Acta Crystallographica Section D
Biological Crystallography
ISSN 1399-0047

The type IV secretion protein TraK from the *Enterococcus* conjugative plasmid pIP501 exhibits a novel fold

Nikolaus Goessweiner-Mohr,[a,b]* Christian Fercher,[a] Karsten Arends,[c] Ruth Birner-Gruenberger,[d] Diana Laverde-Gomez,[e] Johannes Huebner,[e] Elisabeth Grohmann[e]* and Walter Keller[a]*

[a]Institute for Molecular Biosciences, Karl-Franzens-University Graz, Humboldtstrasse 50/III, 8010 Graz, Austria, [b]Institute of Molecular Biotechnology (IMBA), Austrian Academy of Sciences, Dr. Bohr-Gasse 3, 1030 Vienna, Austria, [c]Nordufer 20, 13353 Berlin, Germany, [d]Institute for Pathology and Center of Medical Research, Medical University Graz, Stiftingtalstrasse 24, 8010 Graz, Austria, and [e]Division of Infectious Diseases, University Medical Center Freiburg, Hugstetter Strasse 55, 79106 Freiburg, Germany

Correspondence e-mail: n.goessweiner.mohr@gmail.com, elisabeth.grohmann@uniklinik-freiburg.de, walter.keller@uni-graz.at

Received 19 September 2013
Accepted 22 January 2014

PDB reference: TraK, 4hic

Conjugative plasmid transfer presents a serious threat to human health as the most important means of spreading antibiotic resistance and virulence genes among bacteria. The required direct cell–cell contact is established by a multi-protein complex, the conjugative type IV secretion system (T4SS). The conjugative core complex spans the cellular envelope and serves as a channel for macromolecular secretion. T4SSs of Gram-negative (G−) origin have been studied in great detail. In contrast, T4SSs of Gram-positive (G+) bacteria have only received little attention thus far, despite the medical relevance of numerous G+ pathogens (*e.g.* enterococci, staphylococci and streptococci). This study provides structural information on the type IV secretion (T4S) protein TraK of the G+ broad host range *Enterococcus* conjugative plasmid pIP501. The crystal structure of the N-terminally truncated construct TraKΔ was determined to 3.0 Å resolution and exhibits a novel fold. Immunolocalization demonstrated that the protein localizes to the cell wall facing towards the cell exterior, but does not exhibit surface accessibility. Circular dichroism, dynamic light scattering and size-exclusion chromatography confirmed the protein to be a monomer. With the exception of proteins from closely related T4SSs, no significant sequence or structural relatives were found. This observation marks the protein as a very exclusive, specialized member of the pIP501 T4SS.

1. Introduction

Bacterial conjugation is the major contributor to horizontal gene transfer (Grohmann *et al.*, 2003; Williams & Hergenrother, 2008; Zechner *et al.*, 2012). The conjugative spread of plasmid-encoded antibiotic-resistance and virulence genes presents a serious threat to human health. The process of conjugation involves transport of DNA from a donor to a recipient cell, which requires direct contact between the cells (Cascales & Christie, 2003; Grohmann *et al.*, 2003; Alvarez-Martinez & Christie, 2009). The so-called type IV secretion system (T4SS) is responsible for the transfer. Conjugative T4SSs are plasmid-encoded, multi-protein complexes that are large enough to span the bacterial cell wall (Llosa *et al.*, 2002). A great number of studies have unveiled significant information about T4SSs of Gram-negative (G−) origin, such as the F-plasmid-, R388- and pKM101-encoded T4SSs of *Escherichia coli* and the Ti-plasmid-encoded T4SS of *Agrobacterium tumefaciens* (Llosa *et al.*, 2009; de la Cruz *et al.*, 2010; Hayes *et al.*, 2010; Rêgo *et al.*, 2010; Smillie *et al.*, 2010; Wallden *et al.*, 2010; Thanassi *et al.*, 2012; Zechner *et al.*, 2012; Christie *et al.*, 2014). In contrast, the great majority of our knowledge on T4SSs of Gram-positive (G+) origin is based on similarities to

counterparts from G− systems (Grohmann *et al.*, 2003; Abajy *et al.*, 2007; Goessweiner-Mohr, Arends *et al.*, 2013). Nevertheless, a significant amount of information on G+ type IV secretion (T4S) has become available over the last five years. According to Chen *et al.* (2008), the conjugative transfer of the substrate is initiated by the putative coupling protein PcfC of the *Enterococcus* sex-pheromone plasmid pCF10 (Chen *et al.*, 2008). The protein mediates the NTP-dependent transfer through a pCF10-encoded T4S channel (Li *et al.*, 2012). The Rood group has made a substantial contribution to our understanding of G+ T4S by characterizing the pCW3-encoded T4SS derived from *Clostridium perfringens* (Bannam *et al.*, 2006; Parsons *et al.*, 2007; Teng *et al.*, 2008; Steen *et al.*, 2009; Bantwal *et al.*, 2012). Li and coworkers reported for the first time the horizontal transfer of a G+ pathogenicity island, which was shown to be mediated by a genomic island-type T4SS, and suggested a hypothetical model for T4S in epidemic *Streptococcus suis* strains (Li *et al.*, 2011). Last year, the first structural information on T4SS proteins of G+ origin finally became available (Goessweiner-Mohr *et al.*, 2012; Porter *et al.*, 2012; Walldén *et al.*, 2012; Goessweiner-Mohr, Grumet, Arends *et al.*, 2013; Goessweiner-Mohr, Grumet, Pavkov-Keller *et al.*, 2013).

The multiple antibiotic-resistance plasmid pIP501 (Horodniceanu *et al.*, 1979) isolated from *S. agalactiae* has the broadest known host range in G+ bacteria. Furthermore, Kurenbach and coworkers were able to demonstrate that pIP501 is the first conjugative plasmid of G+ origin which stably replicates in G− bacteria (Kurenbach *et al.*, 2003). The pIP501 T4SS operon encodes 15 putative T4SS (Tra) proteins, of which only four show significant sequence similarity to the Ti-plasmid-encoded T4SS from *A. tumefaciens*. The ATPase TraE (related to VirB4) is most likely to drive the conjugative process by hydrolyzing ATP. Furthermore, it interacts with itself and several other potential pIP501 T4SS proteins (Abajy *et al.*, 2007). The coupling protein TraJ supposedly connects the macromolecular complex of single-stranded plasmid DNA and relaxosome proteins with the secretory conduit. The hexameric protein is related to VirD4; however, it lacks the transmembrane region typical of other coupling proteins (Abajy *et al.*, 2007; E. Grohmann *et al.*, unpublished work). TraJ might be recruited to the cell membrane by the pIP501 T4SS protein TraI (Alvarez-Martinez & Christie, 2009). The muramidase TraG (related to VirB1) is responsible for locally opening the thick peptidoglycan layer of G+ bacteria, a process required for the buildup of the T4SS core complex (Arends *et al.*, 2013). Furthermore, the pIP501-encoded relaxase TraA has been studied in detail (Kopec *et al.*, 2005; Kurenbach *et al.*, 2006). As the first protein of the T4SS operon, it was shown to bind specifically to the origin of transfer (*oriT*) and to auto-regulate the expression of the pIP501 T4SS components.

Recently, we published the 2.5 Å resolution structure of the C-terminal domain of the pIP501 T4SS protein TraMΔ (Goessweiner-Mohr, Grumet, Arends *et al.*, 2013). Despite this first structural data and progress in the functional characterization of some key proteins, the structural and functional characterization of most of the 15 T4SS proteins is still lacking. Moreover, we still need to identify the key components of the proposed conjugative core complex, as was performed for the G− model system from *E. coli* (Fronzes *et al.*, 2009).

Here, we present the biophysical and structural characterization of the N-terminally truncated protein TraK (formerly named ORF11; GenBank CAD44391.1), a 30.6 kDa protein (His-tagged $TraK_{66–307}$, further referred to as TraKΔ) encoded by the *Enterococcus faecalis* conjugative plasmid pIP501. We report the 3.0 Å resolution structure, which has been solved by selenomethionine multiwavelength anomalous dispersion (MAD). The protein localizes to the cell envelope, facing towards the cell exterior, and behaves as a monomer under the tested conditions. With the exception of the T4SS proteins of closely related G+ T4SSs showing high sequence identity, no significant sequence-based or structure-based relationships could be found. These results mark TraK as a highly exclusive T4SS protein that is only found in *Enterococcus* and *Streptococcus* T4SSs. Thus, TraK appears to be an interesting target for further functional studies.

2. Materials and methods

2.1. Expression, purification and buffer optimization

*traK*Δ was cloned into the 7×His-tag expression vector pQTEV (a gift from K. Büssow, Max-Planck-Institute for Molecular Genetics, Berlin, Germany), and *E. coli* BL21-CodonPlus(DE3)-RIL (Stratagene, Amsterdam, The Netherlands) competent cells were transformed with the recombinant construct pQTEV-*traK*Δ. Large-scale expression was performed in 500 ml LB medium supplemented with 100 µg ml^{-1} ampicillin. TraKΔ expression was induced at an OD_{600} of ∼0.6 by the addition of 1 m*M* IPTG, and expression continued for 3 h at 37°C. The cells were harvested and immediately frozen at −20°C. TraKΔ expression levels were monitored by SDS–PAGE.

