FEBS 29969 FEBS Letters 579 (2005) 5353–5360

1.2 Å crystal structure of the *S. pneumoniae* PhtA histidine triad domain a novel zinc binding fold

A. Riboldi-Tunnicliffe[a,b], N.W. Isaacs[b], T.J. Mitchell[a,*]

[a] *University of Glasgow, Division of Infection and Immunity, I.B.L.S. Joseph Black Building, Glasgow G12 8QQ, UK*
[b] *University of Glasgow, Department of Chemistry, Joseph Black Building, Glasgow G12 8QQ, UK*

Received 17 June 2005; revised 30 August 2005; accepted 31 August 2005

Available online 16 September 2005

Edited by Hans Eklund

Abstract The recently described pneumococcal histidine triad protein family has been shown to be highly conserved within the pneumococcus. As part of our structural genomics effort on proteins from *Streptococcus pneumoniae*, we have expressed, crystallised and solved the structure of PhtA-166–220 at 1.2 Å using remote SAD with zinc. The structure of PhtA-166–220 shows no similarity to any protein structure. The overall fold contains 3β-strands and a single short α-helix. The structure appears to contain a novel zinc binding motif. The remaining 4 histidine triad repeats from PhtA have been modelled based on the crystal structure of the PhtA histidine triad repeat 2. From this modelling work, we speculate that only three of the five histidine triad repeats contain the residues in the correct geometry to allow the binding of a zinc ion.

Keywords: Pneumococcal histidine triad protein; PhtA; PhtB; PhtD; PhtE; Zinc binding; SAD

1. Introduction

Structural genomics is flourishing owing to tremendous progress in genome sequencing as well as recent advances in computer and software technology and third generation synchrotron beam-lines for macromolecular crystallography. One of the goals of structural genomics is to map the entire folding space. This can be accomplished by solving the structures of a large number of carefully selected proteins (15–20 000) that show no significant sequence homology to each other, and are therefore likely to include the majority of protein folds [1]. It is anticipated that this effort will expand knowledge of protein structure and will facilitate solving the structures of other proteins. For many proteins, function has not yet been established. It is expected that structural genomics will be able to assign functions to proteins when assignment is not possible from sequence alignment alone [2–8]. This is especially important as thousands of newly identified open reading frames representing putative protein genes became available from genome-sequencing programs. Structural information may provide important functional clues. The selection of proteins for structure determination is key to the structural genomics approach [9]. In this work, we applied the following three criteria: (i) uniqueness of amino acid sequence, to increase chances of finding a new protein fold (ii) unknown function, to aid in assignment of function, (iii) origin from a pathogenic bacterium to provide a basis for future investigation of the protein as a potential drug target.

As part of a pilot structural genomics exercise, we have chosen a subset of proteins of bacterial origin which have an unknown function and/or structure, the initial subset were all proteins from the human pathogen *Streptococcus pneumoniae* as two complete genomes are available in the public domain for this organism. We have chosen to work with the Tigr4 strain (http://www.tigr.org).

The first proteins screened for crystallisation, are those that are specific to the pneumococcus. PhtA is part of a family of uncharacterised proteins that are unique to the pneumococcus and would represent a specific drug target for this organism [10,11]. The proteins in this family are approximately 800 amino acids in length and show very strong sequence conservation within the pneumococcus. PhtA (SP1175) is a member of a novel family of cell surface exposed pneumococcal proteins (Pht family, consisting of PhtA, PhtB, PhtD and PhtE) [10]. This family includes members that can induce antibodies capable of protecting mice against pneumococcal sepsis and death. These proteins were identified from the *S. pneumoniae* genome database and were selected based on their putative hydrophobic leader sequence, which are characteristic of proteins exported across the cytoplasmic membrane [12]. These novel pneumococcal antigens, either alone or in combination with capsular polysaccharides, could serve as effective vaccines against the most prevalent pneumococcal serotypes.

PhtA (SP1175) is a protein of 816 amino acids, and contains five histidine triad repeats which contain the HxxHxH motif. These proteins were therefore termed Pht (**p**neumococcal **h**istidine **t**riad) due to the characteristic repeats [10]. Experiments using surface exposed proteins in immunisation trials in mice [12], revealed that the N-terminus residues of PhtA (18–230) caused the production of antibodies which protected mice against death following intraperitoneal challenge with two strains of *S. pneumoniae* 6B and 6A. Immunization with PhtA resulted in antibodies that were cross reactive to proteins from several strains of *S. pneumoniae* [13] and further work by Hamel et al. [14] has shown protective antibodies are produced by PhtB (BHV3) in immunized mice.

The protein studied here, PhtA (SP1175) from *S. pneumoniae* belongs to a family/class of proteins that are unique to

*Corresponding author. Fax: +44 141 330 4888.
E-mail address: t.mitchell@bio.gla.ac.uk (T.J. Mitchell).

0014-5793/$30.00
doi:10.1016/j.febslet.2005.08.066

the pneumococcus since homologues are not found in other bacterial species. Neither its biological function or catalytic properties (if any) have been characterised. We have determined the high-resolution structure of a single histidine triad repeat domain of PhtA residues (166–220) by SAD analysis of the 'native' protein (shown to contain zinc ions). The protein fragment adopts a simple motif consisting of 3 beta strands and a single alpha helix (linked by short loops).

