Journal of Structural Biology xxx (xxxx) xxx–xxx

Contents lists available at ScienceDirect

Journal of Structural Biology

journal homepage: www.elsevier.com/locate/yjsbi

Structural and functional characterization of TgpA, a critical protein for the viability of *Pseudomonas aeruginosa*

Mónica Uruburu[a,1], Eloise Mastrangelo[a,b,1], Martino Bolognesi[a,c], Silvia Ferrara[a], Giovanni Bertoni[a,*], Mario Milani[a,b,*]

[a] Dipartimento di Bioscienze, Università di Milano, Via Celoria 26, I-20133 Milano, Italy
[b] CNR-IBF, Istituto di Biofisica, Via Celoria 26, I-20133 Milano, Italy
[c] Centro di Ricerca Pediatrica R.E. Invernizzi, Università di Milano, Via Celoria 26, I-20133 Milano, Italy

ARTICLE INFO

Keywords:
Periplasmic proteins
X-ray crystallography
Antibacterials
Point mutations
Heterologous expression
Cystic fibrosis
Peptidoglycans

ABSTRACT

Pseudomonas aeruginosa is an opportunistic pathogen associated with severe diseases, such as cystic fibrosis. During an extensive search for novel essential genes, we identified *tgpA* (*locus* PA2873) in *P. aeruginosa* PAO1, as a gene playing a critical role in bacterial viability. TgpA, the translated protein, is an internal membrane protein with a periplasmic soluble domain, predicted to be endowed with a transglutaminase-like fold, hosting the Cys404, His448, and Asp464 triad. We report here that Cys404 mutation hampers the essential role of TgpA in granting *P. aeruginosa* viability. Moreover, we present the crystal structure of the TgpA periplasmic domain at 1.6 Å resolution as a first step towards structure–activity analysis of a new potential target for the discovery of antibacterial compounds.

1. Introduction

Infections caused by resistant gram-negative bacteria are becoming increasingly prevalent, and a serious threat to public health worldwide, being associated with high morbidity and mortality rates. Colonization with resistant gram-negative bacteria is common among residents in long-term care facilities, particularly for patients with an indwelling device.

Pseudomonas aeruginosa is a gram-negative opportunistic pathogen responsible for chronic lung colonization in cystic fibrosis patients and acute infections in hospitals (Driscoll et al., 2007). *P. aeruginosa* infections are always difficult to treat due to the natural resistance of this bacterium (Obritsch et al., 2005), consequently a therapy based on two or three antimicrobial agents is typically adopted (Lodise et al., 2007; Szaff et al., 1983; Valerius et al., 1991).

Unravelling novel essential genes or pathways that have not yet been targeted by clinical antibiotics can foster the development of new effective antibacterials able to overcome resistance. We previously discovered that the gene coding for the protein TgpA is essential for *P. aeruginosa* viability (Milani et al., 2012). TgpA is a medium size inner membrane protein (668 amino acids, 74.8 kDa) composed of an N-terminal domain with 5 transmembrane (TM) helices (residues 11–321 belonging to the DUF3488 superfamily), a periplasmic domain (residues 180–548), and an additional TM helix (residues 549–567), terminating into a small C-terminal cytoplasmic domain (residues 582–638 belonging to the DUF4129 superfamily). The periplasmic portion of TgpA was predicted to host an eukaryotic-like transglutaminase domain (TG-like domain) (Makarova et al., 1999) of unknown function, characterized by the presence of a Cys, His, Asp triad, proposed to play a catalytic role. In this work, which represents a continuation of our published data (Milani et al., 2012), we validate the available information on the enzyme function both through TgpA point-mutation and through a 3D structure-based analysis. In particular, we prove that the triad's Cys404 residue (Pedersen et al., 1994; Kobayashi et al., 1998) is required for the essential function of TgpA relative to *P. aeruginosa* viability. Furthermore, we present the crystal structure of the TgpA periplasmic domain at 1.6 Å resolution, solved using the SAD method. We show that the periplasmic domain is composed of two subdomains; one hosts the "catalytic triad" of the TG-like domain, while the other is structurally linked to carbohydrate binding domains, and may play a role in substrate binding.

Our results establish *P. aeruginosa* TgpA as a target suitable for development of novel antimicrobial compounds, and provide the structural bases for such discovery activities.

* Corresponding authors CNR-IBF, Istituto di Biofisica, Via Celoria 26, I-20133 Milano, Italy (M. Milani).
E-mail addresses: giovanni.bertoni@unimi.it (G. Bertoni), mario.milani@unimi.it (M. Milani).
[1] These authors have contributed equally to the studies presented.

https://doi.org/10.1016/j.jsb.2018.12.004
Received 29 August 2018; Received in revised form 6 December 2018; Accepted 7 December 2018

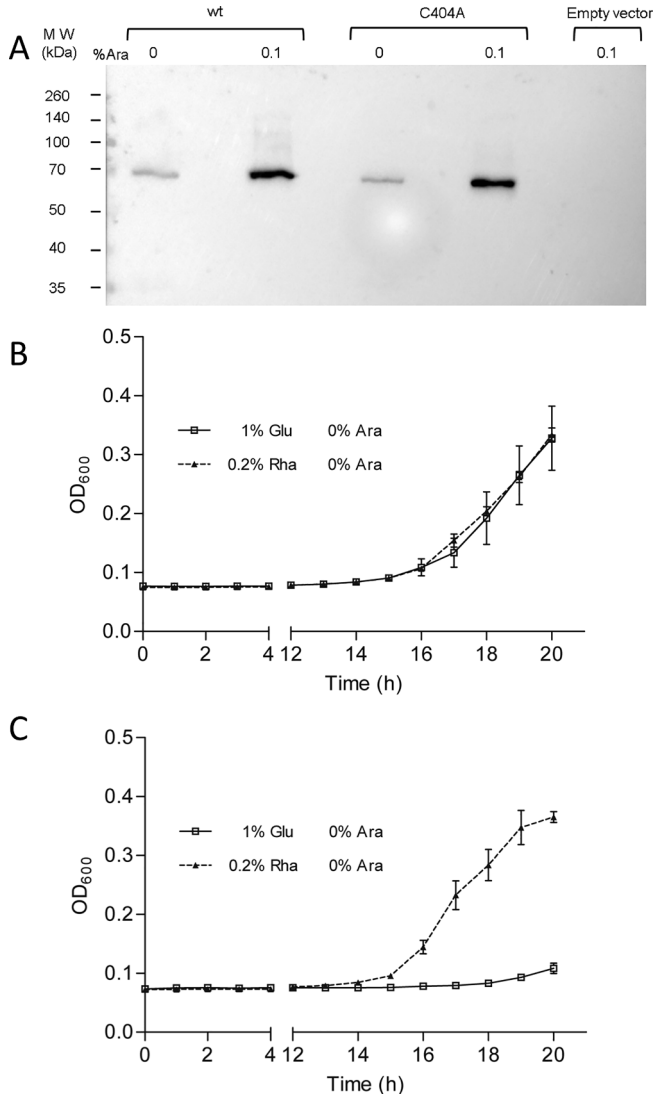

Fig. 1. Ectopic expression of TgpA wt and C404A. A. Western Blot analysis (His-tag antibody) of the ectopic expression of the TgpA wt and C404A full-length proteins in PAO1 with/without arabinose. 10 µl of total membrane fraction were loaded onto 12% SDS-PAGE; 0.1% Ara samples are diluted 1:10 with respect to the samples without arabinose. B. PAO1 P_{rhaB}::*tgpA* strains carrying pHERD28T-TgpA wt or C. pHERD28T-TgpA-C404A, were grown in micro-cultures of 200 µl with the indicated concentrations of rhamnose and glucose. Cultures growth was monitored in real-time by absorbance measurement at 600 nm (OD_{600}).

