

Vahed Oganessian, Paul D.
Adams, Jarmila Jancarik,
Rosalind Kim and Sung-Hou
Kim*

Berkeley Structural Genomics Center, Physical
Biosciences Division, Lawrence Berkeley
National Laboratory, Berkeley,
California 94720, USA

Correspondence e-mail: shkim@lbl.gov

Received 18 August 2006

Accepted 16 April 2007

PDB Reference: O67745_AQUAE, 2hek,
r2heksf.



© 2007 International Union of Crystallography
All rights reserved

Structure of O67745_AQUAE, a hypothetical protein from *Aquifex aeolicus*

Using single-wavelength anomalous dispersion data obtained from a gold-derivatized crystal, the X-ray crystal structure of the protein O67745_AQUAE from the prokaryotic organism *Aquifex aeolicus* has been determined to a resolution of 2.0 Å. Amino-acid residues 1–371 of the 44 kDa protein were identified by *Pfam* as an HD domain and a member of the metal-dependent phosphohydrolase superfamily (accession No. PF01966). Although three families from this large and diverse group of enzymatic proteins are represented in the PDB, the structure of O67745_AQUAE reveals a unique fold that is unlike the others and that is likely to represent a new subfamily, further organizing the families and characterizing the proteins. Data are presented that provide the first insights into the structural organization of the proteins within this clan and a distal alternative GDP-binding domain outside the metal-binding active site is proposed.

1. Introduction

The primary research goal of the Berkeley Structural Genomics Center is to obtain a near-complete structural complement of two closely related pathogens, *Mycoplasma genitalium* and *M. pneumoniae*. Within these constraints, protein targets have been chosen that will most readily yield novel functional and structural information. Within the gene AQ-1910, the 371-amino-acid protein O67745_AQUAE has been identified as a member of the HD-domain enzymatic proteins (Bateman *et al.*, 2004). This protein, which was targeted for its identified motif and predicted activity, was recombinantly expressed as a noncleavable fusion protein bearing an N-terminal His₆ tag *via* a Gly₆ linker, with a total molecular weight of 45 356 Da. Here, we describe in detail the cloning, expression, purification and three-dimensional X-ray structure of the protein O67745_AQUAE. From the 2.0 Å resolution structure of this protein, we propose a novel family of HD-domain phosphohydrolases that bind a metal in their putative active site and bind GDP as a cofactor in a distal alternative binding site.

2. Materials and methods

2.1. Expression and purification

The DNA fragment encoding O67745_AQUAE was amplified by PCR from *Aquifex aeolicus* genomic DNA (American Type Culture Collection) using Deep Vent DNA Polymerase (New England Biolabs, Beverly, MA, USA). The resulting PCR product was purified and prepared for ligation-independent cloning (LIC; Aslanidis & de Jong, 1990) by treatment with T4 DNA polymerase in the presence of 1 mM dTTP for 30 min at 310 K. The DNA fragment was mixed with LIC-prepared pB2 vector for 5 min at room temperature and transformed into DH5 α . The LIC pB2 vector was designed in our laboratory to express the target protein in fusion with a noncleavable N-terminal His₆ tag. The plasmid containing the correct-sized insert was transformed into BL21 (DE3)/pSJS1244 (Kim *et al.*, 1998) for native protein expression. Cells were grown in auto-induction media overnight at 301 K (Studier, 2005). Since the 371-amino-acid protein contains only one methionine residue, which would not produce

Table 1
Data-collection and phasing statistics.

Values in parentheses are for the outermost resolution shell.		
	Gold derivative	Native
Wavelength (Å)	1.04	1.00
Resolution (Å)	2.60	1.92
Space group	$P2_12_12_1$	$C222_1$
Unit-cell parameters (Å)	$a = 94.0, b = 97.8,$ $c = 185.1$	$a = 107.7, b = 144.4,$ $c = 144.7$
Total reflections	469220	383814
Unique reflections	50323	82404
R_{sym}	0.085 (0.60)	0.057 (0.73)
Completeness (%)	94.8 (92.7)	96.8 (72.5)
$I/\sigma(I)$	23.1 (1.9)	24.4 (1.4)
Figure of merit	0.24†	0.74‡

† Before density-modification procedure. ‡ After density-modification procedure.