2. Materials and methods

2.1. Expression and purification

A PCR product containing the coding region for part of the N-terminal domain (residues 18–230) of the PhtA protein was cloned between the *Bam*HI and *Hin*dIII sites of the pQE-10 vector (Qiagen) in frame with an N-terminal His_6 tag. The recombinant protein was purified by nickel affinity chromatography, the protein concentrated using an Amicon pressure cell followed by gel filtration. The protein was concentrated for use in crystal trials to 7 mg/ml^{-1} in 200 mM NaCl, 25 mM Tris and 4 mM imidazole (pH 7.5). *S. pneumoniae* PhtA-166–220 domain crystallised using a mother liquor of 1.6 M ammonium sulfate and 0.1 M HEPES at pH 7.0 at 20 °C using hanging drop vapour diffusion. A single crystal was obtained from 4 μl drops containing equal volumes of protein and reservoir solution, over a period of 18 months [15]. The crystal belongs to the space group C2 with unit cell parameters $a = 62.18$ Å, $b = 35.89$ Å, $c = 72.54$ Å, $\beta = 90.01°$ with three molecules per asymmetric unit.

2.2. Data collection and processing

The crystal was flash cooled at 100 K using dried paraffin oil as the cryoprotectant [16]. Data were collected to 1.2 Å at the XRD-1 Beamline at Elettra using a MarCCD detector. A total of 372 images were collected using a 1° oscillation range per image at a wavelength of 1.0 Å. The data were processed and merged using MOSFLM [17] and SCALA [18] keeping the Bijvoet pairs separate during scaling. Intensities were converted to structure factor amplitudes using the CCP4 suit [19] program TRUNCATE [20].

2.3. Structure solution

The positions of three zinc ions were located using autoSHARP [21]. Position refinement and phase calculation were performed using SHARP [21]. This was followed by a further round to include density modification (solvent flipping) using SOLOMON [22] as implemented in SHARP. This provided a solution that gave excellent phases for model building using Arp/Warp [23]. The remainder of the model was built by hand using Quanta (Accelerys, San Diego, CA, USA) with refinement carried out using REFMAC5 [24]. The model was validated using PROCHECK [25] and WHATCHECK [26].

3. Results and discussion

3.1. Structure solution and zinc sites

Residues 18–220 of *S. pneumoniae* PhtA were expressed with a 6-his tag in *Escherichia coli*. Following affinity purification to immobilised metal ions, the pure protein samples were concentrated to 3 ml and the PhtA domain purified to homogeneity by size-exclusion chromatography. Fractions containing pure PhtA were pooled, concentrated and used in crystal trials. A diffraction quality crystal grew after 18 months. Initial phases were calculated based on the zinc ion sites, located by use of anomalous difference Patterson, to verify the presence of anomalous scatterers [19]. The positions of the Zn ions were determined using autoSHARP with data in the resolution range from 76.0 to 1.2 Å [21], three sites were found. Positional refinement and phase calculations were performed using SHARP [21]. Phases were further improved by solvent flipping using SOLOMON [22] as implemented in SHARP [21]. A free-atom model was built into the electron density calculated from the solvent flipped phases using ARP/wARP [23]. The resulting electron density maps were high quality and allowed auto-tracing of the amino-acid chain using ARP/wARP warpNtrace [23] version 6.0 integrated into CCP4i, which allowed the automatic building of 129 residues in 7 chains. Initial refinement was performed using REFMAC5 [24] from the CCP4 program suite [19]. Rounds of model building using (Quanta) (Accelerys, San Diego, CA, USA) and REFMAC5 [24] allowed all residues to be built into the electron density maps. Application of anisotropic refinement of B factors as implemented in REFMAC5 [24] followed by rounds of refinement using TLS [27] improved both the R factor and the R_{free}. Solvent molecules were then added manually in density peaks greater than 4.0 sigma. Data collection/refinement, phasing and final model statistics are shown in Table 1. Figures were produced with Molscript [28], Bobscript [29] and Raster3D [30]. The final structure contains three independent copies of the histidine triad motif, chains A and B contain 54 residues whilst chain C contains only 53 residues, there are also 3 zinc ions (one associated with each histidine triad motif) and a total of 195 water molecules.