2. Results

2.1. Cys404 in the periplasmic TG-like domain is critical for TgpA cell activity

We previously generated a conditional mutant of *P. aeruginosa*, PAO1 P_{rhaB}::*tgpA*, where genomic TgpA expression is regulated by the P_{rhaB} promoter (Milani et al., 2012). PAO1 P_{rhaB}::*tgpA* requires rhamnose (0.2%) to induce endogenous TgpA expression, and is unable to grow in the presence of glucose due to P_{rhaB} down-regulation. To assess whether bacterial viability could be restored by TgpA exogenous expression, we transformed PAO1 P_{rhaB}::*tgpA* with a plasmid controlled by the P_{BAD} promoter, inducible with arabinose (Ara), expressing either TgpA wt or the C404A mutant.

In the absence of Ara, the P_{BAD} promoter displays a basal activity resulting in ectopic TgpA expression in PAO1 (Fig. 1A; Table 1).

Table 1
Relative mRNA levels of ectopically expressed TgpA wt and C404A mutant.

Strain	OD_{600}	Relative expression [a]	
		0% Ara	0.1% Ara
PAO1-pHERD28T	0.8	1	1
PAO1-pHERD28T-TgpA wt	0.8	16	544
PAO1-pHERD28T-C404A	0.8	11	363

[a] The calculation of the relative expression of the *tgpA* gene from plasmid pHERD28T versus the empty vector was performed as described by the $2^{-\Delta\Delta CT}$ method (Livak and Schmittgen, 2001), normalizing first mRNA amounts to 16S ribosome RNA (ΔCT) and relating the ΔCT in the TgpA wt and C404A mutant ectopically expressed to the chromosomal copy ($\Delta\Delta CT$) (PAO1-pHERD28T Empty vector).

Interestingly, even such basal amount of ectopic wt TgpA, is able to compensate the repression of the chromosomal protein, and PAO1 P_{rhaB}::*tgpA* is able to grow with 1% glucose as well as in the permissive condition of 0.2% rhamnose (Fig. 1B). On the contrary, the ectopic expression of the C404A mutant is not able to counterbalance the repression of the endogenous protein, demonstrating that Cys404 is indispensable for the essential function of TgpA (Fig. 1C). Such result points to a crucial catalytic role of the protein likely involving the catalytic triad Cys404, His448, and Asp464.

2.2. Crystal structure of the periplasmic domain of TgpA

The recombinant putative TG-like periplasmic domain of TgpA (residues 180–544 of the full length protein) was expressed in *E. coli* BL21 CodonPlus (DE3) as a N-terminal His-tagged protein. After Ni-IMAC and size-exclusion chromatography purification steps, SDS page followed by mass spectrometry analysis showed that proteolytic events had removed about 45 amino acids at the C-terminal end. Crystallization of this construct (TG-domain) and structure solution, achieved through the SAD method (based on Hg derivative), produced interpretable electron density (at 1.6 Å resolution; Table 2) for residues Val195 through Leu482, *i.e.* 17 residues upstream of the most probable proteolytic site at Tyr499.

The TG-domain is folded in a tripartite globular structure (with average dimensions 50 × 45 × 30 $Å^3$), resembling a heart built by two sub-domains (*right and left ventricles*) linked by a central antiparallel β-sheet composed of four strands (*the septum*) (Fig. 2A). The *right ventricle* subdomain (mostly formed by the N-terminal portion of the TG-domain) is primarily composed of β strands, with a major solvent exposed antiparallel β sheet (6 strands). On the contrary, the *left ventricle* subdomain that, together with the *septum,* contains the conserved "catalytic triad", is mainly composed of α-helices. Interestingly, the N-and C-terminal ends of the TG-domain are located on the same side of the protein, in agreement with its internal membrane anchoring.

The TgpA residues conserved within the gammaproteobacteria class (Williams et al., 2010) (excluding pseudomonadales order) can be grouped in three clusters (Fig. 2C): residues around the active site likely involved in enzymatic activity (20 amino acids, red sticks), residues with probable structural function (7 amino acids, blue sticks), and isolated residues (3 amino acids, orange sticks).

A structure based homology search, over the whole domain and through the DALI server (Holm and Laakso, 2016) identified two putative cysteine proteases (pdb-id 3ISR, Z = 9.7, id. 18%; and pdb-id 3KD4, Z = 8.7, id. 17%) from gram-negative bacteria, and the transglutaminase coagulation factor XIII (pdb-id 1L9M (Ahvazi et al., 2002), Z = 6.2, id. 21%, Fig. S1 for structural comparison). A new search for similar folds of the C-terminal sub-domain, using the CATH server (http://www.cathdb.info) (Sillitoe et al., 2015), identified the pantoate-β-alanine ligase functional family (Pantothenate synthetase 2, CATH-ID 3.30.1300.10) and the β polymerase domain 2 (class CATH-ID

Table 2
X-ray data-collection and refinement statistics.

Data collection	*TgpA Hg derivative*	*TgpA native*
Beam line & wavelength (Å)	ESRF BM14U 1.0078 Å	ESRF ID30a-3 0.9677
Space group	$P4_32_12$	$P4_32_12$
Unit-cell parameters (Å)	a = b = 72.5; c = 157.5	a = b = 71.6; c = 164.3
Molecules in a.u.	1	1
Resolution (Å)	48.7–1.8 (1.85–1.80)	43.5–1.6 (1.64–1.60)
Mosaicity (°)	0.2	0.2
Unique reflections	74,217 (5,444)[a]	57,370 (4,156)[b]
Completeness (%)	99.9 (99.1)	99.9 (100)
Redundancy	7.1 (3.8)	7.5 (7.8)
Rmeas[†] (%)	9.1 (127.5)	6.4 (146.8)
CC(1/2) (%)	99.9 (42.5)	99.9 (73.9)
Rpim (%)	2.5 (44.6)	3.2 (68.2)
Anomal. completeness (%)	99.9 (98.8)	–
Average I/σ (I)	15.5 (1.1)	17.6 (1.7)
Wilson B-factor (Å^2)	31.08	31.23
Final model		
R factor[‡]/Rfree[§] (%)	15.6/18.7 (31.5/30.2)[a]	16.0/19.4 (30.4/30.8)[b]
r.m.s. bonds (Å)	0.015	0.014
r.m.s. angles (°)	1.58	1.81
n. of waters	332	298
Average protein B fac. (Å^2)	27.4	28.8
Average solvent B fac. (Å^2)	42.0	43.1
Ligand B fac. (Å^2)	42.0	–
Residues in most favored regions (%)	96.9%	96.9%
Residues in additionally allowed regions (%)	2.7%	2.7%
Ramachandran outlier	Glu457	Glu457
PDB-ID	**6G4H**	**6G49**

Values in parentheses are for the highest resolution shell:
[a] (1.85–1.80).
[b] (1.64–1.60).
[†] $R_{meas} = (\Sigma\ (n/(n-1)\ \Sigma\ |I - (I)|)/\Sigma\ I \times 100$, where I is intensity of a reflection and (I) is its average intensity.
[‡] $R_{factor} = \Sigma\ |F_o - F_c|/\Sigma\ |F_o| \times 100$.
[§] R_{free} is calculated on 5% randomly selected reflections, for cross-validation.

