Current Research in Microbial Sciences 2 (2021) 100024

Contents lists available at ScienceDirect

Current Research in Microbial Sciences

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

Phylogenetic analysis of the bacterial Pro-Pro-endopeptidase domain reveals a diverse family including secreted and membrane anchored proteins

Hans C van Leeuwen [a,*], Dick Roelofs [b], Jeroen Corver [c], Paul Hensbergen [d]

[a] Department of CBRN Protection, Netherlands Organization for Applied Scientific Research TNO, Lange Kleiweg 137, 2288 GJ Rijswijk, the Netherlands
[b] KeyGene, Agro Business Park 90, 6708 PW Wageningen, the Netherlands
[c] Department of Medical Microbiology, Leiden University Medical Center, Albinusdreef 2, PO Box 9600, 2300 RC Leiden, the Netherlands
[d] Center for Proteomics and Metabolomics, Leiden University Medical Center, Albinusdreef 2, PO Box 9600, 2300 RC Leiden, the Netherlands

ARTICLE INFO

Keywords:
Endopeptidases
Metalloproteases
Secreted protease
Adhesion
Bacterial virulence
Motility
Metalloprotease
Cell wall
Phylogeny

ABSTRACT

Pro-Pro-endopeptidases (PPEP, EC 3.4.24.89) are secreted, zinc metalloproteases that have the unusual capacity to cleave a peptide bond between two prolines, a bond that is generally less sensitive to proteolytic cleavage. Two well studied members of the family are PPEP-1 and PPEP-2, produced by *Clostridioides difficile,* a human pathogen, and *Paenibacillus alvei*, a bee secondary invader, respectively. Both proteases seem to be involved in mediating bacterial adhesion by cleaving cell surface anchor proteins on the bacterium itself.

By using basic alignment and phylogenetic profiling analysis, this work shows that the complete family of proteins that contain a PPEP domain includes proteins from more than 130 species spread over 9 genera. These analyses also suggest that the PPEP domain spread through horizontal gene transfer events between species within the Firmicutes' classes Bacilli and Clostridia. Bacterial species containing PPEP homologs are found in diverse habitats, varying from human pathogens and gut microbiota to free-living bacteria, which were isolated from various environments, including extreme conditions such as hot springs, desert soil and salt lakes.

The phylogenetic tree reveals the relationships between family members and suggests that smaller subgroups could share cleavage specificity, substrates and functional similarity. Except for PPEP-1 and PPEP-2, no cleavage specificity, specific physiological target, or function has been assigned for any of the other PPEP-family members. Some PPEP proteins have acquired additional domains that recognize and bind noncovalently to various elements of the bacterial peptidoglycan cell-wall, anchoring these PPEPs. Secreted or anchored to the cell-wall surface PPEP proteins seem to perform various functions.

Introduction

Bacteria contain an extensive collection of proteases that have vital roles in viability and virulence. With the secretion of proteases, Gram-positive bacteria can adjust their environment. This varies from non-specific proteases processing nutrient sources and enabling tissue invasion, to proteases involved in defense against attacks by host factors or competing bacteria (Cezairliyan and Ausubel, 2017; Koziel and Potempa, 2013; Lin et al., 2020; Potempa and Potempa, 2012; Seele et al., 2015). A surprising activity is cleavage of adhesion proteins on the bacterial surface by proteases secreted by the same bacteria. The cell-wall envelope of Gram-positives contains proteins which enable the bacteria to anchor to host-animal tissues, plant vascular system, organic material or abiotic surfaces, where they form sessile communities. In response to depletion in nutrients or other stress factors, bacteria have been postulated to secrete proteases which cleave these protein anchors, releasing the bacteria to search for more favorable environments (Hensbergen et al., 2015; McAleese et al., 2001; Pinkston et al., 2011; Tonry et al., 2012).

Proline-proline endopeptidases (PPEPs), named after their unique substrate cleavage site between two prolines, are the latest addition to the clan of M34 metalloendopeptidases of bacterial secreted proteases (Cafardi et al., 2013; Hensbergen et al., 2015, 2014; Klychnikov et al., 2018). Recent advances have resulted in elucidation of the function of two family members, PPEP-1 and PPEP-2, which, although present in

Abbreviations: CBD, Cell-wall-binding-domain; FN3, Fibronectin type III domain; MucBP, Mucin binding protein; PPEP, Pro-pro-endopeptidases; SCWP, Secondary cell-wall polysaccharide; SLH, surface layer homology; TED, thioester domain; VMSP, VWFA-MucBP-SLH-protein; VWFA, Von Willebrand factor type A.

* Corresponding author.
E-mail address: hans.vanleeuwen@tno.nl (H.C. van Leeuwen).

https://doi.org/10.1016/j.crmicr.2021.100024
Received 29 January 2021; Received in revised form 19 February 2021; Accepted 21 February 2021

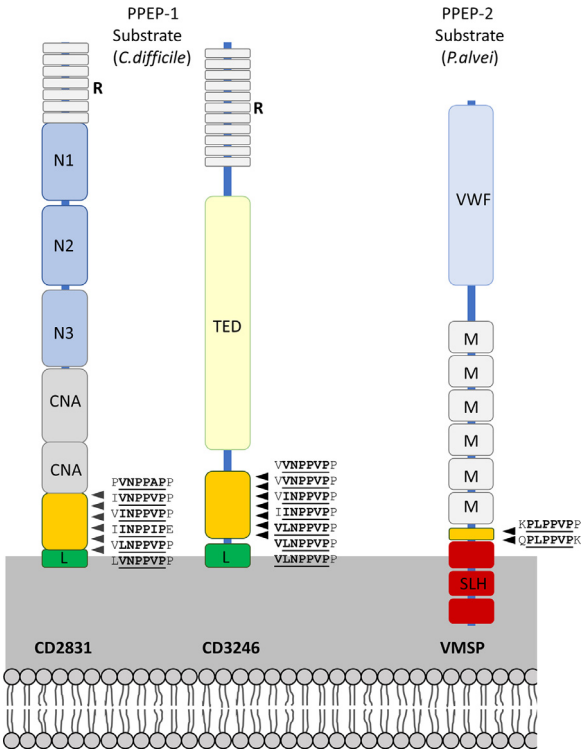

Fig. 1. ***Domain organization of Clostridioides difficile and Paenibacillus alvei adhesion proteins anchored to the peptidoglycan cell wall layer, and their PPEP-1 and PPEP-2 cleavage sites.*** Arrows indicate PPEP-1 and PPEP-2 cleavage sites with bold, underlined specific recognition sequence. Peptidoglycan cell wall layer (gray). *L*= LPXTG-like cell wall anchor, *R* = repetitive sequence, *N* = Amino terminal domain, Orange = region containing PPEP recognition sites. For other domain names see text.(For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

remotely related Firmicutes species, *Clostridioides difficile* and *Paenibacillus alvei*, display similar roles in cleaving cell-wall surface proteins involved in bacterial adhesion (Hensbergen et al., 2015; Klychnikov et al., 2018). Here, we will give a brief overview of current knowledge about the structure and functions of PPEP-1 and PPEP-2, and describe protein homologs with a PPEP-like domain, which have dispersed throughout several bacterial families.

Features and function of PPEP-1

Clostridioides difficile is a pathogenic bacterium that can colonize the colon of individuals whose normal intestinal flora has been disrupted by antibiotic treatment (Schäffler and Breitrück, 2018). Effective colonization depends on adherence of the bacterium to extracellular-matrix components of the epithelium. Two *C. difficile* adhesion proteins, SrtB-anchored collagen-binding adhesin protein (CD2831) and Cys-Gln-thioester bond-forming surface protein (CD3246), are cleaved and released by secreted PPEP-1. Hence, it has been postulated that PPEP-1, by cleaving these anchor proteins, plays a role in the switching from an adhesive to a motile phenotype in *C. difficile* (Hensbergen et al., 2014).

The two substrates of PPEP-1 contain multiple consecutive cleavage sites (see Fig.1), and a consensus sequence analysis showed that in addition to the prolines surrounding the cleavage site, conserved amino acids are present, 3 positions upstream and downstream of the cleavage site (Hensbergen et al., 2015, 2014). This analysis, and *in vitro* cleavage experiments, showed that the preferred PPEP-1 cleavage sites are composed of (ILV)NP↓P(AIV)P (P3 to P3').

