

# SCIENTIFIC REPORTS



OPEN

## Structural and functional probing of PorZ, an essential bacterial surface component of the type-IX secretion system of human oral-microbiomic *Porphyromonas gingivalis*.

Received: 23 August 2016  
Accepted: 03 November 2016  
Published: 24 November 2016

Anna M. Lasica<sup>1,2,\*</sup>, Theodoros Goulas<sup>3,\*</sup>, Danuta Mizgalska<sup>4,\*</sup>, Xiaoyan Zhou<sup>5</sup>, Iñaki de Diego<sup>3,†</sup>, Mirosław Ksiazek<sup>4</sup>, Mariusz Madej<sup>4</sup>, Yonghua Guo<sup>1,4</sup>, Tibisay Guevara<sup>3</sup>, Magdalena Nowak<sup>4</sup>, Barbara Potempa<sup>1</sup>, Apoorv Goel<sup>1,4,‡</sup>, Maryta Sztukowska<sup>1</sup>, Apurva T. Prabhakar<sup>1,4,§</sup>, Monika Bzowska<sup>6</sup>, Magdalena Widziolek<sup>4</sup>, Ida B. Thøgersen<sup>7</sup>, Jan J. Enghild<sup>7</sup>, Mary Simonian<sup>8</sup>, Arkadiusz W. Kulczyk<sup>9</sup>, Ky-Anh Nguyen<sup>5,8</sup>, Jan Potempa<sup>1,4,10</sup> & F. Xavier Gomis-Rüth<sup>3</sup>

*Porphyromonas gingivalis* is a member of the human oral microbiome abundant in dysbiosis and implicated in the pathogenesis of periodontal (gum) disease. It employs a newly described type-IX secretion system (T9SS) for secretion of virulence factors. Cargo proteins destined for secretion through T9SS carry a recognition signal in the conserved C-terminal domain (CTD), which is removed by sortase PorU during translocation. Here, we identified a novel component of T9SS, PorZ, which is essential for surface exposure of PorU and posttranslational modification of T9SS cargo proteins. These include maturation of enzyme precursors, CTD removal and attachment of anionic lipopolysaccharide for anchorage in the outer membrane. The crystal structure of PorZ revealed two  $\beta$ -propeller domains and a C-terminal  $\beta$ -sandwich domain, which conforms to the canonical CTD architecture. We further documented that PorZ is itself transported to the cell surface *via* T9SS as a full-length protein with its CTD intact, independently of the presence or activity of PorU. Taken together, our results shed light on the architecture and possible function of a novel component of the T9SS. Knowledge of how T9SS operates will contribute to our understanding of protein secretion as part of host-microbiome interactions by dysbiotic members of the human oral cavity.

<sup>1</sup>Department of Oral Immunology and Infectious Diseases, University of Louisville School of Dentistry, Louisville, KY, USA. <sup>2</sup>Department of Bacterial Genetics, Institute of Microbiology, Faculty of Biology, University of Warsaw, Warsaw, Poland. <sup>3</sup>Proteolysis Lab, Structural Biology Unit ("Maria-de-Maeztu" Unit of Excellence), Molecular Biology Institute of Barcelona (CSIC), Barcelona Science Park, Barcelona, Catalonia, Spain. <sup>4</sup>Department of Microbiology, Faculty of Biochemistry, Biophysics, and Biotechnology, Jagiellonian University, Krakow, Poland. <sup>5</sup>Department of Oral Biology, Faculty of Dentistry, University of Sydney, Sydney, NSW 2006, Australia. <sup>6</sup>Department of Cell Biochemistry, Faculty of Biochemistry, Biophysics, and Biotechnology, Jagiellonian University, Krakow, Poland. <sup>7</sup>Interdisciplinary Nanoscience Center (iNANO), and the Department of Molecular Biology, Aarhus University, Århus, DK-8000, Denmark. <sup>8</sup>Institute of Dental Research, Westmead Centre for Oral Health and Westmead Institute for Medical Research, Sydney, NSW 2145, Australia. <sup>9</sup>Department of Biological Chemistry and Molecular Pharmacology, Harvard University Medical School, Boston, MA, USA. <sup>10</sup>Małopolska Center of Biotechnology, Jagiellonian University, Krakow, Poland. <sup>†</sup>Present address: ALBA Synchrotron, Cerdanyola del Vallés, Catalonia, Spain. <sup>‡</sup>Present address: Indiana University School of Dentistry, Graduate Periodontics, Indianapolis, IN, USA. <sup>§</sup>Present address: Oral Health Research, School of Dentistry, Virginia Commonwealth University Richmond, VA, USA. \*These authors contributed equally to this work. Correspondence and requests for materials should be addressed to K.A.N. (email: ky-anh.nguyen@sydney.edu.au) or J.P. (email: jan.potempa@louisville.edu) or F.X.G.R. (email: xgrcri@ibmb.csic.es)

As part of host-microbiome interactions, resident bacteria secrete proteins, lipopolysaccharides, and peptidoglycan into the extracellular environment to facilitate antibiotic resistance, deterrence of host immune defenses, attachment, detoxification and nutrient acquisition. This helps them to flourish in a densely populated, highly competitive environment<sup>1,2</sup>. In diderm prokaryotes, represented mostly by Gram-negative bacteria, secreted proteins are synthesized in the cytoplasm and then translocated across two lipid bilayers: the inner (IM) and outer (OM) membranes, which have a periplasmic space between them. To achieve this, diderm bacteria have so far evolved nine known types (I to IX) of secretion systems (TxSS)<sup>1,2,3</sup>. Using such systems, they can assemble cell-surface appendages such as pili (mediated by T2SS, T4SS and T7SS), curli (T8SS), and flagella (T3SS); secrete proteins to the extracellular space (T1SS to T6SS); and inject proteins into eukaryotic host cells (T3SS and T4SS) or the periplasm of other bacteria (T6SS). Proteins to be translocated are either directly secreted from the cytoplasm via T1SS, T3SS, T4SS and T6SS or are first exported to the periplasm across the IM using conserved *Sec* or *Tat* pathways and then translocated through the OM using T2SS, T5SS, T7SS or T8SS<sup>5</sup>. The most recently discovered system of protein secretion is T9SS, also called *Por secretion system*, which operates exclusively in selected species within the Bacteroidetes phylum<sup>4–9</sup>. The Bacteroidetes and Firmicutes phyla accounts for +99% of species and phylotypes of the human gut microbiome<sup>10</sup>. In the oral microbiome, members that colonize dental plaque in periodontal disease such as *Porphyromonas gingivalis* and *Tannerella forsythia* also possess T9SS. These bacteria dominate the dysbiotic biofilm when the commensal microbiome is disrupted. Together with *Treponema denticola*, they give rise to the “red complex” microbial consortium commonly found in severe periodontal disease<sup>11</sup>. T9SS is essential for the secretion of many proteinaceous virulence factors by *P. gingivalis*, and has, thus, attracted considerable attention as a possible target for pharmaceutical intervention to treat severe periodontal disease and restore homeostasis of the oral microbiome<sup>4,12</sup>.

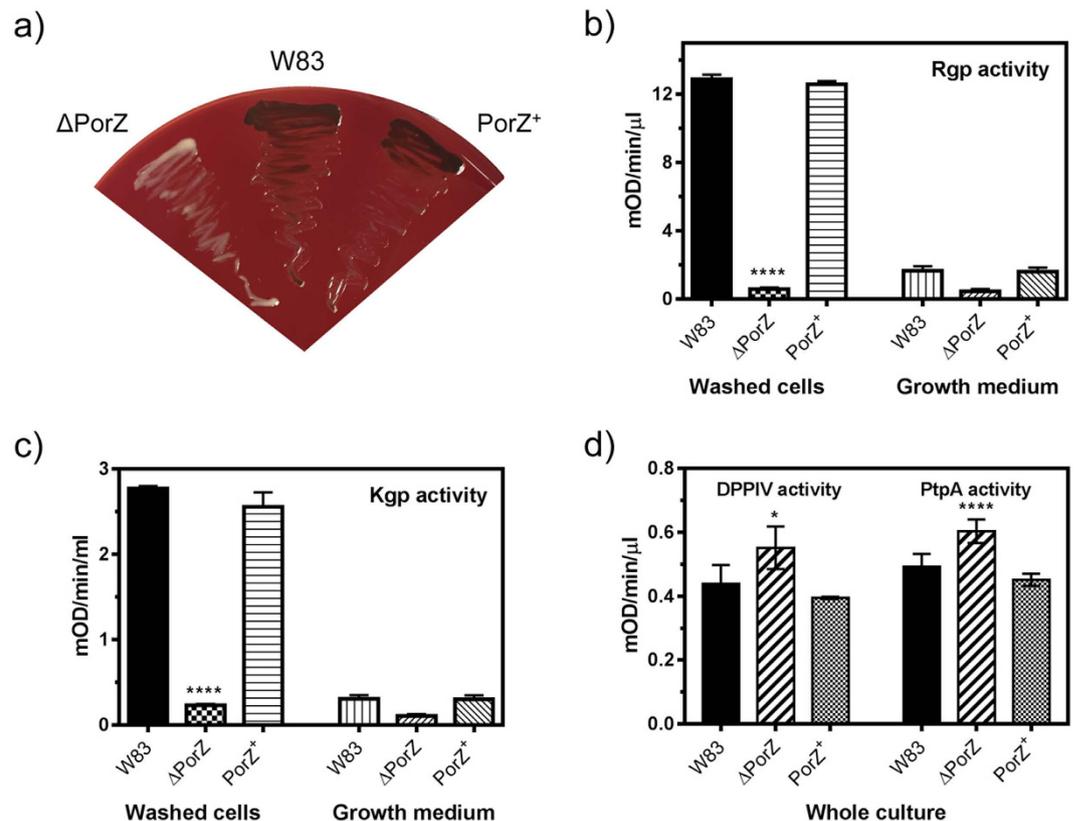
All cargo proteins of *P. gingivalis* T9SS contain a conserved C-terminal domain (CTD) of ~70 residues required for export and attachment to the cell surface<sup>13–15</sup>, which was recently shown to adopt an antiparallel seven-stranded immunoglobulin-like architecture<sup>12</sup>. T9SS cargo proteins carry a typical cleavable N-terminal signal peptide for export across the IM to the periplasm using the *Sec* system. Proteins fold in the periplasm, and are then directed to the T9SS translocon by a targeting signal located in the last two  $\beta$ -strands of CTD<sup>12,15</sup>. During this translocation, CTD is cleaved off<sup>16</sup> and the protein is released extracellularly<sup>17</sup>. At least in selected proteins of *P. gingivalis*, CTD removal occurs concurrently with covalent attachment of anionic lipopolysaccharide (A-LPS) to the newly released C-terminal carboxylate of the processed protein<sup>16</sup>. A-LPS attachment serves to anchor secreted proteins to the OM where they form an electron-dense surface layer, which is characteristic of *P. gingivalis*<sup>18</sup>.

To date, 12 indispensable T9SS genes have been identified in *P. gingivalis*<sup>8</sup>. These are conserved across species that have T9SS, which suggests that they are generally required for assembly or functioning of the secretion apparatus<sup>5</sup>. Among the translated products of these genes, five are predicted to be integral-membrane  $\beta$ -barrels located in the OM (PorP, PorQ, PorT, PorV and Sov); two are type-I inner-membrane proteins (PorM and PorL); and three are putative lipoproteins (PorK, PorW and the unnamed *PG1058* gene product). Recently, lipoproteins PorN and the PorK were shown to interact and form a ring-shaped structure 50-nm in diameter, which is anchored on the periplasmic side of the OM as an integral component of the T9SS machinery<sup>19</sup>. Finally, PorU is found on the bacterial surface with an intact CTD<sup>16,17</sup>. In *P. gingivalis*, inactivation of any of these components leads to the arrest of T9SS cargo proteins in the periplasm, with an intact CTD<sup>4</sup>. Apart from PorU, which is a surface-located cysteine peptidase that functions as a sortase to cleave off the CTD from secreted proteins<sup>18</sup>, little is known about the functional or structural roles of the other T9SS components.

*P. gingivalis* encodes 32 putative CTD-containing proteins, which include PorU and important virulence factors such as the gingipain cysteine peptidases RgpA, RgpB, and Kgp<sup>20,21</sup>; carboxypeptidase D (*alias* Cpg70<sup>22</sup>); 35-kDa heme-binding protein (HBP35<sup>23,24</sup>) and peptidylarginine deiminase (PPAD<sup>25,26</sup>). All of these proteins (apart from PorU which uses LptO/PorV as an anchor) have been shown to be A-LPS-modified after secretion: when they were recovered from the cell envelope, OM or OM vesicles, they migrated in SDS-PAGE with a higher molecular mass than predicted, and reacted to A-LPS-specific antibodies. This indicates that they are secreted and glycosylated via T9SS<sup>27</sup>. In contrast, the protein product of gene *PG1604* (also known as *PG\_RS07070*), tentatively annotated as immunoreactive OM-associated 84-kDa antigen PG93 (see UniProt [UP] database entry Q9S3Q8), was shown in proteomic studies to retain an intact CTD upon secretion<sup>27,28</sup>. Moreover, putative homologues of this protein were found in other Bacteroidetes with T9SS (*Prevotella intermedia*, *Parabacteroides distans* and *Cytophaga hutchinsonii*<sup>27</sup>). An intact CTD is inconsistent with T9SS cargos, but rather reminiscent of intrinsic T9SS component PorU. Thus, the *PG1604* gene product was hypothesized to be a new, conserved component of the T9SS machinery. To verify this hypothesis, we probed the function of the protein by targeted mutagenesis directly in *P. gingivalis*, and analysed the effect of the isogenic *PG1604* gene deletion on T9SS cargo transcription and secretion. In addition, we determined the X-ray crystal structure of the protein to assess the molecular determinants of its function. Cumulatively, the results indicated unambiguously that the *PG1604* product is an essential component of T9SS. To be consistent with the nomenclature of T9SS components, we suggest to call it PorZ.

## Results and Discussion

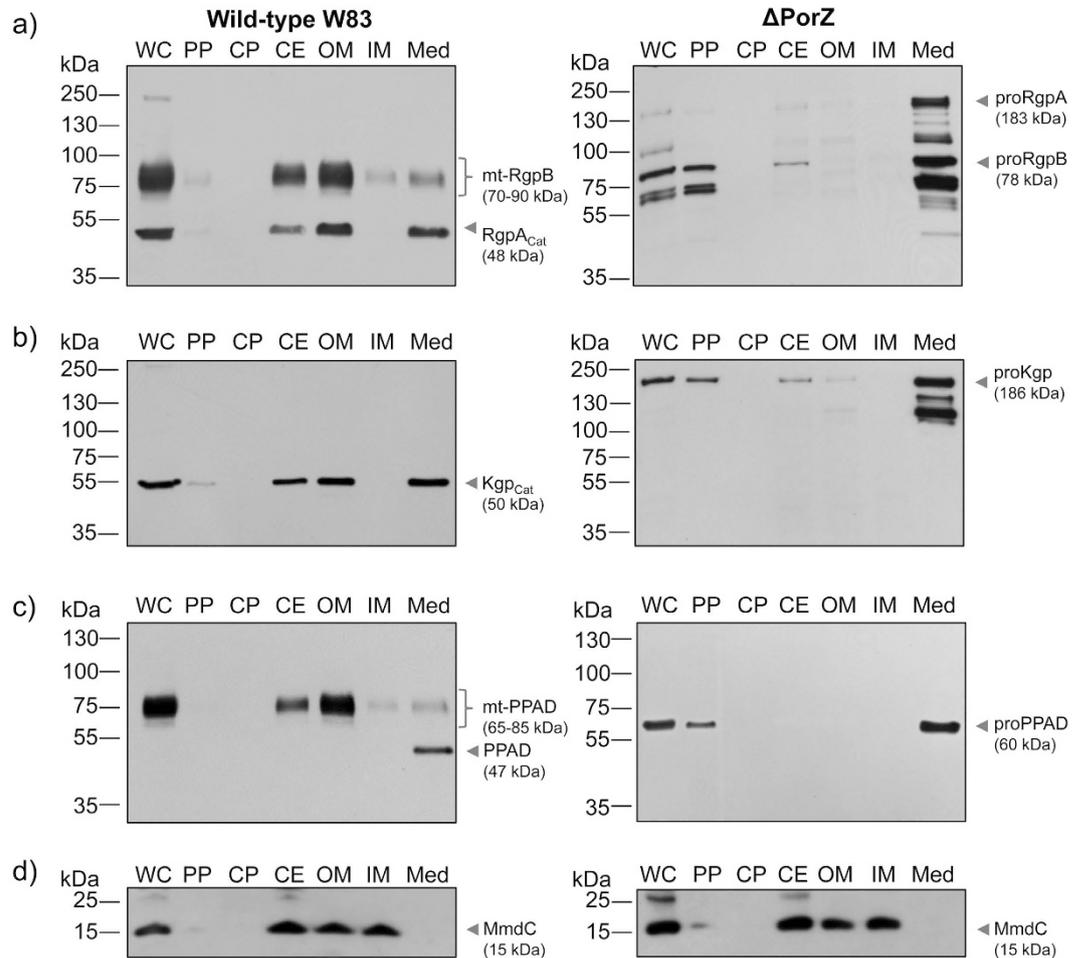
**PorZ is an essential component of T9SS.** An isogenic deletion mutant of the *porZ* gene,  $\Delta$ PorZ, was created by homologous recombination to assess its effect on T9SS cargo secretion and posttranslational processing. Deletion had a negligible effect on the *P. gingivalis* growth rate in complex media (Supplementary Fig. S1). However, on blood agar, the mutant could not accumulate heme on the cell surface and therefore yielded non-pigmented colonies (Fig. 1a). This is attributable to gingipains, which are secreted T9SS cargos that are essential for hemoglobin degradation and heme recruitment<sup>29–31</sup>. Therefore, lack of pigmentation in  $\Delta$ PorZ suggested failure of functional gingipain secretion and activation. Indeed, we found that  $\Delta$ PorZ was deficient in extracellular Kgp and Rgp gingipain activities when compared to the wild-type *P. gingivalis* strain W83 (hereafter, “wild type”; Fig. 1b,c). In contrast, dipeptidyl peptidase IV and prolyl tripeptidyl peptidase A, which are surface enzymes but



**Figure 1. Characterization of the  $\Delta$ PorZ secretory phenotype.** (a) Pigmentation on blood agar of *P. gingivalis* wild type (W83),  $\Delta$ PorZ, and *in trans* *porZ*-complemented  $\Delta$ PorZ (PorZ<sup>+</sup>) strains. Enzymatic activity of (b) Rggs, (c) Kgp, and (d) dipeptidyl peptidase IV (DPPIV) and prolyl tripeptidyl peptidase (PtpA) in whole cultures, fractionated washed cells or growth medium as determined with specific synthetic substrates. Cultures were adjusted to OD<sub>600</sub> = 1.0 prior to testing and processing, and results shown correspond to triplicate experiments. Significant differences between the wild type and mutants are indicated by \* $P < 0.05$  and \*\*\*\* $P < 0.0001$ .

not secreted by T9SS, were produced and transported to the bacterial surface in significantly higher amounts than in the wild type (Fig. 1d). A similar response had been previously observed in an inactivation mutant of an essential T9SS component, PorT<sup>32</sup>. This presumably reflects general upregulation of peptidolytic enzymes as a response to the absence of functional gingipains, which account for 85% of the extracellular proteolytic activity of *P. gingivalis*<sup>33</sup>. Reconstitution *in trans* of the *porZ* gene in the  $\Delta$ PorZ mutant—yielding PorZ<sup>+</sup>—restored both pigmentation and proteolytic activity to wild-type levels (Fig. 1a–d).

