1   **Scaffolding protein GspB/OutB facilitates assembly of the *Dickeya dadantii* type 2 secretion system**

2   **by anchoring the outer membrane secretin pore to the inner membrane and to the peptidoglycan**

3   **cell wall.**

4

5   Shiheng ZHANG[1§], Shuang GU[2§], Piers RYCROFT[2], Florence RUAUDEL[1], Frederic DELOLME[3], Xavier

6   ROBERT[4], Lionel BALLUT[4], Richard W. PICKERSGILL[2*] and Vladimir E. SHEVCHIK[1*]

7

8   1 Univ Lyon, Université Lyon 1, INSA Lyon, CNRS UMR 5240 Microbiologie Adaptation et Pathogénie,

9   F-69622 Villeurbanne, France

10   2 School of Biological and Chemical Sciences, Queen Mary University of London, Mile End Road,

11   London E1 4NS, UK

12   3 Protein Science Facility, SFR BioSciences, CNRS, UMS3444, F-69007 Lyon, France

13   4 Univ Lyon, Molecular Microbiology and Structural Biochemistry, UMR5086 CNRS, F-69007 Lyon,

14   France

15

16   § These authors contributed equally to the work.

17   * Corresponding authors: vladimir.shevchik@insa-lyon.fr and r.w.pickersgill@qmul.ac.uk

18

19   **Keywords**

20   type 2 protein secretion system; GspD secretin; scaffolding protein GspB; peptidoglycan; pathogenic

21   bacteria, membrane proteins, *Dickeya dadantii*

22

23   **Running Title**

24   GspB/OutB facilitates assembly of the secretin pore

25

26   **Summary**

1　The phytopathogenic proteobacterium *Dickeya dadantii* secretes an array of plant cell wall degrading

2　enzymes and other virulence factors via the type 2 secretion system (T2SS). T2SSs are widespread

3　among important plant, animal and human bacterial pathogens. This multi-protein complex spans the

4　double membrane cell envelope and secretes fully folded proteins through a large outer membrane

5　pore formed by 15 subunits of the secretin GspD. Secretins are also found in the type 3 secretion

6　system and the type 4 pili. Usually, specialized lipoproteins termed as pilotins assist the targeting and

7　assembly of secretins into the outer membrane. Here, we show that in *Dickeya*, the pilotin acts in

8　concert with the scaffolding protein GspB. Deletion of *gspB* profoundly impacts secretin assembly,

9　pectinase secretion, and virulence. Structural studies reveal that GspB possesses a conserved

10　periplasmic Homology Region domain that interacts directly with the N-terminal secretin domain. Site-

11　specific photo cross-linking unravels molecular details of the GspB-GspD complex *in vivo*. We show

12　that GspB facilitates outer membrane targeting and assembly of the secretin pores and anchors them

13　to the inner membrane while the C-terminal extension of GspB scaffolds the secretin channel in the

14　peptidoglycan cell wall. Phylogenetic analysis shows that in other bacteria, GspB homologs vary in

15　length and domain composition and act in concert with either a cognate ATPase GspA or a pilotin GspS.

16

17　**Importance**

18　Gram-negative bacteria have two cell membranes sandwiching a peptidoglycan net that form

19　together a robust protective cell envelope. To translocate effector proteins across this multilayer

20　envelope, bacteria have evolved several specialized secretion systems. In the type 2 secretion system

21　and some other bacterial machineries, secretins form large multimeric pores that allow transport of

22　effector proteins or filaments across the outer membrane. The secretins are essential for nutrient

23　acquisition and pathogenicity and constitute a target for development of new antibacterials.

24　Targeting of secretin subunits into the outer membrane is often facilitated by a special class of

25　lipoproteins called pilotins. Here, we show that in *Dickeya* and some other bacteria, the scaffolding

26　protein GspB acts in concert with pilotin, facilitating the assembly of the secretin pore and its

1   anchoring to both the inner membrane and the bacterial cell wall. GspB homologs of varied domain

2   composition are present in many other T2SSs.

3

4   **Introduction**

5   Gram-negative bacteria are surrounded by two membranes, the inner membrane (IM) and the outer

6   membrane (OM) that together delimit a thin (~30 nm) periplasmic space containing the

7   peptidoglycan (PG) layer (1, 2). To exchange proteins, nucleic acids and sugars with the external

8   medium, these bacteria have evolved an array of specialized transport systems (3-7). The type 2

9   secretion system (T2SS) is widespread in proteobacteria and secretes folded proteins that play a

10   pivotal role in colonization of different niches, survival, competition and pathogenicity (8-12). The

11   plant pathogenic γ-proteobacterium *D. dadantii* uses the T2SS, called Out, to secrete pectinases in

12   the infected plant tissues causing soft rot disease in numerous plants and root-vegetables (13, 14).

13   The T2SS is embedded in both the IM and the OM and spans the whole cell envelope of the bacteria.

14   It is composed of 12 core components, generically called GspC to GspM and GspO (OutC to OutO, in

15   *D. dadantii*) as well as some additional components, GspA, B, N and S, that are present in certain

16   T2SSs (8-10, 15, 16).

17   The secretin is an essential T2SS component that forms gated channels in the OM, through

18   which the effector proteins are translocated in the medium or at the cell surface. The secretins are

19   also shared by the type 3 secretion system (T3SS), type 4 pili (T4P) as well as by the competence and

20   the filamentous phage assembly systems (16-19). Secretin homologs have also been identified in

21   mitochondria of some eukaryotes (20). Recent cryo-EM structures have revealed important

22   molecular details of the assembly of the MDa-sized secretin channels consisting of 12 to 16 subunits,

23   with a clear predominance of 15-fold symmetry in more recent, near-atomic resolution structures of

24   the T2SS and T3SS secretins (15, 21-28). Only a small apical segment of the secretin channel

25   consisting of an amphipathic helical loop and a 'β-lip' is embedded into the OM, while the main

26   portion, which is composed of the core secretin domain together with the N-terminal domains, N0 to

1   N3, are located in the periplasm. The conserved C-domains of 15 secretin subunits build up together

2   a double β-barrel channel composed of 60 internal and 60 external β-strands. The N1 to N3 domains

3   also termed as ring-building motifs, adopt mixed α/β fold and form together a cylinder like structure

4   protruding through the periplasm. In the T2SS secretins, the N-terminal N0 domain forms the first

5   gate at the entry of the secretin channel that interacts with the IM portion of the secretion

6   machinery and controls the recruitment of substrates (29-33). This portion of the secretin channel is

7   not seen in the majority of reported structures, or appears as smeared density consistent with its

8   flexibility (21, 23, 34). However, in the recent high-resolution cryo-EM structure of *Klebsiella*

9   *pneumoniae* PulD, the N0-N1 domains have been resolved as a tightly packed ring, stabilized by

10   interaction with the inner membrane PulC component (27). The molecular details of this interaction

11   remain elusive, but this study shows the relevance of attachment of the secretin channel to the

12   periplasmic portion of the IM components.

13   Usually, a small lipoprotein called pilotin guides the cognate secretin subunits through the

14   periplasm to the OM and facilitates their assembly (16, 35). Two groups of sequence and structure-

15   dissimilar pilotins have been identified in various T2SSs. The paradigm pilotins of the OutS/PulS

16   family from *Dickeya, Klebsiella* and enterohemorrhagic *E. coli* bind a short C-terminal region of

17   secretin, termed as S-domain, and pilot the secretin subunit via the Lol system to the inner leaflet of

18   the outer membrane (36-39). Upon binding of pilotin, the disordered S-domain folds into α-helix (40-

19   42). In the assembled secretin channel, the bound pilotins stabilize the external β-barrel by stapling

20   tandemly arranged secretin subunits (15, 43). The T2SS pilotins of the AspS/ExeS family that were

21   identified in *Vibrio*, *Aeromonas*, enterotoxigenic and enteropathogenic *E. coli* seem to fulfil exactly

22   the same functions than OutS/PulS while adopting profoundly different fold (15, 43-45). Targeting

23   and assembly of secretins from the T3SS and T4P systems are also facilitated by specific piloting

24   lipoproteins of different structures and modes of action. For instance, the T3SS pilotins MxiM and

25   ExsB are structurally unrelated neither to one another, nor to PilF/PilW family pilotins from the T4P

26   system, or to the T2SS pilotins (46-49).

1   Beside the pilotins, some other assistance proteins are also involved in the assembly of the

2   secretin channels through the bacterial cell wall. The main portion of the secretin channel is located

3   in the periplasm (18 to 20 nm) and consequently, it crosses through and could interact with the PG

4   mesh (27, 34). Indeed, the distance between the OM and the PG layer is 12 to 25 nm and is

5   controlled by the size of the major OM lipoprotein Lpp, or Braun's lipoprotein, that attaches

6   covalently the PG to the OM (1, 50). In the high-resolution cryo EM tomography of the *Salmonella*

7   *typhimurium* T3SS and the *Myxococcus xanthus* T4P, the PG layer was visualized around the N-

8   terminal portion of the respective secretin channels (25, 51). In the *M. xanthus* T4P, the N0 domain

9   of the secretin PilQ is preceded by a triplet of specialized PG-binding AMIN domains (25, 52). In

10   addition, another T4P component, TsaP, carries a LysM-like PG-binding domain; together they

11   provide anchoring of the secretin channel to the PG (53). The PG-binding modules have been

12   identified in the other trans-envelope machineries of Gram-negative bacteria illustrating that

13   anchoring to the PG is essential for their assembly and function (54-56).

14   In the T2SS, only LspD secretin from *Legionella* is known to possess a specialized PG-binding

15   SPOR domain, located at the N-terminus just prior to the N0 domain (34). Another known example of

16   PG-binding T2SS component is the multimodular ATPase GspA that possesses a periplasmic PG-

17   binding domain of Pfam family PF01471 (57). GspA has been identified in *Vibrio*, *Aeromonas* and

18   some other bacteria, where it acts together with GspB and is thought to be implicated in the

19   insertion of the secretin into or through the PG mesh (58, 59). GspB is also present in some other

20   *Enterobacteriaceae*, such as *Dickeya*, *Klebsiella* and *Pectobacterium*, but in this case, without a GspA

21   counterpart (60). Previous studies have suggested that GspB could interact with the secretin but its

22   precise role has remained unclear and conflicting (60-63).

23   In this study, we explore the structure and function of OutB/GspB from plant pathogen *D.*

24   *dadantii*, we reveal that the GspB periplasmic domain is structurally similar to the homology region

25   (HR) of GspC and depict the molecular details of its interaction with the N0 domain of the secretin.

5

1   We demonstrate that OutB/GspB guides the secretin to the outer membrane and anchors the

2   secretin to the inner membrane and to the cell wall peptidoglycan.

