The Structure and Function of a Novel Glycerophosphodiesterase from *Enterobacter aerogenes*

Colin J. Jackson¹, Paul D. Carr¹, Jian-Wei Liu¹, Stephen J. Watt²
Jennifer L. Beck² and David L. Ollis^{1*}

¹Research School of Chemistry
Australian National University
ACT, 0200, Australia

²Chemistry Department
University of Wollongong
NSW, 2522, Australia

The structure of the glycerophosphodiesterase (GDPD) from *Enterobacter aerogenes*, GpdQ, has been solved by SAD phasing from the active site metal ions. Structural analysis indicates that GpdQ belongs to the α/β sandwich metallo-phosphoesterase family, rather than the $(\alpha/\beta)_8$ barrel GDPD family, suggesting that GpdQ is a structurally novel GDPD. Hexameric GpdQ is generated by interactions between three dimers. The dimers are formed through domain swapping, stabilised by an inter-chain disulfide bond, and β -sheet extension. The active site contains a binuclear metal centre, with a fully occupied α -metal ion site, and partially occupied β -metal ion site, as revealed by anomalous scattering analysis. Using a combination of TLS refinement and normal mode analysis, the dynamic movement of GpdQ was investigated. This analysis suggests that the hexameric quaternary structure stabilises the base of the dimer, which promotes “breathing” of the active site cleft. Comparison with other metallo-phosphodiesterases shows that although the central, catalytic, domain is highly conserved, many of these enzymes possess structurally unrelated secondary domains located at the entrance of the active site. We suggest that this could be a common structural feature of metallo-phosphodiesterases that constrains substrate specificity, preventing non-specific phosphodiester hydrolysis.

Crown Copyright © 2007 Published by Elsevier Ltd. All rights reserved.

Keywords: glycerophosphodiester; phosphodiesterase; quaternary structure; dynamics; binuclear metallo-enzyme

*Corresponding author

Introduction

Glycerophosphodiesterases (GDPDs; EC 3.1.4.46) are a family of enzymes that catalyze the hydrolysis of the 3′–5′ phosphodiester bond of glycerophosphodiester; for instance, hydrolyzing glycerol-3-

phosphoethanolamine into *sn*-glycerol-3-phosphate (G3P) and ethanolamine (Figure 1).¹ GDPDs from bacterial,² yeast,³ plant⁴ and mammalian⁵ cells have been characterised. The primary function of these enzymes is the degradation of glycerophosphodiester produced through the deacylation of phospholipids, thereby maintaining homeostasis of G3P concentration as part of the process of phospholipid remodelling and synthesis.⁶

GDPDs form a large family with an evolutionarily conserved sequence motif (HR(X)ⁿEN(X)ⁿEXD(X)ⁿHD) comprising the active site residues, which is shared by the catalytic domain of the phosphoinositide-specific phospholipase C (PI-PLC) family.^{5,7,8} Structurally, all previously characterised GDPDs, including the GDPD from *Thermotoga maritima*,⁹ the periplasmic (GlpQ)¹⁰ and cytosolic (UgpQ)¹¹ GDPDs from *Escherichia coli*, adopt an $(\alpha/\beta)_8$ barrel tertiary fold. These proteins are also structurally similar to the bac-

Abbreviations used: DALI, distance matrix alignment; ENM, elastic network model; ESI-MS, electrospray ionisation mass spectrometry; G3P, *sn*-glycerol-3-phosphate; GDPD, glycerophosphodiesterase; GPE, glycerophosphoethanolamine; NMA, normal mode analysis; PI-PLC, phosphoinositide-specific phospholipase C; SAD, single-wavelength anomalous diffraction; SCOP, structural classification of proteins; TLS, translation-libration-screw.

E-mail address of the corresponding author:
ollis@rsc.anu.edu.au

Figure 1. The hydrolysis of the glycerophosphodiester sn-glycero-3-phosphoethanolamine. The leaving group will be protonated at physiological pH.

terial¹² and mammalian¹³ PI-PLCs, consistent with the conservation of catalytic residues. In fact, both the GDPD and PI-PLC protein families are classified together as the PLC-like phosphodiesterases by the SCOP database.¹⁴

The related periplasmic (GlpQ) and cytosolic (UgpQ) GDPDs have been characterised in *E. coli* and are expressed in operons that encode proteins that are part of distinct Ugp and Glp glycerophosphodiester uptake pathways.^{2,15} We have recently isolated an operon from *Enterobacter aerogenes*, in which every gene has a homologue in the Ugp pathway other than the GDPD, GpdQ, which is unrelated to UgpQ.¹⁶ More detailed sequence analysis of this enzyme suggested that it might be a structurally unique member of the GDPD family, as it does not contain the canonical GDPD sequence motif. Secondary structure prediction and identification of a conserved sequence motif (DXH(X)ⁿGDXXD(X)ⁿGNHD/E)¹⁷ indicated that GpdQ is likely to be structurally similar to the large and functionally diverse α/β sandwich family of metallo-dependent phosphoesterases that includes purple acid phosphatase,¹⁸ Mre11 nuclease¹⁹ and the Ser/Thr protein phosphatases.²⁰ These enzymes are typically monomeric or dimeric, with active sites containing two transition metals. Members of this family have displayed catalytic activity with a wide range of divalent and trivalent metal ions coordinated in the active site, including Zn^{2+} , Mn^{2+} , Fe^{2+} , and Fe^{3+} .^{19–21}

GpdQ has been cloned and recombinantly expressed in *E. coli*.¹⁶ The purified enzyme displays very broad substrate specificity, efficiently catalysing the hydrolysis of its natural substrate, glycerophosphoethanolamine, in addition to model phosphomonoester, phosphodiester, phosphotriester and phosphorothiolate substrates. We have obtained crystals of apo-GpdQ, into which we have soaked the transition metal ions Zn^{2+} and Co^{2+} , and collected diffraction data to 2.9 Å resolution.²² Herein, we report the structure of a novel bacterial GDPD, GpdQ, as determined by single wavelength anomalous dispersion (SAD) phasing from the active site metals.

Results

Tertiary and quaternary structure

Phases determined from a SAD experiment using the active site metals were used to solve the structures of Zn^{2+} and Co^{2+} substituted GpdQ (Table 1).