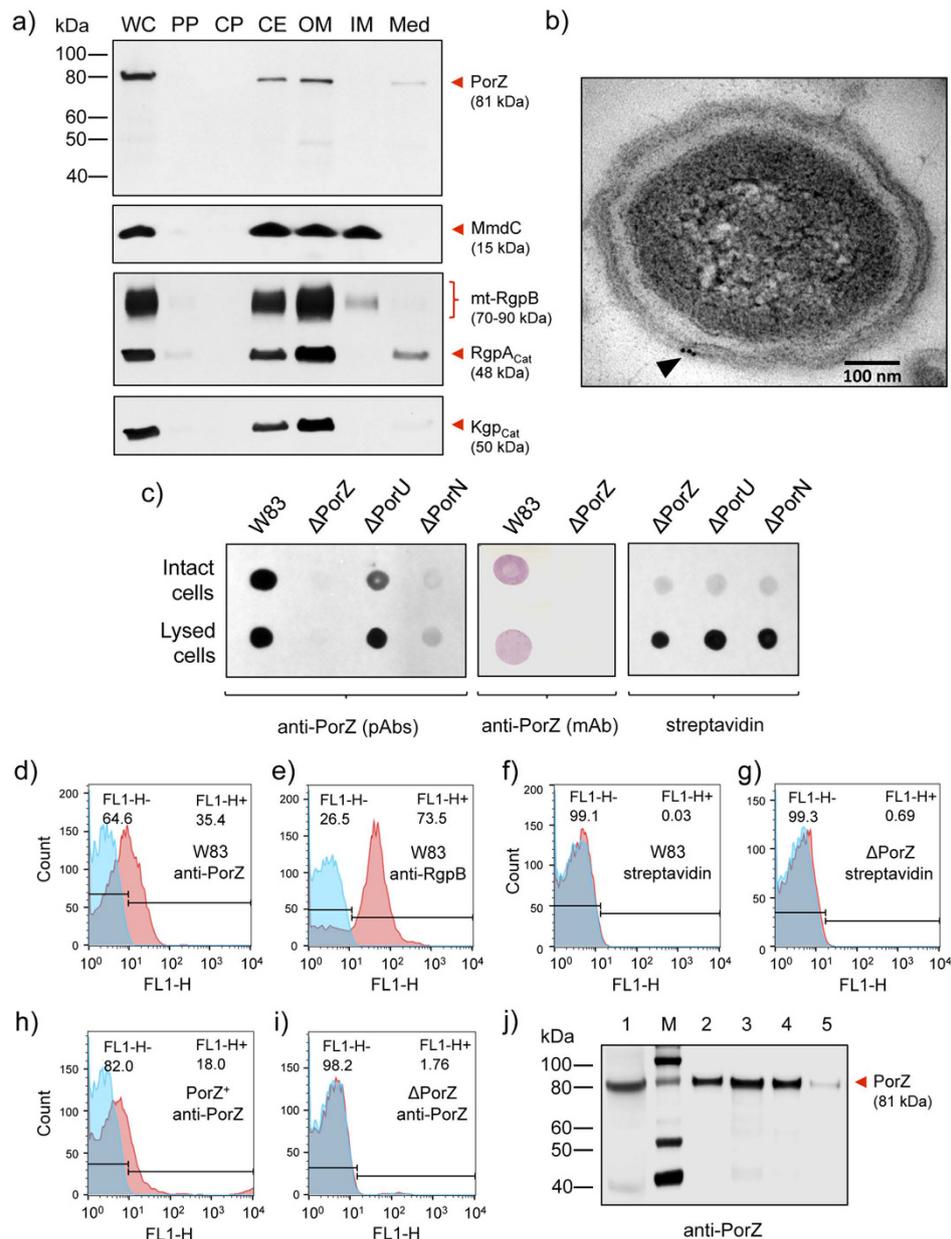
To further investigate the fate of non-secreted T9SS cargos in the absence of PorZ, we performed Western blot analysis of distinct subcellular fractions to detect gingipains (Fig. 2a,b), PPAD (Fig. 2c), and the biotin-containing 15-kDa biotin carboxyl carrier protein (AccB *alias* MmdC or PG1609) as an IM marker (Fig. 2d; see also ref. 32). The latter analysis revealed that the OM fractions obtained from the wild type and the mutant were contaminated with the IM. This is in contrast to undetectable contamination of the IM fraction with OM components, as indicated by the absence of gingipains and PPAD in the IM fraction. In the wild type, gingipains and PPAD were secreted onto the cell surface with CTD removal and proteolytic maturation of their precursors, which led to detectable activity in intact cells<sup>34</sup>. In  $\Delta$ PorZ, they were not processed to the mature forms but rather accumulated as precursors in the periplasmic fraction and in clarified culture media (Fig. 2a–c). Additionally, partially processed gingipain precursors were found with anti-gingipain antibodies. Auto-processing of gingipains has been described in heterologously-overexpressed recombinant proteinases<sup>35</sup>. Moreover, the presence of considerable amounts of PPAD and gingipain precursors in the concentrated growth medium suggested that  $\Delta$ PorZ had a “leaky” OM architecture. This contention was supported by peptide mass fingerprinting of proteins from the growth medium resolved on SDS-PAGE (Supplementary Fig. S2). Although we found in the medium several proteins normally located in the periplasm, including prolyl oligopeptidase family proteins (PG0727 and PG1004), a MEROPS-M16-family peptidase (PG0196), thioredoxin (PG0275), HtrA protease/chaperone (PG0449) and TPR-domain protein (PG0449), CTD-bearing proteins were predominant. Indeed, of the 32 known T9SS cargos of *P. gingivalis*<sup>28</sup>, 12 were found in high abundance and apparently with intact CTDs in the growth medium of  $\Delta$ PorZ, as indicated by high Mascot scores (Supplementary Table S1). These proteins included PorU (*alias* PG0026), carboxypeptidase D (*alias* Cpg70 or PG0232), PPAD (PG1424), internalin-like protein PG0350, putative hemagglutinin PG0411, immunoreactive 47-kDa antigen PG97 (PG1374), immunoreactive 46-kDa antigen PG99 (PG1798), heme-binding protein 30 (PG0616), and proteins PG0495, PG0654, PG1030, and PG2216. In



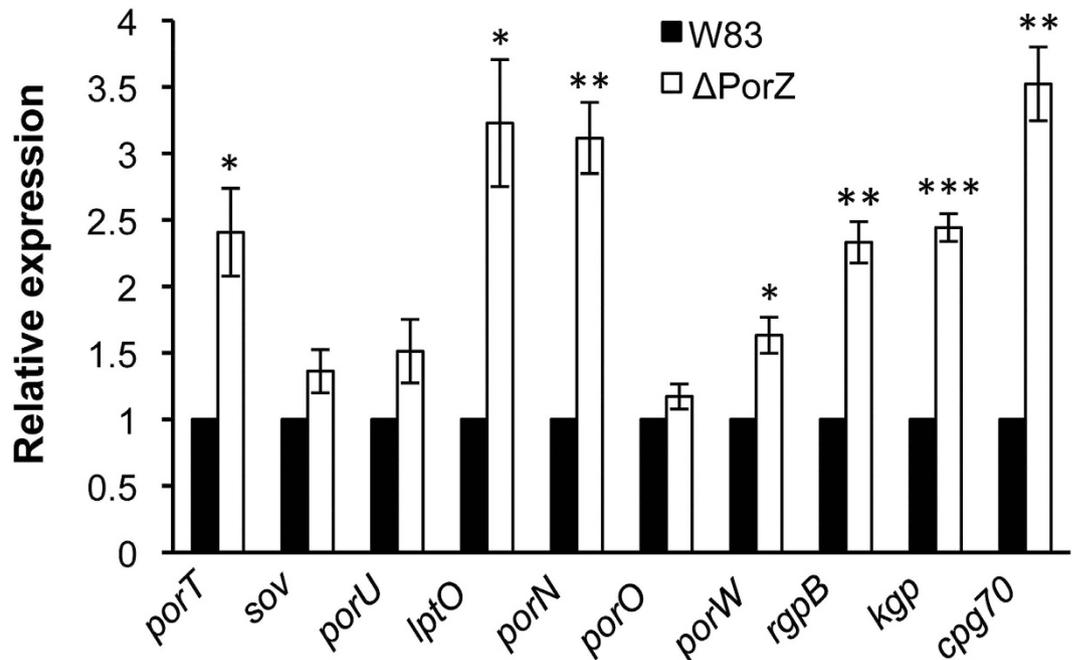
**Figure 2. Subcellular location of gingipains and PPAD.** Whole cells (WC) of wild-type (W83; *left panel*) and  $\Delta$ PorZ (*right panel*) *P. gingivalis* strains were proportionately fractionated into periplasm (PP), cytoplasm (CP), cell envelope (CE), outer membrane (OM), inner membrane (IM) and culture medium fractions (Med; 10-fold concentrated); and probed for (a) Rgps, (b) Kgp, (c) PPAD by Western blotting with specific monoclonal antibodies and (d) biotinylated IM protein (MmdC) through reaction with streptavidin conjugated to horseradish peroxidase. The pinpointed and labeled bands correspond to: (a) catalytic domain of RgpA ( $RgpA_{cat}$ ) and membrane-type RgpB (mt-RgpB) in the wild type (*left panel*) and unprocessed pro-RgpA and pro-RgpB in  $\Delta$ PorZ (*right panel*); (b) catalytic domain of Kgp ( $Kgp_{cat}$ ) in the wild type (*left panel*) and unprocessed pro-Kgp in  $\Delta$ PorZ (*right panel*); (c) mature PPAD and membrane-type PPAD (mt-PPAD) in the wild type (*left panel*) and unprocessed pro-PPAD in  $\Delta$ PorZ (*right panel*); and (d) MmdC in the wild-type (*left panel*) and  $\Delta$ PorZ (*right panel*) strains.

addition, five other potential T9SS cargos were found in the medium, but none of the detected peptides corresponded to their CTDs. When the same proteins were detectable in the growth medium of the wild-type strain, they had much lower Mascot scores and no peptides corresponding to their respective CTDs. This “leaky” OM phenotype, which leads to release of non-cleaved CTDs from CTD-cargo proteins, is similar to that reported by Taguchi *et al.*<sup>36</sup>. These authors reported that the chaperone Skp-like protein (PGN\_0300) is required for OM insertion of PorU sortase, which in turn is necessary for CTD cleavage from CTD-cargo proteins. Consistently, deletion of PGN\_0300 resulted in failure of PorU insertion and, thus, T9SS function<sup>36</sup>.

**PorZ is located on the cell surface of *P. gingivalis*.** Previous proteomics studies identified PorZ in the OM and OM vesicles (OMV) of *P. gingivalis* strain W50<sup>27,28</sup>. To determine the location of the protein more precisely, we performed Western blot analysis on wild-type cultures, quantitatively separated into whole cells and sub-cellular fractions, which included growth medium, periplasm, cytoplasm, and the cell envelope (IM+OM). The latter was additionally fractionated with detergent into OM and IM fractions. Fraction purity was verified by Western blot using Rgp/Kgp and MmdC as markers for the OM and IM, respectively. Again, while the IM fraction was only very slightly contaminated with the OM, the latter fractions contained a notable amount of the IM, as indicated by the presence of MmdC in these fractions (Fig. 3a). In agreement with the predicted localization, PorZ was mostly found associated with the cell envelope and OM fractions (Fig. 3a), as confirmed by immunogold-staining electron microscopy (Fig. 3b). In addition, trace amounts of the PorZ protein were



**Figure 3.** PorZ is located on the cell surface of *P. gingivalis*. (a) Wild-type *P. gingivalis* cells (W83) were proportionately fractionated into whole cell extract (WC), periplasm (PP), cytoplasm (CP), cell envelope (CE), outer membrane (OM), inner membrane (IM) and growth medium (Med), and subsequently analyzed by Western blotting using mouse polyclonal anti-PorZ and anti-Rgp antibodies and mouse monoclonal anti-Kgp antibodies. Streptavidin conjugated to horseradish peroxidase was used to detect MmdC, a biotinylated IM-associated protein. Presence or absence of full-length PorZ (81-kDa band) and other proteins is indicated. (b) Wild-type cells were probed with monoclonal anti-PorZ antibodies and labeled with immunogold to visualize the cellular location of PorZ (black arrowhead) in electron microscopy (bar = 100 nm). (c) Dot blot analysis of intact and lysed wild-type (W83),  $\Delta$ PorZ,  $\Delta$ PorU and  $\Delta$ PorN cells using mouse monoclonal anti-PorZ antibodies (mAb), mouse polyclonal anti-PorZ antibodies (pAb) or streptavidin conjugated to horseradish peroxidase. Flow cytometry analysis showing the surface exposure of (d) PorZ in wild-type cells (W83) with anti-PorZ pAb; (e) RgpB in wild-type cells (W83) with monoclonal anti-RgpB antibodies (positive control); (f) MmdC in wild-type cells (W83) and (g) in  $\Delta$ PorZ cells with streptavidin-Alexa Fluor 488 conjugate; (h) PorZ in *in trans* porZ-complemented  $\Delta$ PorZ (PorZ<sup>+</sup>) cells with pAb and (i) PorZ in  $\Delta$ PorZ cells with anti-PorZ pAb. Isotype negative controls are in blue and immunoprobed cells in red; the histograms shown are representative of three independent experiments. (j) Presence of full-length PorZ (lane 1) detected by Western blot using anti-PorZ pAb on wild-type cells washed with PBS and suspended, respectively, in distilled water (lane 2); in 0.0007% Tween-20 (lane 3); in 0.04% sarcosyl (lane 4); and in 0.02% SDS (lane 5). After 10 min of gentle stirring, cells were removed by centrifugation and the presence of PorZ in the cell pellet (lane 1) and supernatants (lanes 2–5) was checked. The detergent concentrations correspond to one-tenth of the critical micelle concentration (CMC). Lane M, MagicMark™ XP Western Protein Standard.