As the PhtA protein shares no sequence homology to any protein of known structure, it was not expected to be possible to solve the structure by molecular replacement and as the crystal took a long time to grow (over 18 months) it made the growth of new crystals for heavy atom trials unlikely. As the protein was thought to contain several histidine triad repeats we expected that metal ions would have bound to the protein during the expression and subsequent purification, as

Table 1
Data collection and refinement statistics

Data collection	
Resolution (Å)	72.548–1.2 (1.26–1.2)
Number of frames (rotation per frame °)	360 (1)
Total observations	593 315
Unique reflections	49 793
Completeness (%) overall/highest shell	93.9 (93.9)
Mean I/overall/highest shell	8.1/2.6
R_{sym} (%) overall/highest shell	4.7/17.5
Multiplicity overall/highest shell	7.6/7.2
Phasing	
Wavelength (Å)	1.0
Number of zinc sites	3
FOM (centric/acentric)	0.21743/0.55428
Anomalous phasing power (Acentric)	3.15
Refinement	
Space group	C2
Unit cell parameters	$a = 62.18$ Å, $b = 35.89$ Å, $c = 72.54$ Å, $\beta = 90.01°$
Resolution limits	72.5–1.2 Å
R/R_{free}	11.2/13.5
No. atoms	1410
No. water molecules	195
No. zinc ions	3
Mean B value overall ($Å^2$)	8.828
R.m.s. deviations	
Bond length (Å)	0.16
Bond angle (°)	1.93

Values in parentheses are for the outer shell.

these repeats are expected to bind divalent metals, in particular zinc. It is speculated that the binding of zinc ions by the histidine triad motifs would confer the functional conformation on the proteins [10,31]. We decided to try a remote SAD (single wavelength anomalous diffraction) approach and collected the data at a wavelength of 1.0 Å, where a zinc atom would provide three anomalous electrons. We therefore scaled the data keeping the Bijvoet pairs separate. The resulting anomalous Patterson revealed the presence of three peaks which were subsequently refined as Zn ions.

The protein that was used in the crystallisation trials was the full-length construct (residues 18–230), as characterised by Gel Filtration chromatography, SDS–PAGE Gel electrophoresis and dynamic light scattering (data not shown). Calculation of the Matthews coefficient and solvent content [32] suggested that the asymmetric unit would contain a single molecule of approximately 200 amino-acid residues (Matthews coefficient 1.80 solvent content 31.11%). However, on elucidation of the structure it became clear that the asymmetric unit did not contain a single molecule but instead was formed from three repeating units, each of which had a single zinc ion bound. The protein had been degraded over the 18 months it took for the crystal to grow. The protein contained within the crystal consisted of residues 166–220 only (this includes only one of the histidine triad repeats). It is thought that a contaminant protein containing proteolytic activity (either in the 'purified' protein or as a contaminant on the plastic-ware used to setup the screens) had caused the breakdown of the protein over time, however, we cannot rule out the possibility that the protein itself (PhtA-18–220) has some proteolytic activity. Proteolytic activity has been observed for the homologous protein PhtB [14]. Trials of the potential proteolytic activity of this protein have been undertaken and subsequent batches of purified protein have been stored at 4 °C for a period of 2–3 months (crystals were grown at 20 °C). The initial purified protein ran on an SDS–PAGE gel as a single band (with no contaminating bands by Commassie blue staining); running the same sample stored at 4 °C after 3 months does reveal a ladder of breakdown products (data not shown). Although we had initially expected to have grown crystals of the N-terminal residues 18–230 of PhtA from *S. pneumoniae*, we have instead crystallised a fragment of this construct, which contains a single histidine triad repeat region. The final model containing three individual histidine triad repeats of approximately 54 residues 3 zinc ions and 192 water molecules was refined against all of the data (72.55–1.2 Å) to give a final R factor of 11.2% and R_{free} of 13.5% (Fig. 1).

4. Crystal structure of PhtA-166–220

The molecular model fitted well into the electron density (Fig. 2A and B) and was built from Gln-166 to Gly-220 for chain A, Gln-166 to Gly-220 for chain B and Gly-167 to Gly-220 for chain C. The overall structure of PhtA histidine triad repeat 2 (PhtA-166–220) is shown in Fig. 3. It is a single domain molecule of approximately 54 amino acids with an overall topology that consists of a single beta-sheet composed of 3 anti-parallel strands connected by short loops, with a large loop connecting the single alpha helix to the beta-sheet. The secondary structure elements are as follows: β1 residues 181–184; β2 residues 188–193; β3 residues 196–201 and α1 residues 207–220.

In an initial effort to gain clues about the possible biological function of SP1175 (PhtA 166–220), its structure was compared to the proteins of a known spatial structures in the Protein Data Bank (PDB) [33] using DALI [34]. The search was performed against the PDB database in December 2004. This process was attempted using the monomer, the dimer and the full content of the asymmetric unit (a trimer) as the search models. Searches using the monomer resulted in several comparable structures, showing the same overall fold were present in the PDB, however none of which share the zinc binding site.

The zinc ion is chelated (bound) by NE1 of His-194, NE2 of His-197, NE2 of His-199 and OD2 of Asp-173 at distances of 2.02, 2.03, 2.03 and 1.97 Å. The zinc coordinating atoms form a distorted tetrahedral coordination sphere for the zinc ion. The three histidine residues form the base of the triangle, with the zinc ion sitting slightly above the plane of these residues, the apex of the triangle is formed by the side chain of Asp-173. The distances between the side chain atoms and the zinc ion are shorter than expected, typical distances of protein to zinc ions are 2.2 Å between zinc and histidine and 2.1 Å for zinc and aspartic acid residues.