Both CD2831 and CD3246 contain an amino (N)-terminal secretion-signal peptide and a C-terminal cell-wall anchor (L in Fig. 1). The cell-wall anchor is composed of an LPxTG-like motif (S/P)PKTG, which becomes covalently linked to the bacterial cell-wall by a membrane-anchored transpeptidase, sortase B (Paterson and Mitchell, 2004; Peltier et al., 2015; Van Leeuwen et al., 2014).

The closest structural homolog of CD2831, based on the I-TASSER (Iterative Threading ASSEmbly Refinement) structure and function assignment method (Yang et al., 2014a), is the adhesive Surface Protein B (RspB) of *Erysipelothrix rhusiopathiae* (Devi et al., 2012) (suppl. Fig.S1). The N-terminal domains, N1 and N2, of CD2831 (Fig.1) probably form a collagen "hug" domain which is a common theme for many Gram-positive bacteria surface proteins binding to extracellular matrix molecules (Devi et al., 2012; Zong et al., 2005). Several collagen types are bound by CD2831 and no clear preference seems present (Arato et al., 2019). The role of the N3-domain and both two CNA-domains (Fig.1) is unclear, but CNA-domains have been proposed to serve as a 'stalk' that projects the collagen binding domain away from the bacterial cell surface in order to position it for binding and facilitate bacterial adherence (Deivanayagam et al., 2000; Hendrickx et al., 2012; Rich et al., 1998). Both CD2831 and CD3246 contain N-terminal small repeats (R in Fig.1) which can vary in number in *C.difficile* strains with CD3246 encoding 11 to 15 repeats (sequence GQXPDGEKPSD) and CD2831 encoding 7 to 8 repeats (sequence KXTDNKKPEQTPEED). Their function, however, is currently unknown.

The middle domain of CD3246 shows homology to class 2 thioester domains (TEDs) and in particular to the putative adhesin of *Enterococcus faecium* TIE86 (suppl. Fig.S2) and *Bacillus anthracis* BaTIE (Miller et al., 2018). Bacterial surface proteins with an internal thioester have been shown to mediate covalent host binding using a cross-linking reaction, allowing bacterial attachment to human cells (Walden et al., 2015).

Structural analysis of PPEP-1, a complex with a VNPPVP peptide substrate, revealed a compact fold with an active site cleft (Pichlo et al., 2019; Schacherl et al., 2015). In this co-crystal, the substrate-binding cleft of PPEP-1 (Fig.2) is shaped complementary to the double-kinked conformation of the target, proline-rich, sequence. These structured substrates are recognized followed by substrate positioning within the active-site cleft (Pichlo et al., 2019). Upon binding, a flexible loop - the so-called S-loop - closes over the substrate-binding cleft. In this co-crystal structure, the target peptide (P3 to P3') is contacted and stabilized by more than 20 residues (see Fig.2b and 2c). These residues mostly form van der Waals interactions and several aromatic–aliphatic side-chain stacks. The bound conformation also increases the reactivity of the catalytic zinc ion enabling an efficient peptide hydrolysis.

Features and function of PPEP-2

Just as *C. difficile, P. alvei* is an anaerobic, Gram-positive, flagellated and endospore-forming bacterium. *P. alvei* can occur as secondary invader of honeybees, infected with *Melissococcus plutonius,* which cause European foulbrood (Forsgren, 2010). In addition, *P. alvei* is described as a rare causative agent of human infections (DeLeon and Welliver, 2016; Padhi et al., 2013; Reboli et al., 1989). *P. alvei* PPEP-2, despite only 50% amino-acid identity, forms a very similar fold to PPEP-1 (Klychnikov et al., 2018). The endogenous substrate of PPEP-2, VMSP, which contains two optimal cleavage sites PLP↓PVP (Fig.1), was identified in culture medium of *P. alvei strain* DSM29 (Klychnikov et al., 2018). The protein name, VMSP, is an acronym based on the three conserved domains found in the sequence, VWFA, MucBP and SLH (Klychnikov et al., 2018). The consensus cleavage site of PPEP-2 (PLP↓PVP) differs from the optimal PPEP-1 (VNP↓PVP) at positions P2 and P3 (underlined).

The N-terminus of VMSP contains a predicted Von Willebrand factor type A (VWFA) domain fold (Fig.1). VWFA-domains seem to be involved in a wide variety of functions and binding various ligands. In many eukaryotic cells, the VWFA-domain is implicated in extracellular matrix protein interactions (Whittaker and Hynes, 2002). In bacterial adhesins, the VWFA-domain has been shown to interact with dif-

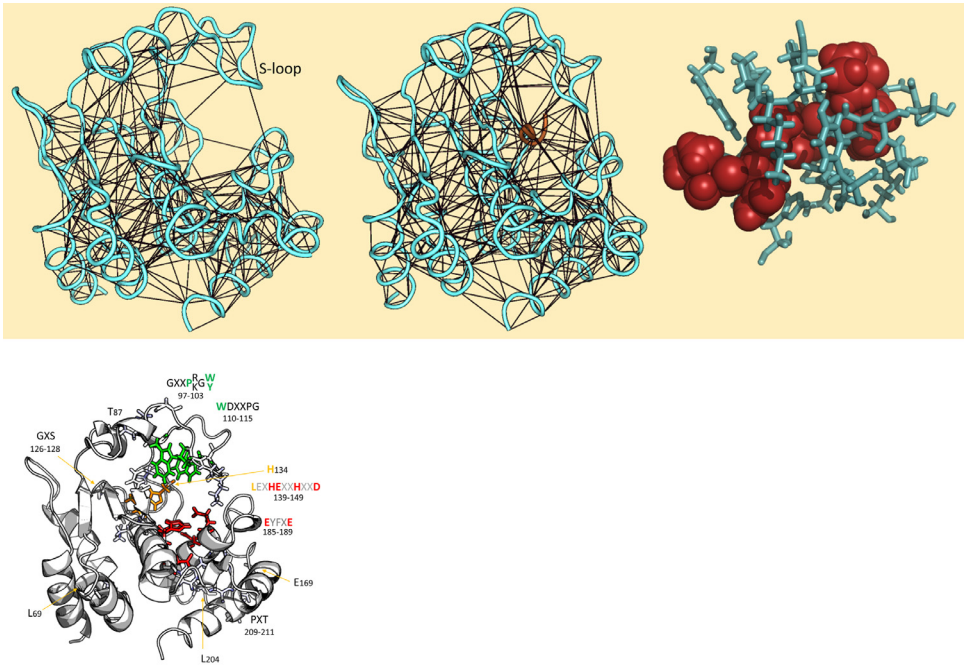

Fig. 2. ***3D structure of PPEP-1 with and without substrate peptide and conserved amino acids of PPEP-family.*** A) without peptide, PDB 5A0P; B) with peptide (orange), PDB 6R57; Black Lines indicate atom-pair contacts with cut-off value of 4 Å (Aydlnkal et al., 2019). C) Ball and stick representation of peptide binding cleft; peptide in red (spheres), contacting residues (sticks) in light blue. D) Conserved residues in PPEP-family alignment (95% threshold) shown as sidechain sticks and indicated in the PPEP-1 structure (PDB 6R57) without peptide (Pichlo et al.2019). Substrate proline P1 contacting residues in green, substrate proline P1' contacts shown in orange. Residues shown in red are involved in zinc coordination and catalytic activity.(For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

ferent elements of the extracellular matrix, namely fibronectin, collagen, and laminin (Vengadesan and Narayana, 2011). The closest homologs of the *P. alvei* VMSP VWFA-domain are the tips of pilins RgrA of *S. pneumoniae* and GBS104 of *Streptococcus agalactiae* (suppl. Fig.S3a). Both these adhesins form stalks to project the collagen-binding VWFA-domain to bind to respiratory- and epithelial cell surfaces (Izoré et al., 2010; Krishnan et al., 2013). Future research should focus on discovering the elements of the extracellular matrix, which are contacted by the *P. alvei* VMSP VWFA-domain.