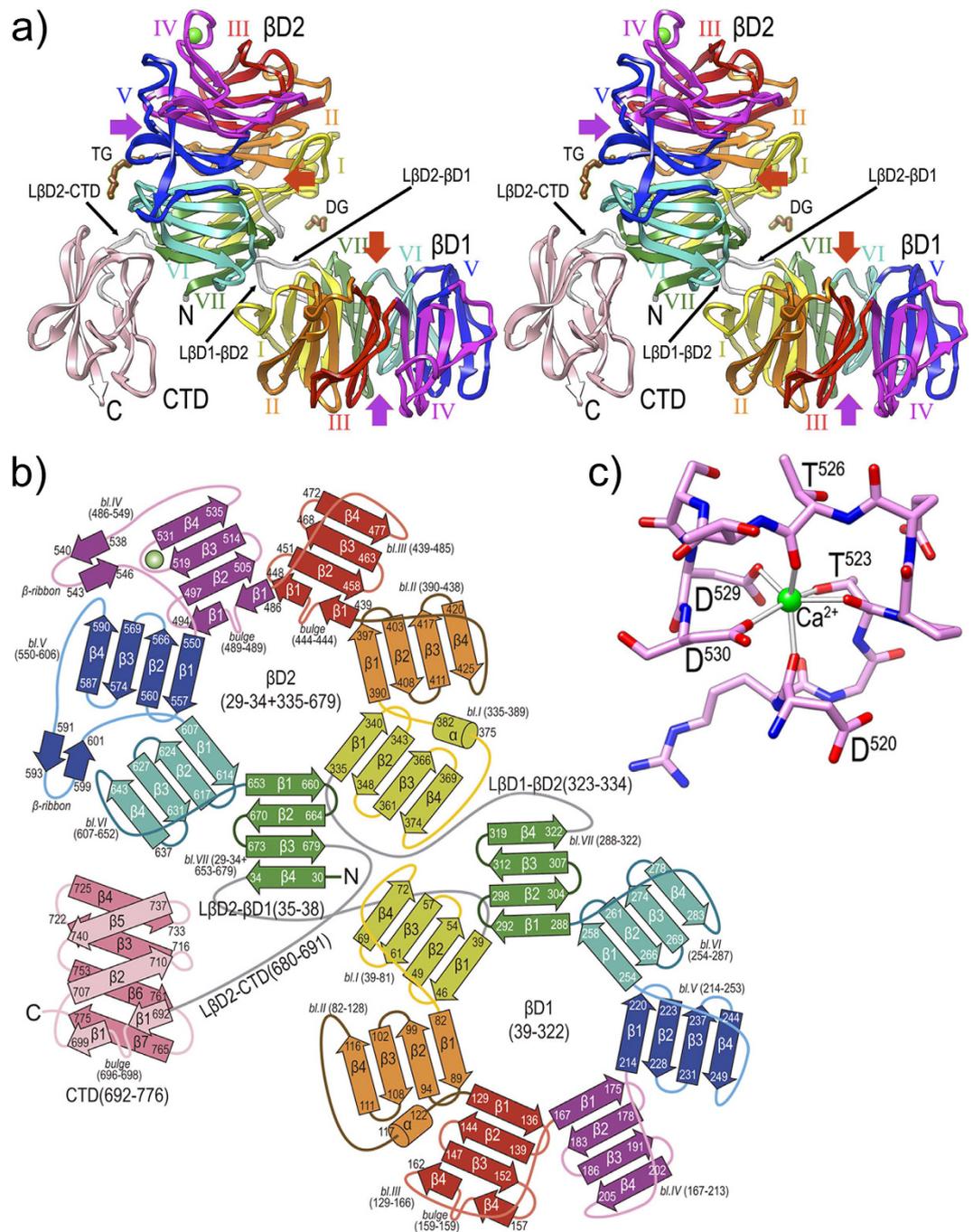
**Figure 4. Effect of PorZ on the expression of T9SS components and cargos.** The respective mRNAs of the genes of T9SS components (*porT*, *sov*, *porU*, *lptO*, *porN*, *porO* and *porW*) and cargos (*rgpB*, *kgp* and *cpq70*) were quantified in mid-logarithmic cultures of wild type (W83; full bars) and  $\Delta$ PorZ (open bars) cells by quantitative RT-PCR. Expression levels were normalized against ribosomal *r16s* expression, and the respective expression levels of the wild type were arbitrarily set to 1.0. The results correspond to four independent experiments. Error bars represent standard deviation as analyzed by Student's t-test. \* $P < 0.05$ ; \*\* $P < 0.01$  and \*\*\* $P < 0.001$ .

detected in concentrated growth medium together with RgpA<sub>cat</sub> and Kgp<sub>cat</sub>, suggesting that, similarly to gingipain catalytic domains, also PorZ can be shed in low amounts from the bacterial surface into the medium.

To determine whether PorZ is located on the periplasmic or the extracellular side of the OM, we performed dot-blot analysis of intact and sonicated wild-type cells plus  $\Delta$ PorZ as a negative control. While no signal was detected for the latter, equivalent staining was observed for both intact and sonicated cells employing two different anti-PorZ antibodies (Fig. 3c). Of note, the integrity of the *P. gingivalis* cell envelope was confirmed by dot-blot analysis with streptavidin-HRP showing the reaction only after cell disruption by sonication. These findings support cell-surface localization of PorZ (Fig. 3c). This location was further assessed by flow cytometry, with surface-exposed gingipain RgpB and intracellular MmdC as positive and negative controls, respectively. Indeed, when anti-PorZ and anti-RgpB antibodies were used, both proteins were identified on the cell surface (Fig. 3d,e), while no staining was detected with streptavidin-Alexa Fluor 488 conjugate (Fig. 3f,g). Specificity of the flow cytometry analysis was checked with  $\Delta$ PorZ, which showed negligible staining, while a strong signal was found in the *in trans* *porZ*-complemented strain, PorZ<sup>+</sup> (Fig. 3h,i). Interestingly, FACS assays with *P. gingivalis* strain HG66, which cannot attach A-LPS to T9SS cargos so that these are subsequently released into the extracellular milieu, revealed that PorZ was still on the cell surface (Supplementary Fig. S3). Finally, when intact *P. gingivalis* cells were gently mixed with pure distilled water or subcritical micellar concentrations of detergents, which are unable to disrupt OM integrity, a sizable fraction of PorZ was released into the liquid phase (Fig. 3j). Significantly, only trace amounts of gingipains were washed out of the bacterial surface under the conditions tested (Supplementary Fig. S3).

**PorZ affects expression of T9SS components and cargos.** To determine whether the absence of gingipain activity in the  $\Delta$ PorZ mutant was also due to their reduced expression, quantitative RT-PCR was used to investigate expression of CTD-cargo proteins and other T9SS components in the  $\Delta$ PorZ mutant. Interestingly, rather than suppression of CTD-cargo protein expression, the mRNA levels of CTD-cargo peptidases RgpB, Kgp and CPG70 in  $\Delta$ PorZ were found to double or triple those of the wild type (Fig. 4). Thus, lack of activity was due to their failure to be secreted and matured in  $\Delta$ PorZ rather than to reduced expression. Further, deletion of PorZ induced significant upregulation of T9SS-components *porT*, *lptO*, and *porN* when compared to the wild type. In particular, the latter two triplicated the expression levels of the wild type (Fig. 4). Other components such as *porO*, *porW*, *sov* and *porU* did not show significant change. These trends were consistent regardless of whether *r16s* (Fig. 4) or *gyrA* (data not shown) was used as a reference gene. Overall, the significance of this variable effect on the expression of T9SS components in  $\Delta$ PorZ is currently unknown.

**Structural analysis of PorZ.** We produced PorZ without its predicted signal peptide (residues Q<sup>26</sup>-R<sup>776</sup>) by recombinant overexpression in *Escherichia coli*, and succeeded in crystallizing and solving its structure by



**Figure 5. Overall crystal structure of PorZ.** (a) Ribbon-type plot in cross-eye stereo of the crystal structure to 2.9 Å resolution of PorZ depicting domains  $\beta D1$ ,  $\beta D2$  and CTD, and the three domain-connecting linkers (white ribbons; labeled  $L\beta D2-\beta D1$ ,  $L\beta D1-\beta D2$ , and  $L\beta D2-CTD$ ). Each of the seven blades of propellers  $\beta D1$  and  $\beta D2$  (labeled counter-clockwise I to VII) is colored in yellow, orange, red, magenta, blue, turquoise and green, respectively; the CTD is in pink. A structural calcium-binding site (green sphere) is found within  $\beta D1$ -blade IV, and a tetraethylene glycol (TG) and a diethylene glycol (DG) were tentatively assigned on the protein surface (brown stick-models). Other (functionally probably irrelevant) ions and ligands were omitted for clarity. The central shafts of  $\beta D1$  and  $\beta D2$  are pinpointed on the entry and exit sides of the propellers by red and purple arrows, respectively. For labels and extension of regular secondary structure elements, see (b). (b) Topology scheme of PorZ, with  $\beta$ -strands as arrows and helices as cylinders, colored as in (a). The polypeptide chain spans residues G<sup>29</sup>–R<sup>776</sup> and the three constituting domains plus the linkers (in grey) are indicated with the residues delimiting each structural element (strands, bulges, helices,  $\beta$ -ribbons, blades and domains). The nomenclature adopted in the text for structure elements is “domain-blade-structural element”, e.g.  $\beta D1$ -VI- $\beta 3$  or  $\beta D2$ -IV- $\beta$ -ribbon. (c) Structural calcium-binding site framed by segment D<sup>520</sup>–D<sup>530</sup> within loop L $\beta 3\beta 4$  of  $\beta D1$ -blade IV. The ion is octahedrally coordinated by D<sup>520</sup>O, T<sup>523</sup>O, T<sup>523</sup>O $\gamma$ , T<sup>526</sup>O, D<sup>529</sup>O $\delta 1$  and D<sup>530</sup>O $\delta 1$ , which are at binding distances of  $\sim 2.4$  Å.

single-wavelength anomalous diffraction with a selenomethionine derivative. The structure was refined with data to 2.9 Å resolution and consists of three domains. The first two are consecutive N-terminal seven-stranded  $\beta$ -propeller or circular-leaflet moieties ( $\beta$ D1: residues K<sup>39</sup>—M<sup>322</sup>, and  $\beta$ D2: G<sup>29</sup>—L<sup>34</sup>+Y<sup>335</sup>—T<sup>679</sup><sup>37–39</sup>; PorZ residue numbering in superscript notation according to UP Q9S3Q8), each featuring a shallow cylinder or thick disk with an “entry side” and an “exit side”<sup>38</sup>. These domains are succeeded by a C-terminal domain (CTD, V<sup>692</sup>—R<sup>776</sup>). The domains are connected by linkers (L): L $\beta$ D2- $\beta$ D1 (L<sup>35</sup>—H<sup>38</sup>), L $\beta$ D1- $\beta$ D2 (P<sup>323</sup>—F<sup>334</sup>), and L $\beta$ D2-CTD (G<sup>680</sup>—G<sup>691</sup>). The two propellers are offset from one another by a  $\sim 90^\circ$  rotation about the intersection axis of the propellers’ planes. This causes the overall molecular structure to be reminiscent of an easy chair of approx. maximal dimensions  $95 \times 80 \times 55$  Å, with  $\beta$ D1 as the seat,  $\beta$ D2 the backrest, and CTD the backrest support (Fig. 5a). The two entry-side surfaces of the PorZ propellers mimic, respectively, the seating and reclining surfaces of the chair.

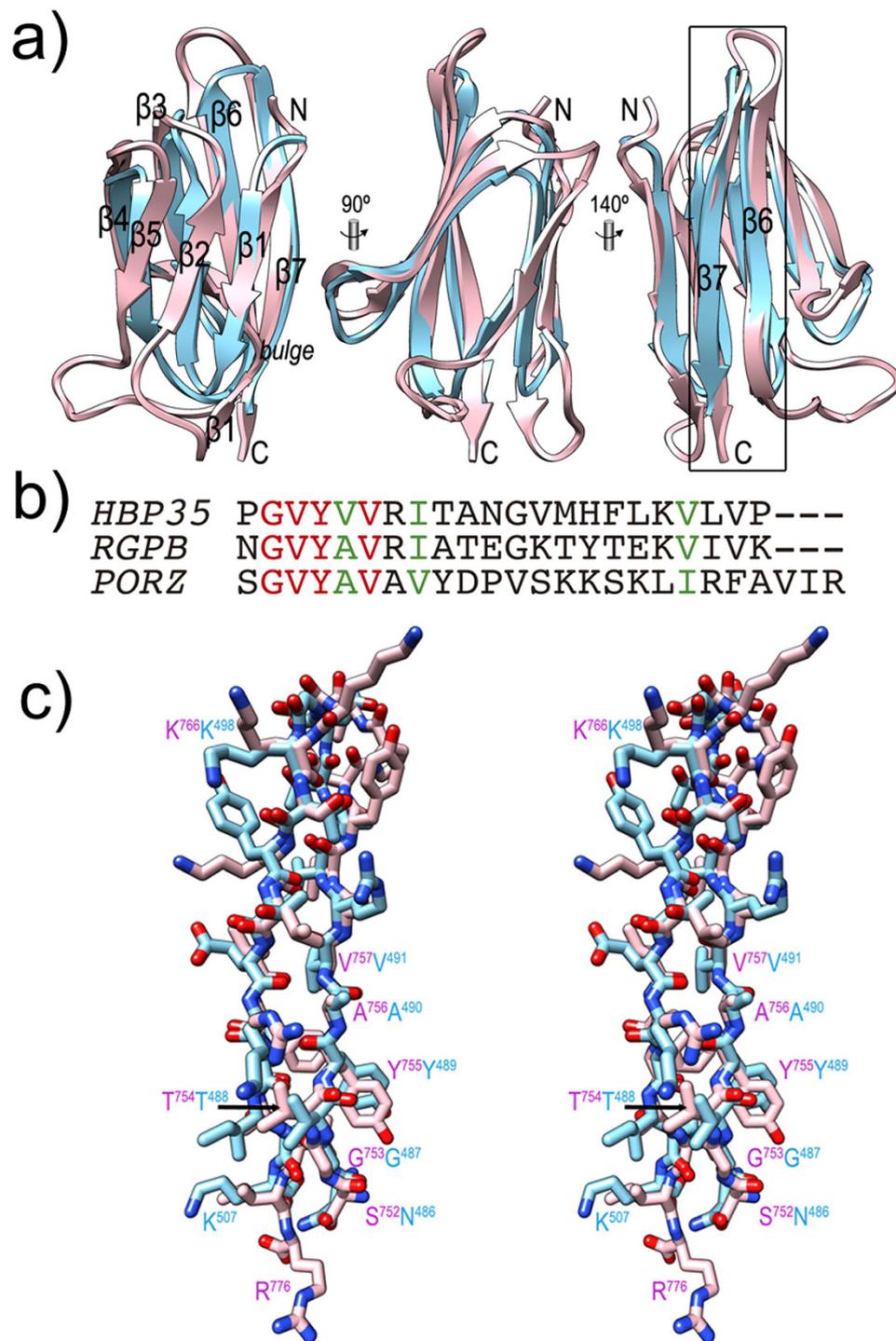
The seven blades of the propellers consist of a four-stranded ( $\beta$ 1— $\beta$ 4) antiparallel  $\beta$ -sheet of simple up-and-down “W” connectivity or  $\beta$ -leaflet topology<sup>40</sup>. Two  $\beta$ 1 strands and one  $\beta$ 4 strand are interrupted by bulges (strands  $\beta$ D2-III- $\beta$ 1,  $\beta$ D2-IV- $\beta$ 1, and  $\beta$ D1-III- $\beta$ 4 Fig. 5b; for structural-element notation, see the legend to Fig. 5b). The blades are radially arranged around a central propeller shaft, which is lined by the respective first strands of each blade ( $\beta$ 1) originating on the entry sides, and the strands of each leaflet are connected by short loops. Exceptions are those connecting  $\beta$ D2-I- $\beta$ 2 with  $\beta$ D2-I- $\beta$ 3, and  $\beta$ D2-IV- $\beta$ 2 with  $\beta$ D2-IV- $\beta$ 3, which span twelve and eight residues respectively, and protrude from the entry side of  $\beta$ D2 (Fig. 5a). Within each sheet and ongoing from  $\beta$ 1 to  $\beta$ 4, the strands accumulate a twist of  $\sim 45^\circ$ , which is right-handed as usual for such structures<sup>41</sup>. Uniquely, blade IV of  $\beta$ D2 contains a metal, which was assigned to a calcium based on the octahedral ligand sphere, chemical nature of the six (oxygen) ligands, and binding distances of the ligands ( $\sim 2.4$  Å; see [http://tanna.bch.ed.ac.uk/newtargs\\_06.html](http://tanna.bch.ed.ac.uk/newtargs_06.html) and ref. 42). The ion is bound by three main-chain oxygens, two aspartate side-chain oxygens, and a threonine side-chain oxygen (Fig. 5c). Each blade further consists of a C-terminal linker, which connects respective strand  $\beta$ 4 with strand  $\beta$ 1 of the downstream blade. Linkers vary in length between four and sixteen residues, and may contain extra regular secondary-structure elements such as  $\beta$ -ribbons (blades IV and V of  $\beta$ D2) and  $\alpha$ -helices (blades I of  $\beta$ D2 and II of  $\beta$ D1; see Fig. 5b). The shafts of the PorZ propellers have an internal diameter spanning  $\sim 5$ – $10$  Å, and, in contrast to other propellers such as the four-fold hemopexin domains<sup>38</sup>, they do not show evidence of tight ion or ligand binding. We only found some loosely bound chemicals from the crystallization and vitrification conditions (data not shown), which could indicate potential binding sites of functional relevance. Further research will be required to verify this assertion, though.