sufficient scattering power for phasing on selenomethionine substitution, selenomethionine labeling of the protein was not attempted. Cells were disrupted using a microfluidizer (Microfluidics, Newton, MA, USA) in 50 mM HEPES pH 7.0, 500 mM NaCl, 5% glycerol, 1 mM PMSF, 10 µg ml⁻¹ DNase, 0.1 µg ml⁻¹ antipain, 1 µg ml⁻¹ chymostatin, 0.5 µg ml⁻¹ leupeptin and 0.7 µg ml⁻¹ pepstatin A. The supernatant was ultracentrifuged in a Ti45 rotor at 35 000 rev min⁻¹ for 30 min at 277 K and then loaded onto a HiTrap Ni²⁺-chelating column (GE Healthcare, Piscataway, NJ, USA). The His-tagged fusion protein was captured on the column in 50 mM HEPES pH 7.0, 100 mM NaCl and eluted with a gradient to the same buffer supplemented with 1 M imidazole. Fractions containing the protein were pooled and dialyzed overnight at 277 K against 50 mM HEPES pH 7.0, 20 mM NaCl. After centrifugation, the protein sample was loaded onto a 5 ml HiTrapQ column and eluted with approximately 500 mM NaCl in a gradient elution. The purity of the target protein was confirmed visually on a Coomassie brilliant blue-stained SDS-PAGE gel and MALDI-TOF mass spectrometry confirmed the protein identity. Dynamic light-scattering experiments (DynaPro 99, Wyatt Technology, Santa Barbara, CA, USA) were performed in the concentration range 0.5–2.5 mg ml⁻¹. Highly homogeneous protein with a hydrodynamic radius of 35 Å was detected in all the concentration ranges. The values of the polydispersity obtained were dependent on the chemistry and pH of the buffer used in the optimum-solubility screen (Jancarik *et al.*, 2004). Citrate buffer at pH 4.0 was determined to be the optimum buffer condition and was used

Table 2
Refinement statistics.

Resolution range (Å)	12.0–2.0
R factor (free R factor)	0.175 (0.220)
R.m.s.d. bonds (Å)	0.015
R.m.s.d. angles (°)	1.452
Residues in most favored region of (ϕ, ψ) space (%)	93.2
Residues in additionally allowed region of (ϕ, ψ) space (%)	6.8
No. of protein atoms	6220
No. of non-protein atoms	487
B factor (model/Wilson) (Å ²)	46.2/44

to exchange the protein buffer. The protein was concentrated to 10 mg ml⁻¹ in a buffer consisting of 20 mM sodium citrate pH 4.0 and 100 mM NaCl and was then used for crystallization studies.

2.2. Crystallization

Crystallization conditions were evaluated using 96 unique conditions each from the three screening applications Crystal Screens I and II (Jancarik & Kim, 1991), Index Screen and Salt Rx (Hampton Research, Aliso Viejo, CA, USA) and also using 48 unique conditions each from the two screening applications Wizard I and Wizard II (deCode Genetics, Bainbridge Island, WA, USA), employing a Phoenix liquid-handling system (Art Robbins Instruments, Sunnyvale, CA, USA). Vapor-diffusion crystallization screening experiments in sitting drops were performed in two sets at 277 and 296 K. Each of the 768 crystallization experiments contained 500 nl crystallization solution and 500 nl concentrated protein in 96-well conical flat-bottom treated nonsterile Crystal EX-3785 plates (Corning Inc., Corning, NY, USA). The plates were imaged daily in the first week and then weekly during the first month using the Crystal Farm Imaging system (Nexus Biosystems, Poway, CA, USA). 25 conditions yielded crystalline precipitation and were further optimized in hanging drops using a 24-well format in which 1 µl reservoir solution and 1 µl protein solution were combined and placed on treated glass cover slides. Screening within the 25 conditions was refined by varying the pH of the buffer and the concentration of the precipitant. Large crystals (0.2 × 0.2 × 0.1 mm) were obtained in four of the 25 conditions. None of the crystals generated could tolerate cryoprotectant or heavy-metal solutions. When 5% (v/v) glycerol was added to the concentrated protein prior to crystallization, the subsequent addition of glycerol required for