Next to the VWFA-domain, Paenibacilli VMSPs contain 60 aa long repeating units of the Mucin binding protein domain (MucBp) (Fig.1 and suppl. Fig.S3b). The closest homologues to these paenibacillus MucBp-repeats are the *Lactobacillus reuteri* mucus-binding protein B1-domain and the β-GF module in *S.aureus* SraP protein (suppl. Fig.S3b). SraP is a surface-exposed protein which promotes *S. aureus* adhesion to host epithelial cells via a specific binding to carbohydrates (Etzold et al., 2014; MacKenzie et al., 2009; Yang et al., 2014b). Unlike the name suggests, the region corresponding to the B1-domains of L.*reuteri* MucBP is not responsible for the mucin binding activity, which requires an additional B2-domain (MacKenzie et al., 2009). This B2-domain is missing in Paenibacillus VMSP MucBP-repeats. As such, we envision that the MucBP-repeats form a stalk that presents the VWFA ligand-binding domain away from the bacterial cell surface, reminiscent of the CNA-domains (see Fig.1).

The LPxTG-like motif containing substrates of PPEP-1 are covalently attached to the cell-wall peptidoglycan, via a sortase mediated reaction. VMSP proteins, in contrast, have a surface-layer homology (SLH) anchor-domain repeat at their C-termini (Fig.1 and suppl. Fig.S3c). Bacterial secreted proteins containing an SLH-domain are tethered to the cell-wall of Gram-positive bacteria through non-covalent interactions with secondary cell-wall polymers (Blackler et al., 2018; Sychantha et al., 2018).

The proposed model suggests that paenibacillus VMPS are flexible filament-like structures composed of an anchor, stalk and matrix-binding domain (Fig.1). Their target could be a cell-surface matrix protein present in the animal gut, promoting adhesion for commensal- or opportunistic Paenibacilli. Alternatively, the VMSP-binding targets may be part of the plant root surface or rhizosphere soil matrix directly associated with the root (Berg et al., 2016).

Phylogenetic analyses of proteins with a PPEP domain

In order to identify PPEP homologs, we screened the NCBI database for homologs of PPEP-1 and PPEP-2 using a probability of 1e-25 expected number of chance as a threshold value. Homologous PPEP domain containing proteins are found in two Firmicutes orders, Clostridiales and Bacillales. The first contains 2 genera, Clostridioides and Clostridium, and the Bacillales contains 7 genera including Paenibacillus (Family Paenibacillaceae), Bacillus, Parageobacillus, Geobacillus, Anoxybacillus (Family Bacillaceae), Salinicoccus and Jeotgalicoccus (Family Staphylococcaceae). In most fully sequenced genomes containing a PPEP homolog, a single gene copy is present, yet in some genomes two or even three PPEP homologs are found.

A phylogenetic tree of species with a PPEP homolog, based on housekeeping proteins commonly used for bacterial taxonomy, FtsZ and RpoD (Beiki et al., 2016; Pal et al., 2019; Parkinson et al., 2011), clearly showed the evolutionary divergence between these genera (suppl. Fig.S4). The species display extremely different lifestyles, varying from enteric pathogens, commensal microbiota, plant roots associated, to free soil-living species. These include halophilic, alkalophilic and thermophilic species. It has been suggested that animal gut- and plant root microbial communities share several characteristics, as both form an active group of microbes that contribute to degradation or modification of nutrients provided by the host surface they depend on (Mendes and Raaijmakers, 2015; Ramírez-puebla et al., 2013). Apparently, PPEP homologs are not restricted to these active group-forming microbes.

A phylogenetic tree of the PPEP homologs revealed multiple aspects of this protein family (Fig.3, table S1). In general, the PPEP domain phylogeny reflects the species distribution. On the other hand, PPEP-1 from *C. difficile* clusters with the PPEP domains from the Paenibacillus group, and not with the PPEP domain from the more closely related Clostridium species. The tree supports three sets of Bacillus branches (Fig.3) labelled Bacillus group A, B and C, with group A the most commonly found. Several bacillus genomes (*e.g. B. ciccensis, B. solani, B.* sp. *FJAT21945*) contain all three PPEP-groups, some group A and B *(e.g. B. methanolicus, B. oceanisediminis*) or only group C (*B.* sp. *COPE52*). Based on this, and the fact that PPEP-like domains in general were not found in many different bacterial orders, we speculate that horizontal transfer of PPEP-like genes, followed by adaptive evolution driven by natural selec-

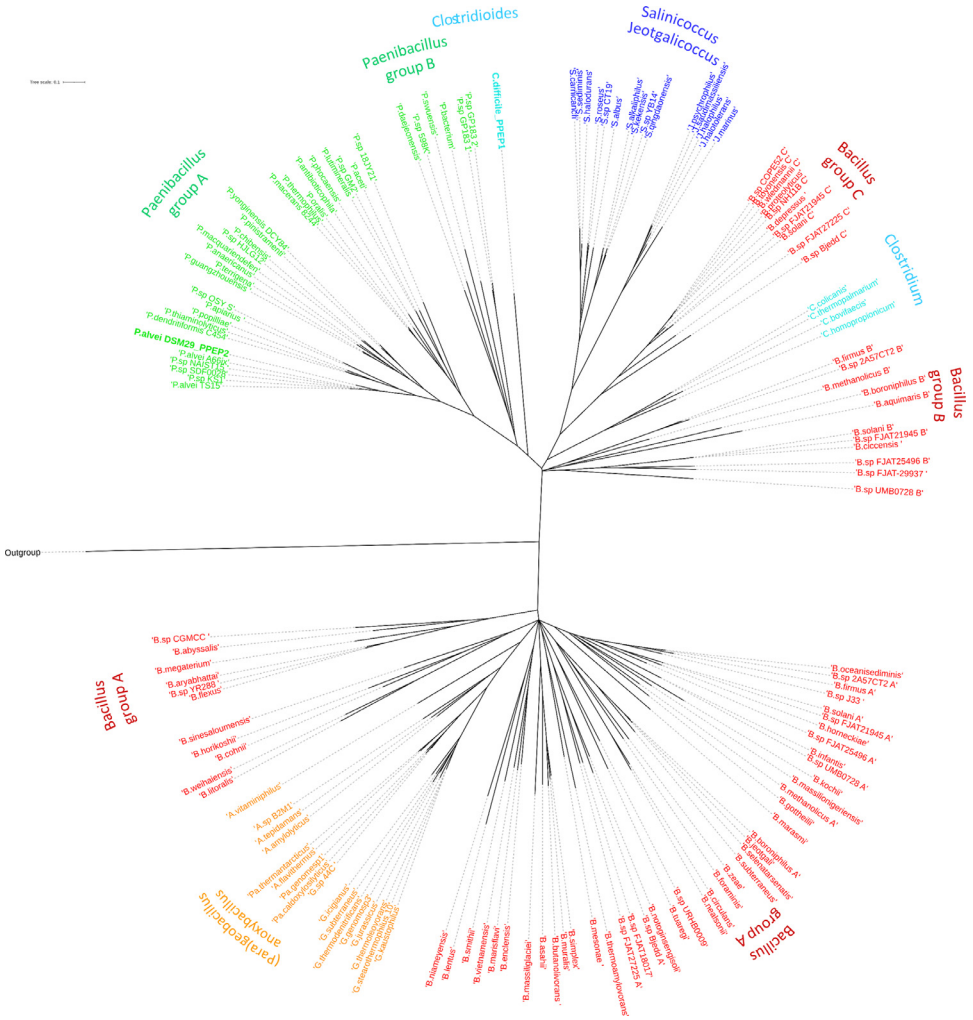

Fig. 3. ***Phylogenetic relationship of PPEP-like proteins.*** A rooted phylogenetic tree was constructed using an alignment of 146 PPEP homologs from 131 species. Tree was built by using the Neighbor-Joining method with the Jukes-Cantor genetic distance model. Major species groups are highlighted by different colors as indicated. Outgroup is M34 peptidase of Vibrio cholera O1 (WP_108257476). Bootstrap values of branches corresponding to indicated groups are all >50 except for the Clostridioides group. Furthest distance in the tree is between B. lentus and J. saudimassiliensis (32%, identities; 52%, similarity; BLAST e-value 7e-28). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

tion on a potentially newly acquired function, may explain the observed species distribution. The driver would be targeting a surface protein, and controlling its function via this cleavage.