In general,  $\beta$ -propeller symmetry ranges from four-fold to eight-fold, the most populated group being the seven-fold propellers<sup>37–39</sup>. This may result from the packing of the blades, which is considered more stable the larger the number of blades<sup>39</sup>. Thus, to circumvent low stability, four-fold and five-fold propellers incorporate additional elements to tether the circular arrangement, such as disulphide bonds between the N- and C-terminal blades<sup>38,43</sup>. Other mechanisms entail that the first blade is made up by the N-terminal part of the polypeptide chain for some of its strands and by the C-terminal part after completion of the entire propeller moiety for the remaining strands, thus featuring a kind of “velcro” mechanism<sup>38</sup>. In the case of PorZ, which consists of two consecutive seven-fold propellers, the N-terminal segment of the polypeptide chain follows such a mechanism and forms strand  $\beta$ D2-VII- $\beta$ 4, while the remaining three strands of this blade are provided by the polypeptide chain after forming both propellers (see Fig. 5b). After strand  $\beta$ D2-VII- $\beta$ 4, the polypeptide enters short, four-residue L $\beta$ D2- $\beta$ D1 and then forms  $\beta$ D1 starting with blade I. After  $\beta$ D1, which ends with strand  $\beta$ D1-VII- $\beta$ 4, L $\beta$ D1- $\beta$ D2 leads to strand  $\beta$ D2-I- $\beta$ 1. Altogether,  $\beta$ D1 and  $\beta$ D2 tightly approach each other through respective blades I and VII and linkers L $\beta$ D2- $\beta$ D1 and L $\beta$ D1- $\beta$ D2 (Fig. 5a). This may explain the stable, nearly perpendicular relative arrangement between propellers, which is reminiscent of that of the two four-fold propellers found in rabbit hemopexin<sup>38,44</sup>.

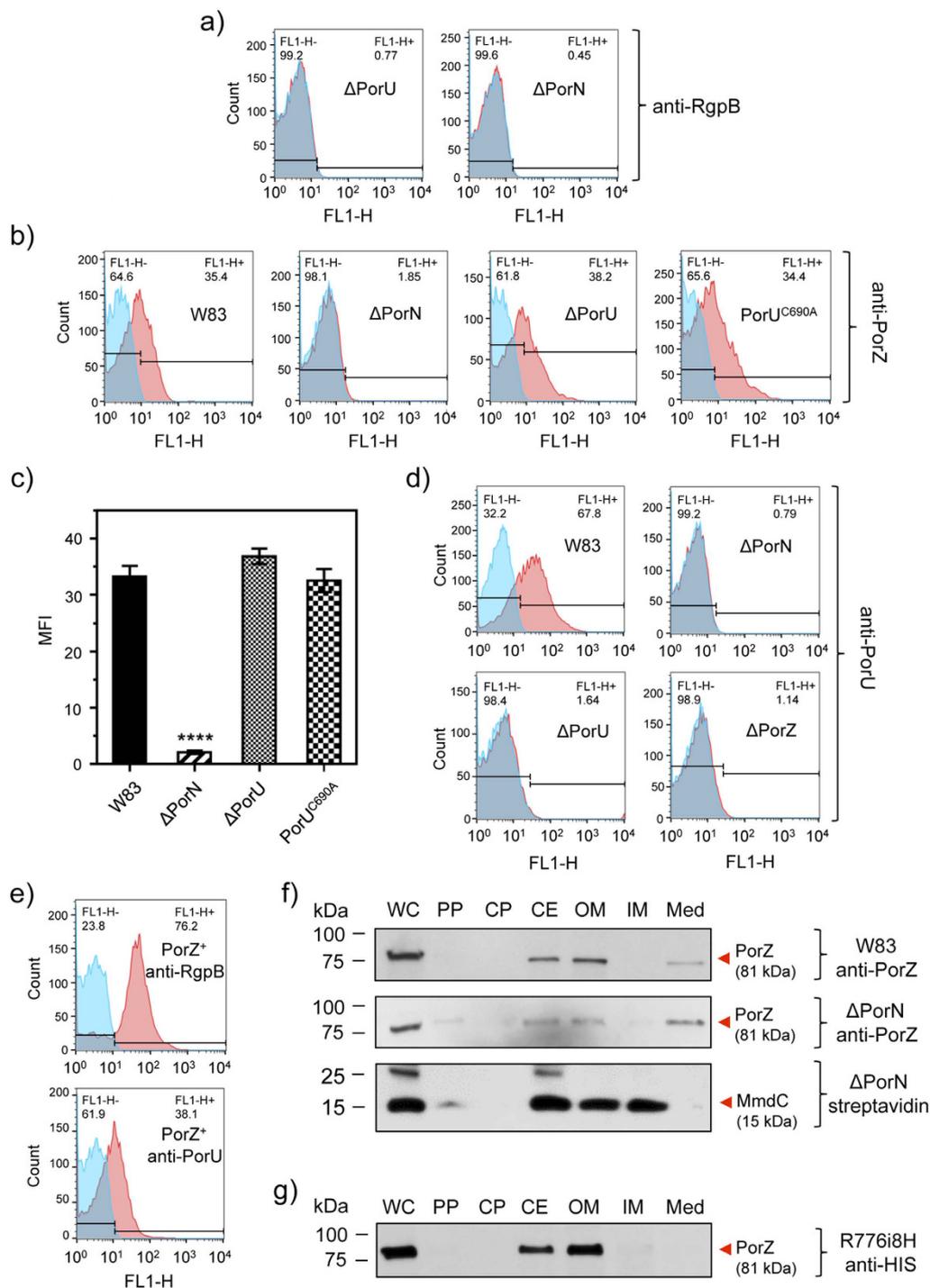
After  $\beta$ D2, L $\beta$ D2-CTD leads to CTD (Figs 5a,b and 6a), which spans 85 residues and is organised as a  $\beta$ -sandwich consisting of a three-fold antiparallel  $\beta$ -sheet (strands CTD- $\beta$ 1, - $\beta$ 2, and - $\beta$ 5; CTD- $\beta$ 1 is interrupted by a bulge, see Fig. 5b) and a four-fold antiparallel  $\beta$ -sheet (CTD- $\beta$ 4, - $\beta$ 3, - $\beta$ 6, and - $\beta$ 7). Despite some general differences, in particular in loops, the architecture and strand connectivity is equivalent to that of the CTD of RgpB, which is the structural paradigm of a functional CTD required for T9SS secretion<sup>12</sup>. Both structures were superposed for 61 topologically equivalent residues, giving rise to an overall core *rmsd* of 1.8 Å and a sequence identity of 20% (see Fig. 6a–c). Although the latter value is rather low<sup>45</sup>, both structural and sequence similarity are particularly significant for the last  $\sim 25$  residues of either molecule (Fig. 6b), a segment that encompasses the apparent signal recognized by T9SS for translocation<sup>12,15</sup>. A common pattern (G-V-Y-V/A-V-X-I/V) arises when the two sequences are aligned based on structural criteria (Fig. 6c) and the sequence of the T9SS cargo HBP35<sup>15</sup> is further included (Fig. 6b). Thus, PorZ possesses a potentially functional CTD for T9SS secretion.

**Possible function of the PorZ propeller domains.** The modular architecture of PorZ is reminiscent of that of the periplasmic sensor-domain moiety of protein BT4663 from *Bacteroides thetaiotaomicron*, which belongs to the human microbiome and is the most prevalent gut colonizer<sup>10</sup>. BT4663 is the transmembrane histidine kinase of a two-component signal transduction system engaged in detection and degradation of complex carbohydrates. BT4663 transits between distinct unbound and bound conformations, and in this way activates the intracellular kinase domain<sup>46</sup>. The general architecture of BT4663 is also found in related potential histidine kinases BT4673 and BT3049 from the same organism<sup>47</sup>. Similarly to PorZ, the three proteins likewise comprise two N-terminal seven-fold  $\beta$ -propeller domains and a C-terminal all- $\beta$  domain, termed the Y<sub>Y</sub>Y domain for BT4663<sup>46</sup>.

However, while BT4663 and related proteins are dimeric<sup>46,47</sup>, PorZ was verified in size-exclusion chromatography to be a monomer in solution (data not shown). In addition, the relative arrangement of the three domains in either the bound or unbound forms, differs from that of PorZ. Differences also arise between the Y<sub>Y</sub>Y domain and PorZ CTD: although they share topology and strand connectivity, the former is much larger (120 vs.



**Figure 6. The C-terminal domain of PorZ.** (a) Ribbon-plot of the CTDs of PorZ (pink) and RgpB (cyan) after optimal superposition facing the three-stranded front  $\beta$ -sheet (left, the seven constituting strands are labelled, see also Fig. 5b); after a vertical  $90^\circ$  rotation (center); and facing the four-stranded back  $\beta$ -sheet (right). C-terminal strands  $\beta 6$  and  $\beta 7$ , which contain the reported molecular determinants for T9SS secretion<sup>12,15</sup>, are framed and labeled in the right panel. (b) Sequence alignment of the 22 C-terminal residues of HBP35 (UniProt Q8G962) and RgpB (UniProt P95493), and the 25 final residues of PorZ after the structural alignment of the structures of the latter two proteins (see also [c]). Identical residues are in red, similar ones in green. G-X-Y sequences are also found in PKD proteins within strands equivalent to CTD- $\beta 6$ . These are also seven-stranded immunoglobulin-like all- $\beta$  domains<sup>84</sup>, although the function of the G-X-Y motif therein is unknown. In addition, the Y\_Y\_Y domain of BT4663 protein contains this signature<sup>46</sup>. (c) Detail in cross-eye stereo showing the strands framed in (a) as full-atom models, i.e. segments S<sup>752</sup>—R<sup>776</sup> of PorZ (with pink carbons and magenta labels) and N<sup>486</sup>—K<sup>507</sup> of RgpB (mature protein numbering as subscripts, add 229 for full-length protein numbering; cyan carbons and labels).



**Figure 7. PorZ is secreted *via* T9SS with an intact CTD independently of PorU.** Flow cytometry analysis using (a) anti-RgpB mAbs in  $\Delta$ PorU and  $\Delta$ PorN; (b) anti-PorZ antibodies in the wild type (W83), PorN-null mutant ( $\Delta$ PorN), PorU-null mutant ( $\Delta$ PorU), and PorU active-site inactivation mutant (PorU<sup>C690A</sup>); (c) PorZ surface exposure as mean fluorescent intensity (MFI) in different strains calculated from flow cytometry analysis (in duplicates) from three different cultures. Significant differences between the wild type and mutants were analyzed by one-way ANOVA with Bonferroni's correction; \*\*\*\* $P < 0.0001$ . Flow cytometry analysis using (d) anti-PorU antibodies in the wild type (W83), PorN-null mutant ( $\Delta$ PorN), PorU-null mutant ( $\Delta$ PorU), and  $\Delta$ PorZ; (e) anti-RgpB and anti-PorU in *porZ* complemented strain (PorZ<sup>+</sup>). The result of using specific antibodies (red surface) and the negative isotype control (blue surface) are shown. (f) Western blot analysis of subcellular locations of PorZ in the  $\Delta$ PorN mutant by probing with anti-PorZ antibodies as compared to the wild type (W83). Streptavidin conjugated to horseradish peroxidase was used to detect MmdC, a biotinylated IM-associated control protein. (g) Same as (f) but using anti-His antibodies to detect the CTD of PorZ in the strain expressing PorZ with an octahistidine tag at the C-terminus (R776i8H). Bacterial cultures were fractionated as described in the Methods section.

85 residues) and has an extra strand. Moreover, while the Y<sub>Y</sub>Y domain functionally acts as a spacer from the cytoplasmic membrane surface further engaged in dimerization, PorZ CTD is a potential signalling domain (see below).

This notwithstanding, BT4663 may yet provide a hint to the molecular function of PorZ, as most of the known propeller domains coordinate ligands or catalyse reactions at or close to the central shafts on their entrance side<sup>37,48</sup>. Among its widespread functions are binding of sugar moieties and related molecules as reported for the five-fold propellers  $\beta$ -fructosidase,  $\alpha$ -L-arabinanase and levan-sucrase. This is also the function of BT4663, which binds disaccharide ligands on the respective entry surfaces of both propeller domains, central to the shafts<sup>46</sup>. Along this line, the crystal structure of PorZ revealed potential polyethylene glycol fragments bound at the hinge between  $\beta$ D1 and  $\beta$ D2 on their entry side, as well as on the exit surface of  $\beta$ D2 (see Fig. 5a). Other smaller molecules were found in or on the shaft entrances (see above). Moreover, the two extended intra-leaflet loops, which protrude from the entry side of  $\beta$ D2 (see above), could also be potentially engaged in ligand binding, thus providing a functional explanation for their exceptional length. In any case, further experiments will be required to verify a potential glycan-binding function of PorZ as part of, or independently from, T9SS secretion.

**PorZ is itself secreted via T9SS with an intact CTD.** The presence of a CTD reminiscent of that found in T9SS cargos and components at the C-terminus of PorZ suggests that the protein is itself secreted to the surface by the same secretion system. In this event, PorZ would be absent from the bacterial surface in *P. gingivalis* mutants with dysfunctional T9SS, as shown for other T9SS cargos. To validate this hypothesis, we investigated the surface location of PorZ by flow cytometry analyses in three secretion-deficient mutants:  $\Delta$ PorN,  $\Delta$ PorU, and a mutant expressing inactive PorU, in which the catalytic cysteine had been mutated to alanine (PorU<sup>C690A</sup>). The defective secretion phenotype of the  $\Delta$ PorN and  $\Delta$ PorU mutants was confirmed by the lack of staining for RgpB (Fig. 7a). In contrast to the wild type, PorZ was absent from the cell surface of mutant  $\Delta$ PorN but surprisingly not of mutants  $\Delta$ PorU and PorU<sup>C690A</sup> (Fig. 7b), which evinced similar levels to the wild type (Fig. 7c). This strongly argues that PorZ is translocated across the OM and anchored at the bacterial surface in a PorU-independent manner.

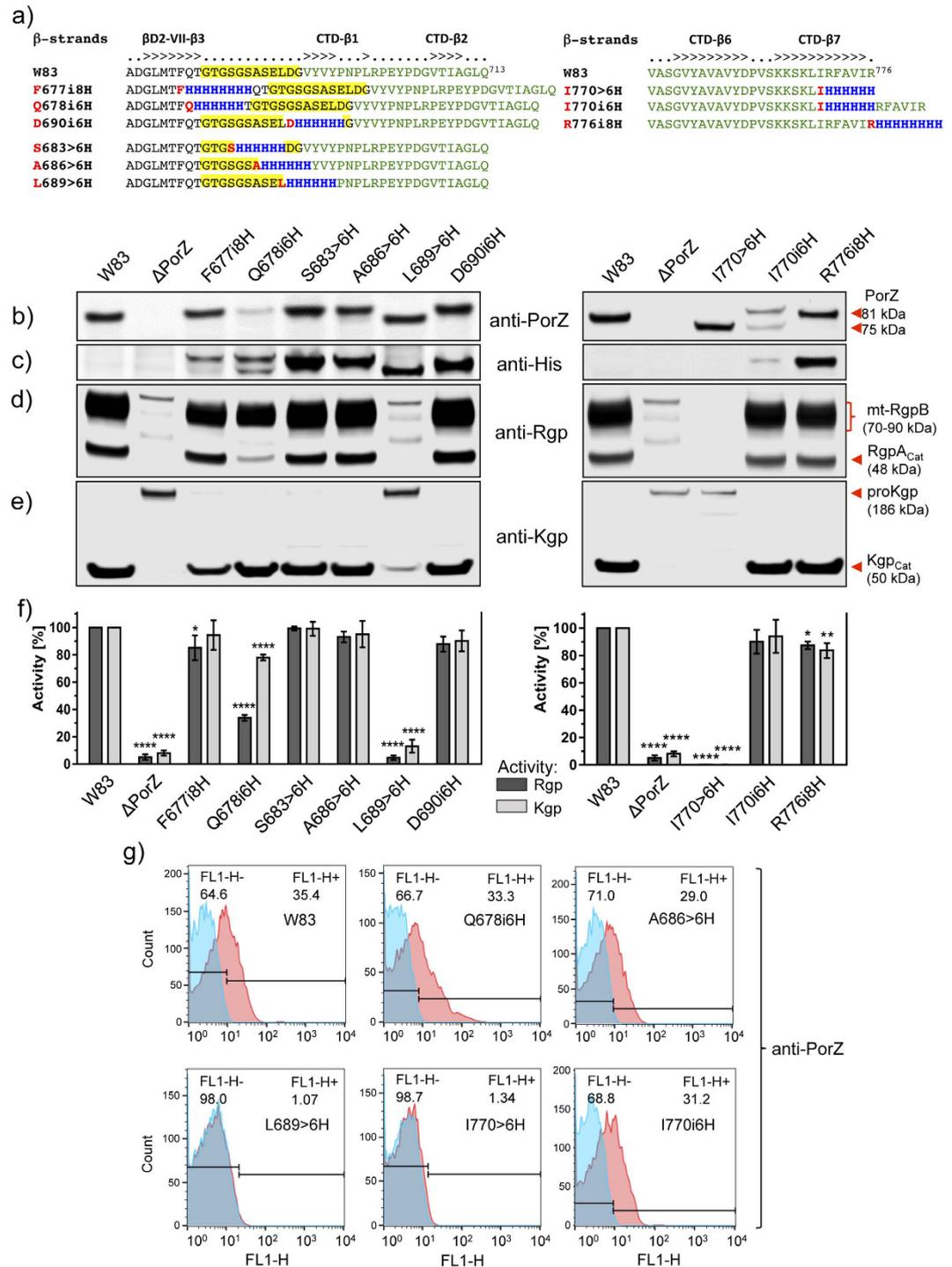
Due to the phenotypic similarity of the  $\Delta$ PorZ mutant with the Skp-like PGN\_0300 mutant, where PorU failed to insert into the OM<sup>36</sup>, we proceeded to investigate the surface exposure of PorU in the  $\Delta$ PorZ and  $\Delta$ PorN mutants as compared to wild type. Flow cytometry using an anti-PorU antibody revealed that PorU was absent from the cell surface of both  $\Delta$ PorZ and  $\Delta$ PorN, in stark contrast to the wild type (Fig. 7d). It was not present in the  $\Delta$ PorU mutant as expected but the complementation of the  $\Delta$ PorZ mutant restored the surface exposure of both PorU and RgpB (Fig. 7e). This result suggests that the export of PorU to the cell surface may be a complex process involving both an intact T9SS pathway as well as the Skp-like chaperone PGN\_0300. It further suggests that PorU may be the last component of T9SS to be secreted to the surface as it requires a functional PorN and PorZ but not vice-versa, as PorZ was exported to the cell surface in the absence of wild-type or inactive PorU (Fig. 7a).