Table 1. Refinement statistics

Data collection ^a	GpdQ-Zn ²⁺	GpdQ-Co ²⁺
Space group	P213	P213
Unit cell parameter <i>a</i> (Å)	164.23	164.46
Temperature (K)	100	100
Wavelength (Å)	1.28269	0.97929
Resolution (Å) ^b	50.0–2.9 (3.1–2.9)	50.0–3.0 (3.2–3.0)
Unique reflections	32,293	29,959
Redundancy	12.5 (11.2)	14.0 (12.3)
$\langle I/\sigma(I) \rangle$	32.2 (4.6)	34.3 (4.6)
Completeness	100 (100)	100 [100]
R_{sym} (%)	7.7 (43.6)	7.3 (42.4)
Refinement		
Resolution range (Å)	50–2.9 (3.0–2.9)	50–3.0 (3.1–3.0)
Reflections (working)	30,501 (2231)	28,390 (2068)
Reflections (test)	1632 (110)	1527 (118)
$R_{\text{work}}/R_{\text{free}}$ (%) (without TLS)	19.5/23.5 (26.1/30.0)	18.7/23.3 (27.2/33.4)
$R_{\text{work}}/R_{\text{free}}$ (%) (with TLS)	16.2/19.5 (22.1/26.0)	16.0/19.4 (21.4/25.4)
B-factors (Å ²)		
Main-chain	76.1	72.6
Side-chain	77.0	73.2
Water molecules	68.9	65.0
Metal ions	77.5	74.8
No. protein atoms	4274	4274
No. water molecules	82	68
No. of metal ions	4	4
r.m.s. deviation from target bonds		
Lengths (Å)	0.01	0.01
Angles (°)	1.35	1.36
Ramachandran (%)		
Favoured	96.5	95.0
Allowed	3.5	5.0
Disallowed	0	0
PDB code	2DXL	2DXN

^a Data collection statistics have been previously published.²²

^b Data collection and refinement statistics for outer shell are in parentheses.

There is one 61.6 kDa dimer in the crystallographic asymmetric unit, containing two, effectively identical, 30.8 kDa subunits of GpdQ related by a non-crystallographic 2-fold axis at $(\theta, \phi) = (141, -17^\circ)$, in polar angles. All but the final three residues (EER) are present in reasonable electron density. Crystallographic 3-fold symmetry axes generate a hexameric enzyme. Each monomer consists of three domains: a classical α/β -sandwich fold where the active site is located, a novel

dimerization domain consisting of a pair of two-stranded antiparallel β -sheets connected by a short β -strand, and a small domain swapped cap of the active site (Figure 2).

The first 195 residues comprise the α/β sandwich domain, which consists of two facing β -sheets ($\beta 1, \beta 2, \beta 3$ and $\beta 4, \beta 5, \beta 6, \beta 7$), surrounded by α -helices ($\alpha A, \alpha B, \alpha C$ and $\alpha D, \alpha E, \alpha F$). This is shown in the topology diagram (Figure 2). A DALI²³ search using the residues 1–195 confirmed the prediction from

Figure 2. (Top) A topology diagram of GpdQ, illustrating the catalytic α/β sandwich domain ($\beta 1$ – $\beta 7/\alpha A$ – αF ; 1–196), the all β -strand dimerisation domain ($\beta 8$ – $\beta 12$; 197–256), and the domain swapped cap domain ($\beta 13, \alpha G$; 257–271) that is stabilised by an inter-chain disulfide bond. The locations of the liganding residues are shown. (Below) A ribbon diagram and carbon- α trace (red) of dimeric GpdQ, showing the location of the active site metals at the center of the α/β sandwich domain and the disulfide bond (yellow).

the earlier study; namely, that despite very low sequence similarity, GpdQ is structurally homologous to the members of a large and functionally diverse α/β sandwich family of metallo-dependent phosphoesterases.^{19,24–26} The closest homology is with purple acid phosphatase from pig,²¹ which is reflected in a reported overall root-mean-square difference (r.m.s.d.) between 168/197 residues of this domain of 2.7 Å (Z-score=21.5). The secondary structure and metal binding residues superimpose well, with the majority of the differences occurring at the surface exposed loop regions.

The second domain (196–255), which we have called the dimerization domain, joins the β -sandwiches from each subunit. In this way, a somewhat disjointed single β -sandwich is formed across the length of the dimer (Figure 2). β -Sheet extension between different polypeptide chains in the formation of quaternary structure, such as we observe in GpdQ, has been reported previously.^{27,28} The fold of this domain consists of a pair of two-stranded β -sheets from each chain, longer than the β -sheets from the α/β sandwich domain, with a shorter third

strand from each chain not directly contributing to the dimer interface (Figure 3). Withdrawal of the hydrophobic surfaces in this domain from aqueous solvent results in the formation of a hydrophobic core, which includes a symmetrical 3.1 Å interaction between the sulfur atoms of M201 from each chain at the centre. The dimer interface is extensive and results in the burial of 1667 Å², or 13.2%, of the total surface area of each chain. In contrast to the DALI search of the α/β sandwich domain, no structures with significant homology to this domain were identified (highest Z-score=2.5). Therefore, this fold may be novel. Additionally, although members of the α/β metallo-phosphatase family are frequently found to be dimeric proteins,²⁹ the β -sheet extension, and the large amount of buried surface area, observed here is, to the best of our knowledge, distinct.

The final 18 residues of each monomer (three of which are disordered) extend approximately 34 Å across the symmetry axis (Figure 2). Domain swapping, and the formation of domain-swapped dimers through C-terminal extensions, is not un-

Figure 3. The dimerisation domain of GpdQ, consisting of a pair of two-stranded antiparallel β -sheets and a short β -strand in the loop between the sheets, from each subunit. Amino acids are coloured by type (basic, blue; polar, yellow; non-polar, grey), illustrating the highly hydrophobic core formed through the dimerization of the two chains (cyan and magenta).

common.^{30,31} This extension forms a short strand, followed by a short helix and results in the burial of a further 1055 Å², or 8.4%, of the total surface area of each chain as a result of dimerization. The domain swapped extension is stabilised by the interactions between three residues of each chain: N53, C54, and R56 from chain A interact with S268, C269 and the backbone oxygen from P270, respectively. The observation of a disulfide bond between C54 and C269 was unexpected as the enzyme was purified from DH5α cells, whose redox potential should not support spontaneous disulfide bond formation; moreover, the formation of a structural disulfide

bond in a cytosolic bacterial enzyme is relatively rare.³² The final helix of the domain swapped extension forms a “cap” at the α/β sandwich subdomain, which significantly narrows the entrance to the active site cavity.

Three dimers interact around a 3-fold crystallographic axis to form hexameric GpdQ. Approximate dimensions of the hexamer are 70 Å across the width of the dimer × 90 Å × 90 Å (Figure 4). The interaction between the three dimers results in the burial of a further 1131 Å², or 9%, of the total surface area of each chain. Thus, the formation of a dimer, domain swapping, and oligomerization of

Figure 4. The D_3 symmetry of hexameric GpdQ. The hexamer is shown down the 3-fold axis (a), and down the perpendicular 2-fold axis (b).