TraKΔ cell pellets were first lysed in 40 ml 25 m*M* HEPES pH 7.6, 75 m*M* Na_2SO_4. 2 U of DNAse (Sigma–Aldrich, St Louis, USA) was added and PMSF and benzamidine were added to final concentrations of 1 and 2 m*M*, respectively. The cell suspension was vigorously mixed (UltraTurrax, IKA, Staufen, Germany) and kept on ice for 30 min. The suspension was sonicated (Sonopuls HD2070, Bandelin; 1 min continuous sonification, ∼80% amplitude) and centrifuged for 30 min at 8°C and 15 000*g*. Pellet and supernatant fractions were analyzed by SDS–PAGE. The pellet was applied to a second extraction step with 20 ml of the above buffer. The TraKΔ-containing supernatant fractions were pooled and loaded onto a HisTrap FF 1 ml column (GE Healthcare, Chalfont St Giles, England) for affinity purification. The purity of TraKΔ was assessed by SDS–PAGE. Imidazole from the HisTrap affinity purification was removed by buffer exchange during concentration (Amicon tubes, 3000 MWCO; Merck, Darmstadt, Germany).

Purified TraKΔ protein with a concentration of 1 mg ml^{-1} was subjected to buffer-optimization screening by differential

scanning fluorimetry (Ericsson *et al.*, 2006) using all crystallization buffers present in the Index, PEG/Ion, MembFac (Hampton Research, Aliso Viejo, California, USA) and Morpheus (Molecular Dimensions, Newmarket, England) screens. For the assay, 10 µl protein sample was mixed with 10 µl of the respective buffer and 5 µl 50× SYPRO Orange (Sigma–Aldrich, St Louis, USA) stock. The resulting thermostability curves were analyzed and a new extraction buffer was designed by combining the buffer components (Collins *et al.*, 2004) which showed a thermostabilizing effect while keeping the composition as simple as possible. The optimized TraKΔ-lysis buffer consisted of 50 m*M* HEPES pH 7.6, 200 m*M* ammonium sulfate and was used for all subsequent TraKΔ extractions, as well as for crystallization.

2.2. Expression of the TraKΔ selenomethionine derivative

For the expression of selenomethionine-labelled (SeMet) TraKΔ protein, pQTEV-*traK*Δ plasmid DNA was isolated and transformed into the methionine-auxotroph *E. coli* strain B834 (DE3) (Novagen, Merck, Darmstadt, Germany) using standard protocols. The cells were resuspended in M9 minimal medium at an OD_{600} of ∼0.6, grown for 1 h at 37°C and induced with 1 m*M* IPTG. At the same time, 25 mg selenomethionine was added and overexpression continued for 3 h. The cells were harvested and immediately frozen at −20°C. SeMet TraKΔ expression levels were monitored by SDS–PAGE. Protein extraction was performed as described above.

2.3. Biophysical analysis of TraKΔ

TraKΔ was extracted and His-affinity purified in 50 m*M* Tris, 100 m*M* ammonium sulfate pH 7.45. TraKΔ-containing fractions were pooled and concentrated *via* centrifugation in Amicon tubes (Millipore Amicon, 3000 MWCO). TraKΔ was further purified by size-exclusion chromatography using a Superdex 200 HR 10/30 column (GE Healthcare, Chalfont St Giles, England). A gel-filtration standard (670, 158, 44, 17 and 1.35 kDa; Bio-Rad, Hercules, California, USA) was used to calculate the molecular weight of TraKΔ. TraKΔ-containing fractions were stored at −80°C for subsequent experiments.

Circular-dichroism (CD) measurements were performed on a Jasco J715 (Jasco Instruments, Gross-Umstadt, Germany) spectropolarimeter equipped with an external thermostat. Spectra were measured from 260 to 190 nm in a 0.01 cm cuvette at a protein concentration of 0.95 mg ml^{-1}. Ten individual spectra were accumulated and the standard deviation was calculated from the repeated measurements. Temperature scans were performed in a 0.02 cm temperature-controlled cuvette in the range from 25 to 95°C using a step-scan procedure with a constant wavelength of 208 nm. Spectra resulting from three accumulated scans were measured every 5°C. The temperature gradient was set to 1°C per minute. TraKΔ was applied at a concentration of 0.46 mg ml^{-1}. The CD data were evaluated using the *DichroWeb* online service (Whitmore & Wallace, 2008) using reference database No. 4.

For the dynamic light-scattering (DLS) measurements, a size-exclusion fraction containing 0.95 mg ml^{-1} TraKΔ was measured directly in a 45 µl cuvette. Ten measurements with constant baseline were merged and the monodispersity was assessed.

For the SAXS measurements on the X33 beamline at DESY, Hamburg, Germany, TraKΔ was suspended in 100 m*M* ammonium sulfate, 100 m*M* NaCl, 50 m*M* HEPES pH 7.0. Size-exclusion purified protein was further concentrated and TraKΔ was measured at three different concentrations: 6.3, 3.0 and 1.64 mg ml^{-1}. The program *PRIMUS* (Konarev *et al.*, 2003) was applied to subtract the buffer from the protein data. Maximum intensity (I_0) and the radius of gyration (R_g) were calculated from the Guinier plot generated from the data at 3.0 mg ml^{-1}. The I_0 was used to calculate the molecular weight of TraKΔ in solution and *GNOME* (Svergun, 1992) was used to generate the output file for subsequent *ab initio* modelling with *GASBOR* (Svergun *et al.*, 2001) and *CRYSOL* (Svergun *et al.*, 1995) to generate a scattering curve from the X-ray-derived model. The *BUNCH* software (Petoukhov & Svergun, 2005) was applied to fit the X-ray data to the SAXS scattering curve of TraKΔ. This was carried out by *ab initio* and rigid-body modelling of TraKΔ and the N-terminal His tag of the construct, which was present in solution but was disordered and thus not visible in the X-ray model. The *GASBOR*-derived model was converted to a volumetric model using *SITUS* (Wriggers, 2010). *CHIMERA* (Pettersen *et al.*, 2004) was used to fit the TraKΔ crystal structure to the converted SAXS model.

2.4. Subcellular fractionation of *E. faecalis* JH2-2 (pIP501) and immunolocalization of TraK

Subcellular fractionation of *E. faecalis* JH2-2 (pIP501) was performed according to Buttaro *et al.* (2000) with modifications. An exponentially growing culture (OD_{600} = 0.5) of *E. faecalis* JH2-2 (pIP501) was chilled on ice for 15 min, washed twice with an equal volume of potassium phosphate buffer (50 m*M*, pH 7.0) and resuspended [1:50(*v*/*v*)] in lysis buffer (50 m*M* KH_2PO_4/K_2HPO_4 pH 7.0, 1 m*M* EDTA, 1 m*M* $MgCl_2$, 100 µg ml^{-1} DNase, 100 µg ml^{-1} RNase). The cells were broken by FastPrep-24 (MP Biomedicals, Illkirch, France) using lysing matrix E (1.4 mm ceramic spheres, 0.1 mm silica spheres, 4 mm glass beads; MP Biomedicals, Illkirch, France). Unlysed cells were removed by low-speed centrifugation. The cell-wall fraction was then harvested by high-speed centrifugation at 17 000*g* for 20 min at 4°C, the membrane fraction was obtained by ultracentrifugation of the supernatant at 163 000*g* for 2 h at 4°C (OTD Combi ultracentrifuge, Thermo Fisher Scientific GmbH, Dreieich, Germany). The remaining supernatant contained the soluble proteins. TraK was localized in the fractions (cell wall, membrane and cytoplasm) by immunostaining with primary polyclonal anti-TraKΔ antibody and a secondary horseradish peroxidase-conjugated anti-rabbit IgG antibody (Promega GmbH, Mannheim, Germany). Polyclonal anti-TraKΔ antibodies were derived from BioGenes (Berlin, Germany) according to the manufacturer's standard immunization protocol for rabbit. The rabbit anti-TraKΔ were subsequently

purified by antigen affinity chromatography according to the manufacturer's protocol.

2.5. TraK start-codon mutation

In order to evaluate whether the TraK double bands arise from proteolytic digestion or the utilization of a second putative ribosomal binding site (RBS) and subsequent start codon, the full open reading frame (ORF) of *traK* including its RBS was cloned into the pQTEV expression vector. We next constructed a double mutant comprising pQTEV and *traK* start-codon mutations (the respective ATG codon was changed to ACG) using the QuikChange Site-Directed Mutagenesis Kit (Agilent Technologies, Santa Clara, California, USA).

The constructs were transformed into *E. coli* BL21-CodonPlus(DE3)-RIL competent cells, which were grown in LB medium supplemented with 100 µg ml^{-1} ampicillin. Expression was induced at an OD_{600} of 0.6 with 0.5 m*M* IPTG at 37°C. The cells were harvested 3 h after induction and samples were loaded onto a 12% SDS polyacrylamide gel and subjected to Western blotting with rabbit polyclonal anti-TraKΔ antibodies followed by a secondary horseradish peroxidase-conjugated anti-rabbit IgG antibody (Promega, Mannheim, Germany).