To determine the fate of PorZ in  $\Delta$ PorN, we performed Western blot analysis on cultures that were quantitatively separated into whole cells and subcellular fractions. When compared with the wild type, the partition of PorZ between the OM and periplasmic fractions was only slightly higher in the periplasm of  $\Delta$ PorN (Fig. 7f). This is in stark contrast with Rgps and Kgp, which are fully processed and associated with the OM in the wild type (Fig. 2a,b) but found in unprocessed and partially processed forms in the periplasm of  $\Delta$ PorN (Supplementary Fig. S5). The same distribution of unprocessed gingipains in the periplasm is apparently a hallmark of all T9SS secretion mutants characterized to date, including  $\Delta$ PorZ (Fig. 2), PorT<sup>49</sup>, PorU<sup>17</sup>, Sov<sup>50</sup>, and LptO<sup>51</sup>. Taken together, our results unambiguously show that PorZ is associated with the OM, and its surface exposure is dependent on PorN but not PorU (Fig. 7b,c). Conversely, PorU exposure on the cell surface requires the presence of PorZ (Fig. 7d,e).

As to which variant of PorZ is found on the cell surface, the relative molecular mass of a PorZ-immunoreactive band in SDS-PAGE was ~80 kDa, which suggests that the protein is full length, without the signal peptide (theoretic molecular mass: 81 kDa). To verify this contention, we constructed *P. gingivalis* mutant strain R776i8H, which expresses PorZ with an octahistidine at the C-terminus (see also the next section). This mutant possesses a secretory phenotype that is indistinguishable from the wild type, as determined by colony pigmentation (data not shown), cellular distribution (Fig. 7e) and gingipain activity (Fig. 8f). Western blot analysis with anti-His-tag antibodies revealed reactivity to a band of ~80 kDa, which confirmed the presence of intact CTD in the mature PorZ protein (Fig. 7g). This observation is consistent with proteomics data reporting that PorZ appears to retain its CTD and does not undergo A-LPS modification as seen in other T9SS cargos<sup>27</sup>.

**An intact CTD is required for cargo translocation across the OM and post translational processing.** It was previously described that an intact CTD was needed for proper secretion, because its removal or C-terminal truncation prevented OM translocation and resulted in accumulation of the cargo protein in the periplasm<sup>32</sup>. However, insertion of additional residues into the linker region between the CTD and the preceding immunoglobulin superfamily domain of RgpB resulted in cleavage of the CTD but prevented A-LPS attachment. In addition, the N-terminal pro-domain responsible for latency maintenance in pro-RgpB<sup>52</sup> was cleaved off, so soluble, fully active gingipain was secreted into the medium<sup>53</sup>.

The crystal structure of PorZ revealed that CTD is preceded by domain  $\beta$ D2, with the inter-domain linker L $\beta$ D2-CTD spanning residues G<sup>680</sup>-G<sup>691</sup> (see above). To study the effect of oligohistidine insertions or replacements on PorZ expression, processing and translocation, we inserted hexa/octahistidines or replaced six consecutive residues with histidines within L $\beta$ D2-CTD (six mutants) and at the C-terminus of the CTD (three mutants) (Fig. 8a). Western blot analysis with either anti-PorZ (Fig. 8b) or anti-His-tag antibodies (Fig. 8c) revealed



**Figure 8. Effects of the introduction of, or replacement with, oligohistidines on PorZ function.** (a) Location of insertions (i) and substitutions (>) of consecutive residues by polyhistidines at the junction (residues G<sup>680</sup>-G<sup>691</sup>, over yellow background) between CTD (green font) and the preceding domain of PorZ (F677i8H, Q678i6H, S683 > 6H, A686 > 6H, L689 > 6H and D690i6H) or at the C-terminus (I770 > 6H, I770i6H, and R776i8H).  $\beta$ -strands are indicated above the alignment. (b–e) Wild-type and mutant strains were grown to OD<sub>600</sub> = 1.0 and whole cultures were subjected to Western blot analysis with anti-PorZ (b), anti-polyhistidine (c), anti-Rgp (d) and anti-Kgp (e) antibodies. (f) The same strains were used for gingipain activity assays. (g) The level of surface exposure of PorZ in various mutants was analyzed by flow cytometry using anti-PorZ antibodies (red) and negative isotype control (blue). Representative histograms are shown from three independent experiments.

~81-kDa (full-length) and/or ~75-kDa (truncated) immunoreactive bands, which were found with the former antibody in all strains tested except  $\Delta$ PorZ (Fig. 8b). Both forms were found in mutant I770i6H with the anti-PorZ antibody, and the shorter fragment was exclusive for the I770 > 6H and L689 > 6H mutants. In turn,

Strain	Relevant genotype	Source
W83	Wild type	Reference strain
HG66	Wild type	Reference strain
$\Delta$ PorZ	<i>porZ</i> (NCBI: PG_RS07070; old locus PG1604)(Em <sup>r</sup> )	This study
PorZ <sup>+</sup>	<i>porZ</i> <sup>+</sup> (Em <sup>r</sup> )	This study
$\Delta$ PorU	<i>porU</i> (NCBI: PG_RS00120; old locus PG0026)(Em <sup>r</sup> )	This study
PorU <sup>C690A</sup>	<i>porU</i> .C690A(Em <sup>r</sup> )	This study
$\Delta$ PorN	<i>porN</i> (NCBI: PG_RS01305; old locus PG0291)(Em <sup>r</sup> )	This study
<i>PorZ</i> modifications		
F677i8H	<i>porZ</i> p.F677_insHHHHHHHH Q678(Em <sup>r</sup> )	This study
Q678i6H	<i>porZ</i> p.Q678_insHHHHHHH_T679(Em <sup>r</sup> )	This study
D690i6H	<i>porZ</i> p.D690_insHHHHHHH_G691(Em <sup>r</sup> )	This study
I770i6H	<i>porZ</i> p.I770_insHHHHHHH_R771(Em <sup>r</sup> )	This study
R776i8H	<i>porZ</i> p.R776insHHHHHHHH(Em <sup>r</sup> )	This study
S683 > 6 H	<i>porZ</i> p.G684H;S685H;A686H;S687H;E688H;L669H(Em <sup>r</sup> )	This study
A686 > 6 H	<i>porZ</i> p.S687H;E688H;L689H;D690H;G691H;V692H(Em <sup>r</sup> )	This study
L689 > 6 H	<i>porZ</i> p.D690H;G691H;V692H;Y693H;V694H;Y695H(Em <sup>r</sup> )	This study
I770 > 6 H	<i>porZ</i> p.R771H;F772H;A773H;V774H;I775H;R776H(Em <sup>r</sup> )	This study

**Table 1.** *P. gingivalis* strains used in this study.

results with anti-His-tag antibodies showed the oligohistidine motif in full-length and/or truncated PorZ in all mutants expressing the protein tagged in the L $\beta$ D2-CTD segment (Fig. 8c, *left panel*). In the case of the C-terminal mutants, the tag was detected only in two mutants expressing, respectively, PorZ C-terminally extended with an octahistidine attached to C-terminal residue R<sup>776</sup> (mutant R776i8H) and when hexahistidine was inserted after I<sup>770</sup> (mutant I770i6H) (Fig. 8c, *right panel*). In contrast, no tag was detected in the ~75-kDa form of PorZ with hexahistidine substituting the six most C-terminal residues (I770 > 6 H) and when inserted in the middle of the last  $\beta$ -strand (CTD- $\beta$ 7) of the protein (I770i6H). The lack of reactivity of the ~75-kDa band with anti-His-tag antibodies in this mutant clearly indicates that mutated PorZ was cleaved at the C-terminus, losing ~5-kDa.

To verify on which side (N- or C-terminal) the protein was cleaved in L $\beta$ D2-CTD linker mutants, we purified the truncated PorZ variant derived from the L689 > 6 H mutant by affinity chromatography, and subjected the protein to N-terminal sequencing analysis. The N-terminus was found to be intact, so cleavage must have occurred at the C-terminus, downstream of the inserted histidine-tag. The same can be assumed for the ~75-kDa PorZ variant derived from mutant Q678i6H.

Next, we determined the function of the PorZ variants in the processing and activation of gingipains. A membrane-bound form of RgpB (mt-RgpB) and the mature RgpA catalytic domain (RgpA<sub>cat</sub>) were clearly visible in the wild type and in mutants F677i8H, Q678i6H, S683 > 6 H, A686 > 6 H, D690i6H, I770i6H and R776i8H (Fig. 8d), which all possessed full-length PorZ (Fig. 8b,c). Conversely, mutants L689 > 6 H and I770 > 6 H, which do not have this PorZ variant, entirely lacked both mt-RgpB and RgpA<sub>cat</sub> (Fig. 8d). Similarly, the mature catalytic domain of Kgp was found only in mutants possessing full-length PorZ, while only the Kgp precursor was present in mutants missing full-length PorZ (Fig. 8e). Lack of gingipain processing in these mutants correlated with substantially lower Rgp and Kgp activities in the culture than in the wild type (Fig. 8f). The only exception was mutant Q678i6H, which seemed to process gingipains, but with much lower yields than the wild type (Fig. 8f). Finally, flow cytometry studies (Fig. 8g) revealed that all mutants with full-length PorZ displayed surface-exposed PorZ at levels comparable to those of the wild type. Conversely, strains with truncated PorZ (L689 > 6 H and I770 > 6 H) had negligible amounts of this protein on the cell surface, despite comparable expression levels of modified PorZ, as shown by Western blot (Fig. 8b,g).

## Conclusions

Members of the dysbiotic oral microbiome, such as *T. forsythia* and *P. gingivalis*, and some environmental Gram-negative bacteria uniquely possess T9SS dedicated to the export of proteins. This multi-component machinery consists of a minimum of twelve Por proteins and is responsible for the secretion of at least 32 proteins in *P. gingivalis*. The export signal is not a flexible peptide but rather a full ~70-residue protein domain located at the C-terminus called the CTD, which is removed upon secretion. Here, we discovered a new component of T9SS, PorZ, and found it loosely associated with the *P. gingivalis* cell surface in a manner independent of A-LPS anchorage. In T9SS cargo enzymes such as PPAD and gingipains, its absence prevented proper secretion to the extracellular medium, cleavage of CTDs, activation from precursor forms, and anchoring in the OM. In this case, cargos accumulated in the periplasmic space and the culture medium, the latter owing to leaky OM architecture. Increased mRNA expression of CTD-cargo proteins in the  $\Delta$ PorZ mutant suggested that absence of these functional proteins on the cell surface induced a feedback response, which augmented expression of these proteins and of some T9SS components, presumably in an effort to secrete more cargo proteins to the surface.

To gain insight into the molecular determinants of PorZ function, we solved its full-length crystal structure, which revealed two N-terminal seven-fold  $\beta$ -propeller domains in tandem. Such domains are widely used in macromolecular recognition and are engaged in protein-protein and protein-substrate interactions, as the

Strain	Relevant genotype	Source
TOP10	F- <i>mcrA</i> $\Delta$ ( <i>mrr-hsdRMS-mcrBC</i> ) $\Phi$ 80 <i>lacZ</i> $\Delta$ M15 $\Delta$ <i>lacX74 recA1 araD139 <math>\Delta</math>(<i>araI</i>)7697 <i>galU galK rpsL</i> (StrR) <i>endA1 nupG</i></i>	Invitrogen
NEB® 5-alpha	<i>fhuA2</i> $\Delta$ ( <i>argF-lacZ</i> )U169 <i>phoA glnV44</i> $\Phi$ 80 $\Delta$ ( <i>lacZ</i> )M15 <i>gyrA96 recA1 relA1 endA1 thi-1 hsdR17</i>	New England Biolabs
BL21 (DE3)	<i>fhuA2</i> [ <i>lon</i> ] <i>ompT gal</i> ( $\lambda$ DE3) [ <i>dcm</i> ] $\Delta$ <i>hdsS</i> $\lambda$ DE3 = $\lambda$ <i>sBamHlo</i> $\Delta$ <i>EcoRI-B int::</i> ( <i>lacI::PlacUV5::T7 gene1</i> ) <i>i21</i> $\Delta$ <i>tnin5</i>	Invitrogen
Plasmid	Relevant features	Source
pUC19	<i>E. coli</i> cloning vector; Ap <sup>r</sup>	Thermo Scientific
pCR-BluntII-TOPO	<i>E. coli</i> cloning vector; Km <sup>r</sup> , Zeo <sup>r</sup>	Invitrogen
pGEX-6P-1	<i>E. coli</i> expression vector, PreScission Protease cleavage site between N-terminal GST tag and target protein; Ap <sup>r</sup>	GE Healthcare
pVA2198	<i>E. coli</i> - <i>Bacteroides</i> shuttle vector, source of <i>ermF-ermAM</i> cassette; Sp <sup>r</sup>	55
pT-COW	<i>E. coli</i> - <i>Bacteroides</i> shuttle vector, Tc <sup>r</sup>	58
p1604AeB-D	Plasmid for <i>porZ</i> deletion mutagenesis, derivative of pUC19	This study
p291AeB-C	Plasmid for <i>porN</i> deletion mutagenesis, derivative of pUC19	This study
pPorU/pUC19/Erm	Plasmid for <i>porU</i> deletion mutagenesis, derivative of pUC19	This study
pT-COW-porZ	Plasmid for <i>porZ</i> complementation, derivative of pT-COW	This study
p1604CeB-H	Master plasmid for PorZ modifications; derivative of pUC19	This study
pAT1	Plasmid for Q678i6H mutagenesis in PorZ, derivative of p1604CeB-H	This study
pAG1	Plasmid for S683 > 6 H mutagenesis in PorZ, derivative of p1604CeB-H	This study
pAL5	Plasmid for I770i6H mutagenesis in PorZ, derivative of p1604CeB-H	This study
pAL11	Plasmid for A686 > 6 H mutagenesis in PorZ, derivative of p1604CeB-H	This study
pAL12	Plasmid for L689 > 6 H mutagenesis in PorZ, derivative of p1604CeB-H	This study
pAL13	Plasmid for D690i6H mutagenesis in PorZ, derivative of p1604CeB-H	This study
p1604M1	Plasmid for I770 > 6 H mutagenesis in PorZ, derivative of p1604CeB-H	This study
p1604M2	Plasmid for F677i8H mutagenesis in PorZ, derivative of p1604CeB-H	This study
p1604M3	Plasmid for R776i8H mutagenesis in PorZ, derivative of p1604CeB-H	This study
pPorU-E	Master plasmid for PorU modification, derivative of pUC19	This study
pPorU/C690A	Plasmid for C690A mutagenesis in PorU, derivative of pPorU-E	This study
pGEX-6P-1/PorZ	Plasmid for PorZ purification from <i>E. coli</i> , derivative of pGEX-6P-1	This study

**Table 2.** *E. coli* strains and plasmids used in this study.

architecture is generally versatile enough to enable binding of small molecules like sugars<sup>43</sup>. Based on structural similarities with sugar-binding proteins from *B. thetaiotaomicron* from the human gut microbiome, we hypothesized that PorZ may likewise have a potential glycan-binding function as part of, or independently from, T9SS secretion. Downstream of the  $\beta$ -propeller domains, PorZ further comprises a C-terminal seven-stranded  $\beta$ -sandwich, which conforms to the canonical CTDs of other T9SS-secreted proteins. This allowed us to further hypothesize that PorZ may itself be a T9SS cargo. Further functional studies showed that PorZ is actually transported to the cell surface *via* T9SS as a full-length protein with an intact CTD- a translocation that was independent of the presence or activity of sortase PorU. Consistently, PorZ was absent from the surface but apparently remained associated with the periplasmic side of the OM in a T9SS mutant lacking protein PorN.

We further studied the effect of oligohistidine insertions or replacements on PorZ expression, processing and translocation to the bacterial surface, and inserted hexa- or octahistidines or replaced six consecutive residues with histidines within linker L $\beta$ D2-CTD and at the C-terminus of CTD. Collectively, these mutations revealed that although CTD can be elongated at the C-terminus without affecting the secretory phenotype of the mutant, alteration of the native C-terminal sequence was absolutely prohibited<sup>26,53</sup>. Interestingly, in both cases truncation/substitution of C-terminal residues resulted in partial cleavage of CTD and accumulation of modified protein in the periplasm. However, modifications of the loop preceding CTD in PorZ had a negligible effect on protein function, as long as they did not affect strand CTD- $\beta$ 1. In T9SS cargos, such mutations decouple CTD removal by sortase PorU from the attachment of A-LPS, thus releasing fully-processed cargo proteins into the culture medium. This accounts for different functions of the linker domain preceding CTD, which is unstructured in cargos and provides a cleavage/A-LPS attachment site for sortase PorU<sup>12</sup>. In stark contrast, the intra-domain loop is well-structured in PorZ and resistant to proteolysis by PorU, thus suggesting an important role of CTD in PorZ function in T9SS.

