opening formed between helix α4 and the loop linking strand β5 and helix α6 (Fig. 5*A*). This small opening is only about 2 Å wide, which is not wide enough for the substrate DAPG to pass through, suggesting that the residues around this opening should undergo conformational changes to allow the substrate access to the active site during catalysis. The residues participating in guarding this opening include His^{114} and Pro^{115} from helix α4 and Asn^{132}, Pro^{133}, and Cys^{134} from the loop linking strand β5 and helix α6 (Fig. 5*B*). In the CRET SMART domain, incorporation of the ligand ceramide is also limited by a small opening, which is guarded by two helices (one corresponds to the helix α4 in PhlG) and a loop (39). In the phosphatidylinositol transfer proteins, there is an additional helix (corresponding to the helix α4 in PhlG) adjacent to the C-terminal helix, which can undergo large conformational changes between a closed, ligand-bound structure and an open, membrane-bound structure, thus enabling the binding and release of large phospholipid molecules (40). All these examples suggest that members of the Bet v1-like superfamily regulate the entry/release of ligand/substrate by controlling an opening of limited size to the interior pocket.

Of the four cysteine residues in one PhlG monomer, Cys^{157} and Cys^{273} are buried and far away from the active site. The other two cysteines, namely Cys^{134} and Cys^{157}, are around the opening of the calculated tunnel. Thus, to confirm the significance of the tunnel opening calculated by CAVER, we attempted to modify Cys^{134} and Cys^{157} with excessive GSSG to sterically block the opening of this substrate access tunnel. Both of these two residues were detected to form mixed disulfide bonds with GSH using liquid chromatography-mass spectrometry (data not shown). As shown in Fig. 3*C*, the GSSG-modified PhlG showed significantly abolished activity, which is in agreement with our inference from the calculated substrate access tunnel. However, we do not have available structural evidence at present to rule out the possibility that the conformation of the finely arranged active site was not also altered upon the GSSG treatment.