2.6. Protease protection assay

E. faecalis JH2-2 (pIP501) was grown in 50 ml Todd Hewitt Broth medium (THB) supplemented with 20 µg ml^{-1} chloramphenicol and 2%(*w*/*v*) glycine to an OD_{650} of 0.4. Cells were chilled on ice for 15 min, harvested by centrifugation at 4000*g* for 10 min at 10°C and resuspended in 10 ml PBS supplemented with 10 mg ml^{-1} BSA, 1 mg ml^{-1} lysozyme and 10 U mutanolysin. *E. faecalis* JH2-2 (pIP501) peptidoglycan was digested for 90 min at 37°C and protoplast formation was confirmed by light and phase-contrast microscopy. In order to evaluate the survival rate of the cells, serial dilutions (10^{-3}–10^{-6}) of 10 µl cell suspension were plated onto modified THB agar plates [THB, 0.5 *M* sodium succinate, 20 µg ml^{-1} chloramphenicol, 0.8%(*w*/*v*) agarose] before and after peptidoglycan digestion. Protoplasts were harvested by centrifugation, washed three times in PBS and 1 mg ml^{-1} BSA and resuspended in THB medium without antibiotics. Protein digestion was induced by adding varying concentrations of proteinase K (0.1–10 µg ml^{-1}) to a 1 ml cell suspension. 100 µl samples were withdrawn after 5 min at 37°C and PMSF was added to a final concentration of 1 m*M*. Protoplasts were harvested by centrifugation as described above, resuspended in up to 30 µl SDS–PAGE sample buffer, denatured and loaded onto an 18% SDS polyacrylamide gel.

To confirm that protection from degradation of the cytoplasmic control protein TraN is owing to the lipid bilayer in the cytoplasmic membrane, Triton X-100 (Roth, Karlsruhe, Germany) was added to a sample containing 5 µg ml^{-1} proteinase K to a final concentration of 1%(*v*/*v*) and protein digestion continued for 5 min. 1 m*M* PMSF was added to stop the reaction and 10 µl cell suspension was mixed with 10 µl SDS–PAGE sample buffer.

Samples were loaded onto an 18% SDS polyacrylamide gel and subjected to Western blotting. Blots were probed with rabbit polyclonal anti-TraKΔ and anti-TraN antibodies (BioGenes, Berlin, Germany) followed by a secondary horseradish peroxidase-conjugated anti-rabbit IgG antibody (Promega, Mannheim, Germany).

2.7. Opsonophagocytosis killing assay (OPA)

The opsonophagocytic assay was performed as described previously (Theilacker *et al.*, 2012; Goessweiner-Mohr, Grumet, Arends *et al.*, 2013). In case of the OPAs including pre-incubation with peptidoglycan-degrading enzymes, a cell-wall enzymatic digestion was performed by co-incubating harvested *E. faecalis* JH2-2 cells harbouring tpIP501 ($OD_{650\ nm}$ = 0.4) with 0.1 mg ml^{-1} lysozyme and 0.01 mg ml^{-1} mutanolysin at 37°C for 15 min. After enzymatic treatment, the samples were washed with TSB medium, readjusted to an $OD_{650\ nm}$ of 0.4 and utilized in the subsequent experiments.

2.8. Crystallization and crystal optimization

Crystallization trials for His-tagged TraKΔ were initially set up with Index screen at a stock concentration of 5.5 mg ml^{-1} using the microbatch method (Chayen *et al.*, 1992). After evaluation of the first plate, the following screens were prepared at different concentrations: Index, Crystal Screen, Crystal Screen 2, MembFac, PEG/Ion (Hampton Research), JCSG and Morpheus (Molecular Dimensions). The drop ratio was 1:1 with a total drop volume of 1 µl. All plates were covered with paraffin oil (a total of ∼4 ml) and stored at 20°C. The formation of crystals was monitored over several weeks. Potential protein crystals were tested for diffraction using a rotating-anode diffractometer (MicroStar; Bruker AXS, Madison, Wisconsin, USA). The most promising of several positive conditions, Morpheus conditions 52 (0.03 *M* each of diethylene, triethylene, tetraethylene and pentaethylene glycols, 0.05 *M* each of imidazole and MES pH 6.5, 12.5% each of MPD, PEG 1K and PEG 3350) and 85 (0.02 *M* each of sodium L-glutamate, alanine, glycine, lysine–HCl and serine, 0.05 *M* each of imidazole and MES pH 6.5, 15% each of PEG MME 550 and PEG 20K) were selected for microbatch precipitant/protein concentration optimization matrices. A constant protein drop volume of 1 µl and different protein stock concentrations were used.

The original conditions showed small, compact crystals with poor diffraction (about 7 Å), which appeared after only a few days. The optimization did not improve the crystal diffraction limit or quality, but the use of the enhanced extraction buffer TraKΔ-lysis led to the formation of larger crystals with improved diffraction behaviour. Optimized conditions were used for the following setup with selenomethionine-containing TraKΔ. Crystals were harvested from condition 52 of the Morpheus screen (protein stock concentration 12.9 mg ml^{-1}). Native crystals were taken from condition 85 of the Morpheus screen (protein stock concentration 14.14 mg ml^{-1}).

2.9. Data collection and processing

Data collection was performed at 100 K on the synchrotron beamline X06DA at SLS, Villigen, Switzerland. No cryoprotectant was needed to preserve the crystals. The data sets were processed and scaled together using *iMosflm* (Battye *et al.*, 2011) and *SCALA* (Evans, 2006). *AutoSol* (McCoy *et al.*, 2007; Terwilliger *et al.*, 2009) and *AutoBuild* (Terwilliger *et al.*, 2008) from the *PHENIX* software suite (Adams *et al.*, 2010) were used to define the selenium heavy-atom sites using a selenium-derivative MAD data set (peak, inflection and high-end remote), as well as to build an initial model. The resulting model was completed manually in *Coot* (Emsley *et al.*, 2010), refined with *REFMAC*5 (Murshudov *et al.*, 2011) and utilized as a template for molecular replacement with *AutoMR* (McCoy *et al.*, 2007) from the *PHENIX* software suite, using a native data set of higher diffraction quality. The resulting model was again completed manually in *Coot* and refined with *REFMAC*5. The refined X-ray model was validated with the online service *MolProbity* (Chen *et al.*, 2010). The secondary-structure elements were determined using *STRIDE* (Heinig & Frishman, 2004). The structural alignment of TraKΔ internal monomers was conducted with *MASS* (Dror *et al.*, 2003). The *DALI* (Holm & Rosenström, 2010) and *MATRAS* (Kawabata, 2003) online structural alignment servers were utilized to search for structural homologues of TraKΔ. *PyMOL* (v.1.3; Schrödinger) was used to prepare structure representations, to calculate the r.m.s.d. of the TraKΔ monomer alignment and to generate the surface representation of the TraKΔ electrostatic potential.

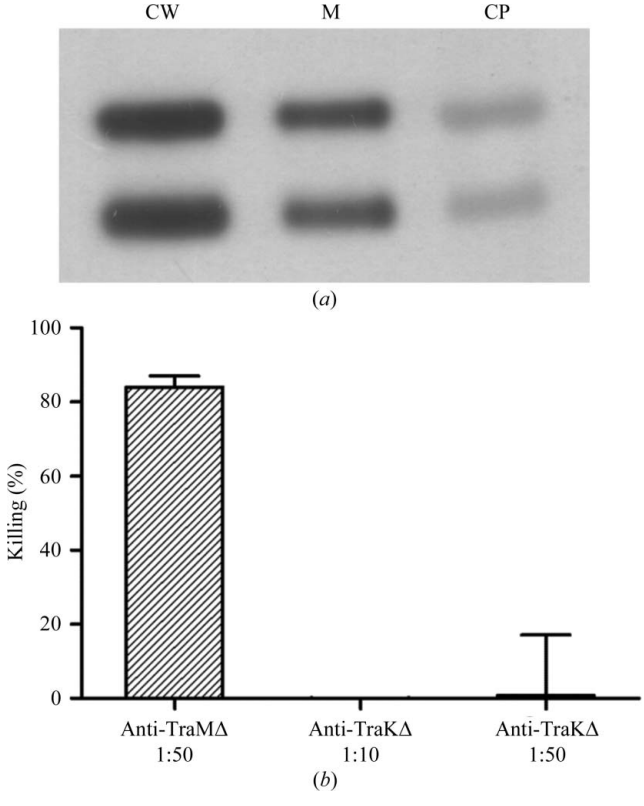

Figure 1
TraK localization and functional characterization. (*a*) TraK localizes to the cell envelope of *E. faecalis* JH2-2 cells harbouring pIP501. The localization of TraK in the cell fractions was detected by Western blot with polyclonal anti-TraKΔ antibodies. CW, cell wall; M, membrane; CP, cytoplasm. (*b*) Opsonophagocytic killing assay using anti-TraKΔ antibodies, as well as anti-TraMΔ antibodies as a positive control.

2.10. Mass spectrometry of TraKΔ crystals

Several crystals of TraKΔ were dissolved in 10 µl pure H_2O and investigated by MALDI–TOF mass-spectrometric analysis (ultrafleXtreme; Bruker, Vienna, Austria).

2.11. Sequence-based comparison and characterization

The following online services were used to search for transmembrane motifs in the TraK sequence and potential homologous proteins: *HMMTOP* (Tusnády & Simon, 2001), *MemsatSVM* (Nugent & Jones, 2009) and *Memsat*3 (Jones *et al.*, 1994).

PSIpred (Jones, 1999) was used to predict the secondary-structure content of TraK and of structurally related proteins, but where known the secondary structure was derived from the crystal structure. General features of the His-tagged TraK construct were assessed with *ProtParam* (Gasteiger *et al.*, 2003).