To sum up, the reported full molecular and functional characterization of PorZ, a novel essential surface component of T9SS, contributes to our understanding of protein secretion as part of host-microbiome interactions by dysbiotic members of the human oral cavity.

## Materials and Methods

**Bacterial strains and general growth conditions.** *Porphyromonas gingivalis* strain W83 (wild type and mutants, listed in Table 1) was grown in enriched tryptic soy broth (eTSB per liter: 30 g trypticase soy broth, 5 g yeast extract, 5 mg hemin, pH 7.5; further supplemented with 0.5 g L-cysteine and 2 mg menadione) or on eTSB

blood agar (eTSB medium *plus* 1.5% agar, further supplemented with 4% defibrinated sheep blood) at 37 °C in an anaerobic chamber (Don Whitley Scientific, UK) with an atmosphere of 90% nitrogen, 5% carbon dioxide and 5% hydrogen. *Escherichia coli* strains (listed in Table 2), used for all plasmid manipulations, were grown in Luria-Bertani (LB) medium and on 1.5% agar LB plates. For antibiotic selection in *E. coli*, ampicillin was used at 100 µg/ml, kanamycin at 50 µg/ml, spectinomycin at 50 µg/ml and erythromycin at 250 µg/ml. *P. gingivalis* mutants were grown in the presence of erythromycin at 5 µg/ml and/or tetracycline at 1 µg/ml.

**Generation of *P. gingivalis* ΔPorZ, ΔPorU and ΔPorN deletion mutants.** All *P. gingivalis* deletion mutants were generated by homologous recombination, as previously described for other genes<sup>54</sup>. Three suicide plasmids were generated analogously. For deletion mutant ΔPorZ, briefly, two 1-kb flanking regions on either side of the *porZ* gene were amplified from genomic DNA by PCR with primer pairs PG1604FrANdeIF/PG1604FrASmaR (for primers and sequences, see Supplementary Table S2) and PG1604FrBXbaIF/PG1604FrBSalIR. The PCR product was cloned into plasmid pUC19 with an inserted erythromycin-resistance (*erm*) cassette (*ermF-ermAM*), amplified from plasmid pVA2198<sup>55</sup> with primers ermFAMsmaIF/ermFAMsalIR. Correct placement and orientation of the DNA segments in resulting plasmid p1604AeB-D was confirmed by sequencing.

A plasmid for mutant ΔPorN, designated p291AeB-C, was generated similarly using vector pUC19 and primer pairs PG291FrBXbaIF/PG291FrBPstIR and PG291FrANdeIF/PG291FrASmaR, as well as primers ermFAMsmaIF/ermFAMXbaIR to introduce cassette *ermF-ermAM* between chromosomal DNA fragments.

In like manner, a plasmid for mutant ΔPorU was obtained using primer pairs PorUFrag5\_F/PorUFrag5\_R and PorUFrag3\_F/PorUFrag3\_R to amplify 1-kb upstream and downstream flanking regions of the *porU* gene respectively. Blunt-end PCR products were cloned into plasmid pCR-BluntII-TOPO using the Zero Blunt TOPO PCR cloning kit (Invitrogen) and then subcloned into a pUC19 plasmid containing a promoterless *erm* resistance gene<sup>55</sup> yielding plasmid pPorU/pUC19/Erm.

The resulting suicide plasmids for generating the ΔPorU, ΔPorZ and ΔPorN mutants were introduced by electroporation into electrocompetent wild-type *P. gingivalis* cells prepared as previously described<sup>56</sup>. Resultant clones were selected on erythromycin plates and double-crossover genomic recombination was confirmed by DNA sequencing of the manipulated region. Isogenic mutant ΔPorU was further verified for inoffensiveness on translation of the downstream gene encoding protein LptO (*alias* PorV and PG0027) through Western blot analysis using rabbit polyclonal anti-LptO antibodies.

**Generation of *P. gingivalis* PorZ and PorU mutants.** Master plasmid p1604CeB-H was created for *porZ* modifications in a pUC19 background through sequential insertion of the following PCR fragments: a 2.3-kb PCR fragment encoding most of the *porZ* gene (primer pair: PG1604FrANdeIF2/PG1604FrCSmaR); an *ermF-ermAM* cassette (primer pair: ermFAMsmaIF/ermFAMsalIR); and a 1-kb downstream fragment of the *porZ* gene (primer pair: PG1604FrBXbaIF/PG1604FrBSalIR). The same method was used to create pPorU-E master plasmid. Here, PCR fragments consisted of a 2.5-kb fragment partially encompassing the *porU* gene (primer pair: Pg26\_AR/Pg26\_AF); an *ermF-ermAM* cassette (primer pair: Pg26\_ER/Pg26\_EF); and a 0.9-kb downstream fragment of the *porU* gene (primer pair: Pg26\_BR/Pg26\_BF).

Similarly to the reported generation of RgpB and Cpg mutants<sup>53</sup>, SLIM mutagenesis<sup>57</sup> was employed to replace the catalytic cysteine of PorU (C690) with alanine within the pPorU-E master plasmid to yield mutant PorU<sup>C690A</sup>.

PorZ mutants including insertion of an oligohistidine-tag or substitution of six consecutive residues with six histidines at various locations at the inter-domain linker between domains βD2 and the CTD (mutants PorZ F677i8H, Q678i6H, S683 > 6H, A686 > 6H, L689 > 6H, and D690i6H) or at the C-terminus (mutants PorZ I770 > 6H, I770i6H, and R776i8H) were generated by the SLIM method within the *porZ* master plasmid (see Supplementary Table S2). Mutated plasmids were verified by DNA sequencing of the pertinent region before being electroporated into wild-type *P. gingivalis* for homologous recombination<sup>54</sup>. Resistant clones were selected using erythromycin-selective media.

**Construction of *porZ* complementation strain (PorZ<sup>+</sup>).** To construct a complementation plasmid (pT-COW-porZ), a 3166-bp fragment containing the entire *porZ* coding sequence *plus* 544 bp upstream and 291 bp downstream was obtained by PCR from *P. gingivalis* genomic DNA with primers pT-COWporZForSalI and pT-COWporZRevNheI. Next, the fragment was cloned into SalI and NheI restriction sites of the pT-COW plasmid<sup>58</sup> by a standard procedure. After pT-COW-porZ sequencing, the plasmid was introduced into ΔPorZ strain by conjugation according to<sup>59</sup>.

**Culture partitioning and subcellular fractionation.** Two-day-old cultures in the stationary phase of wild-type and mutant *P. gingivalis* were adjusted to OD<sub>600</sub> = 1.5 and designated “whole culture” (WC). Bacterial cells were collected from WC by centrifugation (8000 × g, 15 min). The cell pellet was then washed and resuspended in PBS up to the initial volume of WC, and centrifuged. This fraction is referred to as “washed cells”. The collected cell-free culture medium was ultracentrifuged (100,000 × g, 1 h) to remove vesicles and the supernatant was concentrated 10-fold by ultrafiltration with filter devices of 10-kDa cut-off; this fraction was designated “medium” (Med). Subcellular fractionation was carried out as described<sup>60</sup>, with some modifications. Briefly, bacteria were cultured until OD<sub>600</sub> = 1.2 and fraction “cell extract” was obtained by centrifugation of 0.5 ml of WC and a single wash step with PBS. The rest of the culture (7.5 ml) was centrifuged to produce a supernatant and a pellet. The supernatant was filtered through a 0.22-µm syringe filter and concentrated three times with centrifugal filter devices of 10-kDa cut-off to obtain the “medium” fraction. The pellet was rinsed with PBS and suspended in 2.5 ml buffer containing 0.25 M sucrose, 30 mM Tris-HCl, pH 7.6. After a 10-min incubation period, cells were pelleted (12,500 × g, 15 min) and rapidly resuspended in 2.5 ml of cold distilled water to disrupt outer membranes.

After a further 10-min incubation period, spheroplasts were obtained and separated by centrifugation (12,500 × g, 15 min). The supernatant was collected and designated as the “periplasmic” fraction. Next, spheroplasts were washed once with PBS, suspended in 2.5 ml of fresh PBS, and disrupted by brief sonication. The resulting solution was ultracentrifuged (150,000 × g, 1 h) and the soluble fraction was designated as the “cytoplasmic” fraction. The pellet was resuspended in 2.5 ml PBS and briefly sonicated, and was designated as the “cell envelope” fraction. Inner membranes were solubilized by incubation in PBS supplemented with 1% Triton X and 20 mM magnesium chloride (final concentration) for 20 min. Membranes were separated by ultracentrifugation (150,000 × g, 1 h) and the resulting supernatant was dubbed the “inner membrane” fraction. The pellet was then resuspended in PBS and sonicated, which resulted in the “outer membrane” fraction. All fractions were supplemented with peptidase inhibitors (5 mM tosyl-L-lysyl-chloromethane hydrochloride [TLCK], 1 mM 2,2'-dithiodipyridine [DTDP], 1x EDTA-free protein inhibitor cocktail (all from Roche) before storage at −20 °C.

**Antibodies.** Rabbit anti-RgpB antibodies, which also recognize the catalytic domain of RgpA<sup>61</sup>, were used as the primary antibody. Mouse monoclonal antibodies specific for the Kgp catalytic domain and a single epitope on RgpB (18E6), respectively, were developed at the University of Georgia (GA, USA) Monoclonal Antibody Facility using recombinant protein as the antigen<sup>54,62</sup>. A rabbit polyclonal antibody recognizing *P. gingivalis* peptidylarginine deiminase (PPAD) was generated by Cambridge Research Biochemicals, Billingham, UK<sup>63</sup>. Mouse monoclonal antibodies specific for PorU were developed using standard procedures in the Laboratory of Monoclonal Antibodies at the Malopolska Center of Biotechnology, Jagiellonian University (Krakow, Poland) with recombinant PorU as the immunizing antigen.

Mouse polyclonal antibodies against protein PorZ were obtained according to standard procedures. Briefly, six-week old Balb/C mice were injected intraperitoneally with 100 μg of antigen diluted in PBS and mixed 1:1 with Complete Freund's Adjuvant (Sigma Aldrich). Subsequent immunizations were performed with 50 μg of antigen mixed with Incomplete Freund's Adjuvant (Sigma Aldrich). The serum titer of polyclonal antibodies specific to the antigen was determined by ELISA according to standard procedures. After administration of *an adequate anaesthetic* (ketamine and xylazine), blood was collected from immunized animals by cardiac puncture. Polyclonal antibodies were purified from the serum by affinity chromatography with protein G, according to the manufacturer's instructions (Thermo Fisher Scientific), and then dialyzed against sterile PBS. Finally, the antibodies were purified by affinity chromatography employing recombinant PorZ, which was covalently bound to Pierce™ NHS-Activated Agarose Spin Columns (Thermo Fisher Scientific), according to the manufacturer's instructions. This work was carried out in the Laboratory of Monoclonal Antibodies at the Malopolska Center of Biotechnology, Jagiellonian University (Krakow, Poland). Mouse monoclonal antibodies against PorZ were also generated commercially by Abmart (Shanghai, China) using a synthetic peptide of sequence E-K-G-R-K-T-T-Q-F-P. All experiments were performed in accordance with relevant guidelines and regulations.

**Western blot analysis.** Sample preparation, separation by SDS-PAGE, electro-blotting onto polyvinylidene difluoride or nitrocellulose membranes and blocking were performed as described previously<sup>53</sup>. RgpA and RgpB catalytic domains were detected with rabbit polyclonal anti-Rgp antibodies; Kgp with mouse monoclonal anti-Kgp antibodies; PPAD with rabbit polyclonal anti-PPAD antibodies; and PorZ with mouse polyclonal anti-PorZ antibodies (see also above). In addition, oligohistidine-tagged PorZ was detected using mouse monoclonal THE™ His-Tag antibodies (Genscript, USA) at 0.125 μg/ml. Development with polyclonal anti-mouse or anti-rabbit horseradish peroxidase-conjugated secondary antibody (BD Pharmingen and Amersham Pharmacia) was carried out using the ECL Western Blotting substrate kit according to the manufacturer's instructions (Pierce, UK). Streptavidin conjugated to horseradish peroxidase was used to detect MmdC, a biotinylated IM-associated protein.

**Dot blot analysis.** *P. gingivalis* cells were harvested at OD<sub>600</sub> = 1.5, washed and resuspended in cold PBS to yield OD<sub>600</sub> = 1.0. Half of the suspension was sonicated to disrupt cell membranes, and 5 μl of either intact or disrupted cell suspension were spotted onto 0.22-μm nitrocellulose membranes and air dried. Subsequent steps were performed as described for Western blot analysis.

**Flow cytometry analysis.** Wild-type and mutant *P. gingivalis* strains were grown in eTSB until they reached the late exponential or early stationary growth phase (OD<sub>600</sub> ~ 1.2–1.5). Bacterial cells were harvested by centrifugation, washed twice with PBS and adjusted to OD<sub>600</sub> = 1.0 with buffer (PBS supplemented with 1% bovine serum albumin). Then, 100 μl of cell suspension was transferred to a 96-well plate and incubated for 30 min with the previous buffer. Cells were collected by centrifugation (5,000 × g, 5 min) and the pellet was resuspended in the previous buffer containing mouse antiserum specific for PorZ, RgpB or PorU at a total protein concentration of 30 μg/ml, and incubated for 30 min. Thereafter, cells were centrifuged (5,000 × g, 5 min) and the newly obtained pellet was resuspended in the previous buffer containing goat anti-mouse antibody conjugated with fluorescein isothiocyanate (Abcam) at 1:200 dilution and incubated for 30 min. Cells were washed twice with PBS after each incubation with antibodies or streptavidin-Alexa Fluor 488 conjugate. The whole staining procedure was performed on ice. After staining, one-color flow cytometry analyses were performed using a FACSCalibur apparatus (BD Biosciences) operating with CellQuest software (BD Biosciences). Graphs were prepared using the *FLOWJO v.10 program* (Ashland, USA).

**Electron microscopy.** Wild-type *P. gingivalis* cells grown in eTSB were harvested at 4,000 × g, washed once with PBS, and then fixed with 2% [w:v] paraformaldehyde, 0.02% [v:v] glutaraldehyde fixative for 30 min. Cells were washed twice with PBS, then incubated with glycine (2 mg/ml) for 10 min. After a further wash with PBS,

Dataset	PorZ (Se absorption peak)	PorZ (native)
Space group	P4 <sub>3</sub> 2 <sub>1</sub> 2	P4 <sub>3</sub> 2 <sub>1</sub> 2
Cell constants (a, b, c, in Å)	113.4, 113.4, 139.6	115.5, 115.5, 139.9
Wavelength (Å)	0.9793	0.8726
No. of measurements/unique reflections	144,224/31,485 <sup>f</sup>	260,647/21,600
Resolution range (Å) (outermost shell) <sup>a</sup>	44.0–3.10 (3.18–3.10)	48.5–2.90 (3.05–2.90)
Completeness (%)	99.7 (98.6)	99.9 (99.7)
R <sub>merge</sub> <sup>b</sup>	0.146 (0.976)	0.108 (0.913)
R <sub>tr.m.</sub> [=R <sub>meas</sub> ] <sup>c</sup> /CC(1/2) <sup>c</sup>	0.165 (1.206)/0.993 (0.584)	0.113 (0.953)/0.999 (0.837)
Average intensity <sup>d</sup>	10.8 (1.8)	20.0 (3.4)
B-Factor (Wilson) (Å <sup>2</sup> )/Aver. multiplicity	63.0/4.6 (2.8)	63.3/12.1 (12.1)
Number of Se-atom sites used for phasing	10 (out of 12)	
Resolution range used for refinement (Å)		48.5–2.90
No. of reflections used (test set)		20,876 (723)
Crystallographic R <sub>factor</sub> (free R <sub>factor</sub> ) <sup>b</sup>		0.183 (0.238)
No. of protein atoms/solvent molecules/ neutral ligands/ ionic ligands		5,678/43/ 1 tetraethylene glycol, 1 diethylene glycol, 4 glycerol 8 zinc, 4 calcium, 1 chloride, 1 cacodylate
Rmsd from target values <sup>e</sup>		
bonds (Å)/angles (°)		0.010/1.22
Average B-factors (Å <sup>2</sup> )		74.8
All-atom contacts and geometry analysis <sup>e</sup>		
Residues		
in favored regions/outliers/all residues		701 (94.0%)/1/746
with poor rotamers/bad bonds/bad angles		30 (5.0%)/0/0
with Cβ deviations >0.25 Å/clashscore		0/9.13 (97 <sup>th</sup> percentile)
MolProbity score		2.42 (94 <sup>th</sup> percentile)

**Table 3. Crystallographic data.** <sup>a</sup>Values in parenthesis refer to the outermost resolution shell. <sup>b</sup>For definitions, see Table 1 in ref. 85. <sup>c</sup>For definitions, see refs 86 and 87. <sup>d</sup>Average intensity is  $\langle I/\sigma(I) \rangle$  of unique reflections after merging according to the XDS program<sup>70</sup>. <sup>e</sup>According to MOLPROBITY<sup>83,88</sup>. <sup>f</sup>Friedel mates were kept separately.

cells were dehydrated through graded alcohol, embedded in London Resin White (London Resin, UK) in gelatin capsules, and polymerized at 55 °C for 20 hours. Samples were then sectioned to ultra-thin sections (70 μM) on a Leica UC6 Ultramicrotome and picked up on Nickel 200 Square mesh fine bar grids (Gilder grids, UK) coated with 0.12% Formvar and 0.12% Pioloform. Sections were subsequently blocked with 10% [v:v] goat blocking buffer (5% bovine serum albumin [BSA], 10% normal goat serum, 0.2% cold water fish-skin gelatin, 10 mM sodium azide, pH 7.4) for 1 h, washed six times with incubation buffer (0.2% Aurion BSA-c, 10 mM sodium azide), and then incubated with monoclonal anti-PorZ antibodies (5 μg/ml) for 1 h. Sections were then washed six times with incubation buffer and then incubated with goat anti-mouse-5-nm gold conjugate for 1 h (BBI Solutions 1/40). The sections were then washed six times with PBS, post-fixed with 2% [v:v] glutaraldehyde for 5 min, and finally washed twice with PBS and six times with water before drying.