On viewing the three individual histidine triad proteins in the asymmetric unit we have chains A, B and C; the interaction between chain A and B is weak involving only 10% of the total surface area (402 Å^2) whilst the interaction of chains A and C involve a surface area that is almost double, 18.5% of the total surface area is involved in the interaction forming a tight interface. This dimerisation causes the beta sheet to be doubled in length from 3 anti-parallel strands to 6 anti-parallel strands. The third molecule in the asymmetric unit has its 'dimer' partner in the adjacent asymmetric unit (Fig. 4).

5. Molecular modelling of the 4 remaining histidine triad repeats

Using the crystal structure, we have constructed models of the remaining 4 histidine triad repeat regions from PhtA (Fig. 5). Whilst the presence of the HxxHxH motif is obvious for all 5 repeats, modelling has revealed that only 3 of the 5 repeats would be able to bind zinc in the same manner. Repeats 2, 3 and 4 show the zinc ion is held in place in a tetrahedral coordination, by three histidines from the above motif, and an oxygen from the side chain of Aspartic acid 173. In histidine triad repeat 1, this aspartic acid residue has been replaced by an arginine residue, whilst in histidine triad repeat 5 the aspartic acid has been replaced with a valine residue. Attempts to model neighbouring residues to better coordinate the zinc have failed. It is possible that a complete structural rearrangement occurs in this region and that the tetrahedral geometry is maintained by a solvent molecule or the side chain atoms from a residue not located adjacent to the histidine triad repeat site.

6. Proteins identified by DALI

Using the web-based server FAST Alignment and Search Tool (http://biowulf.bu.edu/FAST) it was possible to overlay (based on secondary structure elements) the structures which appear to have a similar fold as they were retrieved after a DALI search

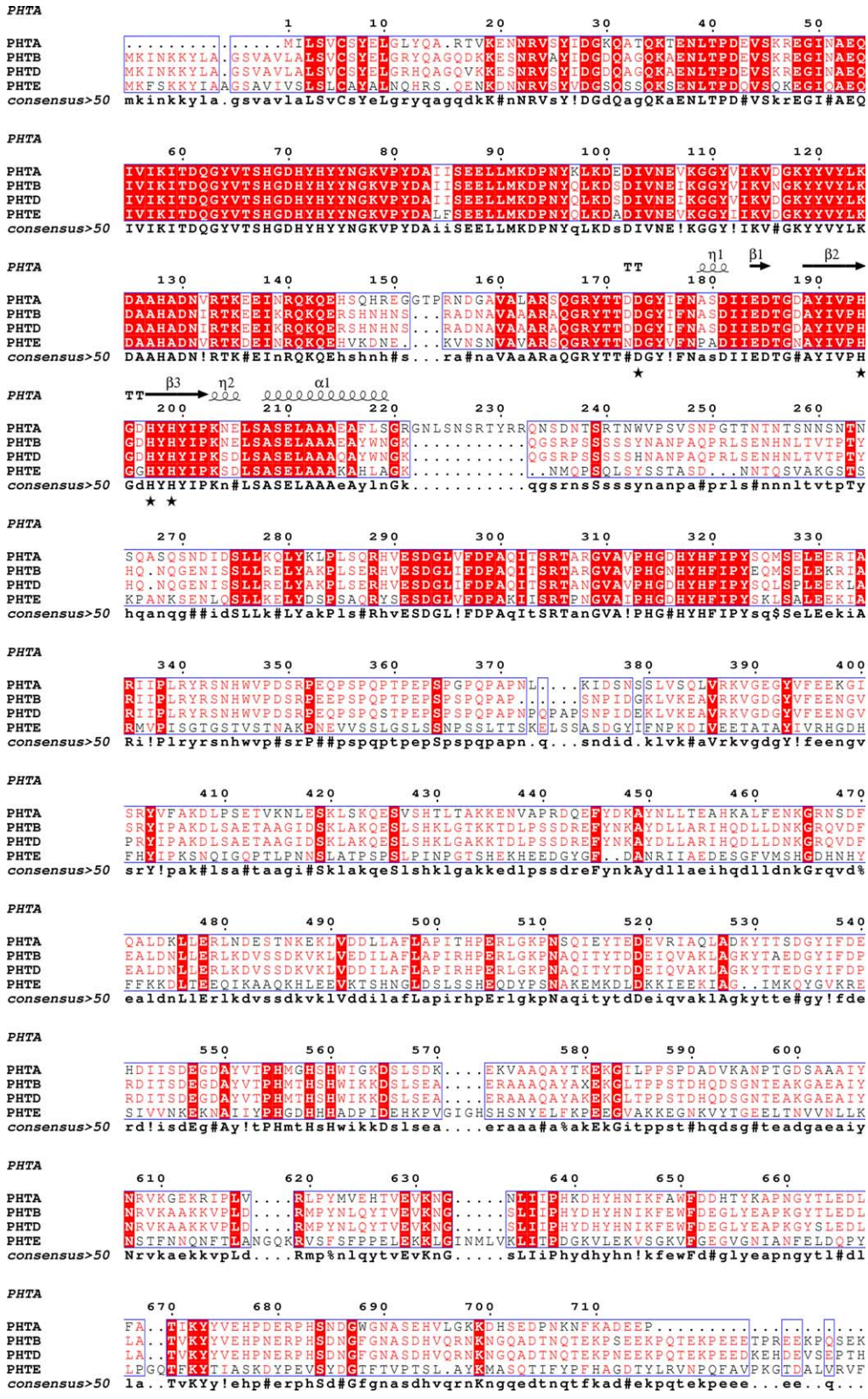

Fig. 1. Sequence alignment of the 4 Pht proteins from *Streptococcus pneumoniae*. Identical residues are in white on a red background.