3

4   **Results**

5   **OutB is required for type 2 secretion and essential for the full virulence of *D. dadantii***

6   Collective action of several pectinases secreted by *D. dadantii* Out T2SS causes soft-rot maceration of

7   infected plants, while mutant bacteria lacking functional Out system are fully or partially avirulent

8   (14, 64). Therefore, to examine the functional relevance of OutB, we assessed its importance in the

9   context of plant infection. Pathogenicity assays with chicory leaves clearly show that the *D. dadantii*

10   *outB* deletion mutant was barely virulent in comparison with the wild-type strain. A very small rotted

11   area formed by *outB* mutant 24 hours post infection did not progress further and was surrounded by

12   dehydrated necrotic plant tissues that form a barrier for bacterial proliferation (Fig. 1A). *D. dadantii*

13   *outD* mutant, which lacks the secretin pore and was used as a negative control, provoked very similar

14   symptoms, with an even smaller infection zone. Since soft-rot symptoms result from the action of

15   pectinases secreted by the Out system, these data indicate that *outB* and *outD* mutants were both

16   unable to efficiently secrete pectinases into the plant tissue. Consistent with this, in plate secretion

17   assays that show pectin and cellulose degradation by the Out-secreted enzymes, the *outB* mutant

18   generated no or very small halos, similar to the *outD* mutant (Fig. 1B). These data show a loss or

19   malfunction of the Out T2SS in the *outB* mutant that causes a striking reduction of bacterial virulence

20   *in planta*.

21

22   **OutB is attached to both the inner and the outer membranes**

23   To examine if the absence of OutB compromises the integrity of the T2SS, the quantity of various

24   T2SS components was assessed by immunoblotting (Fig. S1). The abundance of the inner membrane

25   assembly platform component OutC and the major pseudopilin OutG did not vary significantly, while

26   the quantity of the secretin OutD decreased in the *D. dadantii outB* mutant. Conversely, the amount

1   of OutB was obviously lower in *D. dadantii outD* strain (Fig. S1). These data indicate that OutB is

2   required for secretin channel biogenesis or stability. OutD is localized in the OM while OutB is

3   predicted to reside in the IM, due to its hydrophobic N-terminal segment (Fig. 2A). To test how these

4   proteins could interact within the bacterial envelope, we fractionated membrane vesicles from *D.*

5   *dadantii* wild type cells on a sucrose gradient (Fig. 3). No full-length OutD but an abundant OutD

6   cross-reacting band of ~35 kDa was detected in the OM fractions (Fig. 3A) suggesting OutD

7   degradation during the 60 h centrifugation (probably by the *D. dadantii* metalloproteases (65) since

8   EDTA could not be used in the course of membrane separation). OutB co-localized with both

9   membranes, in contrast to the *bona fide* IM protein TolA and the OM components, OmpA porin and

10   lipopolysaccharide, LPS. This suggests that a fraction of OutB is attached to the OM or to an OM-

11   associated component, for example, the secretin or the peptidoglycan. The treatment of cell extracts

12   with lysozyme, to release the OM vesicles from the PG mesh, improved the separation of the OM

13   components, LPS and OmpA, which moved to the higher density fractions, but OutB remained split

14   between the two membranes (Fig. 3B). However, when the membranes of a *ΔoutD* strain were

15   separated, the effect of the PG on OutB location became clearly visible (Fig. 3C and D). Here, and only

16   after cell-wall cleavage by lysozyme, OutB was detected uniquely in the IM fractions. These data

17   suggest that OutB attaches to both the PG and OutD. We further examined these two possibilities.

18

19   **OutB suppresses PspA induction caused by mislocated secretin**

20   Inefficient targeting of secretins to the OM results in their spontaneous insertion into the IM that

21   causes ion leakage from the cytoplasm. This provokes induction of phage shock protein (Psp) stress

22   response system that acts to restore membrane integrity (37, 66-70). Consequently, an increased

23   level of PspA is an indicator of inefficient targeting of secretins. In *D. dadantii*, PspA is induced in the

24   absence of the pilotin OutS, which promotes OutD targeting to the OM (40, 68).

25   To examine if OutB could also assist the OutD targeting and affect the PspA response, *outD*

26   was expressed in *E. coli* either alone or with *outB* and/or *outS* from constructs designated as D, DB,

7

1 DBS and DS (Fig. S2). As expected, OutS caused a significant increase of the OutD level and reduced

2 that of PspA (Fig. 4A, compare DS with D). In contrast, OutB did not consistently affect the amount of

3 OutD but notably reduced the PspA response, and that, in a manner independent on OutS (Fig. 4A,

4 compare DB and DBS with D and DS). These data suggest that OutB could improve OutD targeting to

5 the OM. Consistent with this possibility, sucrose gradient analysis showed a lower proportion of

6 OutD in the IM fractions of the DBS sample in comparison with DS (Fig. 4B). In addition, the total

7 amount of OutD was visibly lower in the DS gradient, suggesting that in the absence of OutB, OutD is

8 more prone to degradation during the gradient separation (Fig. 4B). Therefore, OutB protects OutD

9 from degradation and facilitates its transport to the OM.

10

11 **The periplasmic OutB region decreases the PspA response caused by OutD mislocation**

12 OutB is composed of a short cytoplasmic region, a transmembrane helix followed in the periplasm by

13 a linker region of variable length, and a conserved periplasmic domain, ended with a 28-residue long

14 C-terminal extension (CTE) (Fig. 2A). We examined which part of OutB is critical for reduction of the

15 PspA response. When the CTE was deleted, the resulting OutB$_{\Delta CTE}$ reduced a PspA level similar to that

16 with the full-length OutB, indicating that this part of OutB is not essential in this context (Fig. 4A,

17 compare DB$_{\Delta C}$S and DBS). In contrast, when the transmembrane segment of OutB was replaced with

18 the cleavable N-terminal signal peptide, the resulting OutB$_{PS}$ reduced the PspA response even slightly

19 better than did OutB (Fig. 4A, compare DB$_{PS}$S and DBS). In the sucrose gradient of DB$_{PS}$S, OutD co-

20 migrated mainly with the OM fractions similarly to that in DBS (Fig. 4B). Notably, a small proportion

21 of the matured, signal peptide-less OutB$_{PS}$ co-migrated with the high-density OM indicating that the

22 periplasmic, membrane segment-free OutB$_{PS}$ binds OutD strongly enough to move together through

23 the gradient.

24

25 **OutD N0 domain is required for secretin targeting and gating the secretin channel**

8

1   In the reverse experiments, we examined which part of OutD is involved in the chaperoning by OutB.

2   Expression of OutD$_{\Delta N0}$ lacking the N0 domain (missing 91 residues of matured OutD, from F7 to I96)

3   elicited high PspA response even in the presence of OutB (Fig. 4A). Membrane separation shows that

4   in contrast to the wild type OutD, more than one-half of OutD$_{\Delta N0}$ co-migrated with the IM fractions,

5   (Fig. 4B, compare D$_{\Delta N0}$BS to DBS). In addition, in the D$_{\Delta N0}$BS gradient, OutB was strictly confined to a

6   few IM fractions while in DBS (OutD$_{WT}$), a significant proportion of OutB co-migrated with the

7   intermediate and the OM fractions. These data show that in the absence of the N0 domain, OutB is

8   no longer able to bind OutD and assist its correct targeting.

9

10 **OutB HR domain interacts with OutD N0 domain**

11 A possible direct interaction between OutB and OutD was first investigated by GST-pull down assay.

12 GST-tagged C-terminal portion of OutB consisting of HR and CTE (residues P112 to K220) bound very

13 efficiently the OutD derivatives comprising N0-N1-N2 domains and to a lone N0 domain but not to

14 the N1-N2 derivative (Fig. S3). Deletion of the 28-residue CTE did not affect the interaction,

15 narrowing the interacting domains down to the HR domain of OutB (residues P112 to G192) and N0

16 domain of OutD (residues A1 to S85) (Fig. S3).

17

18 **NMR characterization of OutB HR/OutD N0 complex**

19 NMR spectroscopy was further used to explore the interaction of the OutB HR and OutD N0 domains.

20 Initial HSQC experiments using ^{15}N-labelled HR domain of OutB (residues 112 to 220) and two OutD

21 derivatives, OutD-N0 (residues 1-85) and OutD-N0,N1,N2 (residues 1-258), showed that both OutD

22 derivatives gave the same pattern of peak shifts confirming that it is the N0 domain of the secretin

23 that interacts with the HR domain of OutB (Fig. S4). Partial assignment of the ^{1}H-^{15}N HSQC spectra for

24 OutB HR domain showed that several residues from different zones on the surface of the HR domain

25 are affected by the interaction with the OutD N0 domain (Fig. S5A). In reciprocal experiments,

26 residues of the OutD N0 domain involved in the interaction surface could not be readily identified

9

1   from peak shifts, because when the OutB HR domain was titrated into ^{15}N-labelled N0 domain there

2   were widespread chemical shift changes (Fig. S5B). Such peak shifts are occasionally seen when the

3   electron distribution is perturbed on forming a complex. Determination of the interdomain interface

4   in the OutB HR/OutD N0 complex by NMR experiments was therefore not straightforward.

5   **Crystal structure of the periplasmic domain of OutB**

6   The high resolution structures of the N0 GspDs domain have been reported for *E. coli* and *P.*

7   *aeruginosa* (31, 71). To define the molecular details of the OutB HR/OutD N0 interaction, we sought

8   to solve the structure of the OutB HR domain. Three constructs were used in crystallization trials:

9   OutB$^{112-220}$, OutB$^{112-202}$ and OutB$^{112-192}$ of which OutB$^{112-202}$ gave crystals used to determine the

10   structure at 2.05 Å resolution (Table S1). The structure of the OutB HR domain was solved using the

11   enterotoxigenic *E. coli* HR GspC domain as a search model in molecular replacement (17% of

12   sequence identity) (31). Residues 115-198 of OutB are clearly defined in the electron density map.

13   The HR domain of GspB consists of two three-stranded antiparallel β-sheets; the β-strands

14   sequentially form the up-down-up β-sheets, so the first sheet comprises strands 1, 2 and 3 and the

15   second, 4, 5 and 6 (Fig. 5A). The two sheets are at approximately 70° to each other, such as the

16   structure forms a β-sandwich with a hydrophobic core. The 3_{10}-helix in the loop between β1 and β2

17   contains a highly conserved R144, which forms a salt bridge to E155, also highly conserved (Fig. 2A

18   and S6A). Perhaps the most striking feature of the structure is the quantity of irregular polypeptide,

19   between the short β-strands, especially but not limited to the β3/ β4 loop, and at the amino- and

20   carboxy-ends of the β-sandwich (Fig. 5A). DSSP algorithm (72) reveals 58% of the residues do not fall

21   into a recognized secondary structure category.

22   A search for structural homologs using the DALI webserver (73) returned the HR GspC domain from

23   enterotoxigenic *E. coli* used as the molecular replacement search model as the most similar structure

24   (PDB entry 3OSS, Z-score 7.8, r.m.s.d. 1.7Å for 58 equivalent CA atoms) (31). The other related

25   architectures are PilP pilot protein from the T4P systems of *Neisseria meningitidis* (PDB entry 2IVW),

26   *P. aeruginosa* (PDB entry 2LC4) and *D. dadantii* OutC HR (PDB entry 2LNV) (74-76) (Fig. S6B).

10

1   ***In vivo* mapping of OutB site interacting with OutD**

2   To characterize the architecture of the OutB-OutD complex *in vivo*, site-specific photo cross-linking

3   was employed. In this approach, a photo-reactive amino acid *para*-benzoyl-phenylalanine (*p*BPA) is

4   incorporated *in vivo* in place of the residue of interest (77, 78). Upon a short UV exposure of cells,

5   *p*BPA can cross-link to any of the carbon-hydrogen bonds within a distance of 3 Å. In this way, we

6   introduced *p*BPA in place of 24 residues of the OutB HR domain and assessed their cross-linked

7   patterns in *D. dadantii* and *E. coli* cells. An abundant adduct of 110 kDa, cross-reacting with OutB and

8   OutD antibodies and compatible with an OutB-OutD complex, was generated by OutB carrying *p*BPA

9   in place of S139, Y137, Q170 and R154 (ranked in a descending order of abundance) (Fig. 5B and S7).