Conserved amino acids in the alignment define this PPEP protein domain family. Among the most conserved residues are the zinc binding site, including the classical metalloprotease HEXXH motif (red in Fig.2d) and residues involved in coordinating the substrate proline - proline cleavage site in PPEP-1 (green and orange in Fig.2d).

Additional domains in several PPEP homologs

PPEP-1 and 2 are secreted proteases and this feature seems to be conserved in many homologs. However, in several homologs extra protein domains were observed. For example, all PPEP-like protein members of bacillus group B (Fig.3) contain a predicted SH3b-domain amino-terminal of the PPEP-domain (Fig.4), whilst group C (Fig.3) is missing this SH3-domain. The PPEP-SH3b-domain has homology to a peptidoglycan cell-wall-binding-domain (CBD) of bacteriophage protein lysPBC5 (suppl. Fig.S5), which is involved in bacterial host-cell surface binding (Lee et al., 2019). This CBD is critical for specific recognition of the *Bacillus cereus'* cell surface by the bacteriophage. SH3b of bacterial cell-wall-binding proteins recognize and bind noncovalently to peptidoglycan of the bacterial cell walls, peptidoglycan cross-bridges in particular (Gonzalez-Delgado et al., 2020; Mitkowski et al., 2019).

Salinicoccus and Jeotgalicoccus are halotolerant or halophilic bacteria that live in saline environments (Schwaiger et al., 2010; Ventosa et al., 1990). Salinicoccus as well as Jeotgalicoccus PPEP-like proteins (Fig.3) have a C-terminal FN3-domain (Fig.4 and suppl.Fig.S6). Bacterial FN3-domain folds are common, suggesting a general ligand-binding module. Several characterized bacterial FN3-domains possess surface-exposed aromatic residues that are thought to make contacts with substrate sugar chains (Alahuhta et al., 2010; Jee et al., 2002; Kataeva et al., 2002). FN3-like domains, capable of binding to carbohydrates, are found mostly in extracellular proteins including bacterial receptors, adhesion proteins or glycosyl-hydrolases. To our knowledge, this is the first FN3-domain described linked to a protease-domain.

Two members of the Paenibacillus group B, *P. swuensis* and *P. spec. GP183* (Fig.3) contain an additional SLH-domain C-terminal of the PPEP-domain (Fig.4, Fig.S7). SLH-domains, consisting of three repeats, have been shown to interact with secondary cell-wall polysaccharide (SCWP) for anchoring proteins to the surfaces of bacteria (Kern et al., 2011; Mesnage et al., 2000). The function of the SLH-domain seems to be accessory and not essential, given the absence of the domain in two Paenibacillus homologs, *P. daejeonensis* and *P.* sp. *598K* (Fig.5a).

Synteny of PPEP proteins

Besides conservation levels of PPEP homologs, similarity in the gene order or distribution between genomes of different species can provide additional support for shared ancestry and function (Junier and Rivoire, 2016). A summary of gene clusters identified around the PPEP homologs is shown in Fig.5.

Within Paenibacillus, the phylogenetic and synteny analysis (Fig.3 and Fig.5a) show that this group can be divided in two subgroups. Group

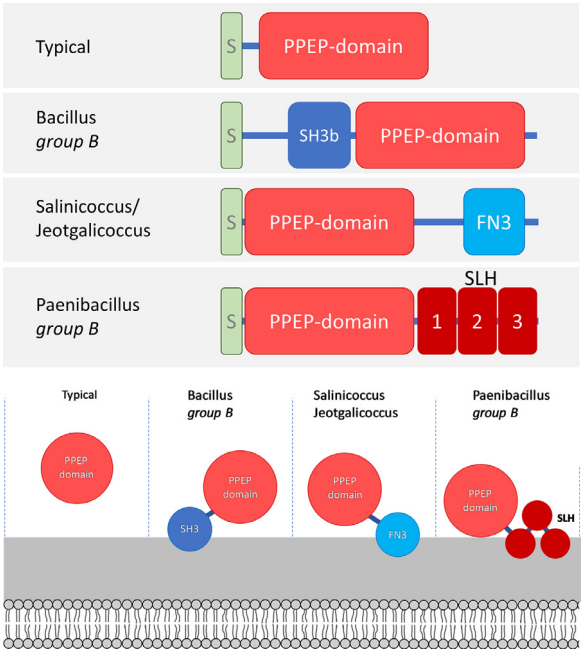

Fig. 4. ***PPEP family proteins domain organization.*** Several members of the PPEP family contain an additional domain, SH3 (Bacillus group B), FN3 (Salinicoccus/Jeotgalicoccus) or SLH (Paenibacillus group B), which are predicted to be involved in anchoring to the Gram-positive bacterium cell wall (gray). *S* = secretion signal. For other domain names see text.

A, containing the VMSP substrate target directly downstream of PPEP (including PPEP-2), and those lacking VMSP, Group B.

Alignment of genome blocks of the Bacillus PPEP group A, which also include Anoxybacillus, Geobacillus and Parageobacillus (Goh et al., 2013; Najar and Thakur, 2020; Zeigler, 2014), shows a largely similar gene order of alignment (Fig.5b). This confirms the evolutionary phylogenetic clustering (Fig.3) and argues for identical function of PPEP-like proteins within this group.

Almost 250 different Clostridium species have been described and at least 40 species' genomes have been completely sequenced. PPEP encoding genes are only found in 4 seemingly unrelated Clostridium species (Fig.5e): *Clostridium colicanis* isolated from canine feces (Greetham et al., 2003), and also identified as a human pathogen in a patient with bacteraemia; *Clostridium thermopalmarium* a moderately thermophilic bacterium isolated from palm wine (Lawson Anani Soh et al., 1991); *Clostridium bovifaecis* isolated from cow manure (Zhu et al., 2018) and *Clostridium homopropionicum* isolated from anaerobic sewage sludge (Beck et al., 2015). The gene organization surrounding the PPEP genes is largely similar confirming the relation of these clostridium species .

Bacteria can acquire new genes by uptake and incorporating of environmental DNA into their genomes (Mell and Redfield, 2014). To investigate whether other gene segments were co-transferred with the PPEP homologs, we also checked the genomic location in closely related species that did not contain a PPEP homolog. Three such genome comparisons are shown in Fig.6. Remarkably, in these examples, no other genes were co-transferred with the PPEP homolog (or these were removed after incorporation). In addition, no transposable elements such as transposons and insertion sequences (Siguier et al., 2014), which catalyze exchanges between genomes, or changes in GC-profiles (Zhang et al., 2014) seem present.

Potential substrates and regulation of orphan PPEPs

Despite belonging to two different Firmicute classes, PPEP-1 and PPEP-2, both involved in cleavage of anchor proteins, share several characteristics: 1. The amino acids in the cleavage sites are P↓P-X-P. 2. Their substrate proteins CD2831 (PPEP-1) and VMSP (PPEP-2) are encoded by the gene directly adjacent to the *ppep* gene, albeit in alternative orientation (see Fig.5a). Genes involved in a common pathway or function are frequently found near each other for miscellaneous reasons such as coregulation, efficiency, gene transfer or survival (Ballouz et al., 2010). PPEP-1 in *C. difficile* strain 630 has an additional target protein, CD3246, which is located ~ 50.000 bp downstream, showing that genetic linking of the protease and this substrate gene is not essential. CD3246, unlike CD2831, is not conserved within *C. difficile* and missing in certain pathogenic strains. 3. Both PPEP-1 and PPEP-2, and their substrates, have a c-di-GMP riboswitch in their 5′-UTR (Fig.5a triangles). By binding to these structures, the second messenger c-di-GMP, which is involved in signal transduction in a wide variety of bacteria, can regulate gene expression upon environmental stimuli (Purcell and Tamayo, 2016).