and the PhtA fragment (residues 116–220). Structural similarities between the PhtA fragment and structures 1AZP (DNA binding protein); 1TUX (neutrophil activating protein) and 3IL8 (interleukin 8) can be seen in Fig. 6A, B and C using only a C-alpha trace. Although the overall fold of these 4 protein structures is remarkably similar, the level of identity at the se-

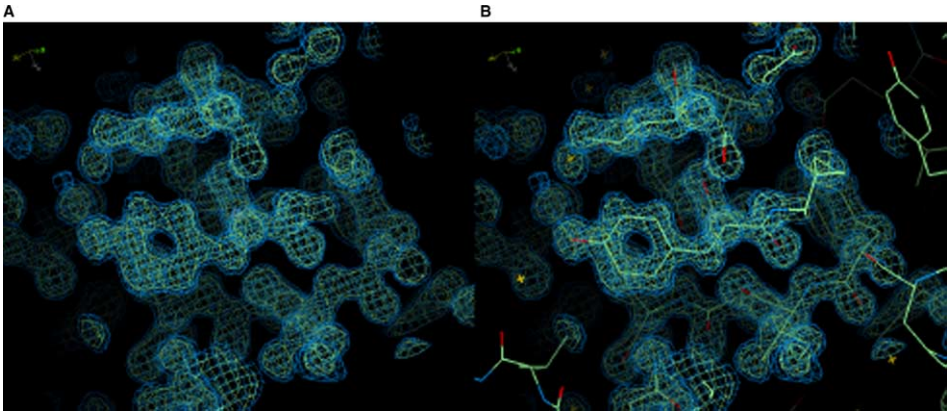

Fig. 2. (A) The experimental electron map after SHARP/SOLOMON. (B) The same area after model building with ARP/wARP.

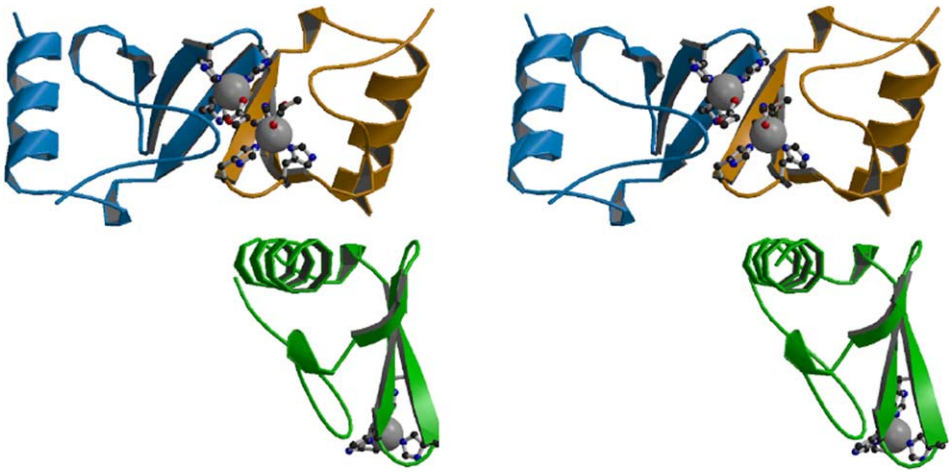

Fig. 3. Stereo view of the asymmetric unit containing three copes of PhtA-166–220. Zinc ions are shown in grey. Chain A in Gold, chain B in Green and chain C in Blue. Residues involved in co-ordination of the zinc ion are shown as ball and stick.

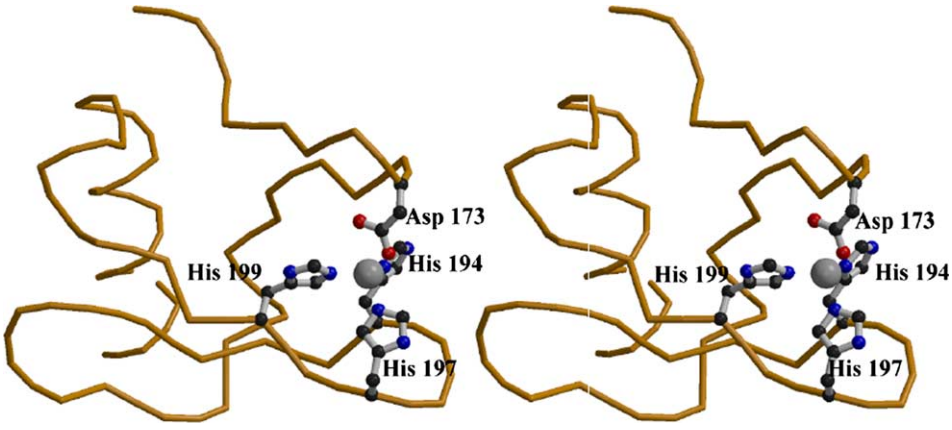

Fig. 4. Stereo view of the zinc binding site of PhtA-166–220. The residues involved in co-ordination of the zinc ion are shown as ball and stick.

quence level is very low (Fig. 6D). Sequence alignments of PhtA-166–220 and each of the three proteins were undertaken (data not shown) which revealed a higher level of sequence identity between PhtA-166–220 and 1TUX and 3IL8, whilst there was virtually no identity between PhtA-166–220 and 1AZP, despite having the best fit structurally.