Grids were then treated with 2% osmium vapour for 5 min, stained first with 1% uranyl acetate and then with Reynolds' lead citrate<sup>64</sup>. Grids were visualized in a 100-kV FEI CM120 electron transmission microscope and digital images were taken with a SIS Morada CCD camera.

**Quantitative RT-PCR analysis.** Quantitative RT-PCR data were obtained from four separate experiments. Briefly, wild-type and ΔPorZ cells were cultured to the mid-exponential phase and collected at OD<sub>600</sub> = 0.8–0.9. Cells were stabilized with RNAProtect Bacteria Reagent (Qiagen, Germany) prior to mRNA extraction with the RNAqueous Total RNA Isolation Kit (Ambion, USA). Reverse transcription was carried out on 2 μl of total RNA using the AffinityScript qPCR cDNA Synthesis Kit (Stratagene, Australia) in 10 μl total volume, according to the manufacturer's instructions. Quantitative RT-PCR was carried out in 25-μl singleplex reactions using 5 μl of a 1:50 dilution of cDNA along with TaqMan probe and primers against genes related to T9SS components (*porT*, *sov*, *porU*, *lptO*, *porN*, *porO* and *porW*) or cargos (*rgpB*, *kpg*, and *cpg70*). A housekeeping DNA gyrase (*gyrA*) gene and ribosomal 16S (*r16s*) were used as calibrator genes (see primers and probes in Supplementary Table S2). Thereafter, thermal cycling with a Stratagene Mx3005 P Real-Time PCR System<sup>®</sup> using the Brilliant<sup>®</sup> II QPCR Master Mix (Stratagene, Australia) entailed activation for 15 min at 95 °C plus 40 cycles of annealing at 95 °C for 20 s and extension at 60 °C for 1 min. Fluorescence intensities were normalized against a passive fluorophore carboxy-X-rhodamine (ROX) present in the Master Mix and converted to absolute quantities using standard curves. The expression of target genes was normalized to the housekeeping *gyrA* and *r16s* genes independently by

expressing the data as the average threshold cycle value of the target gene divided by that of either housekeeping gene. The average ratio in the wild type was arbitrarily set to 1.0 for reference.

### Large-scale expression, purification and oligomeric characterization of recombinant PorZ.

Genomic DNA was isolated from wild-type *P. gingivalis* strain W83 using the Genomic Mini System (A&A Biotechnology, Gdansk, Poland), according to the manufacturer's instructions. The gene encoding residues Q<sup>26</sup>-R<sup>776</sup> (PorZ residue numbers in superscript notation, see UniProt entry Q9S3Q8), i.e. without the predicted signal peptide (M<sup>1</sup>-A<sup>25</sup>), was amplified by PCR, purified, and cloned into the pGEX-6P-1 expression vector using SmaI/NotI sites (for primer sequences, see Supplementary Table S2). The resulting recombinant product encoded an N-terminal glutathione-S-transferase (GST) moiety, a PreScission protease cleavage site, and the cloned protein. The plasmid was verified by DNA sequencing and transformed into *E. coli* expression strain BL21 (DE3) under the control of the T7 promoter. Transformed cells were grown in LB medium at 37 °C until OD<sub>600</sub> = 0.75–1.0, and then for a further 30 min at 20 °C. Protein expression was induced by the addition of 0.25 mM isopropyl-1-thio-β-D-galactopyranoside and allowed to proceed for 8 h at 20 °C. Thereafter, cells were harvested by centrifugation (15 min, 6,000 × g, 4 °C), re-suspended in PBS supplemented with 0.02% sodium azide (15 ml per pellet from 1 liter of culture), and subsequently lysed by sonication. Cell lysates were clarified by centrifugation (50 min, 50,000 × g, 4 °C), filtered through a 0.45-μm syringe filter, and loaded onto a glutathione-Sepharose 4 Fast Flow column equilibrated with PBS supplemented with 0.02% sodium azide at 4 °C. The GST-tag was removed from the recombinant protein bound to the column by intra-column cleavage with PreScission Protease (Amersham Biosciences). The cloning strategy left ten residues (G-P-L-G-S-P-E-F-P-G; confirmed by N-terminal sequencing) attached to the N-terminus of the recombinant protein. The latter was subsequently purified by size-exclusion chromatography using a HiLoad 16/60 Superdex 200 pg (GE Healthcare LifeSciences) column. A selenomethionine variant of PorZ was obtained in the same way except that selenomethionine was used instead of methionine in cell cultures. Size-exclusion chromatography in a calibrated Superdex 200 column further revealed that purified recombinant PorZ eluted as a monomer (Supplementary Fig. S4).

**Purification of histidine-tagged PorZ from *P. gingivalis* mutant.** Among the aforementioned PorZ histidine-insertion mutants (see section *Generation of P. gingivalis PorZ and PorU mutants*), mutant PorZ L689 > 6 H yielded PorZ protein truncated by ~5 kDa. To determine the cleavage site, the protein was purified for N-terminal sequence analysis. Briefly, the mutant cell culture supernatant was clarified by centrifugation (8,000 × g, 30 min, 4 °C). The pellet was resuspended in cold PBS, and peptidase inhibitors TLCK (final concentration 5 mM) and DTDP (1 mM) were added, in addition to 2% protein inhibitor cocktail (Sigma). The sample was then sonicated and ultracentrifuged (150,000 × g, 1 h, 4 °C), and the clarified lysate was dialyzed against buffer (20 mM sodium phosphate, 500 mM sodium chloride, 20 mM imidazole, 0.02% sodium azide, pH 7.4), concentrated by ultrafiltration in a filter device with 10-kDa cut-off, and stirred gently overnight with Nickel Sepharose 6 Fast Flow resin (GE Healthcare), previously equilibrated with the above buffer. The resin was then loaded on a column, washed until OD<sub>280</sub> reached background level, and eluted with 20 mM sodium phosphate, 0.5 M sodium chloride, 500 mM imidazole, pH 7.4 to recover mutant protein PorZ L689 > 6 H. Fractions containing the highest protein concentration - as determined by a bicinchoninic acid assay - were pooled, concentrated, and further analyzed. Purified protein was resolved by SDS-PAGE, transferred onto a polyvinylidene difluoride membrane, and stained with Coomassie Brilliant Blue. The band of protein was excised and subjected to N-terminal sequencing.

**Mass spectrometry of media samples.** Proteins in concentrated (20x) particle-free growth medium were ultracentrifuged and resolved by SDS-PAGE (NuPAGE Novex Bis-Tris System). Gels were stained with SimplyBlue™ SafeStain (Novex) and washed in distilled water. Gel lanes were cut into sections and subjected to mass spectroscopy and follow up analysis, as described earlier<sup>65</sup>.

**Enzyme activity assays.** The extracellular hydrolytic activity of *P. gingivalis* proteases was determined as described previously<sup>66,67</sup>. Briefly, the chromogenic *p*-nitroanilide (*p*NA) substrates benzoyl-Arg-*p*NA, acetyl-Lys-*p*NA, N-Gly-Pro-*p*NA and N-Ala-Phe-Pro-*p*NA (all from Bachem) were used to detect RgpA/B, Kgp, dipeptidyl peptidase IV, and prolyl tripeptidyl peptidase A, respectively. Samples were preincubated in buffer in 96-well plates prior to the addition of substrate to a total volume of 200 μl. The final concentration of each substrate was 1 mM. The rate of substrate hydrolysis, monitored through accumulation of *p*NA, was followed at 405 nm and the activity of each enzyme was given as mOD/min/μl or as a % of wild-type activity.

**Crystallization and diffraction data collection.** Crystallization assays were performed by the sitting-drop vapor diffusion method. Reservoir solutions were prepared by a Tecan robot and 100 nl crystallization drops were dispensed on 96 × 2-well MRC plates (Innovadyne) by a Phoenix nanodrop robot (Art Robbins) or a Cartesian Microsys 4000 XL robot (Genomic Solutions) at the joint IBMB/IRB Automated Crystallography Platform ([www.sbu.csic.es/facilities/automated-crystallographic-plattform](http://www.sbu.csic.es/facilities/automated-crystallographic-plattform)). Plates were stored in Bruker steady-temperature crystal farms at 4 °C or 20 °C. Successful conditions were scaled up to the microliter range in 24-well Cryschem crystallization dishes (Hampton Research). The best crystals of PorZ were obtained at 20 °C in drops containing 2 μl of protein solution (at 6.5 mg/mL in 5 mM Tris-HCl, pH 8.0), 1 μL of reservoir solution (8% polyethylene glycol 10,000, 0.2 M zinc acetate, 0.1 M sodium cacodylate, pH 6.5) and 0.25 μl of additive solution (0.1 M calcium chloride). Selenomethionine-containing crystals were obtained similarly with 1 μl of the same reservoir solution, 1 μl of protein solution (at 5 mg/mL in 5 mM Tris-HCl, pH 8.0) and 0.35 μl of additive solution (7% butanol). Crystals were cryo-protected by rapid passage through drops containing increasing amounts of glycerol (up to 30% [v/v]). A complete dataset was collected at 100 K from a liquid-N<sub>2</sub> flash cryo-cooled (Oxford

Cryosystems 700 series cryostream) native crystal at beam line ID23–2 of ESRF synchrotron (Grenoble, France) using a MAR225 CCD detector. A complete dataset from a selenomethionine-derivatized crystal was similarly collected at the absorption peak wavelength of selenium on a Pilatus 6 M pixel detector (from Dectris) at beam line XALOC of the ALBA synchrotron (Cerdanyola, Barcelona<sup>68</sup>). Crystals were tetragonal and contained one protein molecule in the asymmetric unit (solvent content 57%,  $V_M = 2.9 \text{ \AA}^3/\text{Da}^{69}$ ). Diffraction data were processed with programs XDS<sup>70</sup> and XSCALE<sup>71</sup>, and transformed with XDSCONV to formats suitable for the PHENIX<sup>72</sup> and CCP4 suites<sup>73</sup> of programs. Table 3 provides essential data-processing statistics.

**Structure solution and refinement.** The structure of PorZ was solved by single-wavelength anomalous diffraction with data collected at the selenium absorption peak wavelength from the selenomethionine-derivatized crystal, applying the AUTOSOL protocol from the PHENIX package<sup>74</sup>. The program found 10 of the 12 theoretical selenium sites *plus* seven minor sites corresponding to surface ions, as determined a posteriori, which resulted from the crystallization conditions containing zinc and cacodylate (see the previous section). Phasing with all these sites identified  $P4_32_1$  as the correct enantiomorphic space group and yielded an initial figure of merit of 0.26. Subsequent automatic density modification and model building produced a partial model of 373 (mostly alanine) residues in 72 chains, which showed an overall model-map correlation of 0.34 and an  $R_{\text{free}}$  value of 0.46. The resulting Fourier map enabled us to perform manual model building with the COOT program<sup>75</sup>, which was assisted by a homology model of PorZ obtained by threading with the LOMETS server (<http://zhanglab.ccmb.med.umich.edu/LOMETS><sup>76</sup>). A more complete partial model was obtained and refined against the diffraction data of the selenomethionine derivative to 3.1 Å resolution with the PHENIX<sup>77</sup> and BUSTER/TNT<sup>78</sup> programs. The refined partial model was subsequently used to solve the native structure with data to 2.9 Å resolution by Fourier synthesis with PHENIX. Given the anisomorphism in cell axes a and b between the native and derivative data (see Table 3), these calculations included rigid-body refinement and simulated annealing, in addition to individual coordinate, translation/libration/screw-motion (TLS) and grouped thermal-displacement-parameter refinement. Thereafter, careful manual model building alternated with two cycles of automatic model building and refinement with the AUTOBUILD routine of PHENIX<sup>79</sup>. Finally, additional manual model building was alternated with regular crystallographic refinement using PHENIX and BUSTER/TNT, including TLS refinement, over several cycles until the final model of PorZ was obtained. This consisted of residues G<sup>29</sup>-R<sup>776</sup> *plus* a strongly bound calcium ion. In addition, eight zinc ions, four calcium ions, one chloride ion, one cacodylate ion, one tetraethylene glycol, one diethylene glycol and four glycerol molecules arising from the crystallization and cryo-protection conditions were tentatively assigned on the surface of the molecule in addition to 43 unambiguous solvent molecules. See Table 3 for the final refinement and model quality statistics.

**Bioinformatics and statistical analyses.** Structural similarity searches were performed with DALI<sup>80</sup>, and figures were prepared with the CHIMERA program<sup>81</sup>. Structure superpositions were performed with the SSM routine<sup>82</sup> within COOT. The final structure of PorZ was validated with MOLPROBITY<sup>83</sup> and deposited with the Protein Data Bank (PDB) at [www.pdb.org](http://www.pdb.org) (access code 5M11). Quantitative RT-PCR results were tested for normality distribution, and differences were analysed using Student's t-test with SPSS v.16 software. Differences in enzyme activities and PorZ surface exposure in mutants in comparison to the wild-type parental strain were analyzed with GraphPad Prism 7 software (GraphPad Software, CA, USA) using one-way ANOVA with Bonferroni's correction. *P*-values below 0.05 were considered significant.

## References

- Zückert, W. R. Secretion of bacterial lipoproteins: through the cytoplasmic membrane, the periplasm and beyond. *Biochim. Biophys. Acta* **1843**, 1509–1516 (2014).
- Abby, S. S. *et al.* Identification of protein secretion systems in bacterial genomes. *Sci. Rep.* **6**, 23080 (2016).
- Desvaux, M., Hebraud, M., Talon, R. & Henderson, I. R. Outer membrane translocation: numerical protein secretion nomenclature in question in mycobacteria. *Trends Microbiol.* **17**, 338–340 (2009).
- Sato, K. *et al.* A protein secretion system linked to bacteroidete gliding motility and pathogenesis. *Proc. Natl. Acad. Sci. USA* **107**, 276–281 (2010).
- McBride, M. J. & Zhu, Y. Gliding motility and Por secretion system genes are widespread among members of the phylum Bacteroidetes. *J. Bacteriol.* **195**, 270–278 (2013).
- Chagnot, C., Zorgani, M. A., Astruc, T. & Desvaux, M. Proteinaceous determinants of surface colonization in bacteria: bacterial adhesion and biofilm formation from a protein secretion perspective. *Front. Microbiol.* **4**, 303 (2013).
- Kharade, S. S. & McBride, M. J. *Flavobacterium johnsoniae* PorV is required for secretion of a subset of proteins targeted to the type IX secretion system. *J. Bacteriol.* **197**, 147–158 (2015).
- Nakayama, K. *Porphyromonas gingivalis* and related bacteria: from colonial pigmentation to the type IX secretion system and gliding motility. *J. Periodontol. Res.* **50**, 1–8 (2015).
- Kadowaki, T. *et al.* A two-component system regulates gene expression of the type IX secretion component proteins via an ECF sigma factor. *Sci. Rep.* **6**, 23288 (2016).
- Eckburg, P. B. *et al.* Diversity of the human intestinal microbial flora. *Science* **308**, 1635–1638 (2005).
- Holt, S. C. & Ebersole, J. L. *Porphyromonas gingivalis*, *Treponema denticola*, and *Tannerella forsythia*: the “red complex”, a prototype polybacterial pathogenic consortium in periodontitis. *Periodontol.* **2000** **38**, 72–122 (2005).
- de Diego, I. *et al.* The outer-membrane export signal of *Porphyromonas gingivalis* type IX secretion system (T9SS) is a conserved C-terminal  $\beta$ -sandwich domain. *Sci. Rep.* **6**, 23123 (2016).
- Veith, P. D. *et al.* Major outer membrane proteins and proteolytic processing of RgpA and Kgp of *Porphyromonas gingivalis* W50. *Biochem. J.* **363**, 105–115 (2002).
- Seers, C. A. *et al.* The RgpB C-terminal domain has a role in attachment of RgpB to the outer membrane and belongs to a novel C-terminal-domain family found in *Porphyromonas gingivalis*. *J. Bacteriol.* **188**, 6376–6386 (2006).
- Shoji, M. *et al.* Por secretion system-dependent secretion and glycosylation of *Porphyromonas gingivalis* hemin-binding protein 35. *PLoS ONE* **6**, e21372 (2011).
- Gorasia, D. G. *et al.* *Porphyromonas gingivalis* type IX secretion substrates are cleaved and modified by a sortase-like mechanism. *PLoS Pathog.* **11**, e1005152 (2015).