10   Remarkably, OutB_S139*p*BPA generated a near-quantitative cross-linking to OutD, indicating an

11   important OutB-OutD contact. When OutDΔN0 lacking the N0 domain was used in place of full-length

12   OutD, no OutB-D complex was generated indicating that OutB HR interacts directly with OutD N0

13   (Fig. S8).

14   Notably, the substitutions located at or near the β1 strand of OutB HR generated cross-

15   linking patterns with even/odd alternation, typical for a β-strand addition (79). For instance, an OutB-

16   D complex was generated with *p*BPA substitutions of Y137 and S139 but not with V136, T138 or A140

17   (Fig. 5B and S7). These data suggest that the β1 strand of OutB HR interacts with a β-strand from

18   OutD N0. Since Q170 (β4-β5 turn) is spatially proximal to Y137 and S139 (β1-strand), these three

19   residues seem to constitute the main OutD-interacting zone of OutB. Another residue generating an

20   abundant OutB-OutD complex, R154, is located rather far from this patch of highly reactive residues

21   (Fig. 5A) suggesting formation of an extended HR/N0 inter-domain interface or even two interfaces.

22   Interestingly, some OutB_*p*BPA variants, e.g. D142, K143 and G156, generated two OutB and OutD

23   cross-reactive adducts, one OutB-OutD and another one, of higher mass, compatible with OutB-

24   OutD-OutB complex (Fig. 5B). The latter, putative ternary complex could result from a simultaneous

25   cross-linking of one N0 domain to two HR domains.

26

11

1   **Mapping of OutD site interacting with OutB**

2   To identify the OutD residues interacting with OutB HR, photo cross-linked complexes generated by

3   OutB_pBPA variants S139, R154 and Q170 were purified by Strep-Tactin chromatography, separated

4   by SDS-PAGE, followed by in-gel digestion with chymotrypsin and trypsin. The peptides were then

5   analysed in a high resolution LC-MS/MS Q Exactive HF Mass Spectrometer (Thermo Fisher Scientific)

6   (Fig. 6A and S9). LC-MS/MS data were submitted to StavroX software tools (80), which allowed an

7   accurate identification of OutB_S139pBPA-OutD cross-linked peptides. Precisely, at a 1% False

8   Discovery Rate (FDR), several best spectra with a score above the score cut-off of 48 correspond to

9   the OutD peptide S^{41}YDMMNEGQY^{50}, cross-linked to the OutB peptide YTS$^{139}_{pBPA}$APDKR^{144} (Fig. 6A and

10   S10). The OutD residue, being cross-linked to pBPA, could be assigned from S^{41} to M^{45} with

11   prevalence for D^{43} and M^{44} (Table S2).

12   To further map the OutD binding site, in reverse photo cross-linking experiments, pBPA

13   substitutions were introduced along the β1 and β3 strands of OutD N0, in place of the residues

14   detected by the MS analysis or spatially close to them (Fig. 6B and C). An obvious OutD-OutB

15   complex reactive with both OutD and OutB antibodies was detected with the OutD_pBPA

16   substitutions of E2, S4, S6 (β1), Y42 and M44 (β3) (Fig. 6C). In the OutD N0 structure (Fig. 6B), these

17   residues are located close to each other and hence could interact with the same or proximal sites of

18   OutB HR. Significantly, Y42 and M44 support the OutD cross-linking site identified by the MS analysis

19   (Fig. 6A).

20   The validity of the identified HR-N0 interaction site was further assessed with *in vivo*

21   disulphide-bonding assay. Several residues of OutB HR that generated a pBPA-induced complex were

22   substituted with cysteine and co-expressed with either the wild-type OutD or OutD_Y42C since

23   OutD_M44C variant was barely detectable. Formation of disulphide bonds between spatially

24   proximal cysteine residues were next assessed during bacterial growth. Among the tested

25   combinations, OutB_S139C/OutD_Y42C and OutB_Y137C/OutD_Y42C pairs generated an abundant

26   OutB-OutD complex supporting the proximity of these residues in the functional T2SS (Fig. 6D).

<div align="center">12</div>

1   Therefore, in spite of the different hydrophobicity of the substituted residues, *p*BPA versus cysteine,

2   and the length of the respective cross-linking, 3 Å versus 7 Å, both *in vivo* approaches confirm the MS

3   data and clearly map one OutB-OutD interacting site to S139 of OutB HR and to a short zone around

4   Y42 and M44 of OutD N0.

5

6   **Model of OutD N0-OutB HR interaction**

7   Based on the data of *in vivo* cross-linking experiments, we generated an OutB HR/OutD N0 model by

8   employing the HDOCK server (81). In this docking modelling, the proximity of the residues Y137,

9   S139, R152, R154, Q170 of OutB HR and E2, S4, S6, Y42, M44 of OutD N0 was given as preferable

10   constraints for interacting residues. Among several generated OutB HR/OutD N0 models, one being

11   distinguished by an excellent docking energy of -149.45 and several convincing structural features

12   was retained (Fig. 7A). Firstly, in this model, all the interacting residues of OutB HR and OutD N0,

13   listed above, are placed in a close vicinity. Secondly, a striking aspect of this model is the

14   complementation of several β-strands between the OutB HR and OutD N0 domains. Precisely, the

15   β1-strand of OutB HR defined by the residues S133 to Y137 determines an antiparallel β-sheet with

16   the β1-strand of OutD N0 containing residues S4 to K8; this leads to a complementary five stranded

17   anti-parallel β-sheet implicating the two distinct domains (Fig. 7A). Such a mixed HR/N0 β-sheet is

18   consistent with the even/odd alternation of photo cross-linking patterns observed with the residues

19   of β1-strand of OutB HR and β1-strand of OutD N0 (Fig. 5B and 6C). It seems plausible that formation

20   of this mixed β-sheet together with a rather large HR/N0 interdomain interface could explain high

21   affinity of the OutB HR/OutD N0 interaction observed both *in vivo* and *in vitro* (Fig. 5B, S3 and S12).

22   The *D. dadantii* OutD is highly similar (66% sequence identity) to the secretin PulD of *K.*

23   *pneumoniae* whose pentadecameric structure was recently solved by cryo-EM (27). The N0 domains

24   are notoriously flexible, but have been imaged in this high-resolution structure, as a tightly packed

25   ring stabilized in the presence of PulC (PDB 6HCG). Based on PulD structure we generated the full-

26   length *D. dadantii* secretin model, to validate its ability to bind OutB in such a multimeric

13

1   conformation. The generated model confirms that OutB is indeed able to bind to the 15-meric *D.*

2   *dadantii* secretin with no clashes (Fig. 7B and C). The steric hindrance of the HR/N0 interface area

3   allows the binding of one copy of OutB to one OutD subunit. Our protein/protein docking model

4   between OutD and OutB HR shows that it can be repeated 15 times to form a complete ring with a

5   relative stoichiometry of OutB-OutD complex of 1:1. In this arrangement, OutB HR domains are

6   located radially inward of the N0 ring without any lateral steric contact clashes with neighboring HR

7   domains, forming a cap to the entry of pentadecameric secretin channel.

8

9   **In the functional T2SS, OutC could compete with OutB for the secretin OutD**

10   The HR domain of GspC/OutC components has also been reported to interact with the N0 domain of

11   secretin (31, 32). Interestingly, the OutB HR/OutD N0 complex, that we modelled based on *in vivo*

12   cross-linking data (Fig. 7A), is similar to the crystal structure of GspC HR/GspD N0-N1 complex from *E.*

13   *coli* (PDB entry 3OSS) (31). Particularly, when the N0 domains in these two structures were

14   superimposed, the corresponding HR domains are also well superposed, such as equivalent

15   secondary structure elements are placed in the two HR/N0 interdomain interfaces (Fig. S11). It seems

16   therefore plausible that in the secretion system, OutB and OutC could also take such equivalent

17   arrangements competing for the secretin. To test this hypothesis, NMR HQSC titration experiments

18   were performed with the individual OutB HR, OutC HR and OutD N0 domains (Fig. S12). In these

19   conditions, OutC HR did not compete with OutB HR for OutD N0 indicating that *in vitro*, OutB HR

20   binds to OutD N0 domain with higher affinity than OutC HR. To assess whether these two HR

21   domains compete for OutD *in vivo*, OutB_R154*p*BPA was expressed, either with or without OutC,

22   from CDBS or DBS plasmid, respectively. In *E. coli*, similar amounts of OutB-OutD complex were

23   generated from these two plasmids (Fig. S13), indicating that OutC does not affect the OutB-OutD

24   interaction in this background. In contrast, in *D. dadantii*, the amount of the generated OutB-OutD

25   complex was reproductively lower in the presence of OutC (Fig. S8 and S13), indicating that within

26   the functional T2SS, OutC could compete with OutB for the secretin binding.

14

GspB phylogeny

Search against the UniProtKB database by using the OutB HR domain as template shows the occurrence of GspB homologs in several groups of γ and β-proteobacteria. All these proteins carry a single transmembrane segment and a periplasmic HR domain but vary substantially in length and composition, so that a confident phylogenetic analysis could only be performed with the HR domains of these proteins (Fig. S14). Systematic inspection of *gspB* synteny and domain organization allow us to classify these proteins into seven archetypes (Fig. 2B). In most bacteria, *gspB* gene is co-located with *gspA.* Typically, GspA consists of a cytoplasmic AAA+ ATPase domain fused to a C-terminal transmembrane segment that is followed by a large periplasmic region with a cysteine peptidase-like and a PG-binding domain, such as in *Vibrio* and *Aeromonas* (Fig. 2B). Yet, in *Alteromonas* and *Pseudoalteromonas* groups, GspAs lack the periplasmic region and the cognate GspBs are notably longer than in the other bacteria (Fig. 2B). In addition, GspBs of *Alteromonas* group possess a supplementary cytoplasmic domain of DUF3391 family. Interestingly, in these bacteria, a gene coding for a DUF3391 domain protein is present next to *gspB* (Fig. 2B) indicating a possible origin of DUF3391 domain in GspB. In some bacteria, such as *V. vulnificus*, GspA is fused to the periplasmic portion of GspB in a single GspA-B polypeptide (82). Similar GspA-B fusions of various lengths and domain contents seem to have appeared independently in several groups of γ-proteobacteria (Fig. S14), supporting close functional relationship between GspA and GspB. *E. coli* K-12 and some other close enterobacteria possess an archetypal multidomain GspA in pair with a remarkably small GspB that has no HR domain but carries instead a variable polyampholyte sequence (Fig. 2B). Finally, *Dickeya, Pectobacterium, Klebsiella* and a few other close bacteria lack any *gspA* homolog, and *gspB* gene is co-located with a gene of the OutS/PulS pilotin family. Therefore, depending on bacteria, GspB proteins vary substantially in length and domain composition and act in concert with either a cognate GspA or a pilotin GspS.