The folds of 1AZP and PhtA-166–220 are very similar with and average rmsd of 3.21 Å for 51 C-Alpha atoms. The loop containing the histidine repeat in PhtA is elongated in 1AZP, 1AZP does not contain the residues to enable zinc binding. The residues of 1AZP involved in DNA binding are not conserved in PhtA-166–220. For both 1TUX and 3IL8, the

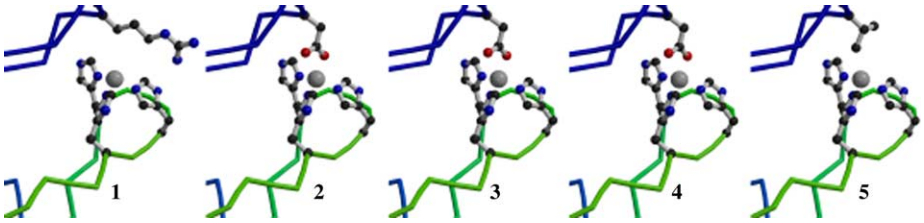

Fig. 5. Homology models of the 5 histidine triad domains found in PhtA. Only the 2nd (crystal structure), 3rd and 4th repeats contain the correct amino acids to bind a zinc ion. In the first repeat the aspartic acid is replaced with an arginine, whilst in the 5th repeat the aspartic acid is replaced with a valine.