3. Results

3.1. TraK localizes to the *Enterococcus* cell envelope

To localize the TraK protein *in vivo*, an exponentially growing culture of *E. faecalis* JH2-2 (pIP501) was fractionated into cell-wall, membrane and cytoplasmic fractions as described by Buttaro *et al.* (2000). TraK was always visualized as a doublet and was mainly found in the cell-envelope fractions (cell wall and membrane; Fig. 1*a*), with only a weak signal in the cytoplasmic fraction. The TraN protein predicted by *PSORTb* to localize to the cytoplasm, which was analyzed in parallel in the same experiment, was exclusively found in the cytoplasmic fraction, consistent with a sound separation of cytoplasmic and cell-envelope proteins (Arends *et al.*, 2013).

A possible explanation for the second signal could be a second start codon with its own ribosomal binding site within the *traK* coding region. As the second gene product was also detected in the cell-envelope fractions, the N-terminal transmembrane motif needs to be present. A start codon at nucleotide position 3923 (GenBank AJ5058232) has an adequate distance to a ribosomal binding site at positions 3908–3912. The mass difference for the larger TraK protein (start codon at nucleotide position 3818) and the gene product starting at position 3923 accounts for 4.1 kDa and is in good agreement with the difference between the two signals on the gel. To evaluate whether the TraK double bands arise from proteolytic digestion or utilization of the second putative ribosomal binding site and the subsequent start codon, we constructed a mutant comprising a mutation in the first TraK start codon and utilized immunodetection to monitor the TraK expression profile. The experiment confirmed that the second

Table 1
Data-collection and processing statistics of scaled data.

Values in parentheses are for the highest resolution shell.

	Se derivative			Native
Data sets	MAD peak	MAD inflection point	MAD high-end remote	Used for molecular replacement and structure refinement
Beamline	X06DA (PXIII), SLS, Villigen, Switzerland	X06DA (PXIII), SLS, Villigen, Switzerland	X06DA (PXIII), SLS, Villigen, Switzerland	X06DA (PXIII), SLS, Villigen, Switzerland
Space group	*I*4	*I*4	*I*4	*I*4
Detector	MAR CCD	MAR CCD	MAR CCD	MAR CCD
Unit-cell parameters (Å, °)	$a = b = 113.59$, $c = 120.84$, $\alpha = \beta = \gamma = 90.00$	$a = b = 113.83$, $c = 121.05$, $\alpha = \beta = \gamma = 90.00$	$a = b = 113.92$, $c = 121.13$, $\alpha = \beta = \gamma = 90.00$	$a = b = 114.04$, $c = 120.52$, $\alpha = \beta = \gamma = 90.00$
Wavelength (Å)	0.9792	0.9797	0.9715	1.0
Resolution range (Å)	46.83–3.50 (3.69–3.50)	41.46–3.50 (3.69–3.50)	41.49–3.50 (3.69–3.50)	41.42–3.00 (3.16–3.00)
R_{meas} (%)	0.127 (0.455)	0.112 (0.434)	0.122 (0.500)	0.036 (0.323)
$\langle I/\sigma(I)\rangle$	15.4 (5.8)	18.8 (6.8)	17.43 (6.0)	12.3 (2.3)
No. of molecules in asymmetric unit	2	2	2	2
Matthews coefficient (Å^3 Da^{-1})	3.18			3.20
Solvent content (%)	61.40			61.60
Measured reflections	122640 (17925)	146350 (21439)	146164 (21317)	117365 (16933)
Unique reflections	9733 (1416)	9802 (1432)	9824 (1428)	15465 (2231)
Multiplicity	12.6 (12.7)	14.9 (15.0)	14.9 (14.9)	7.6 (7.6)
Completeness (%)	99.9 (100)	99.9 (100)	99.9 (100)	99.9 (100)

start codon is indeed being used (Supplementary Fig. S1*a*)[1], thus it is likely that expression of the *traK* gene leads to two gene products *in vivo*.

In order to unambiguously determine the orientation of the TraK protein in the *E. faecalis* membrane, we performed a protease protection assay. To this end, we generated protoplasts from native *E. faecalis* JH2-2 (pIP501) cells from which the peptidoglycan layer had been removed by lysozyme and mutanolysin treatment. As Supplementary Fig. S1(*b*) shows, TraK is proteolytically digested depending on the protease concentration, indicating that the C-terminal domain of TraK is positioned outside the cytoplasmic membrane. In contrast, the cytoplasmic T4SS component $TraN_{pIP501}$ is not digested by the protease, as it is obviously shielded by the intact cytoplasmic membrane of the *E. faecalis* protoplasts and is only degraded when the protoplasts are solubilized by Triton X-100 treatment (Supplementary Fig. S1*c*).

The opsonophagocytic killing assay showed no killing of *E. faecalis* JH2-2 (pIP501) cells with polyclonal antibodies raised against TraKΔ at dilutions of 1:10, 1:50 (Fig. 1*b*) and 1:100 (data not shown), in contrast to the positive control with anti-TraMΔ antibodies. We conclude that while the protein localizes to the bacterial cell envelope and faces the cell exterior, it is not surface-exposed. To evaluate whether the thick peptidoglycan layer of the *E. faecalis* cells was responsible for sterically preventing the anti-TraKΔ antibodies from binding to the protein, we conducted OPA experiments including pre-incubation of the *E. faecalis* cells with varying concentrations of peptidoglycan-degrading enzymes (lysozyme and mutanolysin). None of these setups resulted in increased killing of *E. faecalis* cells harbouring the pIP501 plasmid (D. Laverde-Gomez, T. Sakinc & E. Grohmann, unpublished data), thus surface accessibility of TraK is likely to be prevented by a different mechanism.

3.2. TraKΔ is a monomeric protein

Attempts to overexpress and purify full-length TraK (34.7 kDa) failed owing to solubility problems. Consequently, a stable truncation derivative, TraKΔ (30.6 kDa), was constructed and purified. It lacks the N-terminal putative transmembrane domain, but possesses an N-terminal 7×His tag. TraKΔ eluted from the gel-filtration column as a single peak (Supplementary Fig. S2*a*), indicative of a homogeneous species with an apparent molecular weight of 44.9 kDa. This value compares with the theoretical molecular weight of the His-tagged construct of 30.6 kDa, suggesting that TraKΔ is a monomer in solution.

The monodispersity of TraKΔ was evaluated by DLS. Ten measurements with constant baseline were merged, yielding a single peak with a calculated polydispersity of 13.2% and a hydrodynamic radius (R_h) of 3.6 nm (Supplementary Fig. S2*a*).

Purified TraKΔ is folded in solution and has a mixed α/β composition (Supplementary Fig. S3*a*). The amount of α-helices (29%) exceeds that of β-sheets (20%) (Supplementary Fig. S3*b*). The large proportion of unordered structure (31%) may result from flexible N-terminal or C-terminal parts. Temperature scans revealed that TraKΔ undergoes a transition at 75°C (Supplementary Fig. S3*c*), reaches a plateau at 95°C and is trapped in this state (*i.e.* no refolding during the down-scan).

In order to determine the oligomeric state and shape of TraKΔ in solution, SAXS measurements were performed. The measurements yielded an I_0 of 27.24, a radius of gyration (R_g) of 2.97 nm and a D_{max} of 10 nm, as calculated from the Guinier plot (data at 3.0 mg ml^{-1}) and the $p(r)$ function, respectively. From the I_0, we calculated the apparent molecular weight of TraKΔ in solution using BSA as a molecular-weight standard

[1] Supporting information has been deposited in the IUCr electronic archive (Reference: MN5045).

(Pavkov *et al.*, 2008). The value of 28.1 kDa is in good agreement with the theoretical molecular weight of TraKΔ (30.6 kDa). Calculating *ab initio* models from the scattering function, we observed an elongated particle, which may be owing to the flexible N-terminal end of TraKΔ including the unstructured 7×His tag.

3.3. X-ray data collection

Owing to the lack of structures with significant sequence similarity to TraKΔ, selenomethionine-containing TraKΔ crystals were used for structure solution by multiwavelength anomalous dispersion (MAD). A single selenomethionine-containing crystal showed a non-twinned pattern and diffracted to 3.5 Å resolution at the synchrotron. We performed a fluorescence scan to validate the presence of selenomethionine in the crystal and to define the optimal setup for anomalous data collection at the Se *K* absorption edge. A full MAD data set was collected at the peak (0.9792 Å), inflection (0.9797 Å) and high-end remote (0.9715 Å) wavelengths. A crystal-to-detector distance of 330 mm, an oscillation range of 1.0° and an exposure time of 1 s per image were chosen. 360 frames were taken at each of the three wavelengths.

The selenomethionine-derivative crystal belonged to space group *I*4, with unit-cell parameters a = 113.59, b = 113.59, c = 120.84 Å, $\alpha = \beta = \gamma = 90.00^\circ$ and two molecules per asymmetric unit. Tables 1 and 2 provide an overview of the data-collection and refinement statistics. The Matthews coefficient (Matthews, 1968) was calculated as 3.18 Å^3 Da^{-1}, with a solvent content of 61.4%. The MAD data yielded a preliminary model of TraKΔ, which was used as a template for molecular replacement with native data to 3.0 Å resolution. A data set of 360 frames was collected at a wavelength of 1 Å, with a crystal-to-detector distance of 320 mm, an oscillation range of 0.5° and an exposure time of 1 s per image. The native crystal was found to be isomorphous to the selenomethionine-derivative crystal, belonging to the same space group (*I*4) and with nearly identical unit-cell parameters (a = 114.04, b = 114.04, c = 120.52 Å, $\alpha = \beta = \gamma = 90.00^\circ$) and two molecules per asymmetric unit (Table 1).