Figure 5. Size-exclusion chromatography of GpdQ in 100 mM Hepes (pH 7.0). GpdQ elutes in a single peak and calibration with proteins of known molecular weight (inset) indicated GpdQ forms a large oligomer of approximately 290 kDa.

the dimers to form a hexamer, results in the burial of 31.2% of the surface area of each subunit. Due to the extent of the interaction between dimers, and the relatively high solvent content of these crystals (79%), we do not expect that the formation of a hexamer by the GpdQ dimers is an artefact of crystal packing. To confirm this, and to investigate the oligomeric state of GpdQ in solution, we performed size-exclusion chromatography. GpdQ eluted in a single peak and calibration with molecular weight standards indicated that, assuming GpdQ has a similar shape to the standards used, GpdQ forms a large oligomer in solution of approximately 290 kDa (Figure 5). The homogeneity of GpdQ was confirmed by SDS-PAGE, excluding the possibility that it had co-purified with any other proteins. Size-exclusion chromatography can be inaccurate in the determining molecular weight.³³ Accordingly, the molecular mass, and therefore the oligomeric state of GpdQ, was also determined by electrospray ionisation mass spectrometry (ESI-MS). Under non-denaturing conditions, the observed mass was 185.7 kDa, consistent with the calculated mass of a hexamer as observed in the crystal structure (Table 2). When treated with dithiothreitol (DTT) and diluted into acetonitrile, GpdQ remained associated as a dimer. Although it is surprising that the dimer was stable under these conditions, this result was consistent with the extent of the dimer interface in the crystal struc-

ture, or the possibility that the disulfide bond was not reduced.

TLS refinement

The observation of thermal diffuse scattering and the high Wilson B -factor ($77.2 \text{ \AA}^2$ for GpdQ-Zn) of the crystals indicated that there was considerable molecular disorder.²² Loose crystal packing of protein molecules allows an ensemble of conformations to exist in a crystal lattice, resulting in static and dynamic disorder about the crystallographic average of these structures.³⁴ Accordingly, translation-libration-screw (TLS) parameterization was used to model the collective motions of rigid bodies within the GpdQ structure in an attempt to account for both the large movements between each subunit,³⁵ as well as smaller rigid body movements within each subunit.³⁶ Three TLS models were analyzed: one rigid body containing dimeric GpdQ; two rigid bodies, consisting of the two subunits of the dimer; and six rigid bodies, resulting from the division of each subunit into the α/β sandwich, dimerisation and cap domains. The TLS groups were chosen from inspection of the structure and the main-chain B -factors. The B -factors of all atoms were set to 20 before ten rounds of TLS refinement. Table 3 shows that the six rigid body TLS model allows the best refinement of the data. The division of each GpdQ subunit into three rigid

Table 2. ESI-MS analysis of GpdQ

Conditions	M_w	Theoretical M_w	Multimer
200 mM NH_4OAc	$185,686.3 \pm 23$	185,039.8	Hexamer
50% acetonitrile, 1% formic acid, 2 mM DTT	$61,703.4 \pm 5$	61,679.6	Dimer

Table 3. Refinement statistics of GpdQ using TLS models consisting of one, two, six and eight rigid bodies

TLS Model	TLS groups	R_{work}	R_{free}	FOM
None	0	19.5	23.5	0.835
Dimer	1	18.0	21.2	0.864
Monomers	2	17.7	20.8	0.868
Domains	6	17.3	20.4	0.872

Figure 6. A plot of the total B -factor (B_{TOT} , black), in addition to the residual (B_{RES} , red) and TLS (B_{TLS} , blue) components, of the atoms from one GpdQ subunit. The three rigid bodies used in the refinement are coloured (pink, catalytic domain; green, dimerisation domain; blue, cap domain).

bodies is supported by analysis of the total main-chain B -factor (B_{TOT}) alongside the residual (B_{RES}) and TLS (B_{TLS}) components, shown in Figure 6. Firstly, the variation in B_{TOT} is well modelled by the TLS parameterization, B_{TLS} . Secondly, three different patterns are evident, corresponding to the α/β sandwich domain (1–195), the dimerisation domain (196–255) and the domain swapped active site cap (256–271), indicating that the nature of the atomic displacements within these domains are relatively distinct, supporting the notion that they can be considered to be separate pseudo-rigid bodies. The libration vectors for each rigid body are listed in Table 4 and shown in Figure 7. This demonstrates that the direction and magnitude of the libration axes of the three domains are significantly different from each other, and that the centers of these axes are located very close to the base of the dimer at the hexamer interface. It is clear that the B -factors of the atoms within the catalytic domain increase from the

base to the top of the dimer (Figure 7). This is consistent with stabilisation of the domain at the hexamer interface, allowing it to “rock” from this point, highlighting the important role of the oligomerisation of GpdQ in promoting the enzyme dynamics.

Normal mode analysis

Because the TLS refinement indicated that there was significant variation in the conformation of the molecules with the crystal, normal mode analysis (NMA) of an elastic network model (ENM) of GpdQ was used to investigate the dynamic movement in detail. The ~185 kDa GpdQ hexamer is arguably too large for molecular dynamics simulations, and NMA of an ENM is an appropriate alternative for the analysis of dynamics in large systems. It has been shown to accurately predict low energy conformational changes in protein structures,³⁷ resulting in its increasingly frequent application to the investigation of enzyme dynamics.³⁸ The lowest frequency modes obtained for dimeric GpdQ indicated that the cap domain was highly mobile, which is consistent with the TLS refinement. However, there was no indication of any collective movement that could result in the large increase in B -factors of the atoms within the catalytic and dimerisation domains from the lower to upper halves. In contrast, the two lowest energy modes obtained from an ENM of hexameric GpdQ described a breathing motion in which the base of the catalytic and dimerization domains, located at the hexamer interface, is relatively immobile owing to the quaternary structure interactions (Figure 7). This interface forms a hinge around which these domains can rock, resulting in significantly greater movement of the atoms at the top of the dimer, relative to the base. This motion is consistent with the TLS refinement and B -factor analysis (Figure 6),

Table 4. The directions of the axes of libration, relative to orthogonal axes, and mean-square displacement ($\text{\AA}^2$) of the three pseudo-rigid bodies in each GpdQ subunit

Domain/ rigid body	Angle libration axes make to orthogonal axes ($^\circ$)			Mean-square displacement ($\text{\AA}^2$)	Mean libration ($\text{\AA}^2$)
	X	Y	Z		
Catalytic domain (1–195)	41.45	63.55	60.68	1.74	1.64
	123.18	33.57	85.43	0.55	
	111.87	109.12	27.74	2.63	
Dimerisation domain (196–256)	57.90	54.15	52.26	2.22	1.89
	122.25	37.58	107.17	0.18	
	131.14	99.78	42.80	3.29	
Cap domain (257–296)	39.35	129.20	87.03	2.76	10.57
	53.23	41.02	74.38	22.75	
	102.08	100.21	15.91	6.20

Figure 7. (a) Ribbon diagrams of GpdQ with the three pseudo-rigid bodies/domains coloured red (catalytic domain), green (dimerisation domain) and blue (cap domain). Libration tensors (scaled by 1.5) from the TLS refinement are shown in magenta (catalytic domain), yellow (dimerisation domain) and cyan (cap domain). The centers of libration are all located at the base of the dimer, close to the hexamer interface. (b) GpdQ represented as thermal ellipsoids, obtained through TLS refinement, coloured by B_{TLS} (red, high; blue, low). (c) An ensemble of 11 GpdQ C α conformations obtained from the lowest frequency normal mode. Chains are coloured by r.m.s. deviation from the starting structure (red, highest; blue, lowest). Comparison between (b) and (c) demonstrates the similarity between the NMA and TLS refinement, in the stabilisation of atoms at the hexamer interface.