3.30.460.10), whose members mainly cluster in the phosphotransferase enzyme family EC 2.7. The identification of a potential phosphate binding domain is remarkable, considering that we found, in the derivative crystal, a phosphate ion bound close to the active site (see next section).

Limiting the DALI structural search to the N-terminal portion of the TG-domain (residues Val195 to Ala323) highlighted structural similarity with different carbohydrate binding domains, such as the chitin binding domain (chitinase, pdb-id 5DHE (Hanazono et al., 2016), Z = 2.8, 13% identity), the xylan binding domain (xylanase, pdb-id 1XBD (Simpson et al., 1999), Z = 2.7, 7% identity), and GH52 β-D-xylosidase (pdb-id 4RHH, Z = 3.4, 7% identity), pointing to a possible involvement of this region in binding to the cell wall peptidoglycan. In particular, the chitinase domains (ChBD2 and ChBD3) belong to the carbohydrate binding domain family-II type (IPR001919; CBM2) that is further classified into two subfamilies according to substrate specificities: CBM2a that binds cellulose, and CBM2b interacting specifically with xylan. As for other CBM domains, CBM2 is a β-barrel domain hosting a planar surface able to bind its ligand(s) via a hydrophobic strip of aromatic residues.

2.2.1. Active site

As stated in the previous section, we demonstrated that Cys404 is essential for *P. aeruginosa* viability and therefore for TgpA catalytic function. Cys404 is hosted in the C-terminal portion of the periplasmic domain, forming a typical "catalytic triad" with His448 and Asp464. While His and Asp side chain positions are restrained by a mutual H-bond (2.8 Å), Cys404 side chain is H-bonded with main chain nitrogen of Ala449 (3.4 Å). The environment around Cys404 is characterized by the presence of both hydrophobic and polar residues, *e.g.* Ile207, Tyr380, Phe403, Arg473, and Thr466 (Fig. 2D). Around His448 a network of H-bonds is found involving, besides Asp464, residues Thr466, Arg473 and Tyr380 bridged by two water molecules and linked to the main chains of residues Phe445 and Ala447. In this respect, analysis of the geometry of the "active site" residues Tyr380, Cys404 and His448 using the LabelHash server (Moll et al., 2011) identified structural homology with the active site of a Cys endopeptidase acting on muropeptides (pdb-id 2HBW (Xu et al., 2009)).

The active Cys404 is buried inside the protein domain at about 10 Å from the external surface, but it may become fully accessible when the right substrate is present. The access to the "catalytic triad" cavity is blocked by a small 3_{10}-helix at the C-ter end (residues 478–480; Fig. 2A) in the crystal structure. A rigid body shift of this short helix, or its conformational flexibility, would make the putative active site cavity accessible to substrates. In fact, in the Hg derivative used for SAD phasing structure (where the mercurial compound is bound to Cys404), the C-ter end of the protein appears more mobile and the last C-ter amino acid visible in the electron density is Gln476 (Leu482 in the native structure; Fig. 2B, grey cartoons), yielding a wider entrance path to the "active site". Moreover, in order to accommodate the ethyl-mercury moiety, Phe403 side chain rotates by ~135° around its χ1 angle, widening the "active site" cavity (Fig. 2B). Interestingly, besides the ethyl-mercury moiety, a phosphate ion is hosted within the cavity (Fig. 2B, red sticks), H-bonded to His268 and Tyr200, with electrostatic compensation by Arg473 that is in a different conformation with respect to the native structure (Fig. 2B, grey sticks).

A smaller cavity, lined by N-terminal helix amino acids 205–212, is present in the native structure, where a Cl^- ion (present in the protein buffer) was modelled at 4.9 Å from Cys404 Sγ atom. The putative Cl^- ion interacts with the main chain N-atoms of Gly205, Ile207, Ala208 and Ala405, and to the side chain of His406, partially occupying the place taken by Phe403 in the derivative structure (Fig. 2B). To summarize, we propose a gating mechanism regulating substrate accessibility to the "active site" that would involve: *i*) C-terminal helix, *ii*) N-ter helix (residues 206–211), and *iii*) Phe403 double conformation (Fig. 2B).

2.3. TgpA is able to bind peptidoglycans

As discussed above, structural analysis of the N-terminal portion of the TG-domain (Val195 – Ala323) suggests structural homology with different carbohydrate binding domains. Since the major component of the bacterial cell wall peptidoglycan (PGN) is a complex polymer built by glycan chains cross-linked through peptide stems and isopeptidic bonds, we investigated whether the TgpA TG-domain is able to bind to PGN. To this aim, we performed PGN pull-down assays, where proteins binding to insoluble PGN fragments co-precipitate with them after centrifugation.

Pull-down experiments (Fig. 3A) show that the amount of insoluble TG-domain increases with the concentrations of *P. aeruginosa* PGN. On the contrary, the helicase from Dengue virus, used as a negative control, shows no detectable PGN binding (Fig. 3B), and similar negative results were obtained using BSA or DNase (data not shown).

To further analyse the TG-domain binding to PGN, we used Microscale thermophoresis (MST). In particular, we analysed the binding capability of the TG-domain towards the sacculi from *P. aeruginosa* or *B. subtilis* (gram-positive), showing that the protein is able to bind more efficiently gram-negative PGNs (roughly estimated Kd ~ 40 and 200 μM, respectively; Fig. 3C). The helicase control binds *B. subtilis* PGN with a scarcely significant affinity (Kd ~ 2 mM).