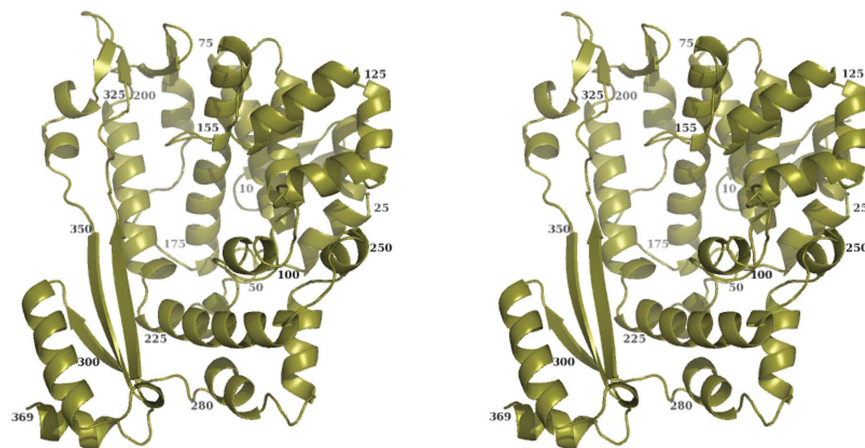


Figure 1
Stereoview of the polypeptide fold of O67745_AQUAE. Approximately every 25th residue is numbered. Residues 1–285 and 305–355 constitute the ‘bow’ and residues 286–304 and 356–369 form the ‘handle’. All figures were prepared using PyMOL (DeLano Scientific LLC, San Carlos, CA, USA).

```

Q9V1T1_PYRAB/57-183 RFEHSLGTTWYLAKKLSQELNLH...YEESLLIQLAALLHDIHGHPFSHTFRI...MIIHGDIDVQLDYITRD
O59446_PYRHO/39-165 RFEHSLGTTWYLARKLSLELQLP...REDSLLIQLAALLHDIHGHPFSHTFRI...MIIHGDIDVQLDYITRD
Y1154_METJA/50-150 RFEHSLGTFMFIASKIAEKIN...ADVELTRVSALLHDIHGHPFSHTLTCIG...KIISGDVDAIRMDYLLRD
YWFO_BACSU/61-181 RFNHSLGVVYEIVRRMVDDVFKG...RPEWDDSERELCLAAALLHDLHGHPFSHSFKV...ISSQIDAIRMDYLQRD
O67745_AQUAE/51-169 RFEHSLGVVYHITERICESLK...VKEKELVKLAGLLHDLHGHPFSHTTEVLL...EITITGEGFSIRMDYLRRD
O28849_ARCFU/51-157 RFEHSLGVVHVTRILQERMG...FDDVVVAALLHDLVGHAPFSHGSRLLL...SVVNGEIDAIRMDYLVRD
O26251_METTH/50-159 RFEHSLGAMYLASRLAEHLGLG...HEKRVLRRLCALHDLVGHGPFSSHVS...QAIISGELDVIRMDYLLRD

```

Figure 2

Amino-acid sequence alignment of several members of the HD superfamily. Conserved residues are shaded red. The fact that all the highly conserved residues in the HD superfamily are histidines or aspartates suggests that coordination of divalent cations could be essential for the activity of these proteins. Conserved residues in contact with the metal ion in the putative active site are marked by an asterisk above the sequence.

proper cryopreservation [up to 25%(v/v)] to the mature crystals did not destroy them. The best crystallization conditions contained 200 mM NaCl, 100 mM phosphate-citrate pH 4.2 and 10% polyethylene glycol 3000. Heavy-atom-containing crystals were obtained by cocrystallizing the protein with 0.1 mM K_2AuCl_4 . Both native and gold-derivatized crystals grew within a week.

2.3. X-ray diffraction and structure determination

Gold-derivatized crystals were washed once with reservoir solution prior to cryoprotection to remove excess unbound gold ions. Both gold-derivatized and native crystals of O67745_AQUAE were cryoprotected with 25% glycerol prior to flash-cooling. The frozen crystals were placed on the goniometer head under a continuous nitrogen stream. In an attempt to collect anomalous data at a bromide absorption-peak wavelength, the native crystal was briefly soaked in potassium bromide (KBr); however, a fluorescence scan did not detect the presence of bromide ions. Interestingly, the KBr-soaked crystal diffracted better than the native crystals and was therefore used to collect the 'native' data.

In data collection from the native crystal, 240 consecutive images with an oscillation range of 0.5° and an exposure time of 3 s were measured. Data were collected to a resolution of 1.92 Å from a single crystal at 100 K using a wavelength of 1.0 Å. The single-wavelength anomalous data for the gold-derivatized crystal consisted of 448 images (224 images for each direct and inverse beam in 5° wedges) with an oscillation range of 0.5° and an exposure time of 5 s. Data were collected to a resolution of 2.6 Å from a single crystal at 100 K using a wavelength of 1.04 Å. Both data sets were collected at the Berkeley Center for Structural Biology (Lawrence Berkeley National Laboratory, Berkeley, CA, USA) at Advanced Light Source beamline 8.2.2 equipped with a Quantum 315 CCD detector (Area Detector System Corporation, Poway, CA, USA). All data were processed with *HKL-2000* (Otwinowski & Minor, 1997). Crystallographic statistics for the data sets from both crystals are given in Table 1.