in which there is a large difference in the magnitude of the displacements of atoms from the base of the catalytic domain, which participates in the hexamer interface, to the atoms of the “free” upper half. The catalytic domain and the cap domain form the two sides of the active site cleft; because the cap domain is not involved in this movement (the TLS analysis indicates that it moves independently) breathing of the active site clefts occurs. Since the active site clefts are located on opposite sides of the dimer, this results in alternate opening of each cleft. Cooperative dynamic movement between monomers in a

dimer has previously been observed in *E. coli* dihydroorotase.³⁹

The active site and substrate binding pocket

The active site cavity of GpdQ is located at the center of the catalytic domain of each subunit (Figure 2), and a diagram of the active site architecture is shown in Figure 8. The seven conserved residues of the phosphodiesterase motif and two metal ions constitute the active site. Given the modest 2.9 Å resolution limit, we could not unambiguously model

Figure 8. The active site architecture of GpdQ. Omit electron density maps were calculated from models refined in the absence of bound ligands. An $mF_o - Df_c$ map is shown, contoured at $+3\sigma$ (red), illustrating the presence of a non-protein ligand. Two water molecules were placed in this density and refined to the positions shown. A Bijvoet difference Fourier map, calculated using anomalous data collected at the Zn K-edge and contoured at 15σ (blue), demonstrates that the Zn^{2+} occupancy of the α site is higher than the β site.

a μ -hydroxide or μ -water bridge through inspection of the electron density. Despite this, an omit electron density map, calculated from models refined in the absence of bound water molecules, reveals positive density at the active site that corresponds to a non-protein ligand. Since bridging water/hydroxide molecules are present in the binuclear metal centers of related enzymes,^{18,19,24,25} a bridging water molecule was modelled into this density. Refinement resulted in the water molecule occupying a position with metal–ligand distances of 2.6 Å to the α -metal and 2.9 Å to the β -metal. A water molecule was also modelled into the second coordination sphere and refined to a position 2.6 Å from the bridging water, 3.2 Å from the amide group of N80, and 3.4 Å from H81. The bond lengths of the bridging water molecule are considerably longer than what would be expected for a μ -hydroxide bridge (2.0–2.1 Å) and are unlikely to be particularly accurate. Accordingly, these water molecules were not included in the final model. Higher resolution structures will be required before the identity of this bridging ligand can be identified with any certainty. The exact coordination of the

metals cannot be determined based on the present model. The α -coordination sphere consists of H10, H197 and D50 as equatorial ligands and D8 as an axial ligand. The β -site coordination sphere consists of H195, N80 and D50 as equatorial ligands and H156 as an axial ligand. The active site residues effectively overlay onto the corresponding residues of other members of this family.

Kinetic analysis demonstrates that GpdQ is catalytically active in the presence of Zn^{2+} and Co^{2+} , although the catalytic efficiency of Zn^{2+} -GpdQ is approximately 30-fold less than Co^{2+} -GpdQ for the model substrate bis(*p*-nitrophenyl) phosphate (Table 5). These metals were soaked into the

Table 5. Kinetic parameters for the hydrolysis of bis(*p*-nitrophenyl) phosphate by Co^{2+} and Zn^{2+} substituted GpdQ

Protein	k_{cat} (s^{-1})	K_{m} (mM)	$k_{\text{cat}}/K_{\text{m}}$ ($\text{s}^{-1} \text{M}^{-1}$)
GpdQ+ Co^{2+}	1.62 ± 0.03	1.4 ± 0.1	1.2×10^3
GpdQ+ Zn^{2+}	0.107 ± 0.004	2.9 ± 0.3	3.7×10^1

apo-enzyme, principally to facilitate phase determination. It was hoped, however, that differences in the coordination of these metals could provide insight into the structural basis for the difference in activity. There were no significant structural differences apparent at this resolution. Using the anomalous data collected at the Zn K-edge, we were able to calculate a Bijvoet difference Fourier map of the active site of the Zn^{2+} -GpdQ enzyme (Figure 8). The peak height of the β metal ion was 80% that of the α -metal ion, which indicates that the more buried α -site has a higher affinity for Zn^{2+} . The Co^{2+} occupancy of the β sites in the Gpdq- Co^{2+} model was also 80% that of the α -site, when the occupancy was set so that the *B*-factors of the metal ions were similar to that of the surrounding amino acid ligands. The observation that the two metal ion binding sites have different affinities is consistent with their different coordination geometries: although both sites contain two histidine ligands and are dually coordinated by D50, the α -site has an additional acidic ligand (D8) whereas the β -site has

an additional polar ligand (N80). This is not unprecedented, as other members of the α/β phosphohydrolase family have been described as containing tight and loose metal ion binding sites.⁴⁰

In order to characterise the substrate binding pocket of GpdQ, the natural substrate glycerophosphoethanolamine (GPE) was manually modelled into the active site (Figure 9). The glycerol group is housed in a cavity formed by, in addition to the metal ligands, Q166, M167, H217, V219, P228, Y230 and the backbone atoms of Q218 and Y229. The leaving group is coordinated in a channel formed by N80, H81, S127 and I170 of chain A, and Y263 and E265 of chain B. S127 is in a position to form a hydrogen bond to the amide group of the ethanolamine leaving group. The phosphate group is coordinated to the β -metal. Coordination of phosphodiester substrates to the binuclear active sites of other metallo-phosphoesterases has previously been observed to occur in a bidentate fashion to the more solvent exposed β -metal, displacing the bridging hydroxide.⁴¹ This binding mode allows the most

Figure 9. The substrate-binding pocket of GpdQ, with the modelled orientation of the natural substrate glycerophosphoethanolamine (GPE). (a) The glycerol-binding pocket. (b) The leaving group channel formed by the domain swapped cap, stabilised by the inter-chain C54–C269 disulfide bond, and the loops surrounding the binuclear active site. (c) The increase in the area of the leaving group channel formed by E265 of chain B, S127 and H81 of chain A. The predicted movement from the lowest frequency normal mode (left to right) results in a significant increase in the area between these side-chains.

reasonable orientation of GPE in the active site of GpdQ, in the absence of additional experimental data. Finally, the oxygen of the leaving group and N^ε of H81 are separated by 3 Å. Another notable feature of the molecular surface around the positively charged binuclear metal center is a large region of electronegative potential.

Discussion

Comparison to structural and functional homologs

The α/β sandwich domain of GpdQ and the architecture of the active site, place GpdQ within the metallo-dependent phosphoesterase structural family (SCOP classification 56300¹⁴), distinct from other members of the glycerophosphodiesterase functional family (SCOP classification 89508,¹⁴ EC 3.1.4.46⁴²). GpdQ is therefore a structurally novel GDPD. Although the structural fold is distinct within the GDPD family, the quaternary structure is not; a search of the Protein Data Bank identified a glycerophosphodiesterase from *Agrobacterium tumefaciens* that, while the fold and active site structures

are of the canonical GDPD type, also exists as a hexamer.⁴³ A two iteration sequence homology search, using PSI-BLAST,⁴⁴ identified over 130 closely related sequences from bacterial proteins of putative phosphodiesterase function with *E*-values less than 1e-50. Considering the biological importance of GDPDs, it is reasonable to suggest that many of these proteins might also be GDPDs.