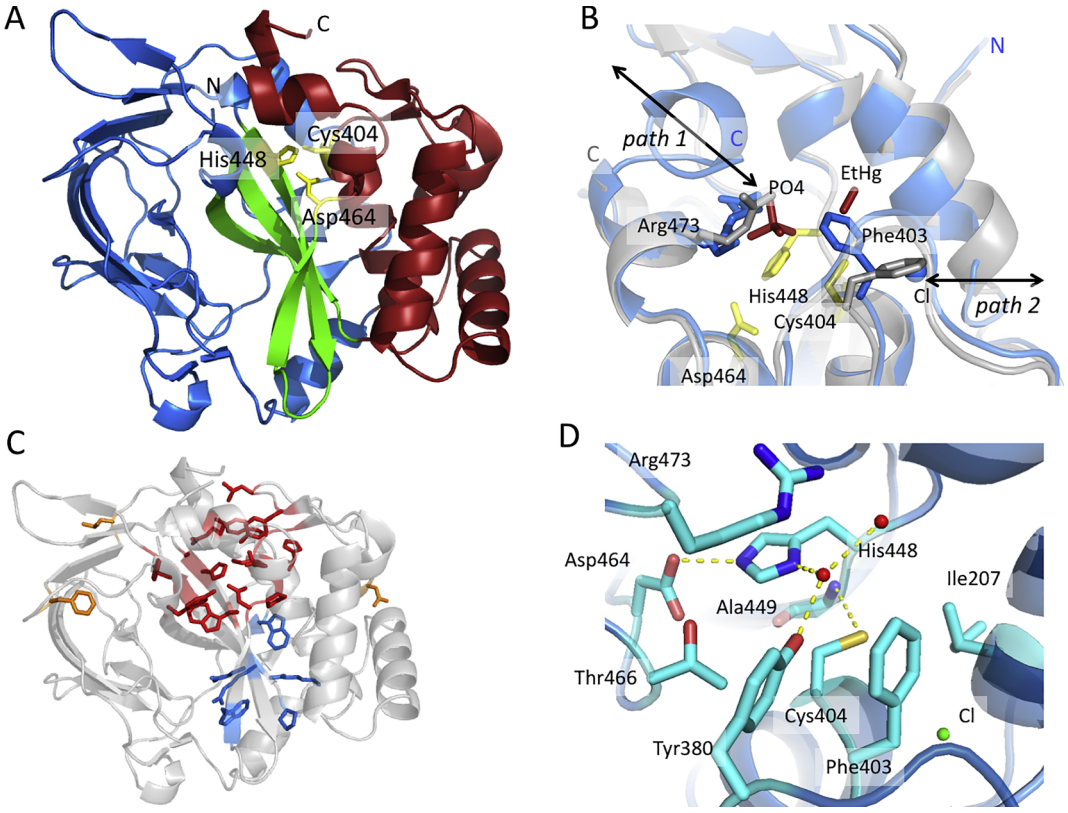

Fig. 2. TgpA structure. (A) TgpA overall periplasmic domain resembles a heart, represented in cartoons coloured with blue for *right ventricle,* green for *septum*, and red for *left ventricle.* The catalytic triad is drawn as yellow sticks. (B) TgpA active site and proposed substrate/product gating mechanism involving Phe403: path 1 between the C-ter helix and Arg473, path 2 toward the N-terminal helix. (C) Strictly conserved residues in an alignment performed with members of the gamma-proteobacteria class (excluding pseudomonadales order; lower sequence identity 45%): residues around the active site represented as red sticks, residues with structural function coloured in blue, and isolated residues in orange. (D) Active site residues around Cys404 (pictures drawn with PyMOL, The PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC).

3. Discussion

P. aeruginosa is presently considered a public health threat due to its ability to develop resistance to multiple drugs. A fruitful strategy for the discovery of new antibiotics relies on the identification of genes that are essential for bacterial viability. In this context we identified TgpA, an inner membrane protein as an essential protein for *P. aeruginosa*, suitable for drug development (Milani et al., 2012). Here, we focus on the analysis of TgpA periplasmic region, predicted to belong to the transglutaminase-like superfamily (TG-like domain), which includes archaeal, bacterial and eukaryotic proteins homologous to animal transglutaminases (Makarova et al., 1999).

Our structural investigation shows that the periplasmic TG-domain is composed of two subdomains linked by a central β-sheet (*the septum*): the N-ter subdomain, rich of β-strands, likely endowed with regulatory and/or substrate-binding/recognition activity, and the α-helical C-ter subdomain, hosting the conserved putative “catalytic triad” Cys404-His448-Asp464, likely responsible for the enzymatic activity.

In our crystallized TG-domain construct, Cys404 is buried inside the protein domain at about 10 Å from the closest external surface, due to a small 3_{10}helix that is accommodated on the “catalytic triad” cavity. Conformational transitions induced by the binding of ethyl-Hg to Cys404 make the C-ter helix more mobile, leaving the site open and presumably fully accessible to the correct substrate (when present). Anyway, we cannot exclude that the location of the C-ter helix may be the result of construct truncation or of the crystal lattice environment.

In order to test whether the hypothetical catalytic activity of TG-domain is essential for bacterial viability, we used a modified *P. aeruginosa* strain with controlled expression of genomic TgpA. We tested complementation of such strain with either wt TgpA or with the C404A mutant, showing that only the former is able to rescue the bacterial growth when the genomic protein is repressed. Such evidences demonstrated the crucial role of Cys404 for TgpA activity, suggesting that the essentiality of TgpA is linked to the TG-domain active site.

The geometry of three residues in the TG-domain active site (Tyr380, Cys404 and His448) indicates similarity with the active site of a Cys endopeptidase acting on muropeptides. Moreover, inspection of the N-terminal portion of the TG-domain (residues Val195 - Ala323) highlights structural homology with carbohydrate binding domains able to bind chitin or cellulose (Hanazono et al., 2016; Simpson et al., 1999). Cellulose, chitin and PGN are major polymers in living organisms, consisting of long-chain carbohydrates linked with β-1,4 glycosidic bonds. Principles and regulatory mechanism of the enzymes involved in the metabolism of these polymers are only partially understood, however, there is growing structural and biochemical information on how synthesis, chemical modifications and hydrolysis processes may occur. In bacteria, there are well-characterized domains that bind cell wall components, tightly related with enzymes that participate in the peptidoglycan degradation and remodelling (Typas et al., 2011). Considering the periplasmic localization of the protein and the evidence that the N-ter subdomain is structurally linked to carbohydrate binding domains, our structural results suggest a possible involvement of the protein in cell wall (re)modelling. Accordingly, we proved that the TG-domain is able to interact with PGN using pull-down experiments. Moreover, with thermophoresis experiments on soluble portion of PGN extracts, we were able to estimate a rough affinity in the μM range. Such degree of affinity is in line with the values measured for the LysM domain (Wong et al., 2014), a well characterized