15

OutB can be efficiently cross-linked to the PG *in vivo*

Separation of *D. dadantii* membrane vesicles in sucrose gradient suggested that OutB is anchored to the PG. Notably, degradation of the PG by lysozyme affects the location of OutB (Fig. 3C and D). In *Aeromonas* and *Vibrio*, GspA carries a specialized PG-binding domain (Fig. 2B) and GspA/GspB complex is involved in the insertion of cognate secretins into or through the PG mesh (57, 83). In *Dickeya* and related bacteria that lack a GspA homolog, OutB has an additional C-terminal polyampholyte sequence (Fig. 2). Therefore, we investigated if in *D. dadantii*, the absence of GspA could be compensated by this C-terminal extension.

The thiol-cleavable cross-linker 3, 3´-dithio-bis (sulfosuccinimidyl) propionate (DTSSP) was used to explore OutB-peptidoglycan interaction. *D. dadantii* Δ*outB* cells harboring a plasmid with *outB*, or *outBΔcte*, or empty vector were treated with DTSSP and PG was extracted by SDS-boiling procedure. The PG-bound proteins were next eluted with 2-mercaptoethanol and analyzed by immunoblotting (Fig. 8A). The PG-associated protein OmpA, used as a positive control, was equally eluted from all three PG preparations. In contrast, only OutB but not OutB$_{ΔCTE}$ was massively eluted from the PG. Therefore, OutB was efficiently cross-linked by DTSSP to the PG via its C-terminal extension suggesting that OutB CTE is proximal to the PG layer and could interact with it.

To further explore the OutB-PG interaction, photo cross-linking with *p*BPA, which has a shorter arm length than DTSSP (3 Å *vs* 12 Å) and a shorter time-lapse (3 min *vs* 20 min) was employed. *p*BPA substitutions were introduced into OutB CTE and their cross-linking patterns were assessed in *D. dadantii* (Fig. 8B). Diffuse multi-band adducts of about 40 to 50 kDa over a less-intense ladder-like background was observed. Treatment of cross-linked cells with lysozyme significantly decreased the amount of diffuse adduct and reduced the ladder-like background indicating that the PG-derived sugars are part of these complexes. Taken together these data suggest that *in vivo*, the C-terminal extension of OutB interacts with the PG. The functional relevance of the CTE was determined using the *D. dadantii outBΔcte* mutant, where the reduced pectinase and cellulase halos indicate diminished secretion (Fig. 1B). The effect was not as substantial as the complete *outB*

16

1 deletion, but was significant, revealing the functional importance of the C-terminal extension of

2 OutB.

3

4 **Discussion**

5 Here we show that OutB is essential for effective assembly and function of the *D. dadantii* T2SS and

6 for expression of full bacterial virulence. We reveal that OutB has three functions linked to distinct

7 regions of the protein: 1) The N-terminal transmembrane helix anchors OutB to the IM; 2) The

8 conserved periplasmic domain of OutB (HR) binds tightly the N0 domain of the secretin OutD and

9 assists its targeting to the outer-membrane, and 3) The C-terminal region (CTE) attaches OutB to the

10 peptidoglycan, anchoring the entire secretion system to the bacterial cell wall.

11   The crystal structure reveals that the conserved periplasmic domain of OutB has the same

12 fold as the homology regions (HR) of the T2SS GspC and the T4P PilP components (31, 74-76). In spite

13 of low sequence conservation (10 to 17% of identical residues), OutB HR structure is highly similar to

14 these HR domains (r.m.s.d. 1.7 Å to 2.8 Å for 58 equivalent CA atoms). This structural similarity

15 coincides with the fact that all these HR domains interact with the N0 domain of cognate secretins

16 (31, 32, 74, 84). By combining structural, *in vivo* and modelling approaches, we unravelled the

17 molecular details of the OutB HR/OutD N0 interaction. Particularly, site-specific photo cross-linking

18 followed by mass spectrometry analysis allowed an unambiguous identification of the OutB-OutD

19 interaction site *in vivo*. The validity of this site was further confirmed with reverse photo cross-linking

20 and *in vivo* disulphide bonding. Consequently, the OutB HR/OutD N0 model constructed after these

21 data reflects a functionally relevant protein contact within the functional secretion system. NMR

22 spectroscopy is roughly consistent with the cross-linking experiments but for various technical

23 reasons failed to give a comprehensive list of residues involved in the OutB HR/OutD N0 interface.

24   The N0 domains constitute the entry of the T2SS secretin channel, which controls the

25 recruitment of substrates and communicates with the IM components of the secretion machinery

26 (29-32). Recent cryo-EM tomography of the *L. pneumophila* T2SS showed high flexibility of the N-

17

1   terminal portion of the secretin channel that could reflect different functional states and interactions

2   with various partners (34). The occurrence of two components, OutB and OutC, carrying structurally

3   similar HR domains that interact with the same N0 domain of the cognate secretin, raises the

4   questions of the hierarchy of these interactions within the T2SS. Noteworthy, the fact that OutB

5   HR/OutD N0 interface modelled after our photo-CL data is strikingly similar to that of the crystal

6   GspC HR/GspD N0-N1 complex (Fig. S11) (31) suggests that these two interactions are spatially

7   incompatible and hence could not occur simultaneously. NMR experiments show that *in vitro* OutB

8   HR binds OutD N0 more tightly than OutC HR does, with the affinities ranging in 1 mM and 10 μM

9   order for OutC and OutB, respectively. (Fig. S12). Similarly, in *E. coli* carrying a partially reconstituted

10   transmembrane complex, OutC does not affect the OutB-OutD complex (Fig. S13). However, in *D.*

11   *dadantii*, OutC seems to interfere with the OutB-OutD interaction (Fig. S8 and S13). Since GspC/OutC

12   components are involved in substrate recruitment (29, 30), it is tempting to speculate that OutC-

13   OutD interactions are prevalent during the course of secretion while OutB-OutD interactions play a

14   role in the assembly, maintenance and scaffolding of the secretin channel.

15   We present a model of the OutD N0/OutB HR interaction in the context of the 15-meric

16   secretin channel (Fig. 7B). In this model, each OutB HR binds one OutD N0 domain in 1:1 ratio, such

17   as OutB HR domains are located radially inward of the N0 ring with no lateral contacts with

18   neighboring HR domains. The model presented here is therefore rather symmetric and constrained

19   compared to the more open and dynamic arrangement, as it can be anticipated in the functional

20   secretion system. Indeed, the functional oligomeric state of the secretin channel within the

21   functional T2SS has not been ultimately determined. Recent cryo-EM studies revealed a C15-fold

22   symmetry of the T2SS secretins (15, 21, 23, 28), while *in vivo* disulfide-bonding analyses are

23   consistent with a C6 symmetry of a hexamer of dimers (32, 71). In agreement with the latter, the N-

24   terminal portion of *P. aeruginosa* secretin XcpQ, comprising the N0-N1-N2 domains, self-assemblies

25   into a hexamer of dimers, forming *in vitro* C6 symmetry ring-shaped channels (85). Lithgow and

26   colleagues proposed an elegant solution to this apparent discrepancy (23) suggesting that the C15

1   symmetry observed in the secretin and N3 domains is further followed by the pseudo 6-fold

2   symmetry for the N0-N1-N2 domains, compatible with C12 or C6 symmetry of the assembly

3   platform, either hexameric as GspE, GspL and GspM, or dodecameric, as GspC (27). The common

4   symmetry element between C6, C12 and C15 symmetries is a three-fold axis. Therefore, C12 and C15

5   complexes could have common three-fold symmetry if every fifth N0 domain is unoccupied. Yet it is

6   not clear how such an arrangement would be enforced; thus it is plausible that the complex overall is

7   asymmetric unless three-fold symmetry is imposed in the periplasm by other components.

8   The role that OutB plays to maintain the secretin channels is closely linked to its ability to

9   bind peptidoglycan. Chemical cross-linking experiments with DTSSP and *p*BPA showed that the C-

10   terminal extension of OutB interacts with the peptidoglycan. The PG is composed of glycan chains

11   connected by short peptides that together form a mesh surrounding bacterial cell. In proteobacteria,

12   the PG layer is located in the periplasm and it is attached to the OM through a covalent linkage to the

13   Braun's lipoprotein Lpp (86). Several multi-protein structures and machineries span the entire cell

14   envelope. The size of the peptidoglycan mesh (about 2 nm) is not compatible with the external

15   dimensions of these large complexes and specialized or housekeeping transglycosylases are recruited

16   for the local rearrangement of the peptidoglycan layer in the course of their assembly through the

17   cell envelope (87, 88). The attachment of these transenvelope machineries to the PG mesh is usually

18   ensured by the specialized PG-binding modules. For instance, the T4P system possesses two

19   components carrying specialized PG-binding domains, secretin PilQ and TsaP, showing the

20   importance of a solid attachment to the PG (25, 52, 53). Generally, the PG-binding components

21   represent a less conserved part of respective transenvelope systems suggesting that anchorage to

22   the PG has evolved and diversified more recently, in accordance with the particular cell wall context

23   (55). For instance, *in silico* analysis of the T6SSs has revealed surprising richness and variability of the

24   associated PG-binding modules (89). The authors have proposed the notion of more or less "evolved"

25   PG-binding modules that have been developed from a few ancestral proteins by gene fusion and loss

26   of some nonessential regions.

1   Phylogenetic analysis shows that in many bacteria, such as in *Vibrio* and *Aeromonas*, GspB

2   acts in concert with a multi-domain ATPase GspA that carries a specialized PG-binding domain, (Fig.

3   2B and S14). In this group, some "more evolved" GspA and GspB are naturally fused into a single

4   polypeptide (59). A C-terminal extension, CTE, similar to that of OutB is only present in the

5   *Dickeya*/*Klebsiella* group that lacks a GspA counterpart (Fig. 2). We showed that the OutB CTE could

6   interact with peptidoglycan. Therefore, it is plausible that in these GspBs, the absence of the PG-

7   binding partner, GspA, is compensated by a polyampholyte C-terminal sequence that interacts with

8   PG. The molecular mechanism of this interaction remains unclear since the OutB CTE does not

9   correspond to any known PG-binding domain. However, the residue combinations of OutB CTE are

10   reminiscent of those in some sugar binding lectin domains. For instance, the bactericidal C-type

11   lectins of RegIII family recognize the PG carbohydrate moiety via an EPN-like motif located at an

12   extended loop region and the variations of this motif, EPN, QPD, EPS, etc. alter its sugar binding

13   specificity (90). It is possible that the OutB C-terminal extension could bind peptidoglycan in a similar

14   way. Some other bacterial proteins, apparently lacking a specialized PG-binding domain, have been

15   shown to bind the PG, e.g. the pilotin InvH from the T3SS of *Salmonella enteritica* (91). Interestingly,

16   in *L. pneumophila* T2SS that lacks any GspB or GspA homolog, the secretin LspD itself carries a SPOR-

17   like PG-binding domain at the N-terminus, prior to the N0 domain (34). Thus, the modes of

18   attachment of T2SS secretins to the PG vary in different bacteria.

19   In *Dickeya*/*Klebsiella* group, *gspB* is co-located with the gene coding for GspS pilotin (Fig. 2B).

20   We propose a model wherein OutB and OutS play complementary functions in targeting, assembly

21   and scaffolding of the secretin (Fig. 9). OutB could act early during OutD export by binding the N0

22   domain, the first part of the secretin released into the periplasm. In this way, OutB could orient the

23   secretin and prevent its misinsertion into the IM. In the later step, pilotin binding to the C-terminal S-

24   domain, with the assistance of the Lol machinery, would target and anchor the C-terminal portion of

25   the secretin into the OM (36), while the N-terminal entry of secretin pore remains connected to the

26   IM via OutB. At the same time, the C-terminal extension of OutB would attach the secretin to the PG.

20

1   Therefore, GspB/OutB has a major role in the assembly of the secretin channel providing it

2   scaffolding with the inner membrane assembly platform and the peptidoglycan layer.