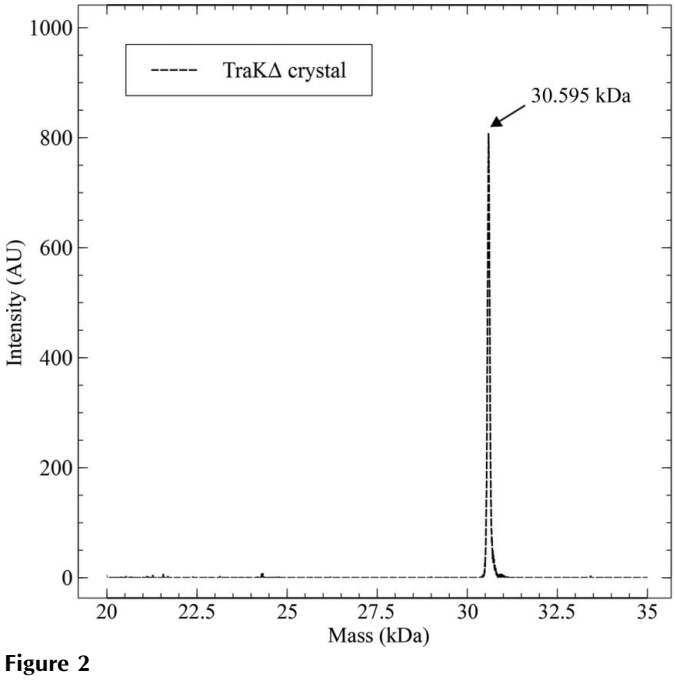

Figure 2
MALDI–TOF analysis of TraKΔ crystals. TraKΔ crystallized as a full-length His-tagged construct (30.6 kDa).

Table 2
Processing statistics of refined data.

Resolution (Å)	40.32–3.00
No. of reflections	15465
R_{work}/R_{free}	0.2050/0.2337
No. of atoms	
Protein	3310
Water	0
B factors (Å^2)	
Protein	59.79
Water	—
R.m.s. deviations	
Bond lengths (Å)	0.011
Bond angles (°)	1.577
Ramachandran outliers	6 (of 404)
Rotamer outliers	18 (of 366)
Ramachandran favoured (%)	94.06
MolProbity score	2.84 (84th percentile)

To confirm the integrity of TraKΔ in the crystals, we analyzed dissolved crystals *via* MALDI–TOF mass spectrometry (MS). The MS analysis showed that the TraKΔ crystals contained the full-length protein (Fig. 2). The N-terminal ends of the monomers, namely the 28-residue His tag and TraK residues 66–102, appeared to be flexible and were not observed in the electron-density maps. Supplementary Fig. S4 provides an overview of the full-length TraK protein, the TraKΔ construct, the predicted and the actual secondary-structure contents and the amino-acid sequence found in the TraKΔ crystals. The final coordinates and structure-factor amplitudes have been deposited in the PDB as entry 4hic.

3.4. The TraKΔ crystal structure

The crystal structure of TraKΔ represents a novel fold, as there were no significantly structurally related proteins found in the PDB as determined by *DALI* and *MATRAS* (Supplementary Table S1). TraKΔ comprises a mixed α/β fold with seven α-helices (h1–h7) and nine β-strands (s1–s9) (Fig. 3*a*). Two antiparallel β-sheets (A, s1–s5; B, s6–s9) are located in the central parts of the protein and are mostly protected from solvent contacts by α-helices and loop regions. Two of the helices (h4 and h5) are positioned in a parallel manner, dominating one side of the molecule. The curved β-sheet B is wrapped around a long α-helix (h2), which is stabilized on the surface of the protein solely through hydrophobic interactions. These interactions occur between the side chains of the helix 2 residues, which face towards the centre of the protein (Ala116, Ala119, Ala122, Phe123 and Trp126), and the side chains of the residues of β-strands 6–9 (Phe234, Trp236, Ile267, Met288, Trp297 and Val299), which are displayed on the concave surface of β-sheet B (Fig. 3*b*). The topology of the TraKΔ crystal structure is depicted in Fig. 3(*c*). The structure

harbours a surface area of 10 597 Å^2. When analyzing the surface electrostatic potential (Fig. 3*d*), we observed a mainly positively charged upper half of the protein (h3 and the loop region between s4 and s5) and a positively charged area in the middle of the protein (the unordered region between h7 and s7), as well as large negatively charged areas in the middle (s6 and the loop region between s7 and s8; the C-terminal residues and h6) and among the lower half of the protein (h2 near the N-terminus). The high degree of surface-exposed charged residues explains the high solubility of the protein which was observed during purification. The alignment of the two TraKΔ monomers showed an r.m.s.d. value of 0.33 Å, suggesting only minor structural differences.

3.5. The TraKΔ crystal structure fits to the SAXS solution data

The crystal structure was compared with the low-resolution solution structure by generating a theoretical SAXS curve from the refined coordinates using *CRYSOL*. The resulting curve showed significant differences from the experimental

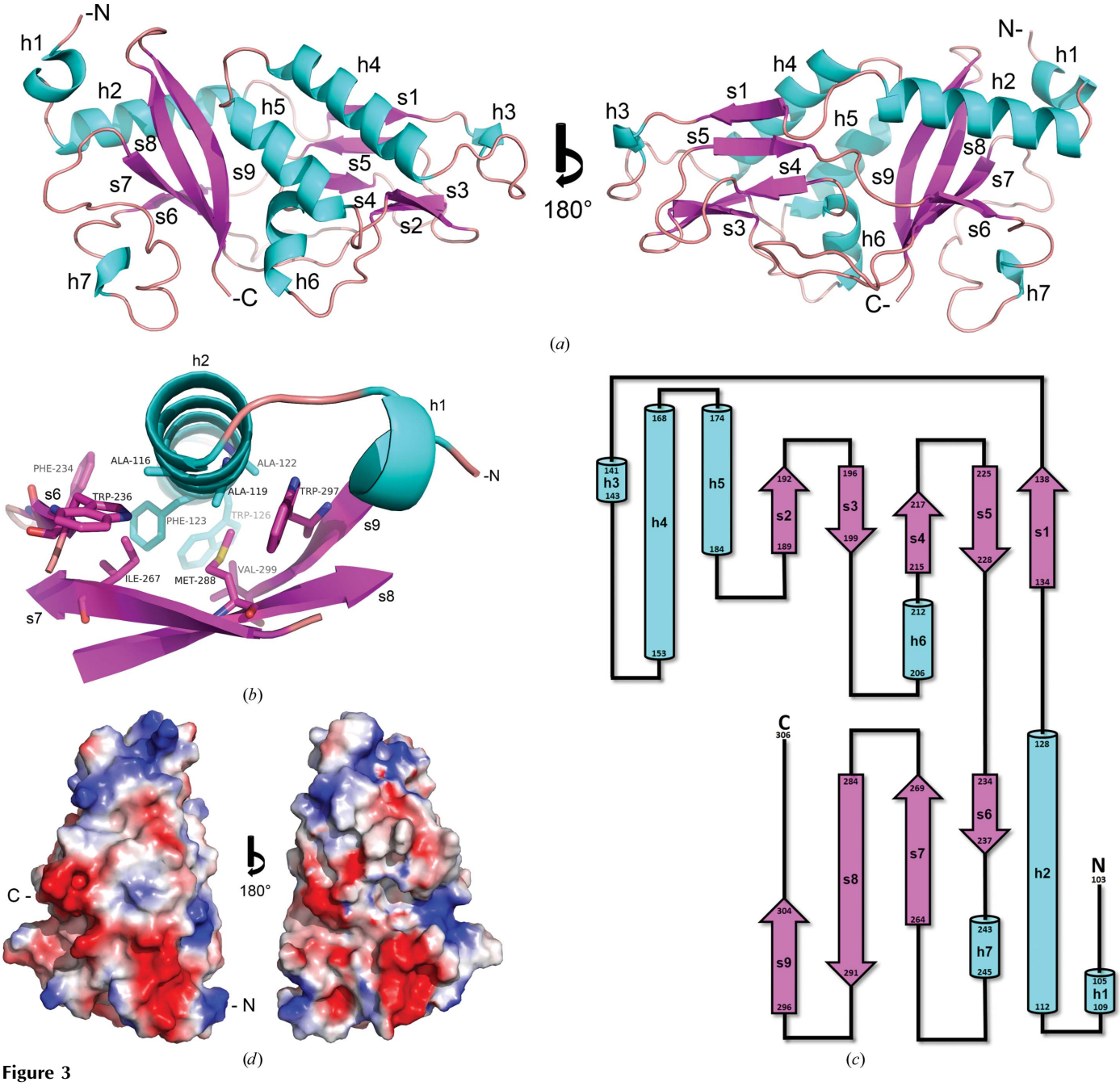

Figure 3
The structure of TraK$_{103–306}$. (*a*) Cartoon representation of TraKΔ with views onto the two parallel helices (h4 and h5) and rotated by 180° around the vertical axis; secondary-structure elements are highlighted (helices in cyan, strands in purple) and numbered. (*b*) Detailed view of the residues involved in the hydrophobic interaction between helix 2 and the antiparallel β-sheet B. (*c*) Topology representation of the TraKΔ fold. (*d*) Surface representation of the electrostatic potential of TraKΔ; front view and rotated by 180°; red, negative charge; blue, positive charge; the C-terminal and N-terminal ends of the protein are indicated.

observations (Fig. 4*a*). As the 7×His tag and TraK residues 66–102 were not assigned in the electron-density map, the refined TraK$\Delta_{103–306}$ structure was significantly smaller than the original construct. To take these differences in protein size into account, we used the program *BUNCH* to generate the missing N-terminal region by *ab initio* modelling and iterative fitting of the resulting model to the experimental data (Fig. 4*a*). 50 *ab initio* models were generated with *GASBOR* and averaged. To compare the size of the crystal structure with the SAXS-derived model of TraKΔ, we fitted the crystal structure to the SAXS model. Although the averaged SAXS model appears to be much more elongated than the crystal structure, TraK$\Delta_{103–306}$ fits very well into the main density (Fig. 4*b*). The N-terminal end of the crystal structure faces towards the empty tail. We conclude that the non-visible N-terminal residues of the TraKΔ construct are responsible for the elongated shape of the SAXS model.