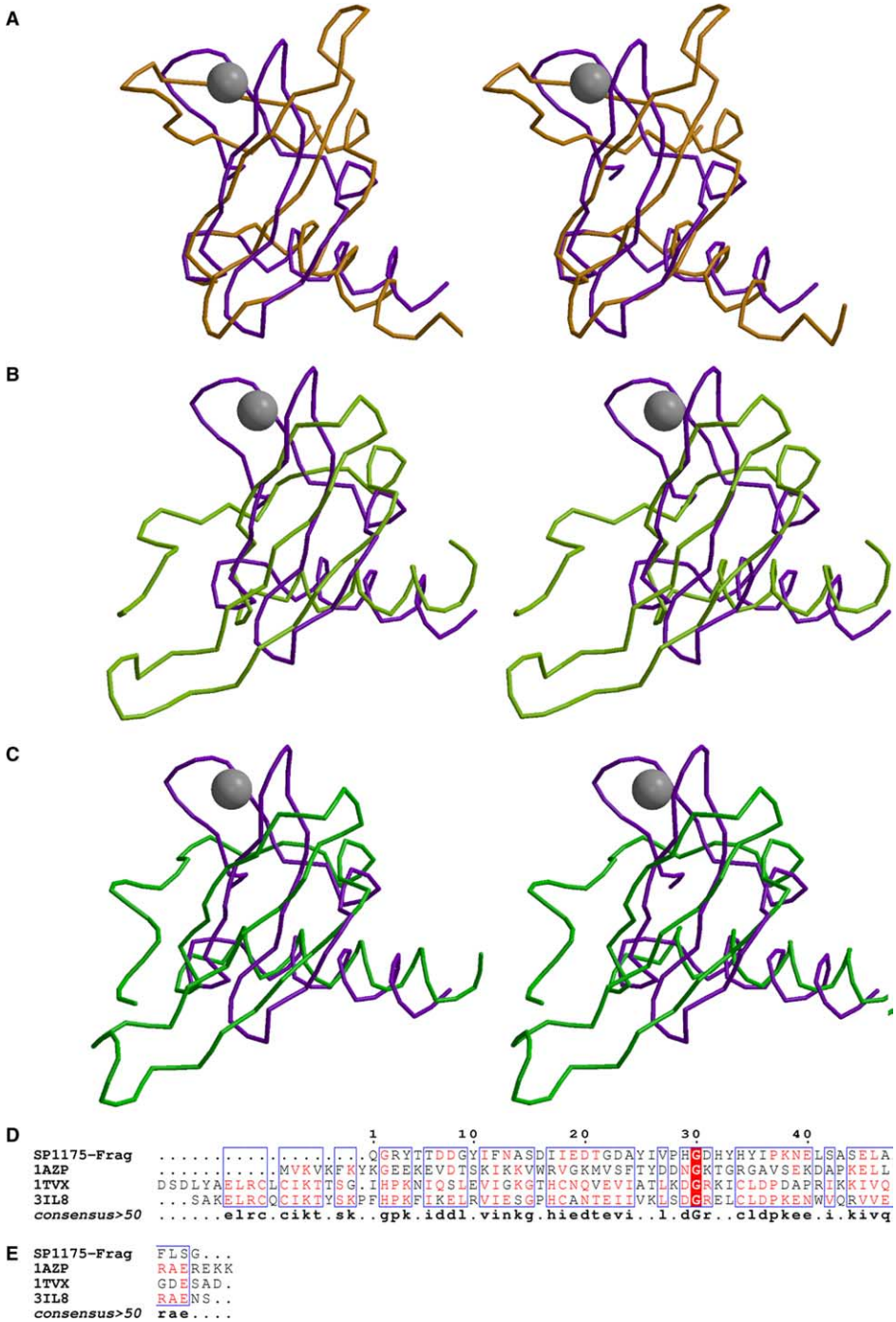

Fig. 6. Stereo views of PhtA-166–220 (purple) and structural homologues. (A) DNA binding protein 1AZP (yellow), (B) neutrophil activating protein 1TUX (light green) and (C) interleukin 8 3IL8 (green). Whilst 1AZP has the best fit, it has the lowest sequence identity (D).

average rmsd are 2.99 and 2.85 Å for 37 and 33 C-Alpha atoms, respectively. Attempts were also made to overlay the structures of other zinc binding proteins belonging to the carbonic anhydrase family (1CAM and 1CSS). This was attempted using both secondary structure elements and residues involved in zinc binding, neither of these resulted in satisfactory overlays.

7. Conclusion

Although on collecting data and solving the structure, we expected to see a single polypeptide of 212 amino acids we have instead solved the crystal structure of a novel zinc binding domain (a degraded version of the N-terminal of PhtA which contains residues 166–220). These residues represent a novel manner for the sequestration of zinc ions contained within a single histidine triad repeat motif (HxxHxH). This motif must be incredibly stable as it has not been digested/degraded along with the rest of the protein. The residues 166–220 contain the final histidine triad motif (of the three present in the original construct). Protein functional analysis suggested that in situations where the bacteria face a zinc-restricted environment, the expression of the Pht proteins would be induced and result in *Streptococcus* adhesion and invasion [31]. Searches using a single histidine triad repeat, the 'dimer' of the histidine triad repeat and the total content of the asymmetric unit have been used as search models gave a number of solutions in DALI [34] in which the overall fold (3 strands and an alpha helix) is preserved.

Although the structure of PhtA-166–220 shows an overall fold similarity to DNA binding proteins, interleukin-8 and neutrophil-activating proteins, further experimental work is necessary to show if PhtA (PhtA-166–220) has any similar activity or not.

A BLAST search of the Swiss-Prot protein database using residues 166–220 revealed the only similarities to be proteins of *Streptococcus* origin. The finding that this sequence is not found within any other organism is not surprising as a search using the complete sequence only produces hits with Pht proteins from other streptococci.

References

[1] Vitkup, D., Melamud, E., Moult, J. and Sander, C. (2001) Completeness in structural genomics. Nat. Struct. Biol. 8, 559–566.

[2] Zarembinski, T.I., Hung, L.W., Mueller-Dieckmann, H.J., Kim, K.K., Yokota, H., Kim, R. and Kim, S.H. (1998) Structure-based assignment of the biochemical function of a hypothetical protein: a test case of structural genomics. Proc. Natl. Acad. Sci. USA 95, 15189–15193.

[3] Hwang, K.Y., Chung, J.H., Kim, S.H., Han, Y.S. and Cho, Y. (1999) Structure-based identification of a novel NTPase from *Methanococcus jannaschii*. Nat. Struct. Biol. 6, 691–696.

[4] Teplova, M., Tereshko, V., Sanishvili, R., Joachimiak, A., Bushueva, T., Anderson, W.F. and Egli, M. (2000) The structure of the yrdC gene product from *Escherichia coli* reveals a new fold and suggests a role in RNA binding. Prot. Sci. 9, 2557–2566.

[5] Schulze-Gahmen, U., Pelaschier, J., Yokota, H., Kim, R. and Kim, S.H. (2003) Crystal structure of a hypothetical protein, TM841 of *Thermotoga maritima*, reveals its function as a fatty acid-binding protein. Proteins 50, 526–530.

[6] Christendat, D., Yee, A., Dharamsi, A., Kluger, Y., Gerstein, M., Arrowsmith, C.H. and Edwards, A.M. (2000) Structural proteomics: prospects for high throughput sample preparation. Prog. Biophys. Mol. Biol. 73, 339–345.

[7] Murzin, A.G. (1999) Structure classification-based assessment of CASP3 predictions for the fold recognition targets. Proteins (Suppl. 3), 88–103.

[8] Oliver, S.G. (1996) From DNA sequence to biological function. Nature 379, 597–600.

[9] Linial, M. and Yona, G. (2000) Methodologies for target selection in structural genomics. Prog. Biophys. Mol. Biol. 73, 297–320.

[10] Adamou, J.E. et al. (2001) Identification and characterization of a novel family of pneumococcal proteins that are protective against sepsis. Infect. Immun. 69, 949–958.