The novel phosphodiesterase from *Methanococcus janacchia*⁴⁵ has an identical active site to GpdQ and activity toward the model GDPD substrate *p*-nitrophenyl phosphocholine. Additionally, both GpdQ and the phosphodiesterase from *M. janacchia* have highly acidic surfaces. This is shown in an electrostatic potential surface (Figure 10), which demonstrates that there is an extensive electronegative region around the active site cleft. The highly charged surface may be an indication of the cellular location of GpdQ, or have a functional role. The *pI* of GpdQ is 5.45, which is very similar to the *pI* of both the periplasmic (GlpQ, 5.38) and cytosolic (UgpQ, 6.09) GDPDs of *E. coli*.² Considering there are no sequence motifs in GpdQ for transport to the periplasm, the acidic surface is unlikely to reflect its cellular location. It is therefore probable that the electronegative surfaces of

Figure 10. Electrostatic potential mapped to the molecular surface of GpdQ. Large regions of negative potential surround the entrance to the active site. Calculation parameters are listed in Materials and Methods.

GpdQ and the phosphodiesterase from *M. janacchia* are functionally significant; for instance, it would make the entrance to the active site repel double-stranded (ds) DNA molecules. This is supported by the structure of a related enzyme, Mre11 nuclease, where the structure suggests that the dsDNA substrate is attracted to an electro-positive groove.¹⁹

The catalytic mechanism

Considering the similarity between the active sites, the catalytic mechanism of GpdQ is most probably very similar to other members of the metallo-phosphoesterase family. There is very strong evidence that phosphate esters are hydrolysed *via* direct in-line attack from a hydroxide ion.⁴⁶ The function of the histidine in the second coordination sphere of these enzymes has been debated, and it has been proposed to be involved in both substrate binding and general acid catalysis.^{20,47} The measurement of enzyme kinetic parameters in this study was performed at a pH value at which the histidine will be uncharged (pH 8.7). It should be noted, however, that H81 is hydrogen bonded to the carbonyl oxygen of N53. Thus, we cannot exclude the possibility that the pK_a of H81 has been altered such that it will be charged at pH 8.7. If it is uncharged, it will not coordinate to the anionic phosphodiester during substrate binding and the delocalised negative charge must be accommodated by the binuclear metal center. This will result in the substrate coordinating in either a bidentate fashion to the loosely bound β -metal, displacing the μ -hydroxide bridge,⁴¹ or in a bridging fashion to both metals.¹⁹ Since the leaving group of the substrate used in these assays (*p*-nitropholate, $pK_a=7.1$) will not require protonation at pH 8.7, the participation of H81 in general acid catalysis cannot be speculated on. However, the ethanolamine leaving group of the natural substrate glycerophosphorylethanolamine has a relatively high pK_a value (9.50), and its protonation through general acid catalysis is likely to be an essential aspect of the catalytic mechanism at physiologically relevant pH values. Further work will be required in order to fully elucidate the catalytic mechanism, especially at physiologically relevant pH values with the natural substrate.

The importance of enzyme dynamics to catalysis

It is interesting that despite the very good fit of the glycerol group into the glycerol binding pocket, the phosphate group at the binuclear metal center, and the ethanolamine group into the leaving group channel, the entrance to the active site in this crystal structure is very narrow and will not be able to accommodate the bulkier leaving group of the substrate assayed here, bis(*p*-nitrophenol phosphate) in the present conformation. In addition to the TLS refinement and NMA, this suggests that there is

considerable breathing of the active site cleft. Analysis of the predicted conformational change demonstrates its functional importance to catalysis: as shown in Figure 9, the independent movement of the catalytic domain and the cap domain results in the area between H81, S127 (catalytic domain) and E265 (cap domain), which form two sides of the cleft, increasing from 17.5 Å² to 28.4 Å². This motion is a direct result of the hexameric organisation of GpdQ, suggesting that the evolution of quaternary structure in GpdQ has had a significant impact on the molecular dynamics, and thus the activity, of the enzyme.

The cap subdomain: a common structural feature of metallo-phosphodiesterases

A significant feature of this structure is the existence of the cap sub-domain. This structural phenomenon, namely, the addition of an unrelated secondary domain to a structurally conserved central domain, has been previously described in other α/β metallo-phosphodiesterases, including a GDPD from *T. maritima*,⁹ Mre11 nuclease,¹⁹ and 5'-nucleotidase.²⁵ In each of these enzymes, the additional domain is larger than the cap in GpdQ, but they are all located at the entrance to the active site. The structural variety of these domains supports the notion that they have evolved to constrain the various substrate specificities of these enzymes, especially as their complexity appears to increase in accordance with the size and nature of the natural substrate. The cap domain of Mre11 is large, contains a positive groove for binding dsDNA, and permits significant opening of the active site cleft. In contrast, the cap domain of GpdQ is very small and narrows the width of the active site cleft to a size that is most appropriate for the leaving groups of glycerophosphodiester ethanoline or choline. Because the catalytic machinery of these proteins is capable of hydrolysing virtually any phosphodiester bond, and non-specific hydrolysis of compounds with phosphodiester bonds would be detrimental to the health of cells, these domains have probably evolved to prevent this.

Summary

The novel GDPD, GpdQ, belongs to a diverse family of metallo-phosphoesterases, for which the catalytic mechanism is still not fully understood. Further analysis of the mechanism of GpdQ, with particular emphasis on the hydrolysis of the natural substrate, could aid in the mechanistic understanding of these enzymes. We present analysis of the enzyme dynamics of GpdQ, highlighting the importance of quaternary structure and independent domain movement to activity. It is hoped that further work could extend this to investigate the importance of enzyme dynamics to other members of this family.

Materials and Methods

SAD phasing

Data collection statistics have been reported elsewhere.²² Phase determination using SAD⁴⁸ from the active site metals of metallo-enzymes has been reported.⁴⁹ Because previous work suggested that the asymmetric unit contained a GpdQ dimer,²² SHELXD,⁵⁰ as implemented in the CCP4 suite of programs,⁵¹ was used to locate the positions of the four Zn²⁺ sites in an anomalous difference Patterson synthesis using the SAD data collected at 1.28269 Å. MLPHARE⁵² was subsequently used to refine the occupancy of the four sites and obtain initial phases with a phasing power of 2.43. SHELXE was used for density modification and phase improvement;⁵³ the high solvent content (79%) undoubtedly contributed to the quality of the phases.⁵⁴

Model building and refinement

The initial phases were used to perform automated model building with ARP-wARP.⁵⁵ This produced an initial model with R_{free} of 38.0%. Several rounds of interactive model building were then carried out using COOT,⁵⁶ followed by structure idealisation as implemented in REFMAC v5.0,⁵⁷ after which R_{free} was 28.7% (Table 1). Restrained refinement and the addition of water molecules reduced R_{free} to 23.5%. The B -factors were then set to 20 Å² and ten rounds of TLS refinement,³⁵ using three rigid bodies comprising residues 1–195, 196–255 and 256–271 of each chain, followed by three rounds of maximum likelihood refinement, further lowered R_{free} to 20.0%, then 19.6%. TLSANL⁵⁸ and ANISOANL⁵⁹ were used to analyse the data produced by TLS refinement. The libration tensors produced by TLSANL were visualised using RIBBONS,⁶⁰ using a scale factor of 1.5. Geometric validation of the structures was made using RAMPAGE,⁶¹ which indicated that all residues were in favourable or allowed regions; PROCHECK,⁶² which indicated all stereochemical parameters were better than or inside normal limits; and SFCHECK,⁶³ which gave an overall error in coordinates by Luzatti plot of 0.45 Å. The GpdQ-Co²⁺ structure was refined using the GpdQ-Zn²⁺ coordinates, after extension of the R_{free} data.