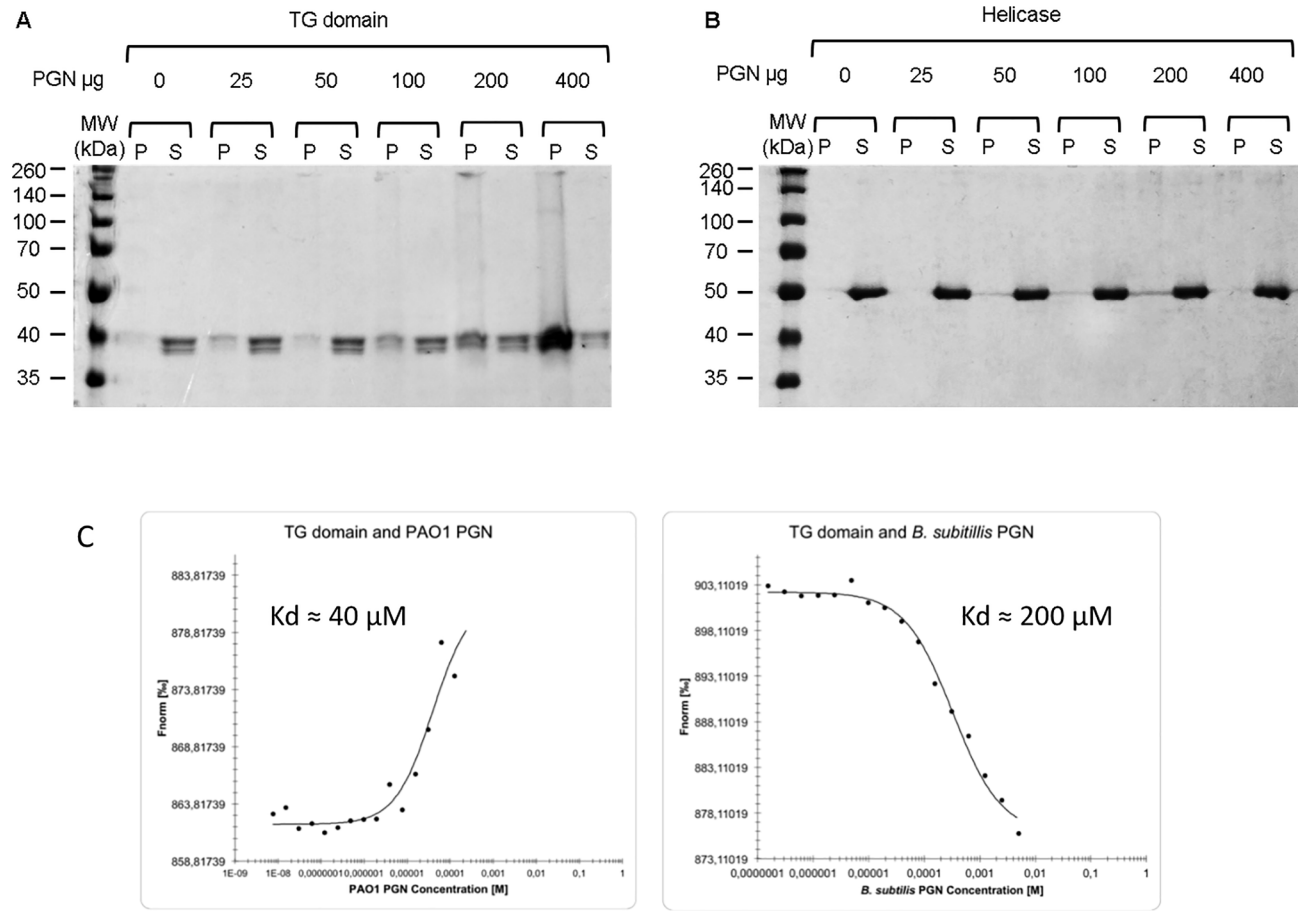

Fig. 3. Pull-down experiments with insoluble PGN. The binding capacities of the (A) TG domain and (B) Dengue virus helicase were assayed by pulldown assay with increasing concentrations of insoluble PGN from *P. aeruginosa*. The lanes are labelled as follows: MW = Molecular weight marker, P = Pellet, S = Supernatant. (C) TG domain-PGN interaction analyzed by MST. The TG domain – PGN interaction was monitored by titrating PGN form *P. aeruginosa* (left) or *B. subtilis* (right) against 75 nM of TG domain labelled with the Kit RED-NHS (NanoTemper technologies, Munich, Germany). The changes of the fluorescent thermophoresis signals were plotted and K_d values were estimated using the NanoTemper analysis software, through non linear curve fitting.

carbohydrate-binding domain with affinity for GlcNac polymers, often associated with proteins involved in PGN synthesis or remodeling. Although the carbohydrate binding module of the TG-domain does not share structural homology with LysM or other known PGN binding domains, the mechanisms supporting PGN recognition are still largely unknown.

The plasticity of the TG-domain C-ter region, suggested by comparison between the native and the Hg-derivative structures, might hint at a tuning of the protein activity via substrate recognition, with opening of the active site cleft only when the right substrate is present. In particular, the structural data suggests a possible gating mechanism regulating the substrate/product permeability to/from the active site involving the membrane anchored N- and C-ter regions of the domain and the double conformation of the side chain of Phe403. Finally, the phosphate ion observed within the "catalytic triad" cavity when Ethyl-Hg is covalently bond to Cys404, might suggest preferential selection of negatively charged substrates.

All the microbiological, structural and biochemical data here reported enrich our views on the function of a new protein essential for *P. aeruginosa* viability, thus representing a valuable target for the development of novel antibacterial leads.

4. Materials and methods

4.1. Construction of plasmids expressing His-tagged TgpA wt and the mutant C404A in the full-length protein

A plasmid expressing a His-tagged full length TgpA was prepared by amplifying the *tgpA* gene (*locus* PA2873) from *P. aeruginosa* PAO1 genomic DNA with primers 5′-TATCATGAACGCGATTCCGCGGGTC-3′ and 5′-AGAAGCTTGTAAAGCTTTCAGTGGTGGTGGTGGTGGTGTGCCT GCTCCTCTC-3′. Amplicons were digested with *Bsp*HI and *Hind*III restriction enzymes and cloned into the broad host range vector pHERD28T (Qiu et al., 2008), digested with *Nco*I and *Hind*III, under the arabinose inducible promoter P_{BAD}, giving rise to the plasmid pHERD28T-TgpA.

The QuikChange II XL Kit (Stratagene) was used to generate the mutant Cys404Ala, using as a template the construct pHERD28T-TgpA, with the complementary primers carrying the mutated codon 5′- GCG CAGCGGCTTCGCCGCGCATTACGCC-3′ and 5′-GGCGTAATGCGCGGC GAAGCCGCTGCGC-3′.

For overexpression, the constructs pHERD28T-TgpA-His and pHERD28T-C404A-His were moved from *E. coli* JM109 to PAO1 strain by triparental mating as described previously (Milani et al., 2012). PAO1 cells carrying pHERD28T-TgpA-His, pHERD28T-TgpA-C404A-His, or the pHERD28T empty vector were grown at 37 °C for 20 h in micro-cultures of 200 µl of LB medium supplemented with 300 µg/ml

trimethoprim (Tmp) and with increasing arabinose concentrations (0, 0.025, 0.05, 0.1 and 0.2%), for the induction of the P_{BAD} promoter. Growth cultures were monitored in real-time by absorbance measurement at 600 nm (OD_{600}).

4.2. Western-Blot analysis

Overnight cultures of PAO1/pHERD28T-TgpA-His, PAO1/pHERD28T-C404A-His and PAO1/pHERD28T in LB medium supplemented with 300 μg/ml Tmp were diluted to OD_{600} of 0.05, incubated at 37 °C until OD_{600} of 0.4–0.6 and induced with 0.1% arabinose. A parallel culture of each strain was grown without the addition of arabinose. After two hours induction, the cells were harvested at 7000 rpm for 10 min at 4 °C and the pellet from 100 ml of culture was resuspended in 1 ml of lysis buffer (10 mM Tris-HCl pH 7.8, 20 μg/ml RNase and DNase, 1 mM PMSF, 0.2 mg/ml lysozyme) and passed through a French press for cell disruption. The lysates were centrifuged at 5000 rpm for 20 min at 4 °C and supernatant was ultracentrifuged at 37,000 rpm for 1 h at 4 °C. The pellets, corresponding to the total membrane fractions, were resuspended in 200 μl of Tris-HCl pH 7.8. Ten microliters of the samples with no arabinose and of the samples induced with 0.1% arabinose diluted 1:10 were analysed on SDS-PAGE. For Western blot analysis, specific monoclonal anti-His antibodies (Roche®) in 0.05% PBS-T were used.