[11] Rigden, D.J., Galperin, M.Y. and Jedrzejas, M.J. (2003) Analysis of structure and function of putative surface-exposed proteins encoded in the *Streptococcus pneumoniae* genome: a bioinformatics-based approach to vaccine and drug design. Crit. Rev. Biochem. Mol. Biol. 38, 143–168.

[12] Wizemann, T.M. et al. (2001) Use of a whole genome approach to identify vaccine molecules affording protection against *Streptococcus pneumoniae* infection. Infect. Immun. 69, 1593–1598.

[13] Zhang, Y., Masi, A.W., Barniak, V., Mountzouros, K., Hostetter, M.K. and Green, B.A. (2001) Recombinant PhpA protein, a unique histidine motif-containing protein from *Streptococcus pneumoniae*, protects mice against intranasal pneumococcal challenge. Infect. Immun. 69, 3827–3836.

[14] Hamel, J., Charland, N., Pineau, I., Ouellet, C., Rioux, S., Martin, D. and Brodeur, B.R. (2004) Prevention of pneumococcal disease in mice immunized with conserved surface-accessible proteins. Infect. Immun. 72, 2659–2670.

[15] Riboldi-Tunnicliffe, A., Bent, C.J., Isaacs, N.W. and Mitchell, T.J. (2004) Expression, purification and X-ray characterization of residues 18–230 from the pneumococcal histidine triad protein A (PhtA) from *Streptococcus pneumoniae*. Acta Crystallogr. D, Biol. Crystallogr. 60, 926–928.

[16] Riboldi-Tunnicliffe, A. and Hilgenfeld, R. (1999) Cryocrystallography with oil – an old idea revived. J. Appl. Cryst. 32, 1003–1005.

[17] A.G.M. Leslie, MOSFLM, Joint CCP4 + ESF-EAMCB Newsletter on Protein Crystallography 26 (1992).

[18] P.R. Evans (1993) in: Proceedings of CCP4 Study Weekend on Data Collection & Processing, pp. 114–122.

[19] C.C.P.N. CCP4 (1994) The CCP4 suite: programs for protein crystallography. Acta Crystallogr., D 50, 760–763.

[20] French, S. and Wilson, K. (1978) On the treatment of negative intensity observations. Acta Crystallogr., A 34, 517–525.

[21] LaFortelle, E.d. and Bricogne, G. (1997) (Sweet, J.R.M. and Carter, C.W., Eds.), Methods in Enzymology, vol. 276, Academic Press, New York.

[22] Abrahams, J.P. and Leslie, A.G.W. (1996) Methods used in the structure determination of bovine mitochondrial F1 ATPase. Acta Crystallogr., D 52, 30–42.

[23] Perrakis, A., Morris, R. and Lamzin, V.S. (1999) Automated protein model building combined with iterative structure refinement. Nat. Struct. Biol. 6, 458–463.

[24] Murshudov, G.N., Vagin, A.A., Lebedev, A., Wilson, K.S. and Dodson, E.J. (1999) Efficient anisotropic refinement of macromolecular structures using FFT. Acta Crystallogr. D, Biol. Crystallogr. 55 (Pt 1), 247–255.

[25] Laskowski, R.A., MacArthur, M.W., Moss, D.S. and Thornton, J.M. (1993) PROCHECK: a program to check the stereochemical quality of protein structures. J. Appl. Cryst. 26, 283–291.

[26] Vriend, G. (1990) WHAT IF: a molecular modeling and drug design program. J. Mol. Graph. 8, 52–56, 29.

[27] Winn, M.D., Isupov, M.N. and Murshudov, G.N. (2001) Use of TLS parameters to model anisotropic displacements in macromolecular refinement. Acta Crystallogr. D, Biol. Crystallogr. 57, 122–133.

[28] Kraulis, P.J. (1991) MOLSCRIPT: A programk to produce both detailed and schematic plots of protein structures. J. Appl. Cryst. 24, 946–950.

[29] Esnouf, R.M. (1999) Further additions to MolScript version 1.4, including reading and contouring of electron-density maps. Acta Crystallogr., D 55, 938–940.

[30] Merrit, E.A. and Bacon, D.J. (1997) (Sweet, J.R.M. and Carter, C.W., Eds.), Methods in Enzymology, vol. 277, Academic Press, New York.

[31] Panina, E.M., Mironov, A.A. and Gelfand, M.S. (2003) Comparative genomics of bacterial zinc regulons: enhanced ion transport, pathogenesis, and rearrangement of ribosomal proteins. Proc. Natl. Acad. Sci. USA 100, 9912–9917.
[32] Matthews, B.W. (1968) Solvent content of protein crystals. J. Mol. Biol. 33, 491–497.
[33] Berman, H.M., Bhat, T.N., Bourne, P.E., Feng, Z., Gilliland, G., Weissig, H. and Westbrook, J. (2000) The Protein Data Bank and the challenge of structural genomics. Nat. Struct. Biol. 7 (Suppl), 957–959.
[34] Holm, L. and Sander, C. (1993) Protein structure comparison by alignment of distance matrices. J. Mol. Biol. 233, 123–138.