Other PPEP-like family members might share these characteristics common to PPEP-1 and PPEP-2. Inspection of the genes 6000 bp upstream and downstream of the PPEP homologs (Fig.5a-e) revealed no secreted proteins with a P↓P-X-P motif directly surrounding the PPEP homolog. However, screening species with a PPEP-like gene for secreted and cell surface proteins with a **PPXP** motif, conserved within the species groups or branches, showed that the broadly conserved MltG protein, a potential terminase for glycan polymerization (Yunck et al., 2016), contains a conserved GL**PPGP** sequence within the Paenibacillus, Salinicoccus and Clostridium group. In addition, a putative polysaccharide deacetylase protein (WP_053475711) with a conserved DL**PPSP** motif is present in the Bacillus B and C group. Moreover, in Bacillus group A, the protein upstream of the PPEP homolog (Fig.5b), annotated as YpjP-like (PF14005), has a secretion signal and contains two conserved di-proline motifs (sites EHPPQD and NTPPKW). This YpjP-like protein, which is functionally uncharacterized, shows homology to *C. difficile* PilJ pilin protein (suppl. Fig.S8) reported to be involved in adhesion (Piepenbrink et al., 2014). In a subgroup of Bacillus group-B, upstream of the *ppep* homolog, a putative secreted collagenase (Caviness et al., 2018; Wilson et al., 2003) is encoded (see Fig.5c and Fig.S9) which also contains a di-proline motif (VLPPDED). Further research is needed to test whether any of the proteins described above are genuine substrates of the corresponding PPEP homolog in these species.

No c-di-GMP riboswitches were found in PPEP-like genes, or their surrounding genes, in other species than *C. difficile* and the Paenibacillus group A. Examining 9 complete genomes in the top half of the phylogenetic PPEP-tree, revealed a total of 41 genes with c-di-GMP riboswitches (data not shown). Only one of those proteins, VWFA-IPT/TIG adhesion protein (Kobayashi et al., 2016) encoded by *P. guangzhouensis*, contains a **PPXP** motif (NA**PPAP**).

Concluding remarks and future prospective

It is unclear why PPEP-1 and 2 are specialized in hard to cleave proline–proline bonds. Possibly, the high number of proline residues leaves their substrates better protected from cleavage by proteases from other, competing, bacteria. Although targets of PPEP-1 and PPEP-2 are involved in adhesion to cell surfaces, other family members, as described in this study, might have evolved other specificities and preferences towards alternative bacterial (surface) proteins or even host proteins. Hence, the incorporation of a *ppep*-like gene in other species may have provided a small benefit unrelated to adhesion. It will therefore be interesting to elucidate the endogenous substrates of these PPEP-homologs, especially the ones that are more remotely related to PPEP-1 and 2. This will also reveal whether they are genuine Pro-Pro-endopeptidases. Techniques such as TAILS, PICS and COFRADIC (Bhagwat et al., 2018) are valuable approaches to study this.

The structures of both PPEP-1 and PPEP-2 have been solved which provide insight in their substrate specificity. More structural data, including substrate bound structures of other family members, will further open avenues for elucidating the binding properties and specificity

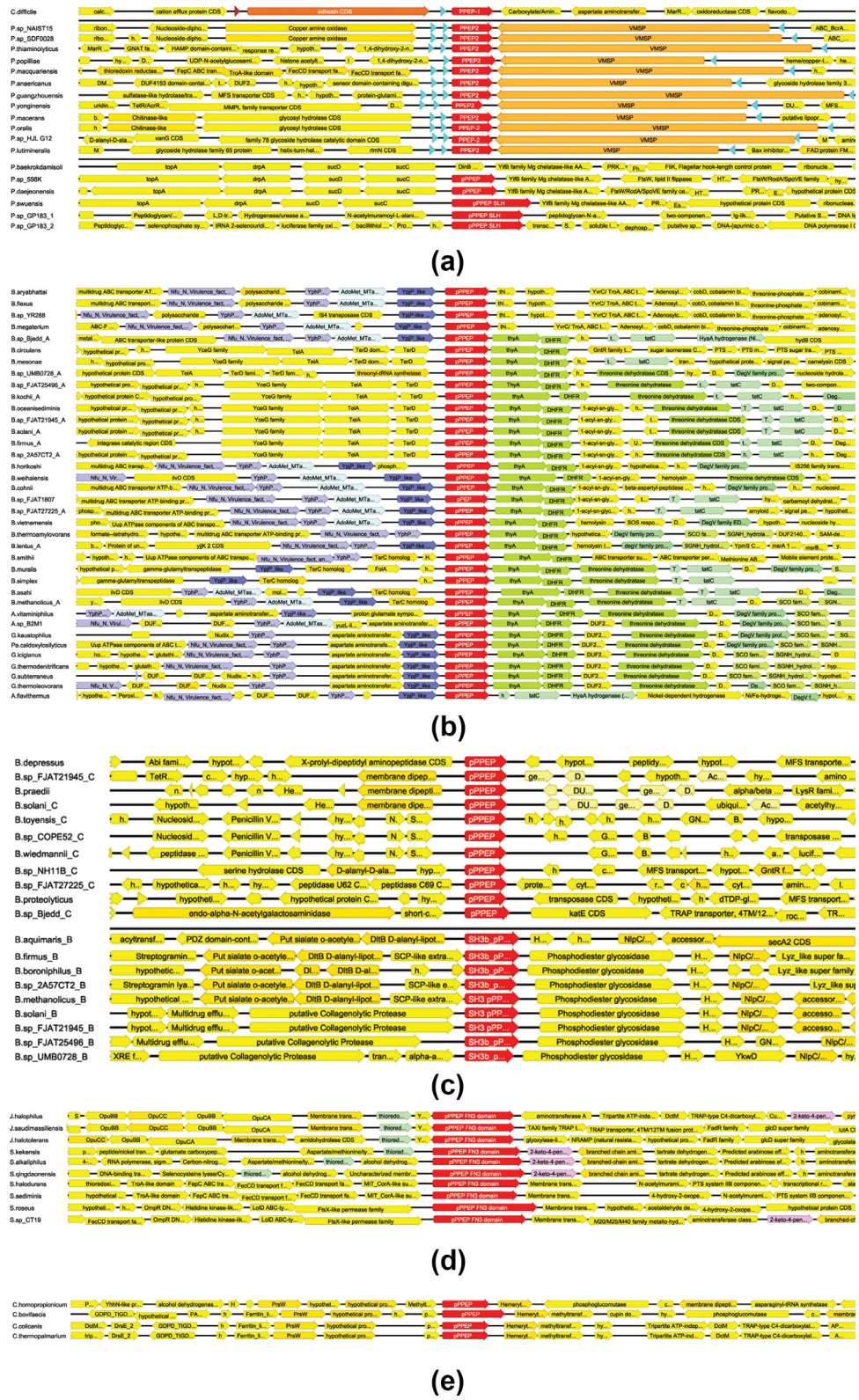

Fig. 5. *Genomic conservation of regions flanking the PPEP homologs (putative PPEP, pPPEP, in red).* A) Genomic conservation of regions flanking PPEP-1, PPEP-2, and homologs in C. difficile and Paenibacillus. Paenibacillus group A containing VMSP (orange) PPEP-2 substrate and group B lacking VMSP. Light blue triangle indicates c-di-GMP type I riboswitch. Top shows C.difficile with its substrate protein (CD2831) in dark orange. Small red triangle indicates c-di-GMP type II riboswitch. SLH, surface-layer homology, domain bind to secondary cell-wall polymers. Paenibacillus (P.); Clostridioides (C.). SLH-domain is predicted to bind surface layer. *B) Bacillus group A. Bacillus (B.), Parageobacillus (Pa.), Geobacillus (G.) and Anoxybacillus (A.). C) Bacillus group B (bottom) and C (top). SH3b-domain is predicted to be a peptidoglycan cell-wall-binding-domain. D) Salinicoccus (S.) and Jeotgalicoccus (J.) group. FN3 domains is predicted to bind to cell wall carbohydrates. E) Clostridium species.*(For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

through identification of the substrate interface forming residues. Engineering these into recognizing specific novel targets seems a possibility leading to the development of new tools for biotechnological, and therapeutic applications. In the example of *C. difficile*, these bacteria cleave their anchors once they are stressed due to nutrient shortage, over-crowding or immune pressure, allowing relocation to places that are more favorable. Blocking release of the bacteria would prevent escape of this environment and result in less survival and less successful dissemination. A knockout of the *C. difficile ppep-1* gene showed attenuated virulence in a hamster infection model (Hensbergen et al., 2015).