4.3. Quantitative RT-PCR analysis

Total RNA was purified with a Total RNA Extraction kit (RBC Bioscience). cDNA was generated by incubating 1 μg of RNA with Superscript II Reverse Transcriptase (RT) (Invitrogen), 100 pg of random primers and buffer supplied by the manufacturer for 50 min at 42 °C. RT was inactivated by incubation at 70 °C for 15 min. As a control of DNA contamination in the subsequent quantitative RT-PCR (qRT-PCR) analysis, reactions were also run without RT. qRT-PCR amplifications were performed in duplicate reactions of 15 μl in tubes filled with iQ™ SYBR Green Supermix (Bio-Rad) and 300 nM of each primer (see below) and run in an iCycler Real-Time PCR machine (Bio-Rad) as follows: 1 cycle at 95 °C for 10 min, 50 cycles at 95 °C for 15 s and 60 °C for 40 s. The calculation of the relative expression of the *tgpA* gene from plasmid pHERD28T versus the empty vector was performed as described by the $2^{-\Delta\Delta CT}$ method (Livak and Schmittgen, 2001), normalizing first mRNA amounts to 16S ribosome RNA (ΔCT) and relating the ΔCT in the TgpA wt and C404A mutant ectopically expressed to the chromosomal copy ($\Delta\Delta CT$) (PAO1-pHERD28T Empty vector). To confirm PCR specificity, the PCR products were subjected to melting curve analysis in a temperature range spanning 55–95 °C with 1 cycle at 55 °C for 50 s and 80 cycles at 55 °C for 10 s set with 0.5 °C increments after the first cycle.

The primer pairs used for qRT-PCR were:

tgpA 5′-CGAAAGCGCTCTGCTGCAA-3′/ 5′-TCTTCGCAGTGGTGGT GGG-3′

16s 5′-tgtcgtcagctcgtgtcgtga-3′/ 5′-atccccaccttcctccggt-3′

4.4. TgpA gene complementation

TgpA wild type (wt) and mutant C404A were tested for their ability to complement a glucose-mediated down-regulation of the chromosomal *tgpA* gene in the PAO1 P_{rhaB}::*tgpA* strain (Milani et al., 2012), when expressed from the ectopic vectors pHERD28T-TgpA or pHERD28T-TgpA-C404A. For the complementation assays, pHERD28T-TgpA and pHERD28T-TgpA-C404A were moved from *E. coli* JM109 to the conditional mutant PAO1 P_{rhaB}::*tgpA* strain by triparental mating as described previously (Milani et al., 2012). Exconjugants P_{rhaB}::*tgpA* clones were selected on M9 minimal medium plates supplemented with 0.2% citrate (M9-citrate), 60 μg/ml gentamicin (Gm), 0.2% rhamnose and 300 μg/ml Tmp. This procedure generated the strains PAO1 P_{rhaB}::*tgpA*/pHERD28T-TgpA and PAO1 P_{rhaB}::*tgpA*/pHERD28T-TgpA-C404.

For both *PrhaB::tgpA*/pHERD28T-TgpA and PAO1 *PrhaB::tgpA*/pHERD28T-TgpA-C404, the conditional growth was assessed after an overnight growth at 37 °C in M9-citrate supplemented with 60 μg/ml Gm, 300 μg/ml Tmp and 0.2% rhamnose. An aliquot of the overnight culture was centrifuged and the bacterial pellet was washed three times with sterile PBS, and diluted to an OD_{600} of 10^{-6} in fresh M9-citrate with appropriated antibiotics and supplemented either with 0.2% rhamnose or 1% glucose. 200 μl of the different cultures were grown at 37 °C with stirring and monitored in real time by OD_{600} measurement in microtiter reader (TECAN) for 20 hrs.

4.5. Expression and purification of TgpA periplasmic domain

TgpA periplasmic domain (residues 180–544) was expressed in *E. coli* BL21CodonPlus (DE3) - RIPL (Agilent; chloramphenicol (Cm) resistance) as N-terminally (His)10-tagged proteins, from plasmid p2N (PRIMM srl.; ampicillin (Amp) resistance). The cells were grown at 37 °C in 2 l of LB medium containing 100 μg/ml Amp and 34 μg/ml Cm, to an OD_{600} of 0.6. After induction with 1 mM isopropyl β-D-thiogalactopyranoside (IPTG), the incubation temperature was lowered to 17 °C and cells were grown for 16 h. Cells, harvested by centrifugation, were resuspended in lysis buffer (1x Phosphate buffer saline (PBS), 1% Triton, 5 mM $MgCl_2$, 0.25 mg/ml lysozyme, 20 μg/ml DNase, 1μg/ml Pepstatin A, 0.5 mM AEBSF and 1 tablet of complete Protease Inhibitor (Roche®)), lysed with a cell disruptor (Constant system) at 27 KPSI, and the soluble fraction was collected after centrifugation at 15,000 rpm for 1 h at 4 °C (Beckman JA20). The protein purification was performed at 4 °C, using a Fast Protein Liquid Chromatography (FPLC) apparatus (ÄKTA system – GE Healthcare). After Ni-IMAC column (HisTrap™ FF crude 5 ml, GE Healthcare) equilibration with Buffer A (20 mM Phosphate buffer pH 7.5, 500 mM NaCl, 0.5 mM AEBSF and 1 tablet of complete Protease Inhibitor Roche®), the soluble cell extract was loaded at 1 ml/min and eluted at 250 mM imidazole. The protein was then loaded (1 ml/min flow-rate) on a Superdex™ 75 column (GE Healthcare), previously equilibrated with Gel Filtration (GF) Buffer (20 mM Tris-HCl pH 7.5, 500 mM NaCl, 20 mM imidazole, 0.5 mM AEBSF and 1 tablet of complete Protease Inhibitor Roche®). During the expression and purification, the recombinant TgpA periplasmic domain was susceptible to proteolytic cleavage. Mass analysis confirmed the loss of more than 40 amino acids at the C-ter end, and the produced peptides indicated different cleavage sites around Tyr499. The purified protein (in GF buffer) was concentrated up to 8 mg/ml using a centrifugal filter unit (10 kDa cut-off, Millipore).

4.6. Crystal structure TgpA periplasmic domain

Microbatch crystallization experiments on TgpA periplasmic domain (6 mg/ml) were assembled with an Oryx-8 crystallization robot (Douglas Instruments, East Garston, UK), with 0.3 μl droplets containing 66/34% of the protein/precipitant solutions, covered with Al's oil. Crystals were obtained after 48 h at 20 °C, in 0.8 M NaH_2PO_4, 0.8 M KH_2PO_4, and 0.1 M Na HEPES pH 7.5. Before collecting crystals for flash freezing in liquid nitrogen, about 0.1 μl glycerol was added to the crystal drop as cryoprotectant. The TgpA native crystals diffracted to a maximum resolution of 1.6 Å using synchrotron radiation at beam-line ID30a3 at the European Synchrotron Radiation Facility (ESRF-Grenoble, France). Additional crystals grown in a different condition (0.7 M $(NH_4)H_2PO_4$, 0.1 M Na HEPES pH 7.2) were selected for soaking with mercury derivatives. The crystals were soaked in a stabilizing solution (0.7 M $(NH_4)H_2PO_4$, 0.1 M Na HEPES pH 7.6) with 5 mM ethylmercurithiosalicylic acid sodium salt for 60 min, then washed (in stabilizing solution) and transferred into a cryoprotectant solution (0.7 M $(NH_4)H_2PO_4$, 0.1 M Na HEPES pH 7.6, glycerol 25%) and flash frozen. TgpA derivative crystal showed a nice absorption peak for

mercury at 12,302 keV (1.0078 Å) and diffracted to a maximum resolution of 1.8 Å at beam-line BM14U (ESRF-Grenoble, France). For both native and derivative crystals, X-ray diffraction data were indexed and scaled using XDS (Kabsch, 2010) (Table 1).