The *HySS* software (Grosse-Kunstleve & Adams, 2003) identified eight gold sites in the four molecules in the asymmetric unit in the gold-derivatized crystal (Matthews parameter $V_M = 2.4 \text{ \AA}^3 \text{ Da}^{-1}$). Phasing, density modification and automatic model building were performed using *SOLVE/RESOLVE* (Terwilliger, 2004). 364 residues of 1484 were built using *RESOLVE*. About 50% of these contained side chains. Manual building with *O* (Jones *et al.*, 1991) of one of the four polypeptide chains in the asymmetric unit in the gold-derivatized crystal yielded 300 consecutive amino acids corresponding to residues 50–351 of the protein O67745_AQUAE. Using this partial structure as a model, molecular replacement was used to resolve the native crystal structure. Two molecules were identified in the asymmetric unit using *MOLREP* (Vagin & Teplyakov, 1997). Electron density with phases from molecular replacement allowed completion of the model before refinement with *REFMAC* (Murshudov *et al.*, 1997)

from the *CCP4* suite (Collaborative Computational Project, Number 4, 1994). During the refinement, 5% of reflections were set aside for R_{free} calculations. Refinement and model statistics are summarized in Table 2. The final model consists of residues 1–369 of both polypeptide chains in the asymmetric unit of the native crystal, 355 solvent molecules, four phosphate ions, two zinc ions, five chloride ions, one bromide ion, two GDP molecules and eight glycerol molecules. The coordinates have been deposited with the PDB (Berman *et al.*, 2000) with code 2hek.

3. Results and discussion

The single polypeptide of O67745_AQUAE has a 'bowl-with-handle' conformation, with a diameter of approximately 60 Å at its widest point and a depth of 35 Å (Fig. 1). It contains 29 secondary-structure elements consisting of 20 α -helices and nine β -strands. The protein exhibits four antiparallel β -sheets. The first antiparallel two-stranded sheet consists of strands β_1 (residues 3–7) and β_2 (residues 10–15) and identifies the beginning of the polypeptide chain. This structure is followed by two nearly perpendicular short α -helices α_1 (residues 17–24) and α_2 (residues 26–33). The third short helix α_3 (residues 39–43) together with two longer helices α_4 (residues 51–69) and α_5 (residues 78–86) constitute the bottom of the 'bowl'. At approximately residue 86, where helix α_5 ends, the polypeptide contains

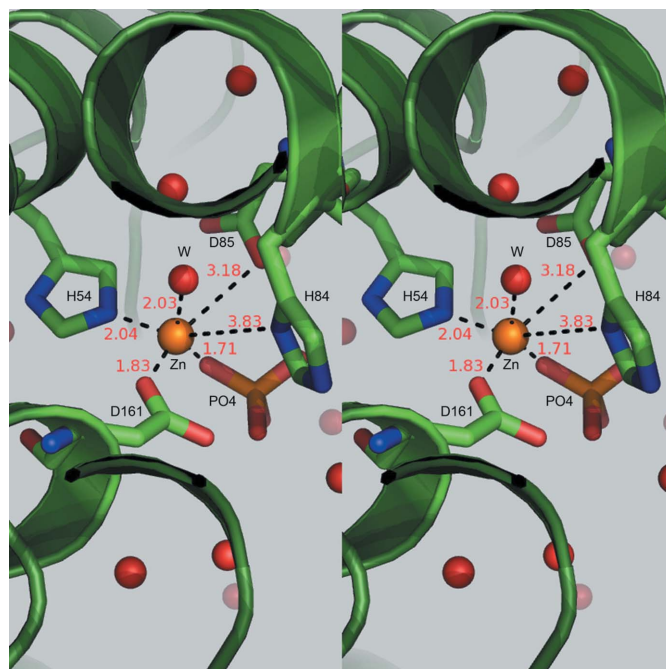


Figure 3

Stereoview of the putative active site of O67745_AQUAE.