On a more biotechnological level, knowledge about PPEPs can be used to develop enzymes that are able to cleave off engineered pro-

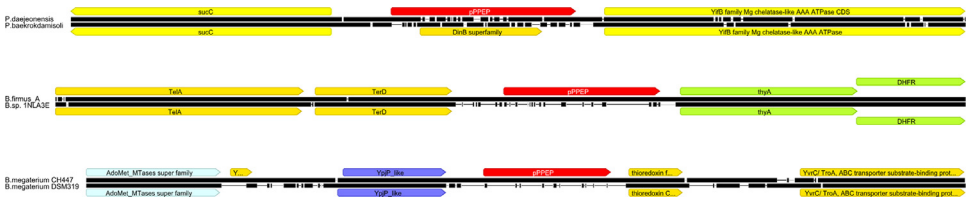

Fig. 6. *Comparison of genomic regions between closely related species with and without a PPEP homolog.* Genome pairs with (top) and without PPEP gene (bottom). 1 Paenibacillus daejeonensis DSM 15491 (438541 – 442364 and Paenibacillus baekrokdamisoli CECT 8890_05 (172529–176227); 2 Bacillus firmus NBRC 15306_18 (2246-4072) and Bacillus sp. 1NLA3E (3111676- 3114991); 3 Bacillus megaterium CH447_03 (510523–515539) and Bacillus megaterium ATCC 14,581_04 (396657–400370). Black lines indicate DNA alignment with the coding sequences (arrow) above and below the respective genes.

tein tags that confer specific binding to standardized affinity resins (Young et al., 2012). Especially in cases where other enzymes that are currently used (for instance factor X) may not work or are impossible to use.

The importance of the PPEP-domain protease and their respective substrates needs to be established, preferably by gene-knockout experiments and phenotypic evaluation of these bacterial knock-outs. This is the first step towards a more comprehensive understanding of their role in the bacterial life cycle.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

CRediT authorship contribution statement

Hans C van Leeuwen: Conceptualization, Methodology, Writing – original draft, Software. **Dick Roelofs:** Validation, Software, Writing – review & editing. **Jeroen Corver:** Validation, Writing – original draft. **Paul Hensbergen:** Conceptualization, Supervision, Writing – original draft.

Acknowledgments

We would like to thank Dustin Dopla for critical reading of the manuscript.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.crmicr.2021.100024.

References

Alahuhta, M., Xu, Q., Brunecky, R., Adney, W.S., Ding, S.Y., Himmel, M.E., Lunin, V.V., 2010. Structure of a fibronectin type III-like module from Clostridium thermocellum. Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun. doi:10.1107/S1744309110022529.

Arato, V., Gasperini, G., Giusti, F., Ferlenghi, I., Scarselli, M., Leuzzi, R., 2019. Dual role of the colonization factor CD2831 in Clostridium difficile pathogenesis. Sci. Rep. doi:10.1038/s41598-019-42000-8.

Ballouz, S., Francis, A.R., Lan, R., Tanaka, M.M., 2010. Conditions for the evolution of gene clusters in bacterial genomes. PLoS Comput. Biol. doi:10.1371/journal.pcbi.1000672.

Beck, M.H., Poehlein, A., Bengelsdorf, F.R., Schiel-Bengelsdorf, B., Daniel, R., Dürre, P., 2015. Draft genome sequence of the strict anaerobe Clostridium homopropionicum LuHBu1 (DSM 5847). Genome Announc. doi:10.1128/genomeA.01112-15.

Beiki, F., Busquets, A., Gomila, M., Rahimian, H., Lalucat, J., García-Valdés, E., 2016. New pseudomonas spp. are pathogenic to citrus. PLoS One doi:10.1371/journal.pone.0148796.

Berg, G., Rybakova, D., Grube, M., Köberl, M., 2016. The plant microbiome explored: implications for experimental botany. J. Exp. Bot. doi:10.1093/jxb/erv466.

Bhagwat, S.R., Hajela, K., Kumar, A., 2018. Proteolysis to identify protease substrates: cleave to decipher. Proteomics doi:10.1002/pmic.201800011.

Blackler, R.J., López-Guzmán, A., Hager, F.F., Janesch, B., Martinz, G., Gagnon, S.M.L., Haji-Ghassemi, O., Kosma, P., Messner, P., Schäffer, C., Evans, S.V., 2018. Structural basis of cell wall anchoring by SLH domains in Paenibacillus alvei. Nat. Commun. doi:10.1038/s41467-018-05471-3.

Cafardi, V., Biagini, M., Martinelli, M., Leuzzi, R., Rubino, J.T., Cantini, F., Norais, N., Scarselli, M., Serruto, D., Unnikrishnan, M., 2013. Identification of a novel zinc metalloprotease through a global analysis of Clostridium difficile extracellular proteins. PLoS One doi:10.1371/journal.pone.0081306.

Caviness, P., Bauer, R., Tanaka, K., Janowska, K., Roeser, J.R., Harter, D., Sanders, J., Ruth, C., Matsushita, O., Sakon, J., 2018. Ca2+-induced orientation of tandem collagen binding domains from clostridial collagenase ColG permits two opposing functions of collagen fibril formation and retardation. FEBS J. doi:10.1111/febs.14611.

Cezairliyan, B., Ausubel, F.M., 2017. Investment in secreted enzymes during nutrient-limited growth is utility dependent. Proc. Natl. Acad. Sci. USA doi:10.1073/pnas.1708580114.

Deivanayagam, C.C.S., Rich, R.L., Carson, M., Owens, R.T., Danthuluri, S., Bice, T., Höök, M., Narayana, S.V.L., 2000. Novel fold and assembly of the repetitive B region of the Staphylococcus aureus collagen-binding surface protein. Structure doi:10.1016/S0969-2126(00)00081-2.

DeLeon, S.D., Welliver, R.C., 2016. Paenibacillus alvei sepsis in a neonate. Pediatr. Infect. Dis. J. doi:10.1097/INF.0000000000001003.

Devi, A.S., Ogawa, Y., Shimoji, Y., Balakumar, S., Ponnuraj, K., 2012. Collagen adhesin-nanoparticle interaction impairs adhesin's ligand binding mechanism. Biochim. Biophys. Acta - Gen. Subj. doi:10.1016/j.bbagen.2012.04.006.

Etzold, S., Kober, O.I., Mackenzie, D.A., Tailford, L.E., Gunning, A.P., Walshaw, J., Hemmings, A.M., Juge, N., 2014. Structural basis for adaptation of lactobacilli to gastrointestinal mucus. Environ. Microbiol. doi:10.1111/1462-2920.12377.

Forsgren, E., 2010. European foulbrood in honey bees. J. Invertebr. Pathol. doi:10.1016/j.jip.2009.06.016.

Goh, K.M., Kahar, U.M., Chai, Y.Y., Chong, C.S., Chai, K.P., Ranjani, V., Illias, R.M., Chan, K.G., 2013. Recent discoveries and applications of Anoxybacillus. Appl. Microbiol. Biotechnol. doi:10.1007/s00253-012-4663-2.

Gonzalez-Delgado, L.S., Walters-Morgan, H., Salamaga, B., Robertson, A.J., Hounslow, A.M., Jagielska, E., Sabała, I., Williamson, M.P., Lovering, A.L., Mesnage, S., 2020. Two-site recognition of Staphylococcus aureus peptidoglycan by lysostaphin SH3b. Nat. Chem. Biol. doi:10.1038/s41589-019-0393-4.

Greetham, H.L., Gibson, G.R., Giffard, C., Hippe, H., Merkhoffer, B., Steiner, U., Falsen, E., Collins, M.D., 2003. Clostridium colicanis sp. nov., from canine faeces. Int. J. Syst. Evol. Microbiol. doi:10.1099/ijs.0.02260-0.

Hendrickx, A.P.A., Poor, C.B., Jureller, J.E., Budzik, J.M., He, C., Schneewind, O., 2012. Isopeptide bonds of the major pilin protein BcpA influence pilus structure and bundle formation on the surface of Bacillus cereus. Mol. Microbiol. doi:10.1111/j.1365-2958.2012.08098.x.