3

4   **Materials and Methods.**

5   ***Strains, plasmids and growth conditions***

6   The bacterial strains and plasmids used in this study are listed in Tables S3 and S4. *D. dadantii* mutant

7   strains carrying chromosomal *outB* mutant alleles were constructed by marker exchange-eviction

8   mutagenesis or by homologous recombination as described (92). The bacteria were grown in Luria-

9   Bertani (LB) or in M9 minimal medium at 28°C with shaking at 120 rpm. If necessary, glucose, glycerol

10   or galacturonate were added at 0.2 % and antibiotics were added at the following final

11   concentrations: ampicillin, 100 μg/ml^{-1}; kanamycin, 50 μg/ml^{-1} and chloramphenicol, 25 μg/ml^{-1}. DNA

12   manipulations were carried out using standard methods. Site-directed mutagenesis was performed

13   with the PrimeSTAR Max DNA Polymerase (TaKaRa) and the primers are listed in Table S5. The

14   sequences of mutant and amplified genes were checked (Eurofins MWG Operon).

15   ***Protein purification***

16   DNA sequence encoding the conserved periplasmic domain of *D. dadantii* OutB was cloned into the

17   pGEX-6P-3 vector (GE Healthcare) encoding a cleavable N-terminal Glutathione S-transferase (GST)

18   affinity tag (Table S4). Three GST-OutB constructs encoding OutB residues 112-220, 112-202 and 112-

19   192 were expressed in *E. coli* BL21 (DE3) cells (Stratagene). The bacteria were grown in Luria broth

20   supplemented with 100 μg/ml ampicillin to an OD_{600} of 0.6 next, induced with 1 mM isopropyl-β-D-

21   thiogalactopyranoside (IPTG) and grown for an additional 16 h at 20°C. The cells were suspended in

22   phosphate buffered saline (PBS) containing 0.01 M phosphate buffer, 0.0027 M KCl and 0.137 M

23   NaCl, pH 7.4 and lysed by sonication. Cell debris was removed by centrifugation at 35,000g for 20

24   min and the supernatant from 1 l of cell culture was incubated with 2.5 ml of Glutathione Sepharose

25   beads (GE Healthcare) at 4°C for 6 h The Glutathione Sepharose beads were then collected in a PD10

26   column and washed with 50 ml of PBS. GST tag was on column cleaved by addition of 20 μl of

1 PreScission protease to 5 ml of Glutathione Sepharose in buffer containing 50 mM Tris-HCl, 150 mM

2 NaCl, 1 mM EDTA, 1 mM DTT, pH 7.0 at 4°C for 6 h. N0 and N0-N1-N2 domains of OutD (residues 1 to

3 85 and 1 to 258 of signal peptide-less OutD, respectively, Table S4) were produced in *E. coli*

4 BL21(DE3) and purified by nickel-affinity chromatography as described previously (33). For NMR

5 spectroscopy, uniformly ^{15}N- and ^{13}C-labeled proteins were produced by growing cell cultures in M9

6 minimal medium that contained 1 g/l ^{15}N-ammonium chloride and 2 g/l ^{13}C-D-glucose (Cambridge

7 Isotope Laboratories, Inc.) as the sources of nitrogen and carbon, respectively. Prior to crystallization

8 trials and NMR spectroscopy, the purified proteins were run on Superdex S200 size-exclusion column

9 and concentrated with Vivaspin devices (Sartorius).

10 ***Crystallization***

11 Crystallization conditions were screened using hanging drop vapour diffusion in 96 well plates set up

12 with a nanodrop Mosquito liquid handling robot (TTP Labtech). Protein and precipitant drops were in

13 the range of 200-400 nl. OutB (residues 112 to 202) crystals suitable for X-ray analysis grew in a solution

14 containing 0.2 M MgCl$_2$, 0.1 M Tris, 30% w/v PEG4000, pH 8.5 (Molecular Dimensions Structure Screen

15 1 condition C9). Selected crystals were transferred to a solution containing mother liquor plus 5%

16 propane-1,2-diol as a cryoprotectant and flash cooled in liquid nitrogen. Diffraction data were

17 collected at 100K using a Pilatus 300M detector at beam line I04 at Diamond Light Source (DLS),

18 Oxfordshire. Data were integrated and scaled by XDS through the Xia2 software package (93) available

19 through DLS. The structure was solved by molecular replacement using Phaser from the CCP4 software

20 package (94) using as a search model *E. coli* GspC-HR structure (PDB code 3OSS) detected using PHYRE2

21 (95). The model was refined using REFMAC v5.8 (96) and 10% of the data were excluded to be used for

22 validation purposes. OutB structure coordinates have been deposited at the Protein Data Bank in

23 Europe (PDBe) database with accession codes 4WFW.

24 ***NMR***

1   NMR spectra were acquired using the shortest OutB HR domain (residues 112-193) at 15°C using

2   Bruker Avance 600 and 700 MHz spectrometers equipped for ^{1}H, ^{15}N, ^{13}C triple resonance

3   experiments. All spectra were processed using NMRPipe/NMRDraw (97) and analyzed using XEASY

4   (98). HNCA, HN(CO)CA, HNCO, HNCACB, and CBCA(CO)NH experiments were used to obtain

5   sequence-specific ^{1}HN, ^{15}N, ^{13}Cα, ^{13}Cβ, and ^{13}C' backbone assignments. The NMR samples had a

6   typical concentration of 0.2 mM in 20 mm Tris-HCl, pH 7.0, and 10% D_2O. NMR titration was achieved

7   by recording a ^{1}H-^{15}N HSQC/ ^{1}H-^{15}N Sofast-HMQC spectra of the labeled protein. Cross titrations were

8   used to eliminate the dilution of the labeled protein. In cross titration, two samples were prepared,

9   one sample was labeled protein only and the second sample had the same concentration of labeled

10   protein and a ten equivalents of unlabeled titration partner. 150 mM NaCl was included in the cross-

11   titration studies.

12   ***Structure modeling***

13   A homology modelling of the *D. dadantii* OutD N0 domain (residues 1 to 78 of the matured, signal

14   peptide-less OutD, UniProtKB entry Q01565) was conducted using the trRosetta server (99) with

15   restraints from both deep learning and homologous templates. The confidence of the predicted model

16   is optimal with a TM-score of 0.912 on a maximum of 1.0. Using this OutD N0 model, we then

17   performed a protein/protein docking with the HR domain of OutB (PDB entry 4WFW). In this order,

18   the HDOCK server (81) was employed with the contacting residues, identified by the photo cross-

19   linking experiments, given as preferable constraints (namely Y137, S139, R152, R154, Q170 of OutB HR

20   and E2, S4, S6, Y42, M44 of OutD N0). In addition, the docking algorithm identified PDB entry 3OSS as

21   template (the crystal structure of enterotoxigenic *E. coli* GspC-GspD complex). This gave results to

22   several distinct models, one being distinguished by an excellent docking energy of -149.45 and several

23   convincing structural features. This last was retained for subsequent steps.

24   A homology modelling of the full-length OutD (residues 1 to 683) was then produced using again the

25   trRosetta server (99). A good model with a TM-score of 0.597 was produced. We then reconstructed

23

1   the full-length OutD/OutB HR domain interface using the previously identified protein/protein docking

2   pose. Finally, this OutD/OutB HR model was employed to reconstruct the *D. dadantii* secretin

3   pentadecameric model. In this goal, we used the cryo-EM structure of the *K. pneumoniae* type II

4   secretion system outer membrane complex (PDB entry 6HCG) as template (27). We superimposed the

5   OutD/OutB HR oligomer model to each of the 15 units of PulD with PyMOL v1.3 (The PyMOL Molecular

6   Graphics System, Schrödinger, LLC) using its 'cealign' function which is very robust for 3D structural

7   alignments of proteins with little to no sequence similarity. We therefore obtained the 15mer model

8   of OutD/OutB HR and we concluded that no further minimization step was needed as its visual

9   inspection demonstrated that no clashes were present in the N-terminal area harboring the

10   OutD/OutB HR interface.

11   ***Pull-down assays***

12   OutB domains of interest fused to GST were co-expressed with various combinations of N0 to N3

13   domains of OutD from the same plasmid (Table S4) in *E. coli* BL21(DE3) and co-purified with batch

14   method as described (100). Briefly, the cells were broken in a French press in TSE buffer (50 mM Tris-

15   HCl, 100 mM NaCl, 1 mM EDTA, pH 7.5), supplemented with 0.1% (v/v) Triton X-100 and incubated

16   with Glutathione Agarose (Machery Nagel). Non-bound proteins were washed away for six times with

17   the same buffer and the bound proteins were eluted with Laemmli sample buffer, separated by SDS-

18   PAGE and stained with Coomassie or probed by immunoblotting with anti-OutD antibodies.

19   ***Site-specific in vivo photo cross-linking***

20   *D. dadantii* or *E. coli* MG1655 cells, carrying pSup-BpaRS-6TRN (77), pREP4 (Qiagen) and a DBS-derived

21   plasmid carrying an *outB_TAG* or *outD_TAG* substitution were grown in BL supplemented with

22   ampicillin, kanamycin and chloramphenicol at 28°C for 16 h. 1 to 1.2 ml cultures were then collected,

23   washed in M9 and diluted (~three-fold) to an OD_{600} of 0.5 into M9 medium containing 1 mM *p*BPA,

24   0.2% glycerol, 0.01% casamino acids and appropriate antibiotics. To induce synthesis of pectinases and

25   Out proteins, *D. dadantii* cultures were additionally supplemented with 0.2% galacturonate. After a 1

24

1 h of growth, isopropyl-β-D-thiogalactopyranoside (IPTG) was added to 1 mM and the cultures were

2 followed for an additional 3 h or 5 h for *E. coli* and *D. dadantii*, respectively. Afterward, a 1.4 ml portion

3 of cells was either chilled on ice (control) or placed in a glass Petri dish and irradiated with an UV lamp

4 (Bio-Link BLX model, 365 nm) for 3 min at 10 cm distance. The cells were chilled on ice, collected by

5 centrifugation, lysed in Laemmli sample buffer and resolved on 9% SDS PAGE. The proteins were

6 detected by immunoblotting using antibodies raised against OutB or OutD.

7 ***Purification of OutB-OutD complex for mass spectrometry analysis***

8 *E. coli* MG1655 cells, carrying pSup-BpaRS-6TRN (77), pREP4 (Qiagen) and DBS-derived plasmid co-

9 expressing an *outB_TAG* variant together with *outD_Strep* were grown and subjected to photo cross-

10 linking exactly as described in the chapter above. Precisely, a series of 10 tubes with 5 ml culture was

11 grown and exposed together to UV in nine 10 cm diameter Petri dishes in the Bio-Link BLX chamber.