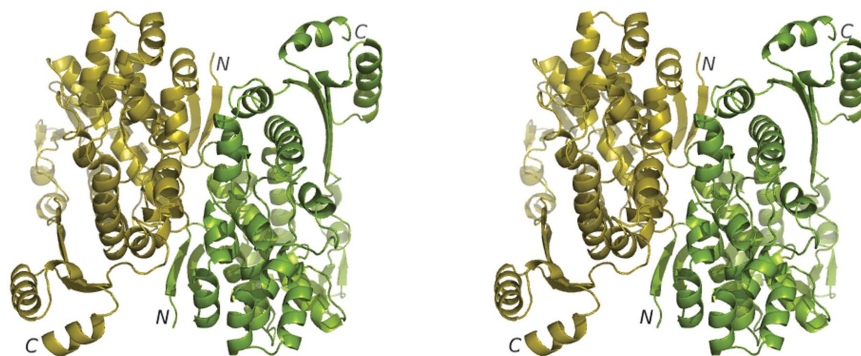


Figure 4
Stereoview of the interaction between two molecules of O67745_AQUAE related by NCS.

residues that form the sides of the ‘bowl’. The secondary-structure elements that participate in building the ‘bowl’ include helices $\alpha 6$ – $\alpha 17$ ($\alpha 6$, residues 95–99; $\alpha 7$, residues 108–115; $\alpha 8$, residues 117–125; $\alpha 9$, residues 128–139; $\alpha 10$, residues 145–155; $\alpha 11$, residues 160–173; $\alpha 12$, residues 183–188; $\alpha 13$, residues 203–222; $\alpha 14$, residues 224–242; $\alpha 15$, residues 249–254; $\alpha 16$, residues 257–266; $\alpha 17$, residues 271–278). The second consecutive antiparallel sheet is formed by strands $\beta 3$ (residues 189–192) and $\beta 4$ (residues 195–199). Strand $\beta 5$ (residues 284–289) and helix $\alpha 18$ (residues 292–305) begin to form the ‘handle’, which is completed by the last helix $\alpha 20$ (residues 358–368). A long stretch of residues from amino acids 311–356 form the other side of the ‘bowl’ and consists of strands $\beta 6$ (residues 311–318), $\beta 7$ (residues 324–328) and $\beta 8$ (residues 331–338), helix $\alpha 19$ (residues 341–346) and strand $\beta 9$ (residues 349–356). Strands $\beta 7$ and $\beta 8$ make up the third short antiparallel sheet, while strands $\beta 5$, $\beta 6$ and $\beta 9$ constitute the fourth and longest antiparallel sheet. Notably, helices $\alpha 10$ and $\alpha 11$ would be combined into one helix if it was not for an intervening four-amino-acid loop which does not conform to the requirements for the φ and ψ angles of α -helices.

The amino-acid sequence of O67745_AQUAE is identified by *Pfam* (Bateman *et al.*, 2004) as belonging to the HD-domain family, a member of a superfamily of enzymes with a predicted or known phosphohydrolase activity. These enzymes appear to be involved in

nucleic acid metabolism, signal transduction and possibly other functions in bacteria, archaea and eukaryotes. Since all the highly conserved residues within the HD superfamily are histidines or aspartates, the coordination of divalent cations could be essential for the activity of these proteins. A sequence alignment of several members of subfamily PF01966 is presented in Fig. 2. A metal ion is found at the bottom of the ‘bowl’ and a phosphate ion is present in the coordination sphere of the ion (Fig. 3). Based on the position of the metal and the phosphate ion, this region of the structure has been identified as the putative active site. The metal ion was modeled as Zn^{2+} based on the tetrahedral nature of its first coordination sphere. The occupancy was adjusted manually to 0.5 according to both the *B* factors ($Zn1$ has a *B* factor of 17 Å² and $Zn2$ 37 Å²) of neighboring atoms and the presence of positive or negative ($F_o - F_c$) electron density.

The asymmetric unit of O67745_AQUAE contains two polypeptides that interact through the bottom of the ‘bowl’ as shown in Fig. 4.

The *PISA* (Krissinel & Henrick, 2005) server finds that the complexation score for the two molecules in the asymmetric unit is only 0.063, which implies that the interface is not significant for complexation and may solely be a result of crystal packing. Only about 11% of the residues from each polypeptide participate in the interface and these residues account for less than 7% of the surface