3.6. TraK-like proteins were exclusively detected in *Enterococcus* and *Streptococcus* T4SSs

We performed an extended search for TraK-like proteins in a broad spectrum of conjugative plasmids, transposons, integrative conjugative elements (ICEs) and genetic islands from G− and G+ bacteria. The candidates were sorted according to their similarity to the template structure (TraK) based on secondary-structure prediction. The results of the prediction-based comparison and sequence alignments of potential TraK-like proteins can be found in detail in Supplementary Figs. S5 and S6. All of the analyzed proteins showing a TraK-like secondary-structure composition belonged to T4SSs of G+ origin (five plasmids and three ICEs). Except for one protein from *S. pyogenes* plasmid pSM19035, all candidate T4SS proteins were found in *Enterococcus* species. With the exception of protein P49 from *E. faecium* conjugative plasmid pVEF3 (261 amino acids), which lacks about 45 amino acids at its N-terminal end, all other TraK-like proteins have a length of approximately 300 residues and a small N-terminal α-helical domain, which is predicted to be cytoplasmic. Interestingly, in the case of the *E. faecalis* plasmid pRE25 (Schwarz *et al.*, 2001) two putative T4SS proteins exhibit high sequence identity to TraK, with ORF34 representing the N-terminal part and ORF35 the C-terminal part of TraK. It appears that the two ORFs have been generated from a single ancestor gene by a frame-shift mutation (Grohmann *et al.*, 2003), as the length and secondary-structure composition of TraK can be reconstructed by merging the amino-acid sequences of T4SS proteins ORF34 and ORF35.

4. Discussion

Prokaryotic genome plasticity is greatly increased by conjugative plasmid transfer. Consequently, the conjugative spread of antibiotic-resistance and virulence genes among pathogens and commensal bacteria has an enormous impact on human healthcare (Zechner *et al.*, 2012). Over the last decades, an increasing interest in the field of bacterial conjugation can be observed.

The T-DNA transfer system from *A. tumefaciens* is the prototype T4SS. It is the best-investigated model system and has been studied since the late 1970s (Gurley *et al.*, 1979). *A. tumefaciens* uses a T4SS to inject tumour-inducing factors into plant cells, a process that involves the transport of proteins and plasmid DNA alike (Cascales & Christie, 2003;

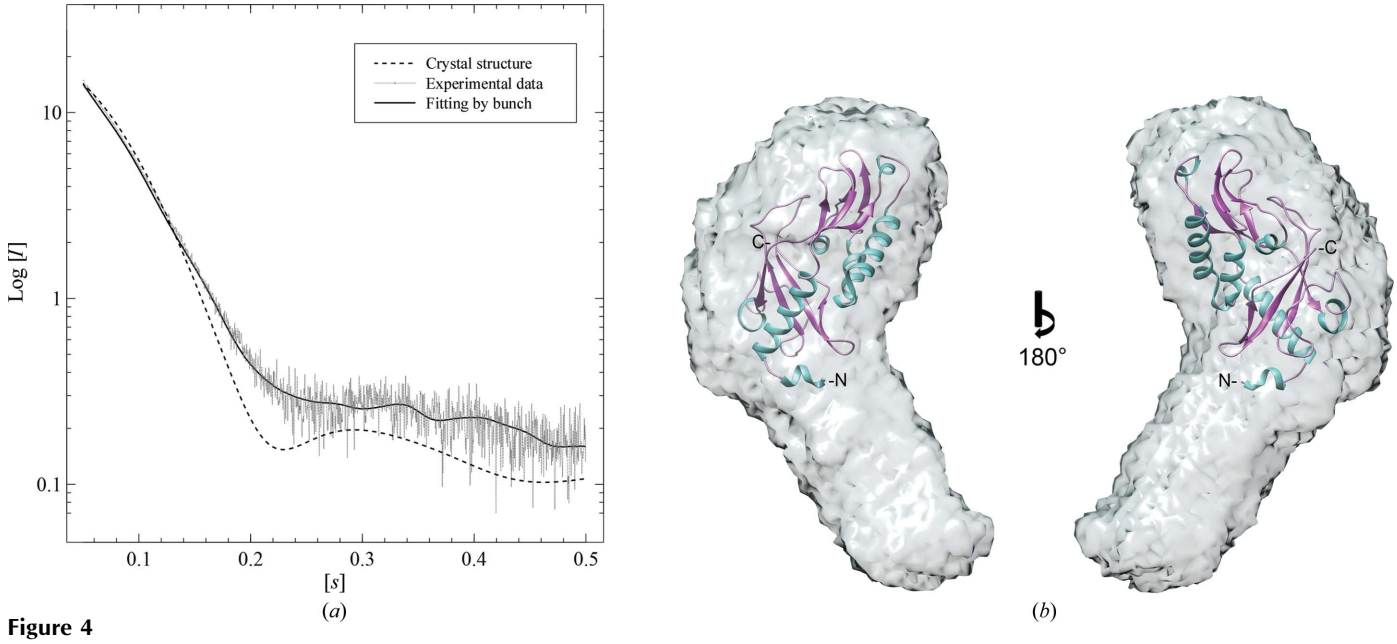

Figure 4
Comparison of the TraKΔ SAXS data and the crystal structure. (*a*) The calculated SAXS curve (dashed black line) from the TraKΔ X-ray structure does not match the experimental SAXS data (grey line); *ab initio* modelling of the missing N-terminal region of the TraMΔ construct allows a significantly improved fit of the data (black line). (*b*) Superposition of the TraKΔ crystal structure with the *GASBOR*-derived SAXS model; 50 *ab initio GASBOR* models were averaged but not filtered.

Nagai & Roy, 2003; Backert & Meyer, 2006). In the last 15 years, more and more T4SSs of G− origin have been studied, leading to a substantial amount of biophysical, functional and structural information. Among the solved protein structures are the VirD4 homologue TrwB from *E. coli* plasmid R388 (Gomis-Rüth *et al.*, 2001) and the VirB5 homologue TraC from *E. coli* plasmid pKM101 (Yeo *et al.*, 2003), as well as VirB11 from *Brucella suis* (Hare *et al.*, 2006) and VirB11 from the Cag pathogenicity island of *Helicobacter pylori* (Yeo *et al.*, 2003). Other examples are the VirB8 proteins from *A. tumefaciens* (Bailey *et al.*, 2006) and *B. suis* (Terradot *et al.*, 2005). Beside these individual proteins, the structure elucidation of the core complex of the pKM101 T4SS from *E. coli* considerably contributed to the knowledge base of G− T4S (Chandran *et al.*, 2009; Fronzes *et al.*, 2009). The combination of electron microscopic and crystallographic methods provided a detailed view of the assembly and partial architecture of the conjugative transfer apparatus.

Equal advances have not yet been achieved for systems originating from G+ bacteria. However, it is promising to see the amount of structural information on G+ T4SSs grow over recent years. In early 2012, the high-resolution structure of the VirB8-like transfer protein TcpC from *C. perfringens* plasmid pCW3 (Porter *et al.*, 2012) became available, shortly followed by the structure of VirB4 from *Thermoanaerobacter pseudethanolicus* (Walldén *et al.*, 2012). In late 2012, Goessweiner-Mohr and coworkers published the VirB8-like structure of the TraM C-terminal domain, the first transfer protein structure solved from the pIP501-encoded enterococcal T4SS (Goessweiner-Mohr, Grumet, Arends *et al.*, 2013). As the cell-wall structures of G− and G+ bacteria differ significantly, reflected by the vastly different makeup of G+ T4SSs, the major components of the conjugative core complex have not yet been identified in G+ T4SSs. Thus, the molecular mechanisms of DNA transfer in G+ bacteria remain largely unknown. Many important human pathogens, such as enterococci, streptococci and staphylococci, belong to this group of prokaryotes. Hence, the lack of knowledge about G+ T4S is particularly concerning (Burns, 2003).