12 After 3 min irradiation with UV at 365 nm, the cells were chilled on ice, collected by centrifugation

13 and stored at -20°C. The frozen cells from each series (50 ml culture) were resuspended in 4 ml TSE

14 buffer (50 mM Tris-HCl pH 8.0, 100 mM NaCl, 1 mM EDTA, 1 mM PMSF) supplemented with 1 mg/ml

15 lysozyme and incubated for 15 min. The cell suspension was next supplemented with SDS to 1% and

16 boiled for an additional 15 min. Non-soluble cell material was removed by centrifugation at 7,000g

17 for 10 min and supernatant was ten-fold diluted with TSE buffer supplemented with 1% Triton X-100

18 and incubated with 2 ml of Step-Tactin XT Superflow (IBA) resin for 2 h at 15°C. The resin was next

19 washed five times with 10 ml of the same buffer and OutD-strep protein and complexes were eluted

20 with TSE buffer supplemented with 1% Triton X-100 and 50 mM biotin. The eluted proteins were

21 precipitated with five volumes of ethanol, resuspended in Laemmli sample buffer and separated on

22 9% SDS-PAGE.

23 ***Liquid chromatography tandem mass spectrometry of OutB_pBPA-OutD complexes***

24 The protein bands corresponding to OutB_*p*BPA-OutD complexes were excised from a stained gel

25 and subjected to digestion with chymotrypsin and trypsin. Precisely, the gel pieces were destained

25

1   with a 50% (v/v) mixture of 50 mM ammonium bicarbonate in acetonitrile, reduced with 10 mM DTT

2   (56°C, 1 h) and alkylated with 55 mM iodacetamide (25°C, 45 min, at darkness). Proteins were first

3   digested with chymotrypsin (Promega) (12.5 ng/µl, 25°C overnight). The 1/10 of the sample was kept

4   for injection and the 9/10 were digested with trypsin (Promega) (12.5 ng/µl, 37°C overnight). After

5   quenching with formic acid (FA) to 5%, the two digests were analyzed with an Ultimate 3000 nano-

6   RSLC (Thermo Fischer Scientific) coupled on line with a Q Exactive HF mass spectrometer via a nano-

7   electrospray ionization source (Thermo Fischer Scientific).

8   The samples were loaded on a Acclaim PepMap100 C18 trap column (20 x 0.075 mm, 100 Å

9   pore size) (Thermo Fischer Scientific) for 3 min at 5 µl/min with 2% acetonitrile, 0.05% trifluoracetic

10   acid (TFA) in H_2O and then separated on a Acclaim PepMap 100 C18 analytical column (50 x 0.075

11   mm, 100 Å pore size) (Thermo Fischer Scientific) linear gradient of A (H_2O, 0.1% FA) and B (100%

12   acetonitrile, 0.1% FA) solvents. Settings were: initial conditions of 4% to 50% of B in 60 min then from

13   50% to 95% of B in 2 min, hold for 10 min and returned to the initial conditions in 1 min for 14 min.

14   The total run duration was set to 90 min at a flow rate of 300 nl/min. The oven temperature was kept

15   at 40°C.

16   ***Mass spectrometry data processing***

17   Samples were analyzed with a TOP20 HCD method. MS data were acquired in a data dependent

18   strategy selecting the fragmentation events based on the 20 most abundant precursor ions in the

19   survey scan (300-1600 Th). The resolution of the survey scan was 60,000 at m/z 200 Th and for

20   MS/MS scan the resolution was set to 15,000 at m/z 200 Th. The Ion Target Value for the survey

21   scans in the Orbitrap and the MS/MS scan were set to 3E6 and 1E5 respectively and the maximum

22   injection time was set to 60 ms for MS scan and for MS/MS scan. Parameters for acquiring HCD

23   MS/MS spectra were as follows: collision energy of 27 and an isolation width of 2.0 m/z. The

24   precursors with unknown charge state, charge state of 1 and 8 or greater than 8 were excluded.

25   Peptides selected for MS/MS acquisition were then placed on an exclusion list for 20 s using the

26   dynamic exclusion mode to limit duplicate.

1   Raw data were first submitted to Proteome Discoverer 2.2 using SEQUEST HT search engine.

2   A FASTA file composed of the OutB and OutD sequences was created. Precursor mass tolerance was

3   set at 10 ppm and fragment mass tolerance was set at 0.02 Da, and up to 2 missed cleavages were

4   allowed. Enzyme parameter was set to a combination of trypsin and chymotrypsin. Oxidation (Met),

5   acetylation (Protein N-terminus) and the $S^{139}p$BPA modification delta mass (+164.0626 Da) were set

6   as variable modifications. Carbamidomethylation (Cys) was set as fixed modification. Validation of

7   the identified peptides were done by a 'Fixed value' approach based on SEQUEST scores.

8   pBPA crosslinked peptides were mapped with the StavroX part of MeroX (2.0.1.4 version

9   http://stavrox.com). Following parameters are set: raw date converted in mgf format; site 1:

10   Ser^{139}BPA (noted as x); site 2: all amino acids with methionine oxidation as variable modification,

11   precursor ion precision at 5 ppm and 0.02 Da on fragments. Validation of the candidates was based

12   on a target-decoy analysis. A False Discovery Rate (FDR) was calculated for each candidate and a FDR

13   score cut-off is applied at the end.

14   ***Chemical cross-linking and PG isolation***

15   Peptidoglycan was extracted from *D. dadantii* cells exponentially grown at BL broth supplemented with

16   ampicillin and galacturonate. The cells (~$5x10^{10}$) were washed with 50 mM sodium phosphate buffer

17   pH 7.2, resuspended in 2 ml of the same buffer and dropped in 2 ml of boiling SDS 10% solution. After

18   1 h of boiling, the suspension was stored overnight at 30°C and peptidoglycan was collected by

19   centrifugation at 48,000 rpm in SW 55 Ti rotor (Beckman Coulter) for 2 h at 30°C. Afterward,

20   peptidoglycan was resuspended in 4 ml of 2% SDS solution, boiled again for 30 min and collected by

21   centrifugation at 48,000 rpm in SW 55 Ti rotor for 1 h. Aliquots of purified peptidoglycan were boiled

22   with Laemmli loading buffer containing 2-mercaptoethanol and the eluted PG-bound proteins were

23   analyzed by immunoblotting with anti-OutB and anti-OmpA antibodies.

24   ***Enzymatic assays***

1   Plate assay for pectinase and cellulase secretion was performed as described (92). Bacteria were

2   patched onto polygalacturonate or carboxymethyl cellulose containing plates, for pectinase and

3   cellulase activities respectively, and grown at 28°C for 12 to 20 h. The plates were next flooded with

4   copper acetate or Congo red solutions, respectively, to reveal degradation halos. The halo size reflects

5   the efficiency of pectinase and cellulase secretion. For β-galactosidase activity assays, *E. coli* MC3

6   carrying pREP4 and a DBS-derived plasmid were grown in LB supplemented with 1 mM IPTG, ampicillin

7   and kanamycin at 28°C for 14 h. Afterward, the cells were permeabilized with toluene and 0.05 % SDS

8   and used in β-galactosidase assay, as described (101).

9   ***Cell fractionation***

10   The inner and outer membrane vesicles were separated on discontinuous sucrose density gradient. *E.*

11   *coli* MC3/pREP4 cells carrying a DBS derived plasmid were grown in LB medium. 100 ml of LB

12   supplemented with ampicillin and kanamycin was inoculated with non-induced overnight culture to

13   an OD_{600} 0.1 and cultivated at 120 rpm and 28°C for 2 h, then 1 mM IPTG was added and culture

14   followed for 7 h. Cells were collected, resuspended in 5 ml of 50 mM HEPES-NaOH pH 7.4 containing

15   protease inhibitor cocktail (Halt EDTA-free, Thermo) and disrupted by French pressure. Unbroken cells

16   and large debris were eliminated at 5,000g for 10 min. The cell-free extracts were supplemented with

17   sucrose to 30% and loaded on a discontinuous sucrose 35 to 65% gradient (with a 5% step). The

18   gradients were centrifuged at 48,000 rpm in SW 55 Ti rotor (Beckman Coulter) for 70 h, then 0.3 ml

19   fractions were collected from the bottom and analyzed by immunoblotting with the indicated

20   antibodies. *D. dadantii* were grown on polygalacturonate containing plates at 28°C for 36 h (3 plates

21   per strain). The plate-grown cells were collected, washed with HEPES-NaOH pH 7.4, broken by French

22   press and separated as above. When indicated, French pressure cell-free extracts were treated with

23   lysozyme (0.1 mg/ml) at 25°C for 30 min prior loading on sucrose gradient.

24   ***Immunoblotting***

1   The protein extracts were separated by electrophoresis on either 7.5, 9, 12 or 4-20 % SDS-PAG (home-

2   made or Bio-Rad), transferred on Immobilon-P membrane (Millipore) and analyzed by immunoblotting

3   with the polyclonal rabbit antibodies raised against OutB, OutC and OutD (39, 60), TolA (provided by

4   J.C. Lazzaroni, MAP Lab), OmpA (generated against recombinant *D. dadantii* OmpA, this study), PspA

5   (provided by Hendrick Osadnik, Leibniz Universität Hannover) or monoclonal mouse Ro-LPS antibody

6   (provided by J.C. Lazzaroni, MAP Lab).

7   ***Phylogenetic analysis***

8   Candidate GspB homologs were searched at a protein level by using BLAST (102) and OutB HR (residues

9   L127 to P191) as a query sequence against the Bacteria UniProtKB database. 1000 best hits were

10   manually inspected to remove incomplete or fragment sequences thus keeping 936 sequences.

11   Sequences were then clustered with CD-HIT software (103) with 30% identity threshold to keep 118

12   representative sequence clusters. Multiple sequence alignments were done with T-Coffee program

13   (104) and the HR domain sequences (correspond to residues L127 to P191 of OutB) were selected and

14   used for phylogenetic analyses. Phylogenetic analyses were performed, with the IQ tree web server in

15   auto mode (105). Consensus tree was constructed from 1000 bootstrap trees. Robinson-Foulds

16   distance between ML tree and consensus tree: 14. Phylogenetic tree was visualized with iTOL software

17   (106).

18   **Acknowledgements**

29

5   **References**

6   1.   Cohen EJ, Ferreira JL, Ladinsky MS, Beeby M, Hughes KT. 2017. Nanoscale-length control of the

7   flagellar driveshaft requires hitting the tethered outer membrane. Science 356:197-200.

8   2.   Miller SI, Salama NR. 2018. The gram-negative bacterial periplasm: Size matters. PLoS Biol

9   16:e2004935.

10 3.   Costa TR, Felisberto-Rodrigues C, Meir A, Prevost MS, Redzej A, Trokter M, Waksman G. 2015.

11   Secretion systems in Gram-negative bacteria: structural and mechanistic insights. Nat Rev

12   Microbiol 13:343-59.

13 4.   Ekiert DC, Bhabha G, Isom GL, Greenan G, Ovchinnikov S, Henderson IR, Cox JS, Vale RD. 2017.

14   Architectures of Lipid Transport Systems for the Bacterial Outer Membrane. Cell 169:273-

15   285.e17.

16 5.   Sherman DJ, Xie R, Taylor RJ, George AH, Okuda S, Foster PJ, Needleman DJ, Kahne D. 2018.

17   Lipopolysaccharide is transported to the cell surface by a membrane-to-membrane protein

18   bridge. Science 359:798-801.

19 6.   Dong C, Beis K, Nesper J, Brunkan-Lamontagne AL, Clarke BR, Whitfield C, Naismith JH. 2006.

20   Wza the translocon for E. coli capsular polysaccharides defines a new class of membrane

21   protein. Nature 444:226-9.