17. Glew, M. D. *et al.* PG0026 is the C-terminal signal peptidase of a novel secretion system of *Porphyromonas gingivalis*. *J. Biol. Chem.* **287**, 24605–24617 (2012).
18. Chen, Y. Y. *et al.* The outer membrane protein LptO is essential for the O-deacylation of LPS and the co-ordinated secretion and attachment of A-LPS and CTD proteins in *Porphyromonas gingivalis*. *Mol. Microbiol.* **79**, 1380–1401 (2011).
19. Gorasia, D. G. *et al.* Structural insights into the PorK and PorN components of the *Porphyromonas gingivalis* Type IX Secretion System. *PLoS Pathog.* **12**, e1005820 (2016).
20. Pike, R., McGraw, W., Potempa, J. & Travis, J. Lysine- and arginine-specific proteinases from *Porphyromonas gingivalis*. Isolation, characterization, and evidence for the existence of complexes with hemagglutinins. *J. Biol. Chem.* **269**, 406–411 (1994).
21. Potempa, J., Pavloff, N. & Travis, J. *Porphyromonas gingivalis*: a proteinase/gene accounting audit. *Trends Microbiol.* **3**, 430–434 (1995).
22. Chen, Y. Y. *et al.* CPG70 is a novel basic metallocarboxypeptidase with C-terminal polycystic kidney disease domains from *Porphyromonas gingivalis*. *J. Biol. Chem.* **277**, 23433–23440 (2002).
23. Shibata, Y. *et al.* A 35-kDa co-aggregation factor is a heme binding protein in *Porphyromonas gingivalis*. *Biochem. Biophys. Res. Commun.* **300**, 351–356 (2003).
24. Shoji, M. *et al.* Characterization of heme-binding protein 35 (HBP35) in *Porphyromonas gingivalis*: its cellular distribution, thioredoxin activity and role in heme utilization. *BMC Microbiol.* **10**, 152 (2010).
25. McGraw, W. T., Potempa, J., Farley, D. & Travis, J. Purification, characterization, and sequence analysis of a potential virulence factor from *Porphyromonas gingivalis*, peptidylarginine deiminase. *Infect. Immun.* **67**, 3248–3256 (1999).
26. Goulas, T. *et al.* Structure and mechanism of a bacterial host-protein citrullinating virulence factor, *Porphyromonas gingivalis* peptidylarginine deiminase. *Sci. Rep.* **5**, 11969 (2015).
27. Veith, P. D. *et al.* Protein substrates of a novel secretion system are numerous in the Bacteroidetes phylum and have in common a cleavable C-terminal secretion signal, extensive post-translational modification, and cell-surface attachment. *J. Proteome Res.* **12**, 4449–4461 (2013).
28. Veith, P. D. *et al.* *Porphyromonas gingivalis* outer membrane vesicles exclusively contain outer membrane and periplasmic proteins and carry a cargo enriched with virulence factors. *J. Proteome Res.* **13**, 2420–2432 (2014).
29. Shi, Y. *et al.* Genetic analyses of proteolysis, hemoglobin binding, and hemagglutination of *Porphyromonas gingivalis*. Construction of mutants with a combination of *rgpA*, *rgpB*, *kgp*, and *hagA*. *J. Biol. Chem.* **274**, 17955–17960 (1999).
30. Okamoto, K. *et al.* Involvement of a lysine-specific cysteine proteinase in hemoglobin adsorption and heme accumulation by *Porphyromonas gingivalis*. *J. Biol. Chem.* **273**, 21225–21231 (1998).
31. Sroka, A., Sztukowska, M., Potempa, J., Travis, J. & Genco, C. A. Degradation of host heme proteins by lysine- and arginine-specific cysteine proteinases (gingipains) of *Porphyromonas gingivalis*. *J. Bacteriol.* **183**, 5609–5616 (2001).
32. Nguyen, K. A. *et al.* Verification of a topology model of PorT as an integral outer-membrane protein in *Porphyromonas gingivalis*. *Microbiology* **155**, 328–337 (2009).
33. Potempa, J., Pike, R. & Travis, J. Titration and mapping of the active site of cysteine proteinases from *Porphyromonas gingivalis* (gingipains) using peptidyl chloromethanes. *Biol. Chem.* **378**, 223–230 (1997).
34. Guo, Y., Nguyen, K. A. & Potempa, J. Dichotomy of gingipains action as virulence factors: from cleaving substrates with the precision of a surgeon's knife to a meat chopper-like brutal degradation of proteins. *Periodontol.* **2000** **54**, 15–44 (2010).
35. Mikolajczyk, J. *et al.* Sequential autolytic processing activates the zymogen of Arg-gingipain. *J. Biol. Chem.* **278**, 10458–10464 (2003).
36. Taguchi, Y. *et al.* Involvement of an Skp-like protein, PGN\_0300, in the Type IX Secretion System of *Porphyromonas gingivalis*. *Infect. Immun.* **84**, 230–240 (2016).
37. Fülöp, V. & Jones, D. T.  $\beta$ -Propellers: structural rigidity and functional diversity. *Curr. Opin. Struct. Biol.* **9**, 715–721 (1999).
38. Gomis-Rüth, F. X. Hemopexin domains. In: *Handbook of metalloproteins*. (ed<sup>^</sup>(eds Messerschmidt, A., Bode, W. & Cygler, M.)). John Wiley & Sons, Ltd. (2004).
39. Murzin, A. G. Structural principles for the propeller assembly of  $\beta$ -sheets: the preference for sevenfold symmetry. *Proteins* **14**, 191–201 (1992).
40. Faber, H. R. *et al.* 1.8 Å crystal structure of the C-terminal domain of rabbit serum haemopexin. *Structure* **3**, 551–559 (1995).
41. Chothia, C., Levitt, M. & Richardson, D. Structure of proteins: packing of  $\alpha$ -helices and pleated sheets. *Proc. Natl. Acad. Sci. USA* **74**, 4130–4134 (1977).
42. Harding, M. M. Small revisions to predicted distances around metal sites in proteins. *Acta Crystallogr. sect. D* **62**, 678–682 (2006).
43. Hunt, L. T., Barker, W. C. & Chen, H. R. A domain structure common to hemopexin, vitronectin, interstitial collagenase, and a collagenase homolog. *Prot. Seq. Data Anal.* **1**, 21–26 (1987).
44. Paoli, M. *et al.* Crystal structure of hemopexin reveals a novel high-affinity heme site formed between two  $\beta$ -propeller domains. *Nat. Struct. Biol.* **6**, 926–931 (1999).
45. Rost, B. Twilight zone of protein sequence alignments. *Prot. Eng.* **12**, 85–94 (1999).
46. Lowe, E. C., Baslé, A., Czjzek, M., Firbank, S. J. & Bolam, D. N. A scissor blade-like closing mechanism implicated in transmembrane signaling in a *Bacteroides* hybrid two-component system. *Proc. Natl. Acad. Sci. USA* **109**, 7298–7303 (2012).
47. Zhang, Z., Liu, Q. & Hendrickson, W. A. Crystal structures of apparent saccharide sensors from histidine kinase receptors prevalent in a human gut symbiont. *FEBS J.* **281**, 4263–4279 (2014).
48. Russell, R. B., Sasieni, P. D. & Sternberg, M. J. Supersites within superfolds. Binding site similarity in the absence of homology. *J. Mol. Biol.* **282**, 903–918 (1998).
49. Sato, K. *et al.* Identification of a new membrane-associated protein that influences transport/maturation of gingipains and adhesins of *Porphyromonas gingivalis*. *J. Biol. Chem.* **280**, 8668–8677 (2005).
50. Saiki, K. & Konishi, K. Identification of a *Porphyromonas gingivalis* novel protein sov required for the secretion of gingipains. *Microbiol. Immunol.* **51**, 483–491 (2007).
51. Ishiguro, I., Saiki, K. & Konishi, K. PG27 is a novel membrane protein essential for a *Porphyromonas gingivalis* protease secretion system. *FEMS Microbiol. Lett.* **292**, 261–267 (2009).
52. Veillard, F. *et al.* Inhibition of gingipains by their profragments as the mechanism protecting *Porphyromonas gingivalis* against premature activation of secreted proteases. *Biochim. Biophys. Acta* **1830**, 4218–4228 (2013).
53. Zhou, X. Y., Gao, J. L., Hunter, N., Potempa, J. & Nguyen, K. A. Sequence-independent processing site of the C-terminal domain (CTD) influences maturation of the RgpB protease from *Porphyromonas gingivalis*. *Mol. Microbiol.* **89**, 903–917 (2013).
54. Nguyen, K. A., Travis, J. & Potempa, J. Does the importance of the C-terminal residues in the maturation of RgpB from *Porphyromonas gingivalis* reveal a novel mechanism for protein export in a subgroup of Gram-Negative bacteria? *J. Bacteriol.* **189**, 833–843 (2007).
55. Fletcher, H. M. *et al.* Virulence of a *Porphyromonas gingivalis* W83 mutant defective in the *prtH* gene. *Infect. Immun.* **63**, 1521–1528 (1995).
56. Smith, C. J. Genetic transformation of *Bacteroides* spp. using electroporation. *Methods Mol. Biol.* **47**, 161–169 (1995).
57. Chiu, J., March, P. E., Lee, R. & Tillett, D. Site-directed, ligase-independent mutagenesis (SLIM): a single-tube methodology approaching 100% efficiency in 4h. *Nucleic Acids Res.* **32**, e174 (2004).

58. Gardner, R. G., Russell, J. B., Wilson, D. B., Wang, G. R. & Shoemaker, N. B. Use of a modified *Bacteroides-Prevotella* shuttle vector to transfer a reconstructed  $\beta$ -1,4-*D*-endoglucanase gene into *Bacteroides uniformis* and *Prevotella ruminicola* B(1)4. *Appl. Environ. Microbiol.* **62**, 196–202 (1996).
59. Bélanger, M., Rodrigues, P. & Progulske-Fox, A. Genetic manipulation of *Porphyromonas gingivalis*. *Curr. Protoc. Microbiol.* Chapter 13, **Unit13C**, 12 (2007).
60. Parker, A. C. & Smith, C. J. Genetic and biochemical analysis of a novel Ambler class A  $\beta$ -lactamase responsible for cefoxitin resistance in *Bacteroides* species. *Antimicrob. Agents Chemother.* **37**, 1028–1036 (1993).
61. Potempa, J., Pike, R. & Travis, J. The multiple forms of trypsin-like activity present in various strains of *Porphyromonas gingivalis* are due to the presence of either Arg-gingipain or Lys-gingipain. *Infect. Immun.* **63**, 1176–1182 (1995).
62. Sztukowska, M. *et al.* The C-terminal domains of the gingipain K polyprotein are necessary for assembly of the active enzyme and expression of associated activities. *Mol. Microbiol.* **54**, 1393–1408 (2004).
63. Quirke, A. M. *et al.* Heightened immune response to autocitrullinated *Porphyromonas gingivalis* peptidylarginine deiminase: a potential mechanism for breaching immunologic tolerance in rheumatoid arthritis. *Ann. Rheum. Dis.* **73**, 263–269 (2014).
64. Reynolds, E. S. The use of lead citrate at high pH as an electron-opaque stain in electron microscopy. *J. Cell Biol.* **17**, 208–212 (1963).
65. Ksiazek, M. *et al.* Miropin, a novel bacterial serpin from the periodontopathogen *Tannerella forsythia*, inhibits a broad range of proteases by using different peptide bonds within the reactive center loop. *J. Biol. Chem.* **290**, 658–670 (2015).
66. Banbula, A. *et al.* Rapid and efficient inactivation of IL-6 gingipains, lysine- and arginine-specific proteinases from *Porphyromonas gingivalis*. *Biochem. Biophys. Res. Commun.* **261**, 598–602 (1999).
67. Banbula, A. *et al.* Emerging family of proline-specific peptidases of *Porphyromonas gingivalis*: purification and characterization of serine dipeptidyl peptidase, a structural and functional homologue of mammalian prolyl dipeptidyl peptidase IV. *Infect. Immun.* **68**, 1176–1182 (2000).
68. Juanhuix, J. *et al.* Developments in optics and performance at BL13-XALOC, the macromolecular crystallography beamline at the ALBA synchrotron. *J. Synchrotron Radiat.* **21**, 679–689 (2014).
69. Matthews, B. W. Solvent content of protein crystals. *J. Mol. Biol.* **33**, 491–497 (1968).
70. Kabsch, W. XDS. *Acta Crystallogr. sect. D* **66**, 125–132 (2010).
71. Kabsch, W. Integration, scaling, space-group assignment and post-refinement. *Acta Crystallogr. sect. D* **66**, 133–144 (2010).
72. Adams, P. D. *et al.* PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. sect. D* **66**, 213–221 (2010).
73. Winn, M. D. *et al.* Overview of the CCP4 suite and current developments. *Acta Crystallogr. sect. D* **67**, 235–242 (2011).
74. Terwilliger, T. C. *et al.* Decision-making in structure solution using Bayesian estimates of map quality: the PHENIX AutoSol wizard. *Acta Crystallogr. sect. D* **65**, 582–601 (2009).
75. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. sect. D* **66**, 486–501 (2010).
76. Wu, S. & Zhang, Y. LOMETS: a local meta-threading-server for protein structure prediction. *Nucleic Acids Res.* **35**, 3375–3382 (2007).
77. Afonine, P. V. *et al.* Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr. sect. D* **68**, 352–367 (2012).
78. Smart, O. S. *et al.* Exploiting structure similarity in refinement: automated NCS and target-structure restraints in BUSTER. *Acta Crystallogr. sect. D* **68**, 368–380 (2012).
79. Terwilliger, T. C. *et al.* Iterative model building, structure refinement and density modification with the PHENIX AutoBuild wizard. *Acta Crystallogr. sect. D* **64**, 61–69 (2008).
80. Holm, L. & Rosenström, P. Dali server: conservation mapping in 3D. *Nucleic Acids Res.* **38**, W545–W549 (2010).
81. Pettersen, E. F. *et al.* UCSF Chimera—A visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).
82. Krissinel, E. & Henrick, K. Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. *Acta Crystallogr. sect. D* **60**, 2256–2268 (2004).
83. Chen, V. B. *et al.* MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr. sect. D* **66**, 12–21 (2010).
84. Bycroft, M. *et al.* The structure of a PKD domain from polycystin-1: implications for polycystic kidney disease. *EMBO J.* **18**, 297–305 (1999).
85. García-Castellanos, R. *et al.* Three-dimensional structure of MecI: Molecular basis for transcriptional regulation of staphylococcal methicillin resistance. *J. Biol. Chem.* **278**, 39897–39905 (2003).
86. Weiss, M. S. Global indicators of X-ray quality. *J. Appl. Cryst.* **34**, 130–135 (2001).
87. Karplus, P. A. & Diederichs, K. Linking crystallographic model and data quality. *Science* **336**, 1030–1033 (2012).
88. Davis, I. W. *et al.* MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. *Nucleic Acids Res.* **35**, W375–W383 (2007).

## Acknowledgements

We thank Joan Pous and Xandra Kreplin from the IBMB/IRB Protein Crystallography Platform for their excellent technical assistance during the crystallization experiments. This study was supported in part by grants from European, Spanish, Polish, Australian, and Catalan agencies (grant references FP7-HEALTH-2012-306029-2 “TRIGGER”; BFU2015-64487R, UMO-2015/17/B/NZ1/00666, UMO-2015/19/N/NZ1/00322, 1306/MOB/IV/2015/0, MDM-2014-0435; BIO2013-49320-EXP; 2014SGR9; 2014/15/D/NZ6/02546, 2012/04/A/NZ1/00051, 2975/7.PR/13/2014/2, and NIDCR (DE 09761 and DE 022597)). The Structural Biology Unit (www.sbu.csic.es) of IBMB is a “María de Maeztu” Unit of Excellence from the Spanish Ministry of Economy and Competitiveness. We acknowledge the help provided by ESRF and ALBA synchrotron local contacts. Funding for travelling and synchrotron data collection was provided in part by ESRF. The Faculty of Biochemistry, Biophysics and Biotechnology is a partner of the Leading National Research Center (KNOW) supported by the Ministry of Science and Higher Education of Poland.

## Author Contributions

F.X.G.R., K.A.N. and J.P. conceived the work; A.G., A.M.L., A.T.P., D.M., K.A.N., M.Sz., X.Z., and Y.G. carried out the molecular biology manipulations and mutant analysis; A.W.K. and M.Si. were responsible for electron microscopy; M.B. and M.W. provided antibodies; M.M. was responsible for flow cytometry; I.B.T. and J.J.E. performed mass spectrometry experiments; B.P., M.K., and M.N. produced recombinant PorZ; F.X.G.R., I.d.D., T.Go., and T. Gu. performed the structural studies; F.X.G.R., K.A.N. and J.P. wrote the paper.

## Additional Information

**Supplementary information** accompanies this paper at <http://www.nature.com/srep>

**Competing financial interests:** The authors declare no competing financial interests.

**How to cite this article:** Lasica, A. M. *et al.* Structural and functional probing of PorZ, an essential bacterial-surface component of the type-IX secretion system of human oral-microbiomic *Porphyromonas gingivalis*. *Sci. Rep.* **6**, 37708; doi: 10.1038/srep37708 (2016).

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>

© The Author(s) 2016