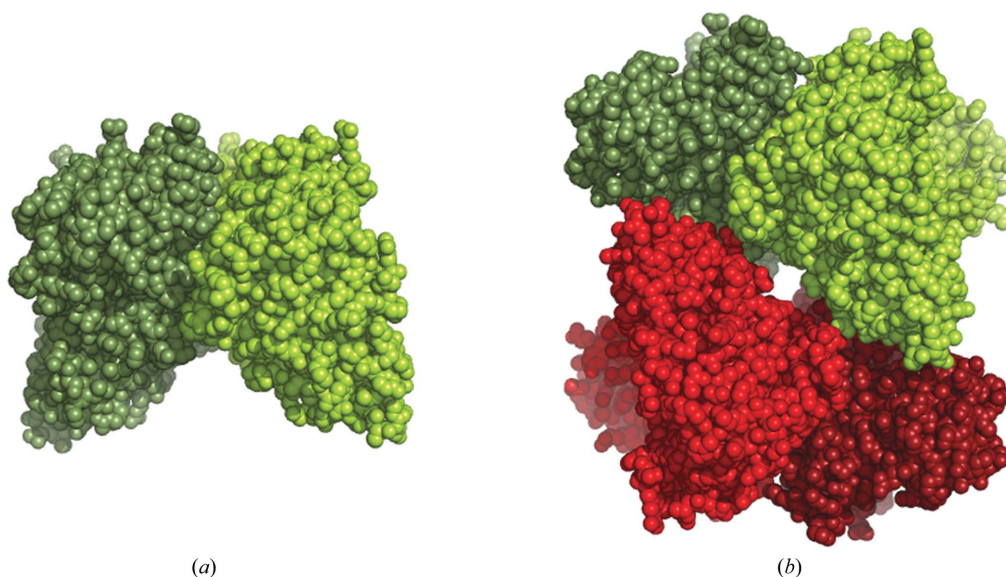
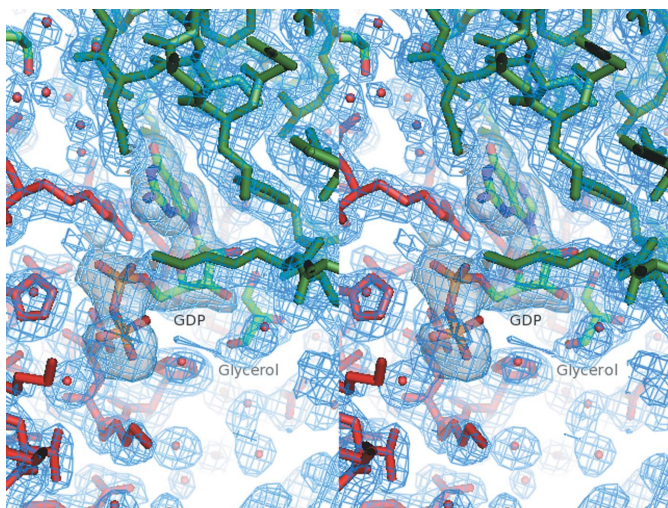


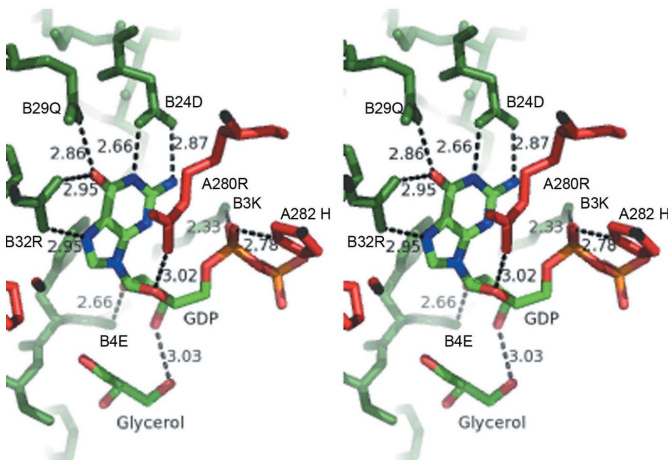
Figure 5
Dimerization (a) and tetramerization (b) of O67745_AQUAE.

area (the total surface area of a single polypeptide is about 17 200 Å²). The PISA server also suggests that the most probable complexation state for O67745_AQUAE would be a tetramer (Fig. 5), in which the buried area would account for nearly 40% of the total surface area of the tetramer (56 136 Å² total surface area compared with 22 413 Å² buried surface area).

Experimental data on the oligomerization of O67745_AQUAE is limited to dynamic light scattering which, across a wide range of protein concentrations, consistently showed a monodisperse peak with a hydrodynamic radius of 35 Å corresponding to a molecular weight of about 50 kDa. Size-exclusion chromatography or sedimentation and static light-scattering experiments may have provided a more reliable and definite determination of the oligomerization state of O67745_AQUAE in solution. Although the dynamic light-scattering data suggest no complexation in solution and the PISA server predicts that the interface between the two molecules in the asymmetric unit is not significant for complexation, we still hypothesize that O67745_AQUAE is a dimer in solution. Such a hypothesis is based on the presence of two GDP molecules detected at the interface between two polypeptides of the asymmetric unit (Fig. 6).