Normal mode analysis of GpdQ using an elastic network model

In order to examine the conformational flexibility of GpdQ, NMA was performed using an ENM of dimeric and hexameric GpdQ. The structures were submitted to the elNemo server[†] using default parameters. Details of the theory can be found elsewhere.⁶⁴ The output consisted of the five lowest frequency mode, the lowest two of which were effectively identical and displayed the most reasonable conformational change, when compared to experimentally determined B -factors and data from TLS refinement. Movies of the predicted movement, shown as the molecular surface around the active site cleft and as a ribbon diagram, were made using PYMOL, and are available in the supporting information.

Size-exclusion chromatography

Size-exclusion chromatography was performed at 4 °C using an FPLC system (AKTA, Amersham Biosciences, USA) on a Superdex 200 (Amersham Biosciences) column equilibrated with elution buffer (50 mM Hepes (Sigma, Australia, pH 7.5), 100 mM NaCl (Ajax Finechem, Australia). Fractions of 0.5 ml were collected at a flow rate of 0.1 ml/min, and absorbance was measured at 280 nm. The partition coefficient K_{av} was calculated using the formula $K_{\text{av}} = (V_e - V_0)/(V_t - V_0)$, where V_e is the elution volume of the sample, V_0 is the excluded volume of the column, V_t is the total volume of the column. The excluded volume, V_0 (7.75 ml) was measured by calibration with Blue Dextran 2000. On the assumption of similar shape factors, the calibration plot of $\log K_{\text{av}}$ versus M_w allowed determination in first approximation of the molecular weight of GpdQ. The calibration plot was constructed using a high molecular weight calibration kit (Amersham Biosciences) containing Blue Dextran 2000, thyroglobulin (669 kDa), ferritin (440 kDa), catalase (232 kDa) and aldolase (158 kDa).

Mass spectrometry

For mass spectrometry analysis GpdQ was exchanged into 200 mM CH₃COONH₄ (Ajax Finechem, pH 7.0), using a Millipore Biomax centrifugal filter (USA) with the application of 6 × 400 µl aliquots of this solvent. For spectra of the dimer, GpdQ (in 200 mM CH₃COONH₄) was treated with 2 mM DTT (ICN Biomedicals, USA), before dilution into a solution giving 1% (v/v) formic acid and 50% (v/v) CH₃CN. All mass spectra were acquired using a Waters QToF UltimaTM mass spectrometer using equipped with a Z-spray ESI source. Typically, the capillary voltage was 2.5 kV, cone 120 V, RF lens 160 V, RF 1 lens pressure 2.6 × 10⁻³ mbar, collision cell 4 eV, transport and aperture 5 V and desolvation temperature 120 °C. For the analysis of the intact hexamer, the cone, RF lens, and RF 1 lens pressure were increased to 249 V, 184 V, and 1.0 × 10⁻¹ mbar, respectively. Typically, 20 scans were acquired for mass spectra.

Enzyme assays

Enzyme concentration was determined by measuring absorbance at 280 nm in 6 M urea, using the extinction coefficient 39,880 M⁻¹ cm⁻¹. Determination of the steady state kinetic parameters of purified GpdQ was achieved by monitoring the production of *p*-nitrophenolate at 410 nm ($\epsilon_{410} = 16,600 \text{ M}^{-1} \text{ cm}^{-1}$) at various concentrations of bis-*p*NPP (Sigma), ranging from 50 µM to 10.2 mM. The reaction mixture contained 50 mM KCl (Ajax Finechem), 100 mM Ches (Sigma, pH 8.7), and 100 µM of either CoCl₂ or ZnCl₂ (Ajax Finechem). All measurements were made in duplicate. The values for V_{max} and K_m were determined from a fit of the data to the following equation:

$$v = V_{\text{max}} A / (K_m + A)$$

where v is the initial velocity, V_{max} is the maximum velocity, K_m is the Michaelis constant, and A is the substrate concentration. The k_{cat} was calculated by dividing V_{max} by the concentration of enzyme used in the reaction ($V_{\text{max}} = k_{\text{cat}}[E]$).

[†] <http://www.igs.cnrs-mrs.fr/elNemo/>

Calculation of electrostatic potential

Electrostatic potential was calculated by the program Swiss PDB Viewer using partial and full charges, computed using simple Coulomb interaction.⁶⁵ The protein was taken to be at pH 7.0, with default protonation states for all residues, i.e. histidine residues are singly protonated. Default dielectric constants were used. The electrostatic potential shown in **Figures 8 and 9** was calculated with a -5 kT/e cutoff for negative potential and a 0 kT/e cutoff for positive potential.

Protein Data Bank accession codes

The coordinates of Zn^{2+} and Co^{2+} containing GpdQ have been deposited in the RCSB Protein Data Bank under accession codes 2DXL and 2DXN.

Acknowledgements

This work was supported by an Australian Research Council grant to D.O., and the Australian Synchrotron Research Program. We thank C.-S. Chan and the staff of the Structural Biology Centre at Argonne National Laboratory for expert assistance at the beamline. We thank A. Oakley for assistance in data collection. C. J. J. was supported by ANU Graduate School and A. Sargeson scholarships.