Hensbergen, P.J., Klychnikov, O.I., Bakker, D., Van Winden, V.J.C., Ras, N., Kemp, A.C., Cordfunke, R.A., Dragan, I., Deelder, A.M., Kuijper, E.J., Corver, J., Drijfhout, J.W., Van Leeuwen, H.C., 2014. A novel secreted metalloprotease (CD2830) from clostridium difficile cleaves specific proline sequences in LPXTG cell Surface Proteins. Mol. Cell. Proteomics 13. doi:10.1074/mcp.M113.034728.

Hensbergen, P.J., Klychnikov, O.I., Dennis Bakker, D., Dragan, I., Kelly, M.L., Minton, N.P., Corver, J., Kuijper, E.J., Drijfhout, J.W., van Leeuwen, H.C., 2015. Clostridium difficile secreted Pro-Pro endopeptidase PPEP-1 (ZMP1/CD2830) modulates adhesion through cleavage of the collagen binding protein CD2831. FEBS Lett. doi:10.1016/j.febslet.2015.10.027.

Izoré, T., Contreras-Martel, C., El Mortaji, L., Manzano, C., Terrasse, R., Vernet, T., Di Guilmi, A.M., Dessen, A., 2010. Structural basis of host cell recognition by the pilus adhesin from streptococcus pneumoniae. Structure doi:10.1016/j.str.2009.10.019.

Jee, J.G., Ikegami, T., Hashimoto, M., Kawabata, T., Ikeguchi, M., Watanabe, T., Shirakawa, M., 2002. Solution structure of the fibronectin type III domain from Bacillus circulans WL-12 chitinase A1. J. Biol. Chem. doi:10.1074/jbc.M109726200.

Junier, I., Rivoire, O., 2016. Conserved units of co-expression in bacterial genomes: an evolutionary insight into transcriptional regulation. PLoS One doi:10.1371/journal.pone.0155740.

Kataeva, I.A., Seidel, R.D., Shah, A., West, L.T., Li, X.L., Ljungdahl, L.G., 2002. The fibronectin type 3-like repeat from the Clostridium thermocellum cellobiohydrolase CbHa promotes hydrolysis of cellulose by modifying its surface. Appl. Environ. Microbiol. doi:10.1128/AEM.68.9.4292-4300.2002.

Kern, J., Wilton, R., Zhang, R., Binkowski, T.A., Joachimiak, A., Schneewind, O., 2011. Structure of surface layer homology (SLH) domains from Bacillus anthracis surface array protein. J. Biol. Chem. doi:10.1074/jbc.M111.248070.

Klychnikov, O.I.O.I., Shamorkina, T.M.T.M., Weeks, S.D.S.D., van Leeuwen, H.C.H.C., Corver, J., Drijfhout, J.W.J.W., van Veelen, P.A.P.A., Sluchanko, N.N.N.N., Strelkov, S.V.S.V., Hensbergen, P.J.P.J., 2018. Discovery of a new Pro-Pro endopeptidase, PPEP-2, provides mechanistic insights into the differences in substrate specificity within the PPEP family. J. Biol. Chem. 293. doi:10.1074/jbc.RA118.003244.