22 7.   Wu T, Malinverni J, Ruiz N, Kim S, Silhavy TJ, Kahne D. 2005. Identification of a multicomponent

23   complex required for outer membrane biogenesis in Escherichia coli. Cell 121:235-45.

24 8.   Gu S, Shevchik VE, Shaw R, Pickersgill RW, Garnett JA. 2017. The role of intrinsic disorder and

25   dynamics in the assembly and function of the type II secretion system. Biochim Biophys Acta

26   Proteom 1865:1255-1266.

9. Korotkov KV, Sandkvist M, Hol WG. 2012. The type II secretion system: biogenesis, molecular architecture and mechanism. Nat Rev Microbiol 10:336-51.

10. Thomassin JL, Santos Moreno J, Guilvout I, Tran Van Nhieu G, Francetic O. 2017. The trans-envelope architecture and function of the type 2 secretion system: new insights raising new questions. Mol Microbiol 105:211-226.

11. Cianciotto NP, White RC. 2017. Expanding Role of Type II Secretion in Bacterial Pathogenesis and Beyond. Infect Immun 85.

12. Denise R, Abby SS, Rocha EPC. 2019. Diversification of the type IV filament superfamily into machines for adhesion, protein secretion, DNA uptake, and motility. PLoS Biol 17:e3000390.

13. Hugouvieux-Cotte-Pattat N, Condemine, G., Gueguen, E. and Shevchik, V.E. . 2020. Dickeya Plant Pathogens, p 1-10, eLS doi:10.1002/9780470015902.a0028932. John Wiley & Sons, Ltd (Ed.).

14. Hugouvieux-Cotte-Pattat N, Condemine G, Shevchik VE. 2014. Bacterial pectate lyases, structural and functional diversity. Environ Microbiol Rep 6:427-40.

15. Howard SP, Estrozi LF, Bertrand Q, Contreras-Martel C, Strozen T, Job V, Martins A, Fenel D, Schoehn G, Dessen A. 2019. Structure and assembly of pilotin-dependent and -independent secretins of the type II secretion system. PLoS Pathog 15:e1007731.

16. Silva YRO, Contreras-Martel C, Macheboeuf P, Dessen A. 2020. Bacterial secretins: Mechanisms of assembly and membrane targeting. Protein Sci 29:893-904.

17. Korotkov KV, Gonen T, Hol WG. 2011. Secretins: dynamic channels for protein transport across membranes. Trends Biochem Sci 36:433-43.

18. Majewski DD, Worrall LJ, Strynadka NC. 2018. Secretins revealed: structural insights into the giant gated outer membrane portals of bacteria. Curr Opin Struct Biol 51:61-72.

19. Denise R, Abby SS, Rocha EPC. 2020. The Evolution of Protein Secretion Systems by Co-option and Tinkering of Cellular Machineries. Trends Microbiol 28:372-386.

31

20. Horváthová L, Žárský V, Pánek T, Derelle R, Pyrih J, Motyčková A, Klápšťová V, Vinopalová M, Marková L, Voleman L, Klimeš V, Petrů M, Vaitová Z, Čepička I, Hryzáková K, Harant K, Gray MW, Chami M, Guilvout I, Francetic O, Franz Lang B, Vlček Č, Tsaousis AD, Eliáš M, Doležal P. 2021. Analysis of diverse eukaryotes suggests the existence of an ancestral mitochondrial apparatus derived from the bacterial type II secretion system. Nat Commun 12:2947.

21. Yan Z, Yin M, Xu D, Zhu Y, Li X. 2017. Structural insights into the secretin translocation channel in the type II secretion system. Nat Struct Mol Biol 24:177-183.

22. Worrall LJ, Hong C, Vuckovic M, Deng W, Bergeron JRC, Majewski DD, Huang RK, Spreter T, Finlay BB, Yu Z, Strynadka NCJ. 2016. Near-atomic-resolution cryo-EM analysis of the Salmonella T3S injectisome basal body. Nature 540:597-601.

23. Hay ID, Belousoff MJ, Lithgow T. 2017. Structural Basis of Type 2 Secretion System Engagement between the Inner and Outer Bacterial Membranes. mBio 8.

24. D'Imprima E, Salzer R, Bhaskara RM, Sanchez R, Rose I, Kirchner L, Hummer G, Kuhlbrandt W, Vonck J, Averhoff B. 2017. Cryo-EM structure of the bifunctional secretin complex of Thermus thermophilus. Elife 6.

25. Chang YW, Rettberg LA, Treuner-Lange A, Iwasa J, Sogaard-Andersen L, Jensen GJ. 2016. Architecture of the type IVa pilus machine. Science 351:aad2001.

26. Koo J, Lamers RP, Rubinstein JL, Burrows LL, Howell PL. 2016. Structure of the Pseudomonas aeruginosa Type IVa Pilus Secretin at 7.4 Å. Structure 24:1778-1787.

27. Chernyatina AA, Low HH. 2019. Core architecture of a bacterial type II secretion system. Nat Commun 10:5437.

28. Hay ID, Belousoff MJ, Dunstan RA, Bamert RS, Lithgow T. 2018. Structure and Membrane Topography of the Vibrio-Type Secretin Complex from the Type 2 Secretion System of Enteropathogenic Escherichia coli. J Bacteriol 200.

29. Pineau C, Guschinskaya N, Robert X, Gouet P, Ballut L, Shevchik VE. 2014. Substrate recognition by the bacterial type II secretion system: more than a simple interaction. Mol Microbiol 94:126-40.

30. Douzi B, Ball G, Cambillau C, Tegoni M, Voulhoux R. 2011. Deciphering the Xcp Pseudomonas aeruginosa type II secretion machinery through multiple interactions with substrates. J Biol Chem 286:40792-801.

31. Korotkov KV, Johnson TL, Jobling MG, Pruneda J, Pardon E, Heroux A, Turley S, Steyaert J, Holmes RK, Sandkvist M, Hol WG. 2011. Structural and functional studies on the interaction of GspC and GspD in the type II secretion system. PLoS Pathog 7:e1002228.

32. Wang X, Pineau C, Gu S, Guschinskaya N, Pickersgill RW, Shevchik VE. 2012. Cysteine scanning mutagenesis and disulfide mapping analysis of arrangement of GspC and GspD protomers within the type 2 secretion system. J Biol Chem 287:19082-93.

33. Login FH, Fries M, Wang X, Pickersgill RW, Shevchik VE. 2010. A 20-residue peptide of the inner membrane protein OutC mediates interaction with two distinct sites of the outer membrane secretin OutD and is essential for the functional type II secretion system in Erwinia chrysanthemi. Mol Microbiol 76:944-55.

34. Ghosal D, Kim KW, Zheng H, Kaplan M, Truchan HK, Lopez AE, McIntire IE, Vogel JP, Cianciotto NP, Jensen GJ. 2019. In vivo structure of the Legionella type II secretion system by electron cryotomography. Nat Microbiol 4:2101-2108.

35. Koo J, Burrows LL, Howell PL. 2012. Decoding the roles of pilotins and accessory proteins in secretin escort services. FEMS Microbiol Lett 328:1-12.

36. Collin S, Guilvout I, Nickerson NN, Pugsley AP. 2011. Sorting of an integral outer membrane protein via the lipoprotein-specific Lol pathway and a dedicated lipoprotein pilotin. Mol Microbiol 80:655-65.

37. Hardie KR, Lory S, Pugsley AP. 1996. Insertion of an outer membrane protein in Escherichia coli requires a chaperone-like protein. Embo j 15:978-88.

38. Korotkov KV, Hol WG. 2013. Crystal structure of the pilotin from the enterohemorrhagic Escherichia coli type II secretion system. J Struct Biol 182:186-91.

39. Shevchik VE, Robert-Baudouy J, Condemine G. 1997. Specific interaction between OutD, an Erwinia chrysanthemi outer membrane protein of the general secretory pathway, and secreted proteins. Embo j 16:3007-16.

40. Gu S, Rehman S, Wang X, Shevchik VE, Pickersgill RW. 2012. Structural and functional insights into the pilotin-secretin complex of the type II secretion system. PLoS Pathog 8:e1002531.

41. Rehman S, Gu S, Shevchik VE, Pickersgill RW. 2013. Anatomy of secretin binding to the Dickeya dadantii type II secretion system pilotin. Acta Crystallogr D Biol Crystallogr 69:1381-6.

42. Tosi T, Nickerson NN, Mollica L, Jensen MR, Blackledge M, Baron B, England P, Pugsley AP, Dessen A. 2011. Pilotin-secretin recognition in the type II secretion system of Klebsiella oxytoca. Mol Microbiol 82:1422-32.

43. Yin M, Yan Z, Li X. 2018. Structural insight into the assembly of the type II secretion system pilotin–secretin complex from enterotoxigenic Escherichia coli. Nature Microbiology 3:581-587.

44. Dunstan RA, Heinz E, Wijeyewickrema LC, Pike RN, Purcell AW, Evans TJ, Praszkier J, Robins-Browne RM, Strugnell RA, Korotkov KV, Lithgow T. 2013. Assembly of the type II secretion system such as found in Vibrio cholerae depends on the novel Pilotin AspS. PLoS Pathog 9:e1003117.

45. Strozen TG, Li G, Howard SP. 2012. YghG (GspSβ) is a novel pilot protein required for localization of the GspSβ type II secretion system secretin of enterotoxigenic Escherichia coli. Infect Immun 80:2608-22.

46. Okon M, Moraes TF, Lario PI, Creagh AL, Haynes CA, Strynadka NC, McIntosh LP. 2008. Structural characterization of the type-III pilot-secretin complex from Shigella flexneri. Structure 16:1544-54.

47. Izoré T, Perdu C, Job V, Attree I, Faudry E, Dessen A. 2011. Structural characterization and membrane localization of ExsB from the type III secretion system (T3SS) of Pseudomonas aeruginosa. J Mol Biol 413:236-46.

48. Koo J, Tammam S, Ku SY, Sampaleanu LM, Burrows LL, Howell PL. 2008. PilF is an outer membrane lipoprotein required for multimerization and localization of the Pseudomonas aeruginosa Type IV pilus secretin. J Bacteriol 190:6961-9.

49. Trindade MB, Job V, Contreras-Martel C, Pelicic V, Dessen A. 2008. Structure of a widely conserved type IV pilus biogenesis factor that affects the stability of secretin multimers. J Mol Biol 378:1031-9.

50. Asmar AT, Ferreira JL, Cohen EJ, Cho SH, Beeby M, Hughes KT, Collet JF. 2017. Communication across the bacterial cell envelope depends on the size of the periplasm. PLoS Biol 15:e2004303.

51. Hu B, Lara-Tejero M, Kong Q, Galan JE, Liu J. 2017. In Situ Molecular Architecture of the Salmonella Type III Secretion Machine. Cell 168:1065-1074.e10.

52. Carter T, Buensuceso RN, Tammam S, Lamers RP, Harvey H, Howell PL, Burrows LL. 2017. The Type IVa Pilus Machinery Is Recruited to Sites of Future Cell Division. mBio 8.

53. Siewering K, Jain S, Friedrich C, Webber-Birungi MT, Semchonok DA, Binzen I, Wagner A, Huntley S, Kahnt J, Klingl A, Boekema EJ, Sogaard-Andersen L, van der Does C. 2014. Peptidoglycan-binding protein TsaP functions in surface assembly of type IV pili. Proc Natl Acad Sci U S A 111:E953-61.