(a)



(b)

Figure 6

(a) Fragment of the electron density along with a model of the protein with GDP and glycerol fitted. Residues in green and in red belong to different polypeptide chains of the asymmetric unit. Eight sites within the asymmetric unit were fitted with glycerol molecules. (b) Residues from both polypeptide chains (shown in different colors) contact each GDP molecule. Short contacts are shown.

Upon dimerization, the two polypeptides in the asymmetric unit create two similar and highly positively charged deep clefts that accommodate the GDP molecules (Fig. 7). Amino acids from both polypeptides in the asymmetric unit coordinate the GDP molecules. Arg280 and His282 from one polypeptide unit and Lys3, Glu4, Asp24, Glu29 and Arg32 from the other polypeptide unit are within hydrogen-bonding distance of the docked GDP molecule. The second GDP molecule of the asymmetric unit has nearly identical coordination.

3.1. Comparison to structural homologues

In an attempt to establish the identity and subsequently assign the polypeptide fold of O67745_AQUAE, a spatial similarity search was carried out using the DALI server (Holm & Sander, 1999). Four structures were found with Z scores above 6. The highest score was with the exopolyphosphatase from *Escherichia coli* (PDB code 1u6z; Hasson *et al.*, unpublished work). 147 residues of the C-terminal second domain of exopolyphosphatase match the N-terminal portion of O67745_AQUAE with a root-mean-square deviation (r.m.s.d.) value of 3.2 Å. Within the aligned regions of these two polypeptides the sequence identity is only 12% owing to 17 gaps in the alignment spread over 320 residues of O67745_AQUAE.

The second highest Z score was found with the enteropathogenic *E. coli* type III secretion-system protein EscJ (PDB code 1yj7; Yip *et al.*, 2005). The N-terminal 54 residues of EscJ, equivalent to one third of the protein, very closely match (r.m.s.d. value of 1.6 Å) the fold of the 'handle' of O67745_AQUAE, although the sequence identity is only 13%. In the supermolecular complex of the type III secretion system, the N-termini of all the EscJ molecules are localized towards the broad face of the complex and are believed to be involved in the anchoring of the protein to the inner membrane. Whether or not the 'handle' in O67745_AQUAE plays a similar anchoring role is currently unknown.

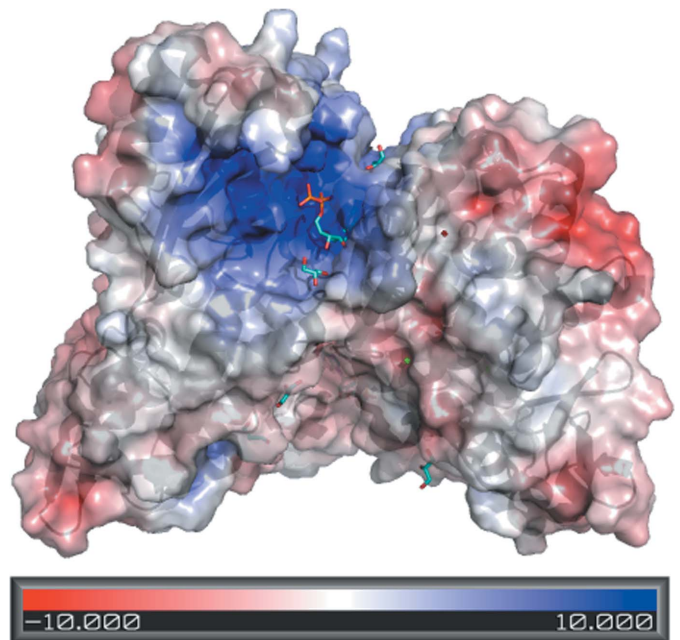


Figure 7

The positively charged cleft formed on the interface of two polypeptides of the asymmetric unit accommodates a GDP molecule. The surface electrostatic potential was calculated using the Adaptive Poisson-Boltzmann Solver software package (Baker *et al.*, 2001).

The third highest *Z* score of 6.7 was with the hypothetical protein AF1432 (PDB code 1ynb; Dong *et al.*, unpublished), where 133 residues of the 167-amino-acid protein match O67745_AQUAE with a sequence identity of 11%, the insertion of 15 gaps and an r.m.s.d. value of 4.6 Å.