Supplementary Data

Supplementary data associated with this article can be found, in the online version, at [doi:10.1016/j.jmb.2007.01.032](https://doi.org/10.1016/j.jmb.2007.01.032)

References

- Larson, T. J., Ehrmann, M. & Boos, W. (1983). Periplasmic glycerophosphodiester phosphodiesterase of *Escherichia coli*, a new enzyme of the glp regulon. *J. Biol. Chem.* **258**, 5428–5432.
- Tommassen, J., Eiglmeier, K., Cole, S. T., Overduin, P., Larson, T. J. & Boos, W. (1991). Characterization of two genes, glpQ and ugpQ, encoding glycerophosphoryl diester phosphodiesterases of *Escherichia coli*. *Mol. Gen. Genet.* **226**, 321–327.
- Fisher, E., Almaguer, C., Holic, R., Griac, P. & Patton-Vogt, J. (2005). Glycerophosphocholine-dependent growth requires Gde1p (YPL110c) and Git1p in *Saccharomyces cerevisiae*. *J. Biol. Chem.* **280**, 36110–36117.
- van der Rest, B., Boisson, A. M., Gout, E., Bligny, R. & Douce, R. (2002). Glycerophosphocholine metabolism in higher plant cells. Evidence of a new glyceryl-phosphodiester phosphodiesterase. *Plant Physiol.* **130**, 244–255.
- Zheng, B., Berrie, C. P., Corda, D. & Farquhar, M. G. (2003). GDE1/MIR16 is a glycerophosphoinositol phosphodiesterase regulated by stimulation of G protein-coupled receptors. *Proc. Natl Acad. Sci. USA*, **100**, 1745–1750.
- Raetz, C. R. (1986). Molecular genetics of membrane phospholipid synthesis. *Annu. Rev. Genet.* **20**, 253–295.
- Meldrum, E., Parker, P. J. & Carozzi, A. (1991). The PtdIns-PLC superfamily and signal transduction. *Biochim. Biophys. Acta*, **1092**, 49–71.
- Zheng, B., Chen, D. & Farquhar, M. G. (2000). MIR16, a putative membrane glycerophosphodiester phosphodiesterase, interacts with RGS16. *Proc. Natl Acad. Sci. USA*, **97**, 3999–4004.
- Santelli, E., Schwarzenbacher, R., McMullan, D., Biorac, T., Brinen, L. S., Canaves, J. M. *et al.* (2004). Crystal structure of a glycerophosphodiester phosphodiesterase (GDPD) from *Thermotoga maritima* (TM1621) at 1.60 Å resolution. *Proteins: Struct. Funct. Genet.* **56**, 167–170.
- Zhang, R., Kim, Y., Dementieva, I., Duck, N., Stols, L., Donnelly, M. & Joachimiak, A. (2004). Structural genomics, crystal structure of glycerophosphoryl diester phosphodiesterase from *E. coli*. *RCSB Protein Data Bank 1TQ8*.
- Malashkevich, V. M., Fedorov, E. & Almo, S. C. (2005). Crystal structure of periplasmic glycerophosphodiesterase from *Escherichia coli*. *RCSB Protein Data Bank 1YDY*.
- Heinz, D. W., Ryan, M., Bullock, T. L. & Griffith, O. H. (1995). Crystal structure of the phosphatidylinositol-specific phospholipase C from *Bacillus cereus* in complex with myo-inositol. *EMBO J.* **14**, 3855–3863.
- Essen, L. O., Perisic, O., Cheung, R., Katan, M. & Williams, R. L. (1996). Crystal structure of a mammalian phosphoinositide-specific phospholipase C delta. *Nature*, **380**, 595–602.
- Andreeva, A., Howorth, D., Brenner, S. E., Hubbard, T. J., Chothia, C. & Murzin, A. G. (2004). SCOP database in 2004: refinements integrate structure and sequence family data. *Nucl. Acids Res.* **32**, 226–229.
- Brzoska, P. & Boos, W. (1988). Characteristics of a ugp-encoded and phoB-dependent glycerophosphoryl diester phosphodiesterase which is physically dependent on the ugp transport system of *Escherichia coli*. *J. Bacteriol.* **170**, 4125–4135.
- McLoughlin, S. Y., Jackson, C., Liu, J. W. & Ollis, D. L. (2004). Growth of *Escherichia coli* coexpressing phosphotriesterase and glycerophosphodiester phosphodiesterase, using paraoxon as the sole phosphorus source. *Appl. Environ. Microbiol.* **70**, 404–412.
- Koonin, E. V. (1994). Conserved sequence pattern in a wide variety of phosphoesterases. *Protein Sci.* **3**, 356–358.
- Klabunde, T., Strater, N., Frohlich, R., Witzel, H. & Krebs, B. (1996). Mechanism of Fe(III)-Zn(II) purple acid phosphatase based on crystal structures. *J. Mol. Biol.* **259**, 737–748.
- Hopfner, K. P., Karcher, A., Craig, L., Woo, T. T., Carney, J. P. & Tainer, J. A. (2001). Structural biochemistry and interaction architecture of the DNA double-strand break repair Mre11 nuclease and Rad50-ATPase. *Cell*, **105**, 473–485.
- Rusnak, F. & Mertz, P. (2000). Calcineurin: form and function. *Physiol. Rev.* **80**, 1483–1521.
- Guddat, L. W., McAlpine, A. S., Hume, D., Hamilton, S., de Jersey, J. & Martin, J. L. (1999). Crystal structure of mammalian purple acid phosphatase. *Struct. Fold. Des.* **7**, 757–767.
- Jackson, C. J., Carr, P. D., Kim, H. K., Liu, J. W. & Ollis, D. L. (2006). The purification, crystallization and preliminary diffraction of a glycerophosphodiesterase from *Enterobacter aerogenes*. *Acta Crystallog. sect. F*, **62**, 659–661.