Kobayashi, K., Kanesaki, Y., Yoshikawa, H., 2016. Genetic analysis of collective motility of paenibacillus sp. NAIST15-1. PLoS Genet. doi:10.1371/journal.pgen.1006387.
Koziel, J., Potempa, J., 2013. Protease-armed bacteria in the skin. Cell Tissue Res. doi:10.1007/s00441-012-1355-2.
Krishnan, V., Dwivedi, P., Kim, B.J., Samal, A., MacOn, K., Ma, X., Mishra, A., Doran, K.S., Ton-That, H., Narayana, S.V.L., 2013. Structure of Streptococcus agalactiae tip pilin GBS104: a model for GBS pili assembly and host interactions. Acta Crystallogr. Sect. D Biol. Crystallogr. doi:10.1107/S0907444913004642.
Lawson Anani Soh, A., Ralambotiana, H., Ollivier, B., Prensier, G., Tine, E., Garcia, J.L., 1991. Clostridium thermopalmarium sp. nov., a moderately thermophilic butyrate-producing bacterium isolated from palm wine in senegal. Syst. Appl. Microbiol. doi:10.1016/S0723-2020(11)80291-2.
Lee, K.O., Kong, M., Kim, I., Bai, J., Cha, S., Kim, B., Ryu, K.S., Ryu, S., Suh, J.Y., 2019. Structural basis for cell-wall recognition by bacteriophage PBC5 endolysin. Structure doi:10.1016/j.str.2019.07.001.
Lin, H.H., Yu, M., Sriramoju, M.K., Hsu, S.T.D., Liu, C.Te, Lai, E.M., 2020. A high-throughput interbacterial competition screen identifies ClpAP in enhancing recipient susceptibility to type VI secretion system-mediated attack by agrobacterium tumefaciens. Front. Microbiol. doi:10.3389/fmicb.2019.03077.
MacKenzie, D.A., Tailford, L.E., Hemmings, A.M., Juge, N., 2009. Crystal structure of a mucus-binding protein repeat reveals an unexpected functional immunoglobulin binding activity. J. Biol. Chem. doi:10.1074/jbc.M109.040907.
McAleese, F.M., Walsh, E.J., Sieprawska, M., Potempa, J., Foster, T.J., 2001. Loss of clumping factor B fibrinogen binding activity by staphylococcus aureus involves cessation of transcription, shedding and cleavage by metalloprotease. J. Biol. Chem. doi:10.1074/jbc.M102389200.
Mell, J.C., Redfield, R.J., 2014. Natural competence and the evolution of DNA uptake specificity. J. Bacteriol. doi:10.1128/JB.01293-13.
Mendes, R., Raaijmakers, J.M., 2015. Cross-kingdom similarities in microbiome functions. ISME J. doi:10.1038/ismej.2015.7.
Mesnage, S., Fontaine, T., Mignot, T., Delepierre, M., Mock, M., Fouet, A., 2000. Bacterial SLH domain proteins are non-covalently anchored to the cell surface via a conserved mechanism involving wall polysaccharide pyruvylation. EMBO J. doi:10.1093/emboj/19.17.4473.
Miller, O.K., Banfield, M.J., Schwarz-Linek, U., 2018. A new structural class of bacterial thioester domains reveals a slipknot topology. Protein Sci. doi:10.1002/pro.3478.
Mitkowski, P., Jagielska, E., Nowak, E., Bujnicki, J.M., Stefaniak, F., Niedziałek, D., Bochtler, M., Sabała, I., 2019. Structural bases of peptidoglycan recognition by lysostaphin SH3b domain. Sci. Rep. doi:10.1038/s41598-019-42435-z.
Najar, I.N., Thakur, N., 2020. A systematic review of the genera Geobacillus and Parageobacillus: their evolution, current taxonomic status and major applications. Microbiology doi:10.1099/mic.0.000945.
Padhi, S., Dash, M., Sahu, R., Panda, P., 2013. Urinary tract infection due to paenibacillus alvei in a chronic kidney disease: a rare case report. J. Lab. Phys. doi:10.4103/0974-2727.119872.
Pal, A., Saha, B.K., Saha, J., 2019. Comparative in silico analysis of ftsZ gene from different bacteria reveals the preference for core set of codons in coding sequence structuring and secondary structural elements determination. PLoS One doi:10.1371/journal.pone.0219231.
Parkinson, N., Bryant, R., Bew, J., Elphinstone, J., 2011. Rapid phylogenetic identification of members of the Pseudomonas syringae species complex using the rpoD locus. Plant Pathol. doi:10.1111/j.1365-3059.2010.02366.x.
Paterson, G.K., Mitchell, T.J., 2004. The biology of Gram-positive sortase enzymes. Trends Microbiol. doi:10.1016/j.tim.2003.12.007.
Peltier, J., Shaw, H.A., Couchman, E.C., Dawson, L.F., Yu, L., Choudhary, J.S., Kaever, V., Wren, B.W., Fairweather, N.F., 2015. Cyclic diGMP regulates production of sortase substrates of clostridium difficile and their surface exposure through ZMPI protease-mediated cleavage. J. Biol. Chem. doi:10.1074/jbc.M115.665091.
Pichlo, C., Juetten, L., Wojtalla, F., Schacherl, M., Diaz, D., Baumann, U., 2019. Molecular determinants of the mechanism and substrate specificity of Clostridium difficile proline-proline endopeptidase-1. J. Biol. Chem. doi:10.1074/jbc.RA119.009029.
Piepenbrink, K.H., Maldarelli, G.A., De La Peña, C.F.M., Mulvey, G.L., Snyder, G.A., De Masi, L., Von Rosenvinge, E.C., Günther, S., Armstrong, G.D., Donnenberg, M.S., Sundberg, E.J., 2014. Structure of clostridium difficile pilj exhibits unprecedented divergence from known type iv pilins. J. Biol. Chem. doi:10.1074/jbc.M113.534404.
Pinkston, K.L., Gao, P., Diaz-Garcia, D., Sillanpää, J., Nallapareddy, S.R., Murray, B.E., Harvey, B.R., 2011. The Fsr quorum-sensing system of Enterococcus faecalis modulates surface display of the collagen-binding MSCRAMM Ace through regulation of gelE. J. Bacteriol. doi:10.1128/JB.05026-11.
Potempa, M., Potempa, J., 2012. Protease-dependent mechanisms of complement evasion by bacterial pathogens. Biol. Chem. doi:10.1515/hsz-2012-0174.
Purcell, E.B., Tamayo, R., 2016. Cyclic diguanylate signaling in Gram-positive bacteriaa. FEMS Microbiol. Rev. doi:10.1093/femsre/fuw013.
Ramírez-puebla, S.T., Servín-Garcidueñas, L.E., Jiménez-marín, B., Bolaños, L.M., Rosenblueth, M., Martínez, J., Rogel, M.A., Ormeño-orrillo, E., Martínez-romero, E., 2013. Gut and root microbiota commonalities. Appl. Environ. Microbiol. doi:10.1128/AEM.02553-12.
Reboli, A.C., Bryan, C.S., Farrar, W.E., 1989. Bacteremia and infection of a hip prosthesis caused by Bacillus alvei. J. Clin. Microbiol. doi:10.1128/jcm.27.6.1395-1396.1989.
Rich, R.L., Demeler, B., Ashby, K., Deivanayagam, C.C.S., Petrich, J.W., Patti, J.M., Narayana, S.V.L., Höök, M., 1998. Domain structure of the Staphylococcus aureus collagen adhesin. Biochemistry doi:10.1021/bi981773r.
Schacherl, M., Pichlo, C., Neundorf, I., Baumann, U., 2015. Structural basis of proline-proline peptide bond specificity of the metalloprotease Zmp1 implicated in motility of clostridium difficile. Structure doi:10.1016/j.str.2015.06.018.
Schäffler, H., Breitrück, A., 2018. Clostridium difficile - From colonization to infection. Front. Microbiol. doi:10.3389/fmicb.2018.00646.
Schwaiger, K., Hölzel, C., Mayer, M., Bauer, J., 2010. Notes on the almost unknown genus Jeotgalicoccus. Lett. Appl. Microbiol. doi:10.1111/j.1472-765X.2010.02811.x.
Seele, J., Beineke, A., Hillermann, L.M., Jaschok-Kentner, B., Von Pawel-Rammingen, U., Valentin-Weigand, P., Baums, C.G., 2015. The immunoglobulin M-degrading enzyme of Streptococcus suis, Ide Ssuis, is involved in complement evasion. Vet. Res. doi:10.1186/s13567-015-0171-6.
Siguier, P., Gourbeyre, E., Chandler, M., 2014. Bacterial insertion sequences: their genomic impact and diversity. FEMS Microbiol. Rev. doi:10.1111/1574-6976.12067.
Sychantha, D., Chapman, R.N., Bamford, N.C., Boons, G.J., Howell, P.L., Clarke, A.J., 2018. Molecular basis for the attachment of S-layer proteins to the cell wall of bacillus anthracis. Biochemistry doi:10.1021/acs.biochem.8b00060.
Tonry, J.H., Mcnichol, B.A., Ramarao, N., Chertow, D.S., Kim, K.S., Stibitz, S., Schneewind, O., Kashanchi, F., Bailey, C.L., Popov, S., Chung, M.C., 2012. Bacillus anthracis protease InhA regulates BslA-mediated adhesion in human endothelial cells. Cell. Microbiol. doi:10.1111/j.1462-5822.2012.01791.x.
Van Leeuwen, H.C., Klychnikov, O.I., Menks, M.A.C., Kuijper, E.J., Drijfhout, J.W., Hensbergen, P.J., 2014. Clostridium difficile sortase recognizes a (S/P)PXTG sequence motif and can accommodate diaminopimelic acid as a substrate for transpeptidation. FEBS Lett. doi:10.1016/j.febslet.2014.09.041.
Vengadesan, K., Narayana, S.V.L., 2011. Structural biology of Gram-positive bacterial adhesins. Protein Sci. doi:10.1002/pro.613.
Ventosa, A., Márquez, M.C., Ruiz-Berraquero, F., Kocur, M., 1990. Salinicoccus roseus gen. nov., sp. nov., a New Moderately Halophilic Gram-Positive Coccus. Syst. Appl. Microbiol. doi:10.1016/S0723-2020(11)80177-3.
Walden, M., Edwards, J.M., Dziewulska, A.M., Bergmann, R., Saalbach, G., Rohde, M., Schwarz-Linek, U., Banfield, M.J., 2015. Covalent host-targeting by thioester domains of gram-positive pathogens. Acta Crystallogr. Sect. A Found. Adv. doi:10.1107/s2053273315099520.
Whittaker, C.A., Hynes, R.O., 2002. Distribution and evolution of von Willebrand/integrin A domains: widely dispersed domains with roles in cell adhesion and elsewhere. Mol. Biol. Cell. doi:10.1091/mbc.E02-05-0259.
Wilson, J.J., Matsushita, O., Okabe, A., Sakon, J., 2003. A bacterial collagen-binding domain with novel calcium-binding motif controls domain orientation. EMBO J. doi:10.1093/emboj/cdg172.
Yang, J., Yan, R., Roy, A., Xu, D., Poisson, J., Zhang, Y., 2014a. The I-TASSER suite: protein structure and function prediction. Nat. Methods doi:10.1038/nmeth.3213.
Yang, Y.H., Jiang, Y.L., Zhang, J., Wang, L., Bai, X.H., Zhang, S.J., Ren, Y.M., Li, N., Zhang, Y.H., Zhang, Z., Gong, Q., Mei, Y., Xue, T., Zhang, J.R., Chen, Y., Zhou, C.Z., 2014b. Structural insights into SraP-mediated staphylococcus aureus adhesion to host cells. PLoS Pathog. doi:10.1371/journal.ppat.1004169.
Young, C.L., Britton, Z.T., Robinson, A.S., 2012. Recombinant protein expression and purification: a comprehensive review of affinity tags and microbial applications. Biotechnol. J. doi:10.1002/biot.201100155.
Yunck, R., Cho, H., Bernhardt, T.G., 2016. Identification of MltG as a potential terminase for peptidoglycan polymerization in bacteria. Mol. Microbiol. doi:10.1111/mmi.13258.
Zeigler, D.R., 2014. The Geobacillus paradox: why is a thermophilic bacterial genus so prevalent on a mesophilic planet? Microbiol. (United Kingdom) doi:10.1099/mic.0.071696-0.
Zhang, R., Ou, H.-Y., Gao, F., Luo, H., 2014. Identification of horizontally-transferred genomic islands and genome segmentation points by using the GC profile method. Curr. Genom. doi:10.2174/1389202915999140328163125.
Zhu, H., Fu, B., Lu, S., Liu, Hongbo, Liu, He, 2018. Clostridium bovifaecis sp. nov., a novel acetogenic bacterium isolated from cow manure. Int. J. Syst. Evol. Microbiol. doi:10.1099/ijsem.0.002928.
Zong, Y., Xu, Y., Liang, X., Keene, D.R., Höök, A., Gurusiddappa, S., Höök, M., Narayana, S.V.L., 2005. A “Collagen Hug” model for staphylococcus aureus CNA binding to collagen. EMBO J. doi:10.1038/sj.emboj.7600888.