In this study, structural and biophysical approaches were used to characterize TraK, a putative T4SS protein from the *E. faecalis* conjugative model plasmid pIP501. This task was especially demanding, as no sequence similarities of TraK to T4SS components of G− origin have been detected. We showed that TraK is a cell envelope-located protein (Fig. 1*a*) with its C-terminal domain facing the cell exterior (Supplementary Fig. S1*b*). However, antibodies directed against TraKΔ were not able to recruit macrophages to pIP501-harbouring *E. faecalis* cells. We conclude that although TraK is located in the bacterial cell envelope, it is not surface-exposed, in contrast to the previously reported pIP501 T4SS protein TraM (Goessweiner-Mohr, Grumet, Arends *et al.*, 2013). A possible explanation for these findings is that TraK might not be positioned near the opening in the peptidoglycan layer locally generated by the pIP501-encoded muramidase TraG for the assembly of the conjugative T4SS core complex (Arends *et al.*, 2013). In this scenario, the anti-TraKΔ antibodies would not be able to access the protein surface and thus would not be able to recruit macrophages to the *E. faecalis* cells. As OPAs including peptidoglycan-degrading enzymes were negative, we conclude that the surface accessibility of TraK is likely to be prevented by a different mechanism. A role as an exclusion factor or its involvement in cell-to-cell attachment, as suggested for the *E. faecalis* pCF10-encoded PrgA and PrgB (Alvarez-Martinez & Christie, 2009), is unlikely, since these tasks would clearly require a surface-exposed location.

A second explanation would be that TraK might be an integral component of the core complex. In the case of the pKM101 core complex, three proteins homologous to VirB7, VirB9 and VirB10 are part of the 172 Å diameter ring-shaped structure (Fronzes *et al.*, 2009; Rivera-Calzada *et al.*, 2013). One of the components, VirB9, is covered by the elongated molecule VirB7, which substantially limits its accessible surface area. Similarly, as yet undetected interactions of TraK with other components of the putative pIP501 core complex might also restrict its accessible surface and could potentially exclude stable binding of the antibodies to their respective surface epitopes. Previous studies using yeast two-hybrid and pull-down assays have not shown any interactions of TraK with other pIP501 T4SS components (Abajy *et al.*, 2007). Nonetheless, the elucidation of the pKM101 core complex buildup, *i.e.* the interaction of its components, was only made possible by their co-expression (Fronzes *et al.*, 2009; Rivera-Calzada *et al.*, 2013). We expect a similar behaviour for the pIP501 core complex proteins; co-expression tests have recently been started.

It is surprising that TraK only shows a very limited number of structural relatives. Furthermore, only some of the known enterococcal T4SSs encode a putative transfer protein with a TraK-like secondary-structure composition, *e.g.* the well characterized enterococcal T4SS encoded by the conjugative plasmid pCF10 lacks a corresponding gene product (Bhatty *et al.*, 2013). This suggests an exclusive role for TraK-like proteins in the respective T4SSs, in contrast to the prevalence of the major T4SS proteins: relaxases (Kopec *et al.*, 2005; Garcillán-Barcia *et al.*, 2009; Lang *et al.*, 2010), coupling proteins (Llosa *et al.*, 2003; Gomis-Rüth *et al.*, 2004; Parsons *et al.*, 2007), transglycosylases (Koraimann, 2003; Zahrl *et al.*, 2005; Bantwal *et al.*, 2012; Arends *et al.*, 2013) and ATPases (Savvides, 2007), as well as VirB8-like proteins (Baron, 2006; Porter *et al.*, 2012; Goessweiner-Mohr, Grumet, Arends *et al.*, 2013). One explanation for the exclusive distribution of TraK-like proteins might be the specific host range of the T4SS. pIP501, as well as pRE25, owing to their nearly identical buildup (Schwarz *et al.*, 2001), exhibit a particular broad host range, with pIP501 even able to transfer and stably replicate in bacteria of G− origin (Kurenbach *et al.*, 2003).

We suggest that the TraK-like proteins might play an important role in the adaptation of their respective T4SSs to new hosts. Most likely, this adaptation involves adjustment to the distinct cell-envelope compositions (Vollmer & Seligman, 2010) of the various types of G+ bacteria. Interestingly, close homologues of the pIP501 T4SS protein TraN

were only found in the same range of enterococcal and streptococcal T4SSs (N. Goessweiner-Mohr & W. Keller, unpublished data), which suggests a connected function for the two proteins.

Despite the growing structural and functional information on T4SSs in general, further efforts are needed to identify the function of TraK-like proteins, as well as to reveal the components of the G+ T4SS core complex. Nevertheless, we are confident that the structural elucidation of TraK will prove to be a keystone in the growing understanding of T4S, as well as in deciphering the structural differences and adaptations in T4SSs among G+ bacteria.

This work was supported by Austrian Science Fund (FWF) projects P19794 and F4604. Staff support during data collection on the X06DA beamline at the SLS synchrotron and the X33 SAXS beamline at DESY is gratefully acknowledged. EG was supported by the European Union Sixth Framework Program 'Approaches to Control Multi-resistant Enterococci ACE' (LSHE-CT-2007-037410).

References

Abajy, M. Y., Kopeć, J., Schiwon, K., Burzynski, M., Döring, M., Bohn, C. & Grohmann, E. (2007). *J. Bacteriol.* **189**, 2487–2496.

Adams, P. D. *et al.* (2010). *Acta Cryst.* D**66**, 213–221.

Alvarez-Martinez, C. E. & Christie, P. J. (2009). *Microbiol. Mol. Biol. Rev.* **73**, 775–808.

Arends, K., Celik, E. K., Probst, I., Goessweiner-Mohr, N., Fercher, C., Grumet, L., Soellue, C., Abajy, M. Y., Sakinc, T., Broszat, M., Schiwon, K., Koraimann, G., Keller, W. & Grohmann, E. (2013). *J. Bacteriol.* **195**, 4436–4444.

Backert, S. & Meyer, T. F. (2006). *Curr. Opin. Microbiol.* **9**, 207–217.

Bailey, S., Ward, D., Middleton, R., Grossmann, J. G. & Zambryski, P. C. (2006). *Proc. Natl Acad. Sci. USA*, **103**, 2582–2587.

Bannam, T. L., Teng, W. L., Bulach, D., Lyras, D. & Rood, J. I. (2006). *J. Bacteriol.* **188**, 4942–4951.

Bantwal, R., Bannam, T. L., Porter, C. J., Quinsey, N. S., Lyras, D., Adams, V. & Rood, J. I. (2012). *Plasmid*, **67**, 139–147.

Baron, C. (2006). *Biochem. Cell Biol.* **84**, 890–899.

Battye, T. G. G., Kontogiannis, L., Johnson, O., Powell, H. R. & Leslie, A. G. W. (2011). *Acta Cryst.* D**67**, 271–281.

Bhatty, M., Laverde Gomez, J. A. & Christie, P. J. (2013). *Res. Microbiol.* **164**, 620–639.

Burns, D. L. (2003). *Curr. Opin. Microbiol.* **6**, 29–34.

Buttaro, B. A., Antiporta, M. H. & Dunny, G. M. (2000). *J. Bacteriol.* **182**, 4926–4933.

Cascales, E. & Christie, P. J. (2003). *Nature Rev. Microbiol.* **1**, 137–149.

Chandran, V., Fronzes, R., Duquerroy, S., Cronin, N., Navaza, J. & Waksman, G. (2009). *Nature (London)*, **462**, 1011–1015.

Chayen, N. E., Shaw Stewart, P. D. & Blow, D. M. (1992). *J. Cryst. Growth*, **122**, 176–180.

Chen, V. B., Arendall, W. B., Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S. & Richardson, D. C. (2010). *Acta Cryst.* D**66**, 12–21.

Chen, Y., Zhang, X., Manias, D., Yeo, H.-J., Dunny, G. M. & Christie, P. J. (2008). *J. Bacteriol.* **190**, 3632–3645.

Christie, P. J., Whitaker, N. & González-Rivera, C. (2014). *Biochim. Biophys. Acta*, doi:10.1016/j.bbamcr.2013.12.019.

Collins, B. K., Tomanicek, S. J., Lyamicheva, N., Kaiser, M. W. & Mueser, T. C. (2004). *Acta Cryst.* D**60**, 1674–1678.

Cruz, F. de la, Frost, L. S., Meyer, R. J. & Zechner, E. L. (2010). *FEMS Microbiol. Rev.* **34**, 18–40.

Dror, O., Benyamini, H., Nussinov, R. & Wolfson, H. (2003). *Bioinformatics*, **19**, i95–i104.

Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. (2010). *Acta Cryst.* D**66**, 486–501.

Ericsson, U. B., Hallberg, B. M., DeTitta, G. T., Dekker, N. & Nordlund, P. (2006). *Anal. Biochem.* **357**, 289–298.

Evans, P. (2006). *Acta Cryst.* D**62**, 72–82.

Fronzes, R., Schäfer, E., Wang, L., Saibil, H. R., Orlova, E. V. & Waksman, G. (2009). *Science*, **323**, 266–268.

Garcillán-Barcia, M. P., Francia, M. V. & de la Cruz, F. (2009). *FEMS Microbiol. Rev.* **33**, 657–687.

Gasteiger, E., Gattiker, A., Hoogland, C., Ivanyi, I., Appel, R. D. & Bairoch, A. (2003). *Nucleic Acids Res.* **31**, 3784–3788.

Goessweiner-Mohr, N., Arends, K., Keller, W. & Grohmann, E. (2013). *Plasmid*, **70**, 289–302.

Goessweiner-Mohr, N., Fercher, C., Abajy, M. Y., Grohmann, E. & Keller, W. (2012). *Acta Cryst.* F**68**, 1402–1405.