23. Holm, L. & Sander, C. (1993). Protein structure comparison by alignment of distance matrices. *J. Mol. Biol.* **233**, 123–138.
24. Kissinger, C. R., Parge, H. E., Knighton, D. R., Lewis, C. T., Pelletier, L. A., Tempczyk, A. *et al.* (1995). Crystal structures of human calcineurin and the human FKBP12-FK506-calcineurin complex. *Nature*, **378**, 641–644.
25. Knofel, T. & Strater, N. (1999). X-ray structure of the *Escherichia coli* periplasmic 5'-nucleotidase containing a dimetal catalytic site. *Nature Struct. Biol.* **6**, 448–453.
26. Strater, N., Klabunde, T., Tucker, P., Witzel, H. & Krebs, B. (1995). Crystal structure of a purple acid phosphatase containing a dinuclear Fe(III)-Zn(II) active site. *Science*, **268**, 1489–1492.
27. Bullock, T. L., Roberts, T. M. & Stewart, M. (1996). 2.5 Å resolution crystal structure of the motile major sperm protein (MSP) of *Ascaris suum*. *J. Mol. Biol.* **263**, 284–296.
28. Carr, P. D., Cheah, E., Suffolk, P. M., Vasudevan, S. G., Dixon, N. E. & Ollis, D. L. (1996). X-ray structure of the signal transduction protein from *Escherichia coli* at 1.9 Å. *Acta Crystallog. sect. D*, **52**, 93–104.
29. Schenk, G., Gahan, L. R., Carrington, L. E., Mitic, N., Valizadeh, M., Hamilton, S. E. *et al.* (2005). Phosphate forms an unusual tripodal complex with the Fe-Mn center of sweet potato purple acid phosphatase. *Proc. Natl Acad. Sci. USA*, **102**, 273–278.
30. Pasta, S. Y., Raman, B., Ramakrishna, T. & Rao Ch, M. (2002). Role of the C-terminal extensions of alpha-crystallins. Swapping the C-terminal extension of alpha-crystallin to alphaB-crystallin results in enhanced chaperone activity. *J. Biol. Chem.* **277**, 45821–45828.
31. Rousseau, F., Schymkowitz, J. W. & Itzhaki, L. S. (2003). The unfolding story of three-dimensional domain swapping. *Structure*, **11**, 243–251.
32. Kadokura, H., Katzen, F. & Beckwith, J. (2003). Protein disulfide bond formation in prokaryotes. *Annu. Rev. Biochem.* **72**, 111–135.
33. Richter, W. (2005). Determining the subunit structure of phosphodiesterases using gel filtration and sucrose density gradient centrifugation. *Methods Mol. Biol.* **307**, 167–180.
34. Burgi, H. B. (2000). Motion and disorder in crystal structure analysis: measuring and distinguishing them. *Annu. Rev. Phys. Chem.* **51**, 275–296.
35. Winn, M. D., Isupov, M. N. & Murshudov, G. N. (2001). Use of TLS parameters to model anisotropic displacements in macromolecular refinement. *Acta Crystallog. sect. D*, **57**, 122–133.
36. Painter, J. & Merritt, E. A. (2006). Optimal description of a protein structure in terms of multiple groups undergoing TLS motion. *Acta Crystallog. sect. D*, **62**, 439–450.
37. Tama, F. & Sanejouand, Y. H. (2001). Conformational change of proteins arising from normal mode calculations. *Protein Eng.* **14**, 1–6.
38. Bahar, I. & Rader, A. J. (2005). Coarse-grained normal mode analysis in structural biology. *Curr. Opin. Struct. Biol.* **15**, 586–592.
39. Lee, M., Chan, C. W., Mitchell Guss, J., Christopherson, R. I. & Maher, M. J. (2005). Dihydroorotase from *Escherichia coli*: loop movement and cooperativity between subunits. *J. Mol. Biol.* **348**, 523–533.
40. White, D. J., Reiter, N. J., Sikkink, R. A., Yu, L. & Rusnak, F. (2001). Identification of the high affinity Mn²⁺ binding site of bacteriophage lambda phosphoprotein phosphatase: effects of metal ligand mutations on electron paramagnetic resonance spectra and phosphatase activities. *Biochemistry*, **40**, 8918–8929.
41. Hamdan, S., Carr, P. D., Brown, S. E., Ollis, D. L. & Dixon, N. E. (2002). Structural basis for proofreading during replication of the *Escherichia coli* chromosome. *Structure*, **10**, 535–546.
42. Fleischmann, A., Darsow, M., Degtyarenko, K., Fleischmann, W., Boyce, S., Axelsen, K. B. *et al.* (2004). IntEnz, the integrated relational enzyme database. *Nucl. Acids Res.* **32**, D434–D437.
43. Rao, K. N., Bonanno, J. B., Burley, S. K. & Swaminathan, S. (2006). Crystal structure of glycerophosphodiester phosphodiesterase from *Agrobacterium tumefaciens* by SAD with a large asymmetric unit. *Proteins*, **65**, 514–518.
44. Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W. & Lipman, D. J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucl. Acids Res.* **25**, 3389–3402.
45. Chen, S., Yakunin, A. F., Kuznetsova, E., Busso, D., Pufan, R., Proudfoot, M. *et al.* (2004). Structural and functional characterization of a novel phosphodiesterase from *Methanococcus jannaschii*. *J. Biol. Chem.* **279**, 31854–31862.
46. Mueller, E. G., Crowder, M. W., Averill, B. A. & Knowles, J. R. (1993). Purple acid phosphatase: a diiron enzyme that catalyzes a direct phospho group transfer to water. *J. Am. Chem. Soc.* **115**, 2974–2975.
47. Funhoff, E. G., Wang, Y., Andersson, G. & Averill, B. A. (2005). Substrate positioning by His92 is important in catalysis by purple acid phosphatase. *FEBS J.* **272**, 2968–2977.
48. Dauter, Z., Dauter, M. & Dodson, E. (2002). Jolly SAD. *Acta Crystallog. sect. D*, **58**, 494–506.
49. Liu, D., Lepore, B. W., Petsko, G. A., Thomas, P. W., Stone, E. M., Fast, W. & Ringe, D. (2005). Three-dimensional structure of the quorum-quenching N-acyl homoserine lactone hydrolase from *Bacillus thuringiensis*. *Proc. Natl Acad. Sci. USA*, **102**, 11882–11887.
50. Schneider, T. R. & Sheldrick, G. M. (2002). Substructure solution with SHELXD. *Acta Crystallog. sect. D*, **58**, 1772–1779.
51. Collaborative-Computational-Research-Project-4. (1994). The CCP4 suite: programs for protein crystallography. *Acta Crystallog. sect. D*, **50**, 760–763.
52. Otwinowski, Z. (1991). *Daresbury Study Weekend proceedings*.
53. Sheldrick, G. M. (2002). Macromolecular phasing with SHELXE. *Z. Kristallor.* **217**, 644–650.
54. Terwilliger, T. C. (2001). Map-likelihood phasing. *Acta Crystallog. sect. D*, **57**, 1763–1775.
55. Perrakis, A., Harkiolaki, M., Wilson, K. S. & Lamzin, V. S. (2001). ARP/wARP and molecular replacement. *Acta Crystallog. sect. D*, **57**, 1445–1450.
56. Emsley, P. & Cowtan, K. (2004). Coot: model-building tools for molecular graphics. *Acta Crystallog. sect. D*, **60**, 2126–2132.
57. Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997). Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallog. sect. D*, **53**, 240–255.
58. Howlin, B., Butler, S. A., Moss, D. S., Harris, G. W. & Driessen, H. P. C. (1993). TLSANL: TLS parameter-analysis program for segmented anisotropic refinement of macromolecular structures. *J. Appl. Crystallog.* **26**, 622–624.
59. Winn, M. (2001). ANISOANL - analysing anisotropic displacement parameters. *CCP4 Newsletter*, **39**, 3.

60. Carson, M. (1991). RIBBONS 2.0. *J. Appl. Crystallog.* **24**, 958–961.
61. Lovell, S. C., Davis, I. W., Arendall, W. B., 3rd, de Bakker, P. I., Word, J. M., Prisant, M. G. *et al.* (2003). Structure validation by Calpha geometry: phi,psi and Cbeta deviation. *Proteins: Struct. Funct. Genet.* **50**, 437–450.
62. Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallog.* **26**, 283–291.
63. Vaguine, A. A., Richelle, J. & Wodak, S. J. (1999). SFCHECK: a unified set of procedures for evaluating the quality of macromolecular structure-factor data and their agreement with the atomic model. *Acta Crystallog. sect. D*, **55**, 191–205.
64. Suhre, K. & Sanejouand, Y. H. (2004). ElNemo: a normal mode web server for protein movement analysis and the generation of templates for molecular replacement. *Nucl. Acids Res.* **32**, W610–W614.
65. Kaplan, W. & Littlejohn, T. G. (2001). Swiss-PDB viewer (Deep View). *Brief. Bioinform.* **2**, 195–197.

Edited by I. Wilson

(Received 6 September 2006; received in revised form 10 January 2007; accepted 11 January 2007)
Available online 20 January 2007