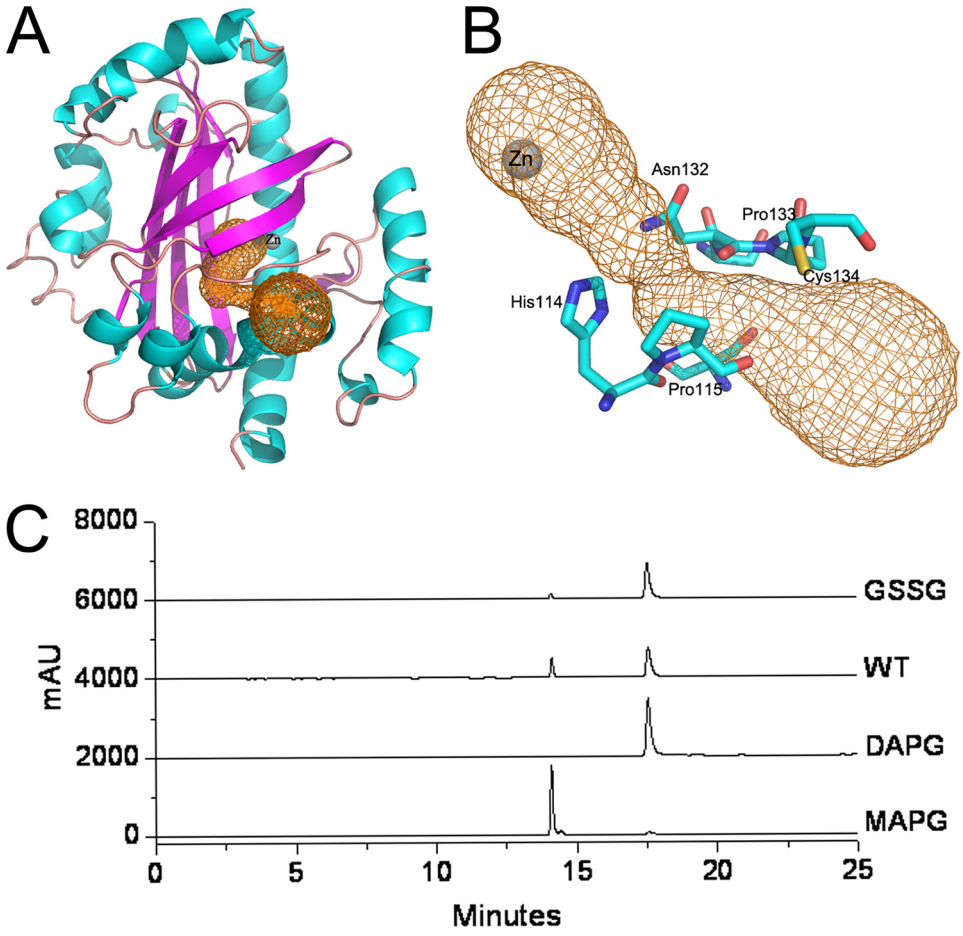

FIGURE 5. **Putative substrate access tunnel calculated by CAVER.** *A,* front view of the substrate access tunnel denoted in *mesh*. *B,* close-up side view of the substrate access tunnel; residues in the vicinity are shown as *sticks* and labeled. *C,* overlay of HPLCs of glutathionylated PhlG enzyme and untreated PhlG wild-type (*WT*) enzyme: from *top* to *bottom,* PhlG wild-type enzyme treated with 10 mM GSSG, wild-type enzyme, DAPG, and MAPG. *mAU*, milliabsorbance unit.

DISCUSSION

PhlG Defines a Novel Family of C–C Bond-cleaving Hydrolase—We showed that the overall structure of a C–C bond hydrolase, PhlG, uses the Bet v1-like fold as the catalytic domain, which is distinct from the classical α/β-fold hydrolases. To the best of our knowledge, this is the first identified hydrolytic enzyme with the Bet v1-like fold.

The penta-coordinated catalytic zinc ion bound in the pocket of PhlG is a unique feature among the proteins with the Bet v1-like fold and other C–C bond hydrolases. For those Bet v1-like fold proteins such as plant PR-10, human START5, and tetracenomycin aromatase/cyclase, although one or more metal-binding sites are located in the conserved internal pocket, none of these bound metal ions play catalytic roles or have the similar coordination geometry to the zinc ion found in PhlG. Rather, the reported C–C bond hydrolases either hexa-coordinate Ca^{2+} to stabilize the reaction intermediates or use the strictly conserved triad (His, Asp, and Ser) of α/β fold for catalysis. Using the protein metal site data base MESPEUS, we searched for the hydrolases containing catalytic metal ions with similar coordination residues and geometry as the zinc ion of PhlG. The phosphodiesterase was identified to coordinate a Zn^{2+} with two histidines, two aspartates, and a solvent molecule in pentagonal bipyramidal configuration, quite similar to PhlG (36, 41). Because it has been widely accepted that phosphodiesterases as well as other Zn^{2+}-dependent hydrolases, such as alkaline phosphatase (42) and carboxypeptidase A (43), possess a zinc-bound water or hydroxyl ion to execute the nucleophilic attack, we hypothesized that this metal ion bound in PhlG may play an essential role in catalysis and consequently performed the computational docking studies.

Structural Basis of Strict Substrate Specificity—The main feature of the Bet v1-like fold is a deep interior pocket formed between a seven- or eight-stranded antiparallel β-sheet and a long C-terminal α-helix, yet the volume of this pocket shows remarkable variability in different members of this superfamily. For instance, a plant pathogenesis-related protein of class 10 (PR-10) (PDB code 2QIM) that binds plant hormones such as zeatin and cytokinin has an internal pocket of 4500 Å^3 in volume (44), START5 (the human star-related lipid transfer protein 5 (PDB code 2R55)) that binds cholesterol or other sterols has a pocket of 800 Å^3, and tetracenomycin aromatase/cyclase (PDB code 2REZ) that catalyzes the cyclization of polyketide has a pocket of 600 Å^3 (45). In contrast, PhlG has a much smaller pocket of less than 500 Å^3 in volume, which can barely accommodate a DAPG molecule but not a triacetylphloroglucinol molecule or other larger analogs of DAPG, providing a structural interpretation of the strict substrate specificity of the enzyme.

Based on the results from computational docking, site-directed mutagenesis, and enzymatic assays, we propose that cleavage of the C–C bond in DAPG proceeds via nucleophilic attack at the carbonyl group of one of the two acetyl groups by a water molecule coordinated by a zinc ion. This zinc ion can effectively activate the coordinated water molecule by dramatically lowering its pK_a. The two zinc ion ligands Glu^{160} and Glu^{270}, both of which are hydrogen-bonded to the coordinated water, can be great candidates for conveying the released proton from water to the keto-O of the departing acetyl group. Aromatic residues such as Phe^{111}, Trp^{124}, Phe^{173}, and Phe^{267} can be essential in shaping the substrate binding pocket and positioning the phenyl group of DAPG via aromatic π-π stacking interactions and hydrophobic interactions. Residues Tyr^{121}, Tyr^{229}, and Asn^{132}, which are predicted to be hydrogen-bonded to the hydroxyl groups and unhydrolyzed acetyl group, can

finely tune and position the bound substrate in a reactive orientation. It is noteworthy that Asn^{132} is not preserved in phloretin hydrolase (Fig. 2*B*), which is consistent with the fact the phloretin molecule does not have an acetyl group (12). Although multiple sequence alignment of distantly related PhlG homologs reveals that the His^{214} is strictly conserved and the docking simulation also suggests a catalytic/substrate binding role for His^{214}, site mutagenesis analysis indicates that this residue does not play a key role in the catalysis of the enzyme.