Goessweiner-Mohr, N., Grumet, L., Arends, K., Pavkov-Keller, T., Gruber, C. C., Gruber, K., Birner-Gruenberger, R., Kropec-Huebner, A., Huebner, J., Grohmann, E. & Keller, W. (2013). *J. Biol. Chem.* **288**, 2018–2028.

Goessweiner-Mohr, N., Grumet, L., Pavkov-Keller, T., Birner-Gruenberger, R., Grohmann, E. & Keller, W. (2013). *Acta Cryst.* F**69**, 178–183.

Gomis-Rüth, F. X., Moncalián, G., Pérez-Luque, R., González, A., Cabezón, E., de la Cruz, F. & Coll, M. (2001). *Nature (London)*, **409**, 637–641.

Gomis-Rüth, F. X., Solà, M., de la Cruz, F. & Coll, M. (2004). *Curr. Pharm. Des.* **10**, 1551–1565.

Grohmann, E., Muth, G. & Espinosa, M. (2003). *Microbiol. Mol. Biol. Rev.* **67**, 277–301.

Gurley, W. B., Kemp, J. D., Albert, M. J., Sutton, D. W. & Callis, J. (1979). *Proc. Natl Acad. Sci. USA*, **76**, 2828–2832.

Hare, S., Bayliss, R., Baron, C. & Waksman, G. (2006). *J. Mol. Biol.* **360**, 56–66.

Hayes, C. S., Aoki, S. K. & Low, D. A. (2010). *Annu. Rev. Genet.* **44**, 71–90.

Heinig, M. & Frishman, D. (2004). *Nucleic Acids Res.* **32**, W500–W502.

Holm, L. & Rosenström, P. (2010). *Nucleic Acids Res.* **38**, W545–W549.

Horodniceanu, T., Bougueleret, L., El-Solh, N., Bouanchaud, D. H. & Chabbert, Y. A. (1979). *Plasmid*, **2**, 197–206.

Jones, D. T. (1999). *J. Mol. Biol.* **292**, 195–202.

Jones, D. T., Taylor, W. R. & Thornton, J. M. (1994). *Biochemistry*, **33**, 3038–3049.

Kawabata, T. (2003). *Nucleic Acids Res.* **31**, 3367–3369.

Konarev, P. V., Volkov, V. V., Sokolova, A. V., Koch, M. H. J. & Svergun, D. I. (2003). *J. Appl. Cryst.* **36**, 1277–1282.

Kopec, J., Bergmann, A., Fritz, G., Grohmann, E. & Keller, W. (2005). *Biochem. J.* **387**, 401–409.

Koraimann, G. (2003). *Cell. Mol. Life Sci.* **60**, 2371–2388.

Kurenbach, B., Bohn, C., Prabhu, J., Abudukerim, M., Szewzyk, U. & Grohmann, E. (2003). *Plasmid*, **50**, 86–93.

Kurenbach, B., Kopeć, J., Mägdefrau, M., Andreas, K., Keller, W., Bohn, C., Abajy, M. Y. & Grohmann, E. (2006). *Microbiology*, **152**, 637–645.

Lang, S., Gruber, K., Mihajlovic, S., Arnold, R., Gruber, C. J., Steinlechner, S., Jehl, M. A., Rattei, T., Fröhlich, K. U. & Zechner, E. L. (2010). *Mol. Microbiol.* **78**, 1539–1555.

Li, F., Alvarez-Martinez, C., Chen, Y., Choi, K.-J., Yeo, H.-J. & Christie, P. J. (2012). *J. Bacteriol.* **194**, 4041–4051.

Li, M., Shen, X., Yan, J., Han, H., Zheng, B., Liu, D., Cheng, H., Zhao, Y., Rao, X., Wang, C., Tang, J., Hu, F. & Gao, G. F. (2011). *Mol. Microbiol.* **79**, 1670–1683.

Llosa, M., Gomis-Rüth, F. X., Coll, M. & de la Cruz, F. (2002). *Mol. Microbiol.* **45**, 1–8.

Llosa, M., Roy, C. & Dehio, C. (2009). *Mol. Microbiol.* **73**, 141–151.

Llosa, M., Zunzunegui, S. & de la Cruz, F. (2003). *Proc. Natl Acad. Sci. USA*, **100**, 10465–10470.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. (2007). *J. Appl. Cryst.* **40**, 658–674.
Murshudov, G. N., Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F. & Vagin, A. A. (2011). *Acta Cryst.* D**67**, 355–367.
Nagai, H. & Roy, C. R. (2003). *Cell. Microbiol.* **5**, 373–383.
Nugent, T. & Jones, D. T. (2009). *BMC Bioinformatics*, **10**, 159.
Parsons, J. A., Bannam, T. L., Devenish, R. J. & Rood, J. I. (2007). *J. Bacteriol.* **189**, 7782–7790.
Pavkov, T., Egelseer, E. M., Tesarz, M., Svergun, D. I., Sleytr, U. B. & Keller, W. (2008). *Structure*, **16**, 1226–1237.
Petoukhov, M. V. & Svergun, D. I. (2005). *Biophys. J.* **89**, 1237–1250.
Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C. & Ferrin, T. E. (2004). *J. Comput. Chem.* **25**, 1605–1612.
Porter, C. J., Bantwal, R., Bannam, T. L., Rosado, C. J., Pearce, M. C., Adams, V., Lyras, D., Whisstock, J. C. & Rood, J. I. (2012). *Mol. Microbiol.* **83**, 275–288.
Rêgo, A. T., Chandran, V. & Waksman, G. (2010). *Biochem. J.* **425**, 475–488.
Rivera-Calzada, A., Fronzes, R., Savva, C. G., Chandran, V., Lian, P. W., Laeremans, T., Pardon, E., Steyaert, J., Remaut, H., Waksman, G. & Orlova, E. V. (2013). *EMBO J.* **32**, 1195–1204.
Savvides, S. N. (2007). *Structure*, **15**, 255–257.
Schwarz, F. V., Perreten, V. & Teuber, M. (2001). *Plasmid*, **46**, 170–187.
Smillie, C., Garcillán-Barcia, M. P., Francia, M. V., Rocha, E. P. C. & de la Cruz, F. (2010). *Microbiol. Mol. Biol. Rev.* **74**, 434–452.
Steen, J. A., Bannam, T. L., Teng, W. L., Devenish, R. J. & Rood, J. I. (2009). *J. Bacteriol.* **191**, 2926–2933.
Svergun, D. I. (1992). *J. Appl. Cryst.* **25**, 495–503.
Svergun, D., Barberato, C. & Koch, M. H. J. (1995). *J. Appl. Cryst.* **28**, 768–773.
Svergun, D. I., Petoukhov, M. V. & Koch, M. H. J. (2001). *Biophys. J.* **80**, 2946–2953.
Teng, W. L., Bannam, T. L., Parsons, J. A. & Rood, J. I. (2008). *J. Bacteriol.* **190**, 5075–5086.
Terradot, L., Bayliss, R., Oomen, C., Leonard, G. A., Baron, C. & Waksman, G. (2005). *Proc. Natl Acad. Sci. USA*, **102**, 4596–4601.
Terwilliger, T. C., Adams, P. D., Read, R. J., McCoy, A. J., Moriarty, N. W., Grosse-Kunstleve, R. W., Afonine, P. V., Zwart, P. H. & Hung, L.-W. (2009). *Acta Cryst.* D**65**, 582–601.
Terwilliger, T. C., Grosse-Kunstleve, R. W., Afonine, P. V., Moriarty, N. W., Zwart, P. H., Hung, L.-W., Read, R. J. & Adams, P. D. (2008). *Acta Cryst.* D**64**, 61–69.
Thanassi, D. G., Bliska, J. B. & Christie, P. J. (2012). *FEMS Microbiol. Rev.* **36**, 1046–1082.
Theilacker, C., Kropec, A., Hammer, F., Sava, I., Wobser, D., Sakinc, T., Codée, J. D. C., Hogendorf, W. F. J., van der Marel, G. A. & Huebner, J. (2012). *J. Infect. Dis.* **205**, 1076–1085.
Tusnády, G. E. & Simon, I. (2001). *Bioinformatics*, **17**, 849–850.
Vollmer, W. & Seligman, S. J. (2010). *Trends Microbiol.* **18**, 59–66.
Wallden, K., Rivera-Calzada, A. & Waksman, G. (2010). *Cell. Microbiol.* **12**, 1203–1212.
Walldén, K., Williams, R., Yan, J., Lian, P. W., Wang, L., Thalassinos, K., Orlova, E. V. & Waksman, G. (2012). *Proc. Natl Acad. Sci. USA*, **109**, 11348–11353.
Whitmore, L. & Wallace, B. A. (2008). *Biopolymers*, **89**, 392–400.
Williams, J. J. & Hergenrother, P. J. (2008). *Curr. Opin. Chem. Biol.* **12**, 389–399.
Wriggers, W. (2010). *Biophys Rev.* **2**, 21–27.
Yeo, H.-J., Yuan, Q., Beck, M. R., Baron, C. & Waksman, G. (2003). *Proc. Natl Acad. Sci. USA*, **100**, 15947–15952.
Zahrl, D., Wagner, M., Bischof, K., Bayer, M., Zavecz, B., Beranek, A., Ruckenstuhl, C., Zarfel, G. E. & Koraimann, G. (2005). *Microbiology*, **151**, 3455–3467.
Zechner, E. L., Lang, S. & Schildbach, J. F. (2012). *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **367**, 1073–1087.