The fourth highest *Z* score of 6.6 was found by a comparison of the protein with phosphodiesterase 5 (PDE5; PDB code 1rkp; Huai *et al.*, 2004). In PDE5, 157 residues of the 311-amino-acid polypeptide match O67745_AQUAE with an r.m.s.d. value of 4.2 Å. Superimposition of the structures of O67745_AQUAE and PDE5 shows the distance between zinc ions in the two polypeptides to be only 1.5 Å, with most of the identical residues shared by the proteins occupying very close positions in the vicinity of the metal ion. Histidines 54, 84 and 88 of O67745_AQUAE align with histidines 617, 653 and 657 of PDE5, and aspartic acids 85 and 151 of O67745_AQUAE align with aspartic acids 654 and 764 of PDE5. The only mismatch of residues is Arg51 of O67745_AQUAE and His613 of PDE5. Notably, the NH1 atom of Arg51, which is within hydrogen-bonding distance of the zinc-bound phosphate ion, is only 1.2 Å away from the NE2 atom of His613. In addition, the glycerol molecule found in the putative active site of O67745_AQUAE aligns well with the inhibitor IBMX found in the active site of PDE5.

All the structures to which *DALI* found some degree of similarity belong to different *Pfam* families.

Based on the presence of two GDP molecules and the similarity to phosphodiesterase PDE5 found by *DALI*, we would like to hypothesize that O67745_AQUAE is a homodimeric enzyme belonging to one of the 23 known phosphoesterases/phosphodiesterase families, making the fold space of phosphodiesterases even broader.

The authors thank Dr John-Marc Chandonia for target selection, Hisao Yokota and Barbara Gold for cloning, Marlene Henriquez and Bruno Martinez for expression and cell-paste preparation and Dr Alexander Iakounine of the University of Toronto for functional assays. We are indebted to M. Damschroder for reading the manuscript. We are also reliant on other members of BSGC as well as staff members for user support at beamline 8.2.2 at the Advanced Light

Source. The Advanced Light Source is supported by the Director, Office of Science, Office of Basic Energy Sciences and Material Sciences Division of the US Department of Energy under contract No. DE-AC03-76SF00098 at Lawrence Berkeley National Laboratory. The research presented here was supported by a Protein Structure Initiative grant from the National Institutes of Health (Grant No. 62412).

References

- Aslanidis, C. & de Jong, P. J. (1990). *Nucleic Acids Res.* **18**, 6069–6074.
- Baker, N. A., Sept, D., Joseph, S., Holst, M. J. & McCammon, J. A. (2001). *Proc. Natl Acad. Sci. USA*, **98**, 10037–10041.
- Bateman, A., Coin, L., Durbin, R., Finn, R. D., Hollich, V., Griffiths-Jones, S., Khanna, A., Marshall, M., Moxon, S., Sonnhammer, E. L., Studholme, D. J., Yeats, C. & Eddy, S. R. (2004). *Nucleic Acids Res.* **32**, D138–D141.
- Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N. & Bourne, P. E. (2000). *Nucleic Acids Res.* **28**, 235–242. Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* **D50**, 760–763.
- Grosse-Kunstleve, R. W. & Adams, P. D. (2003). *Acta Cryst.* **D59**, 1966–1973.
- Holm, L. & Sander, C. (1999). *Nucleic Acids Res.* **27**, 244–247.
- Huai, Q., Liu, Y., Francis, S. H., Corbin, J. D. & Ke, H. (2004). *J. Biol. Chem.* **279**, 13095–130101.
- Jancarik, J. & Kim, S.-H. (1991). *J. Appl. Cryst.* **24**, 409–411.
- Jancarik, J., Pufan, R., Hong, C., Kim, S.-H. & Kim, R. (2004). *Acta Cryst.* **D60**, 1670–1673.
- Jones, T. A., Zou, J.-Y., Cowan, S. W. & Kjeldgaard, M. (1991). *Acta Cryst.* **A47**, 110–119.
- Kim, R., Sandler, S. J., Goldman, S., Yokota, H., Clark, A. J. & Kim, S.-H. (1998). *Biotechnol. Lett.* **20**, 207–210.
- Krissinel, E. & Henrick, K. (2005). *CompLife 2005*, edited by M. R. Berthold, R. Glen, K. Diederichs, O. Kohlbacher & I. Fischer, pp. 163–174. Berlin/Heidelberg: Springer-Verlag.
- Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997). *Acta Cryst.* **D53**, 240–255.
- Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
- Studier, F. W. (2005). *Protein Expr. Purif.* **41**, 207–234.
- Terwilliger, T. (2004). *J. Synchrotron Rad.* **11**, 49–52.
- Vagin, A. & Teplyakov, A. (1997). *J. Appl. Cryst.* **30**, 1022–1025.
- Yip, C. K., Kimbrough, T. G., Felize, H. B., Vuckovic, M., Thomas, N. A., Pfuetzner, R. A., Frey, E. A., Finlay, B. B., Miller, S. I. & Strynadka, N. C. J. (2005). *Nature (London)*, **435**, 702–707.