Structural Implication of the Substrate Entry—The entrance to the active site appears to be rather restricted in the PhlG structure, as indicated by the dumbbell-shaped substrate access tunnel. A small opening to the active site is defined by the finely arranged His^{114} and Pro^{115} on one side and Asn^{132}, Pro^{133}, and Cys^{134} on the other side, separated by distances of ~2 Å when hydrogen atoms are added to the protein structure (Fig. 5*B*). Considering the size of DAPG and MAPG, this opening is not large enough to allow substrate entry and product release. Therefore, conformational changes near the opening must happen in order for the substrate to access the active site. Because proline residues usually have exceptional rigidity compared with other amino acids due to their locked φ backbone dihedral angle at approximately $-75°$ (46), it is more likely residues His^{114} and Asn^{132} serve as tunnel gatekeepers that control the substrate entry through conformational changes of their side chains. It is also possible that the substrate can induce conformational changes near the opening of the substrate access tunnel. However, because DAPG and MAPG are molecules of relatively small sizes, we speculate that the side-chain conformational changes of His^{114} and Asn^{132} are sufficient for substrate entry and product release.

Acknowledgments—Diffraction data were collected at the Beijing Synchrotron Radiation Facility. We gratefully acknowledge the support of the K. C. Wong Education Foundation, Hong Kong. We thank Dr. Yuhui Dong at Beijing Synchrotron Radiation Facility and Jiang Yu at University of Science and Technology of China for their help during data collection. We are grateful to Dr. Jie Nan and Prof. Xiaodong Su at Peking University for help with solving the PhlG structure.