The three-dimensional structure of TgpA was solved using the SAD method based on the mercury derivative (X-ray wavelength 1.0078 Å) using the program package SHELXC/D/E (Sheldrick, 2010) within the CCP4i interface (Potterton et al., 2003). Briefly, SHELXC detected an anomalous signal up to ~ 1.9 Å; SHELXD (at 2 Å resolution) found 6 possible heavy atoms sites (CC All/Weak 42.8/25.1, CFOM 67.9, best 67.9, PATFOM 16.28), and SHELXE calculated the initial phases and performed density modification (FOM = 0.582, Pseudo-free CC = 63.71%). The obtained experimental electron density was then used to trace an initial protein model with ARP/WARP (Lamzin and Wilson, 1997). The single TgpA molecule present in the crystal asymmetric unit was then subjected to manual (program COOT (Emsley et al., 2010)) and constrained refinements using REFMAC5 (Steiner et al., 2003), and the partially refined model was used for molecular replacement to solve the structure of the native dataset (Vagin and Teplyakov, 1997). Additional refinements with BUSTER (Smart et al., 2012) and REFMAC5 were subsequently performed for both datasets. Data collection, refinement statistics as well as stereochemical quality of the two models (checked also using the wwPDB Validation Server, https://validate-rcsb-1.wwpdb.org/) are summarized in Table 1 (Read et al., 2011). Atomic coordinates and structure factors have been deposited in the PDB (Berman et al., 2000), with accession codes 6G49 and 6G4H for the native and mercury derivative datasets, respectively.

4.7. Cell wall peptidoglycans isolation

A published method (Desmarais et al., 2014) with minimum modifications was followed for the isolation of cell wall peptidoglycans (PGN) from *P. aeruginosa*. Briefly, cells were harvested from an exponential phase culture (OD_{600} of 0.8) by centrifugation, washed two times with phosphate-buffered saline (PBS). The final pellet was resuspended in 3 ml of PBS, slowly added to boiling 6% SDS solution and boiled for 2 h; the mixture was then stirred overnight at room temperature. PNGs were recovered by ultra-centrifugation at 400,000 g, washed six times with ultrapure water at room temperature and then digested with 100 µg/ml of Pronase E for 2 h at 60 °C. The reaction was stopped with the addition of 200 µl of 6% SDS and incubation for 30 min at 100 °C. Additional washes by ultra-centrifugation at 400,000 g were performed until the excess of SDS was removed. The final PNGs were resuspended in 50 mM MOPS, pH 7.0 and stored at −20 °C.

4.8. Insoluble peptidoglycan pull down assay

For PGN binding assay, we slightly modified a procedure already described (Wong et al., 2014). In a 50 µl reaction mixture, 50 µg of purified TG domain was added to increasing amounts of PGN (25, 50, 100, 200 and 400 µg) in the assay buffer (50 mM Tris-HCl, pH 8, 1 M NaCl, 5 mM β-mercaptoethanol, 1% tween-20). The reaction was incubated for 30 min at 25 °C, and then centrifuged at 10,500 g for 45 min. The supernatant was collected, and the insoluble PGN pellet was washed three times in assay buffer. Ten microliters of supernatant and resuspended pellet solution were analysed on a 10% SDS-PAGE gel. As a reaction control, three different proteins (BSA, DNAse (Sigma) and helicase from Dengue virus) were used in reaction with the PGN at the same conditions as TG domain. A negative control was performed with 50 µg of each protein under the same conditions but without PGN, to ensure that the proteins were not precipitating on their own.

4.9. Cell wall binding affinity

To evaluate the binding ability of the TG domain to cell wall components, we used PGN isolated from both *Bacillus subtilis* (Sigma – Aldrich) and *P. aeruginosa*. Protein interactions with PGN were measured through MicroScale thermophoresis (MST). The soluble fraction of the sacculi from *P. aeruginosa* and *B. subtilis*, was separated from the insoluble portion by ultracentrifugation (15,000 *g*), maintaining ~50% of the total PGN in the soluble fraction. It is possible to estimate a hypothetical dissociation constant (K_d) assuming that the binding is due to the muropeptide glycans N-Acetylglucosamine [GlcNac] and N-acetylmuramic acid [MurNAc].

Proteins were labelled with the monolith NT™ Protein labelling Kit RED-NHS (Amine Reactive; NanoTemper technologies, Munich, Germany) to achieve a 1:1 M ratio of labelled protein to dye. Serial dilutions of the PGN were prepared (from 250 µM to 100 nM) keeping constant the concentration of labelled TG domain (150 nM), in thermophoresis buffer (50 mM Tris/HCl pH 7.4, 150 mM NaCl, 10 mM $MgCl_2$, 0.05% Tween-20). MST measurements (MST power 60%) were performed at 24 °C on a Monolith NT.115 instrument (NanoTemper Technologies) using Premium coated capillary. As negative control, helicase from Dengue virus was evaluated for its ability to bind to PGN from *B. subtilis*.

The K_d of protein-ligand complexes was then calculated with the following equation:

$$F = U + \frac{(B-U)\bullet([ligand] + [protein] + K_d - \sqrt{([ligand] + [protein] + K_d)^2 - 4\bullet[ligand]\bullet[protein]})}{2[protein]}$$

where F indicates the measured fluorescence, while U and B represent the response values of the unbound and bound states, respectively.

5. Accession numbers

Coordinates and structure factors have been deposited in the Protein Data Bank with accession number **6G4H** and **6G49**.

Acknowledgements

We are grateful to the beam line scientists from the ESRF facility for support during data collections, to C. Rotella for technical assistance, to D. Vecchietti and D. Tarantino for technical assistance and helpful discussion.

Funding information

This work was funded by European Commission (grant NABATIVI, EU-FP7-HEALTH-2007-B contract number 223670). This funder had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jsb.2018.12.004.

References

Ahvazi, B., Kim, H.C., Kee, S.H., Nemes, Z., Steinert, P.M., 2002. Three-dimensional structure of the human transglutaminase 3 enzyme: binding of calcium ions changes structure for activation. EMBO J. 21, 2055–2067.
Berman, H.M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T.N., Weissig, H., Shindyalov, I.N., Bourne, P.E., 2000. The Protein Data Bank. Nucleic Acids Res. 28, 235–242.
Desmarais, S.M., Cava, F., de Pedro, M.A., Huang, K.C., 2014. Isolation and preparation of bacterial cell walls for compositional analysis by ultra performance liquid chromatography. J. Vis. Exp., e51183.
Driscoll, J.A., Brody, S.L., Kollef, M.H., 2007. The epidemiology, pathogenesis and treatment of Pseudomonas aeruginosa infections. Drugs 67, 351–368.
Emsley, P., Lohkamp, B., Scott, W.G., Cowtan, K., 2010. Features and development of Coot. Acta Crystallogr. Sect. D 66, 486–501.
Hanazono, Y., Takeda, K., Niwa, S., Hibi, M., Takahashi, N., Kanai, T., Atomi, H., Miki, K.,

2016. Crystal structures of chitin binding domains of chitinase from Thermococcus kodakarensis KOD1. FEBS Lett. 590, 298–304.
Holm, L., Laakso, L.M., 2016. Dali server update. Nucleic Acids Res. 44, W351–W355.
Kabsch, W., 2010. Xds. Acta Crystallogr. D Biol. Crystallogr. 66, 125–132.
Kobayashi, K.H.K., Yokozeki, K., Yamanaka, S., 1998. Molecular cloning of the transglutaminase gene from Bacillus subtilis and its expression in Escherichia coli. Biosci. Biotechnol. Biochem. 62, 1109–1114.
Lamzin, V.S., Wilson, K.S., 1997. Automated refinement for protein crystallography. Methods Enzymol. 277, 269–305.
Livak, K.J., Schmittgen, T.D., 2001. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods 25, 402–408.
Lodise Jr., T.P., Lomaestro, B., Drusano, G.L., 2007. Piperacillin-tazobactam for Pseudomonas aeruginosa infection: clinical implications of an extended-infusion dosing strategy. Clin. Infect. Dis. 44, 357–363.
Makarova, K.S., Aravind, L., Koonin, E.V., 1999. A superfamily of archaeal, bacterial, and eukaryotic proteins homologous to animal transglutaminases. Protein Sci. 8, 1714–1719.
Milani, A., Vecchietti, D., Rusmini, R., Bertoni, G., 2012. TgpA, a protein with a eukaryotic-like transglutaminase domain, plays a critical role in the viability of Pseudomonas aeruginosa. PLoS One 7, e50323.
Moll, M., Bryant, D.H., Kavraki, L.E., 2011. The LabelHash server and tools for substructure-based functional annotation. Bioinformatics 27, 2161–2162.
Obritsch, M.D., Fish, D.N., MacLaren, R., Jung, R., 2005. Nosocomial infections due to multidrug-resistant Pseudomonas aeruginosa: epidemiology and treatment options. Pharmacotherapy 25, 1353–1364.
Pedersen, L.C., Yee, V.C., Bishop, P.D., Le Trong, I., Teller, D.C., Stenkamp, R.E., 1994. Transglutaminase factor XIII uses proteinase-like catalytic triad to crosslink macromolecules. Protein Sci. 3, 1131–1135.
Potterton, E., Briggs, P., Turkenburg, M., Dodson, E., 2003. A graphical user interface to the CCP4 program suite. Acta Crystallogr. D Biol. Crystallogr. 59, 1131–1137.
Qiu, D., Damron, F.H., Mima, T., Schweizer, H.P., Yu, H.D., 2008. PBAD-based shuttle vectors for functional analysis of toxic and highly regulated genes in Pseudomonas and Burkholderia spp. and other bacteria. Appl. Environ. Microbiol. 74, 7422–7426.
Read, R.J., Adams, P.D., Arendall 3rd, W.B., Brunger, A.T., Emsley, P., Joosten, R.P., Kleywegt, G.J., Krissinel, E.B., Lutteke, T., Otwinowski, Z., Perrakis, A., Richardson, J.S., Sheffler, W.H., Smith, J.L., Tickle, I.J., Vriend, G., Zwart, P.H., 2011. A new generation of crystallographic validation tools for the protein data bank. Structure 19, 1395–1412.
Sheldrick, G.M., 2010. Experimental phasing with SHELXC/D/E: combining chain tracing with density modification. Acta Crystallogr. D Biol. Crystallogr. 66, 479–485.
Sillitoe, I., Lewis, T.E., Cuff, A., Das, S., Ashford, P., Dawson, N.L., Furnham, N., Laskowski, R.A., Lee, D., Lees, J.G., Lehtinen, S., Studer, R.A., Thornton, J., Orengo, C.A., 2015. CATH: comprehensive structural and functional annotations for genome sequences. Nucleic Acids Res. 43, D376–D381.
Simpson, P.J., Bolam, D.N., Cooper, A., Ciruela, A., Hazlewood, G.P., Gilbert, H.J., Williamson, M.P., 1999. A family IIb xylan-binding domain has a similar secondary structure to a homologous family IIa cellulose-binding domain but different ligand specificity. Structure 7, 853–864.
Smart, O.S., Womack, T.O., Flensburg, C., Keller, P., Paciorek, W., Sharff, A., Vonrhein, C., Bricogne, G., 2012. Exploiting structure similarity in refinement: automated NCS and target-structure restraints in BUSTER. Acta Crystallogr. Sect. D 68, 368–380.
Steiner, R.A., Lebedev, A.A., Murshudov, G.N., 2003. Fisher's information in maximum-likelihood macromolecular crystallographic refinement. Acta Crystallogr. Sect. D 59, 2114–2124.
Szaff, M., Hoiby, N., Flensborg, E.W., 1983. Frequent antibiotic therapy improves survival of cystic fibrosis patients with chronic Pseudomonas aeruginosa infection. Acta Paediatr. Scand. 72, 651–657.
Typas, A., Banzhaf, M., Gross, C.A., Vollmer, W., 2011. From the regulation of peptidoglycan synthesis to bacterial growth and morphology. Nat. Rev. Microbiol. 10, 123–136.
Vagin, A., Teplyakov, A., 1997. MOLREP: an automated program for molecular replacement. J. Appl. Crystallogr. 30, 1022–1025.
Valerius, N.H., Koch, C., Hoiby, N., 1991. Prevention of chronic Pseudomonas aeruginosa colonisation in cystic fibrosis by early treatment. Lancet 338, 725–726.
Williams, K.P., Gillespie, J.J., Sobral, B.W., Nordberg, E.K., Snyder, E.E., Shallom, J.M., Dickerman, A.W., 2010. Phylogeny of gammaproteobacteria. J. Bacteriol. 192, 2305–2314.
Wong, J.E., Alsarraf, H.M., Kaspersen, J.D., Pedersen, J.S., Stougaard, J., Thirup, S., Blaise, M., 2014. Cooperative binding of LysM domains determines the carbohydrate affinity of a bacterial endopeptidase protein. FEBS J. 281, 1196–1208.
Xu, Q., Sudek, S., McMullan, D., Miller, M.D., Geierstanger, B., Jones, D.H., Krishna, S.S., Spraggon, G., Bursalay, B., Abdubek, P., Acosta, C., Ambing, E., Astakhova, T., Axelrod, H.L., Carlton, D., Caruthers, J., Chiu, H.J., Clayton, T., Deller, M.C., Duan, L., Elias, Y., Elsliger, M.A., Feuerhelm, J., Grzechnik, S.K., Hale, J., Han, G.W., Haugen, J., Jaroszewski, L., Jin, K.K., Klock, H.E., Knuth, M.W., Kozbial, P., Kumar, A., Marciano, D., Morse, A.T., Nigoghossian, E., Okach, L., Oommachen, S., Paulsen, J., Reyes, R., Rife, C.L., Trout, C.V., van den Bedem, H., Weekes, D., White, A., Wolf, G., Zubieta, C., Hodgson, K.O., Wooley, J., Deacon, A.M., Godzik, A., Lesley, S.A., Wilson, I.A., 2009. Structural basis of murein peptide specificity of a gamma-D-glutamyl-l-diamino acid endopeptidase. Structure 17, 303–313.