REFERENCES

1. Haas, D., and Keel, C. (2003) *Annu. Rev. Phytopathol.* **41,** 117–153
2. Haas, D., and Défago, G. (2005) *Nat. Rev. Microbiol.* **3,** 307–319
3. Broadbent, D., Mabelis, R. P., and Spencer, H. (1976) *Phytochemistry* **15,** 1785
4. Keel, C., Wirthner, P., Oberhänsli, T., Voisard, C., Burger, U., Haas, D., and Défago, G. (1990) *Symbiosis* **9,** 327–341
5. Keel, C., Schnider, U., Maurhofer, M., Voisard, C., Laville, J., Burger, U., Wirthner, P., Haas, D., and Défago, G. (1992) *Mol. Plant-Microbe Interact.* **5,** 4–13
6. Bangera, M. G., and Thomashow, L. S. (1999) *J. Bacteriol.* **181,** 3155–3163
7. Abbas, A., McGuire, J. E., Crowley, D., Baysse, C., Dow, M., and O'Gara, F. (2004) *Microbiology* **150,** 2443–2450
8. Abbas, A., Morrissey, J. P., Marquez, P. C., Sheehan, M. M., Delany, I. R., and O'Gara, F. (2002) *J. Bacteriol.* **184,** 3008–3016
9. Delany, I., Sheehan, M. M., Fenton, A., Bardin, S., Aarons, S., and O'Gara, F. (2000) *Microbiology* **146,** 537–543
10. Schnider-Keel, U., Seematter, A., Maurhofer, M., Blumer, C., Duffy, B., Gigot-Bonnefoy, C., Reimmann, C., Notz, R., Défago, G., Haas, D., and Keel, C. (2000) *J. Bacteriol.* **182,** 1215–1225
11. Bottiglieri, M., and Keel, C. (2006) *Appl. Environ. Microbiol.* **72,** 418–427
12. Schoefer, L., Braune, A., and Blaut, M. (2004) *Appl. Environ. Microbiol.* **70,** 6131–6137
13. Nardini, M., and Dijkstra, B. W. (1999) *Curr. Opin. Struct. Biol.* **9,** 732–737
14. Grogan, G. (2002) *J. Mol. Catal. B Enzym.* **19,** 73–82
15. Grogan, G. (2005) *Biochem. J.* **388,** 721–730
16. Bateman, R. L., Bhanumoorthy, P., Witte, J. F., McClard, R. W., Grompe, M., and Timm, D. E. (2001) *J. Biol. Chem.* **276,** 15284–15291
17. Schleberger, C., Sachelaru, P., Brandsch, R., and Schulz, G. E. (2007) *J. Mol. Biol.* **367,** 409–418
18. Otwinowski, Z., and Minor, W. (1997) *Macromol. Crystallogr. A* **276,** 307–326
19. Usón, I., and Sheldrick, G. M. (1999) *Curr. Opin. Struct. Biol.* **9,** 643–648
20. Terwilliger, T. C., and Berendzen, J. (1999) *Acta Crystallogr. D Biol. Crystallogr.* **55,** 849–861
21. Morris, R. J., Perrakis, A., and Lamzin, V. S. (2002) *Acta Crystallogr. D Biol. Crystallogr.* **58,** 968–975
22. Collaborative Computational Project Number 4 (1994) *Acta Crystallogr. D Biol. Crystallogr.* **50,** 760–763
23. Emsley, P., and Cowtan, K. (2004) *Acta Crystallogr. D Biol. Crystallogr.* **60,** 2126–2132
24. Davis, I. W., Leaver-Fay, A., Chen, V. B., Block, J. N., Kapral, G. J., Wang, X., Murray, L. W., Arendall, W. B., 3rd, Snoeyink, J., Richardson, J. S., and Richardson, D. C. (2007) *Nucleic Acids Res.* **35,** W375–W383
25. Laskowski, R. A., Macarthur, M. W., Moss, D. S., and Thornton, J. M. (1993) *J. Appl. Crystallogr.* **26,** 283–291
26. Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994) *Nucleic Acids Res.* **22,** 4673–4680
27. Gouet, P., Robert, X., and Courcelle, E. (2003) *Nucleic Acids Res.* **31,** 3320–3323
28. DeLano, W. L. (2002) *The PyMOL Molecular Graphics System,* DeLano Scientific LLC, San Carlos, CA
29. Hermann, J. C., Ghanem, E., Li, Y., Raushel, F. M., Irwin, J. J., and Shoichet, B. K. (2006) *J. Am. Chem. Soc.* **128,** 15882–15891
30. Adams, P. D., Grosse-Kunstleve, R. W., Hung, L. W., Ioerger, T. R., McCoy, A. J., Moriarty, N. W., Read, R. J., Sacchettini, J. C., Sauter, N. K., and Terwilliger, T. C. (2002) *Acta Crystallogr. D Biol. Crystallogr.* **58,** 1948–1954
31. Hermann, J. C., Ridder, L., Höltje, H. D., and Mulholland, A. J. (2006) *Org. Biomol. Chem.* **4,** 206–210
32. Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., and Ferrin, T. E. (2004) *J. Comput. Chem.* **25,** 1605–1612
33. Moustakas, D. T., Lang, P. T., Pegg, S., Pettersen, E., Kuntz, I. D., Brooijmans, N., and Rizzo, R. C. (2006) *J. Comput. Aided Mol. Des.* **20,** 601–619
34. Iyer, L. M., Koonin, E. V., and Aravind, L. (2001) *Proteins* **43,** 134–144
35. Radauer, C., Lackner, P., and Breiteneder, H. (2008) *BMC Evol. Biol.* **8,** 286
36. Xu, R. X., Hassell, A. M., Vanderwall, D., Lambert, M. H., Holmes, W. D., Luther, M. A., Rocque, W. J., Milburn, M. V., Zhao, Y., Ke, H., and Nolte, R. T. (2000) *Science* **288,** 1822–1825
37. Gomis-Rüth, F. X. (2003) *Mol. Biotechnol.* **24,** 157–202
38. Liu, S., Mansour, M. N., Dillman, K. S., Perez, J. R., Danley, D. E., Aeed, P. A., Simons, S. P., Lemotte, P. K., and Menniti, F. S. (2008) *Proc. Natl. Acad. Sci. U.S.A.* **105,** 13309–13314
39. Kudo, N., Kumagai, K., Tomishige, N., Yamaji, T., Wakatsuki, S., Nishijima, M., Hanada, K., and Kato, R. (2008) *Proc. Natl. Acad. Sci. U.S.A.* **105,** 488–493
40. Cockcroft, S., and Carvou, N. (2007) *Biochim. Biophys. Acta* **1771,** 677–691
41. Houslay, M. D., Schafer, P., and Zhang, K. Y. (2005) *Drug Discov. Today* **10,** 1503–1519
42. Coleman, J. E. (1992) *Annu. Rev. Biophys. Biomol. Struct.* **21,** 441–483
43. Makinen, M. W., Wells, G. B., and Kang, S. O. (1984) *Adv. Inorg. Biochem.* **6,** 1–69
44. Fernandes, H., Pasternak, O., Bujacz, G., Bujacz, A., Sikorski, M. M., and Jaskolski, M. (2008) *J. Mol. Biol.* **378,** 1040–1051
45. Ames, B. D., Korman, T. P., Zhang, W., Smith, P., Vu, T., Tang, Y., and Tsai, S. C. (2008) *Proc. Natl. Acad. Sci. U.S.A.* **105,** 5349–5354
46. Li, S. C., Goto, N. K., Williams, K. A., and Deber, C. M. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93,** 6676–6681