54. Roujeinikova A. 2008. Crystal structure of the cell wall anchor domain of MotB, a stator component of the bacterial flagellar motor: implications for peptidoglycan recognition. Proc Natl Acad Sci U S A 105:10348-53.

55. Aschtgen MS, Gavioli M, Dessen A, Lloubès R, Cascales E. 2010. The SciZ protein anchors the enteroaggregative Escherichia coli Type VI secretion system to the cell wall. Mol Microbiol 75:886-99.

56. Veith PD, Glew MD, Gorasia DG, Reynolds EC. 2017. Type IX secretion: the generation of bacterial cell surface coatings involved in virulence, gliding motility and the degradation of complex biopolymers. Mol Microbiol 106:35-53.

57. Li G, Howard SP. 2010. ExeA binds to peptidoglycan and forms a multimer for assembly of the type II secretion apparatus in Aeromonas hydrophila. Mol Microbiol 76:772-81.

58. Vanderlinde EM, Strozen TG, Hernández SB, Cava F, Howard SP. 2017. Alterations in Peptidoglycan Cross-Linking Suppress the Secretin Assembly Defect Caused by Mutation of GspA in the Type II Secretion System. J Bacteriol 199.

59. Strozen TG, Stanley H, Gu Y, Boyd J, Bagdasarian M, Sandkvist M, Howard SP. 2011. Involvement of the GspAB complex in assembly of the type II secretion system secretin of Aeromonas and Vibrio species. J Bacteriol 193:2322-31.

60. Condemine G, Shevchik VE. 2000. Overproduction of the secretin OutD suppresses the secretion defect of an Erwinia chrysanthemi outB mutant. Microbiology 146 (Pt 3):639-647.

61. Vanderlinde EM, Zhong S, Li G, Martynowski D, Grochulski P, Howard SP. 2014. Assembly of the type two secretion system in Aeromonas hydrophila involves direct interaction between the periplasmic domains of the assembly factor ExeB and the secretin ExeD. PLoS One 9:e102038.

62. D'Enfert C, Pugsley AP. 1989. Klebsiella pneumoniae pulS gene encodes an outer membrane lipoprotein required for pullulanase secretion. J Bacteriol 171:3673-9.

63. Jahagirdar R, Howard SP. 1994. Isolation and characterization of a second exe operon required for extracellular protein secretion in Aeromonas hydrophila. J Bacteriol 176:6819-26.

64. Reverchon S, Nasser W. 2013. Dickeya ecology, environment sensing and regulation of virulence programme. Environ Microbiol Rep 5:622-36.

65. Létoffé S, Delepelaire P, Wandersman C. 1990. Protease secretion by Erwinia chrysanthemi: the specific secretion functions are analogous to those of Escherichia coli alpha-haemolysin. Embo j 9:1375-82.

36

66. Flores-Kim J, Darwin AJ. 2016. The Phage Shock Protein Response. Annu Rev Microbiol 70:83-101.

67. Guilvout I, Chami M, Engel A, Pugsley AP, Bayan N. 2006. Bacterial outer membrane secretin PulD assembles and inserts into the inner membrane in the absence of its pilotin. Embo j 25:5241-9.

68. Shevchik VE, Condemine G. 1998. Functional characterization of the Erwinia chrysanthemi OutS protein, an element of a type II secretion system. Microbiology 144 (Pt 11):3219-28.

69. Brissette JL, Russel M, Weiner L, Model P. 1990. Phage shock protein, a stress protein of Escherichia coli. Proc Natl Acad Sci U S A 87:862-6.

70. Junglas B, Huber ST, Heidler T, Schlösser L, Mann D, Hennig R, Clarke M, Hellmann N, Schneider D, Sachse C. 2021. PspA adopts an ESCRT-III-like fold and remodels bacterial membranes. Cell 184:3674-3688.e18.

71. Van der Meeren R, Wen Y, Van Gelder P, Tommassen J, Devreese B, Savvides SN. 2013. New insights into the assembly of bacterial secretins: structural studies of the periplasmic domain of XcpQ from Pseudomonas aeruginosa. J Biol Chem 288:1214-25.

72. Kabsch W, Sander C. 1983. Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. Biopolymers 22:2577-637.

73. Holm L, Laakso LM. 2016. Dali server update. Nucleic Acids Res 44:W351-5.

74. Gu S, Kelly G, Wang X, Frenkiel T, Shevchik VE, Pickersgill RW. 2012. Solution structure of homology region (HR) domain of type II secretion system. J Biol Chem 287:9072-80.

75. Tammam S, Sampaleanu LM, Koo J, Sundaram P, Ayers M, Chong PA, Forman-Kay JD, Burrows LL, Howell PL. 2011. Characterization of the PilN, PilO and PilP type IVa pilus subcomplex. Mol Microbiol 82:1496-514.

76. Golovanov AP, Balasingham S, Tzitzilonis C, Goult BT, Lian LY, Homberset H, Tønjum T, Derrick JP. 2006. The solution structure of a domain from the Neisseria meningitidis lipoprotein PilP reveals a new beta-sandwich fold. J Mol Biol 364:186-95.

77. Ryu Y, Schultz PG. 2006. Efficient incorporation of unnatural amino acids into proteins in Escherichia coli. Nat Methods 3:263-5.

78. Farrell IS, Toroney R, Hazen JL, Mehl RA, Chin JW. 2005. Photo-cross-linking interacting proteins with a genetically encoded benzophenone. Nat Methods 2:377-84.

79. Remaut H, Waksman G. 2006. Protein-protein interaction through beta-strand addition. Trends Biochem Sci 31:436-44.

80. Götze M, Pettelkau J, Schaks S, Bosse K, Ihling CH, Krauth F, Fritzsche R, Kühn U, Sinz A. 2012. StavroX--a software for analyzing crosslinked products in protein interaction studies. J Am Soc Mass Spectrom 23:76-87.

81. Yan Y, Tao H, He J, Huang SY. 2020. The HDOCK server for integrated protein-protein docking. Nat Protoc 15:1829-1852.

82. Martynowski D, Grochulski P, Howard PS. 2013. Structure of a periplasmic domain of the EpsAB fusion protein of the Vibrio vulnificus type II secretion system. Acta Crystallogr D Biol Crystallogr 69:142-9.

83. Howard SP, Gebhart C, Langen GR, Li G, Strozen TG. 2006. Interactions between peptidoglycan and the ExeAB complex during assembly of the type II secretin of Aeromonas hydrophila. Mol Microbiol 59:1062-72.

84. Balasingham SV, Collins RF, Assalkhou R, Homberset H, Frye SA, Derrick JP, Tønjum T. 2007. Interactions between the lipoprotein PilP and the secretin PilQ in Neisseria meningitidis. J Bacteriol 189:5716-27.

85. Douzi B, Trinh NTT, Michel-Souzy S, Desmyter A, Ball G, Barbier P, Kosta A, Durand E, Forest KT, Cambillau C, Roussel A, Voulhoux R. 2017. Unraveling the Self-Assembly of the Pseudomonas aeruginosa XcpQ Secretin Periplasmic Domain Provides New Molecular Insights into Type II Secretion System Secreton Architecture and Dynamics. mBio 8.

86. Asmar AT, Collet JF. 2018. Lpp, the Braun lipoprotein, turns 50-major achievements and remaining issues. FEMS Microbiol Lett 365.

87. Santin YG, Cascales E. 2017. Domestication of a housekeeping transglycosylase for assembly of a Type VI secretion system. EMBO Rep 18:138-149.

88. Scheurwater E, Reid CW, Clarke AJ. 2008. Lytic transglycosylases: bacterial space-making autolysins. Int J Biochem Cell Biol 40:586-91.

89. Aschtgen MS, Thomas MS, Cascales E. 2010. Anchoring the type VI secretion system to the peptidoglycan: TssL, TagL, TagP... what else? Virulence 1:535-40.

90. Lehotzky RE, Partch CL, Mukherjee S, Cash HL, Goldman WE, Gardner KH, Hooper LV. 2010. Molecular basis for peptidoglycan recognition by a bactericidal lectin. Proc Natl Acad Sci U S A 107:7722-7.

91. Pucciarelli MG, García-del Portillo F. 2003. Protein-peptidoglycan interactions modulate the assembly of the needle complex in the Salmonella invasion-associated type III secretion system. Mol Microbiol 48:573-85.

92. Bouley J, Condemine G, Shevchik VE. 2001. The PDZ domain of OutC and the N-terminal region of OutD determine the secretion specificity of the type II out pathway of Erwinia chrysanthemi. J Mol Biol 308:205-19.

93. Winter G, Lobley CM, Prince SM. 2013. Decision making in xia2. Acta Crystallogr D Biol Crystallogr 69:1260-73.

94. McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ. 2007. Phaser crystallographic software. J Appl Crystallogr 40:658-674.

95. Kelley LA, Sternberg MJ. 2009. Protein structure prediction on the Web: a case study using the Phyre server. Nat Protoc 4:363-71.

96. Murshudov GN, Skubák P, Lebedev AA, Pannu NS, Steiner RA, Nicholls RA, Winn MD, Long F, Vagin AA. 2011. REFMAC5 for the refinement of macromolecular crystal structures. Acta Crystallogr D Biol Crystallogr 67:355-67.

97. Delaglio F, Grzesiek S, Vuister GW, Zhu G, Pfeifer J, Bax A. 1995. NMRPipe: a multidimensional spectral processing system based on UNIX pipes. J Biomol NMR 6:277-93.

1   98.   Bartels C, Xia TH, Billeter M, Güntert P, Wüthrich K. 1995. The program XEASY for computer-

2   supported NMR spectral analysis of biological macromolecules. J Biomol NMR 6:1-10.

3   99.   Yang J, Anishchenko I, Park H, Peng Z, Ovchinnikov S, Baker D. 2020. Improved protein

4   structure prediction using predicted interresidue orientations. Proc Natl Acad Sci U S A

5   117:1496-1503.

6   100.   Login FH, Shevchik VE. 2006. The single transmembrane segment drives self-assembly of OutC

7   and the formation of a functional type II secretion system in Erwinia chrysanthemi. J Biol Chem

8   281:33152-62.

9   101.   Lallemand M, Login FH, Guschinskaya N, Pineau C, Effantin G, Robert X, Shevchik VE. 2013.

10   Dynamic interplay between the periplasmic and transmembrane domains of GspL and GspM

11   in the type II secretion system. PLoS One 8:e79562.

12   102.   Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. 1997. Gapped

13   BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids

14   Res 25:3389-402.

15   103.   Li W, Jaroszewski L, Godzik A. 2001. Clustering of highly homologous sequences to reduce the

16   size of large protein databases. Bioinformatics 17:282-3.

17   104.   Notredame C, Higgins DG, Heringa J. 2000. T-Coffee: A novel method for fast and accurate

18   multiple sequence alignment. J Mol Biol 302:205-17.

19   105.   Trifinopoulos J, Nguyen LT, von Haeseler A, Minh BQ. 2016. W-IQ-TREE: a fast online

20   phylogenetic tool for maximum likelihood analysis. Nucleic Acids Res 44:W232-5.

21   106.   Letunic I, Bork P. 2019. Interactive Tree Of Life (iTOL) v4: recent updates and new

22   developments. Nucleic Acids Res 47:W256-w259.

23   107.   Robert X, Gouet P. 2014. Deciphering key features in protein structures with the new ENDscript

24   server. Nucleic Acids Res 42:W320-4.

25

26   **Figures**

40

1
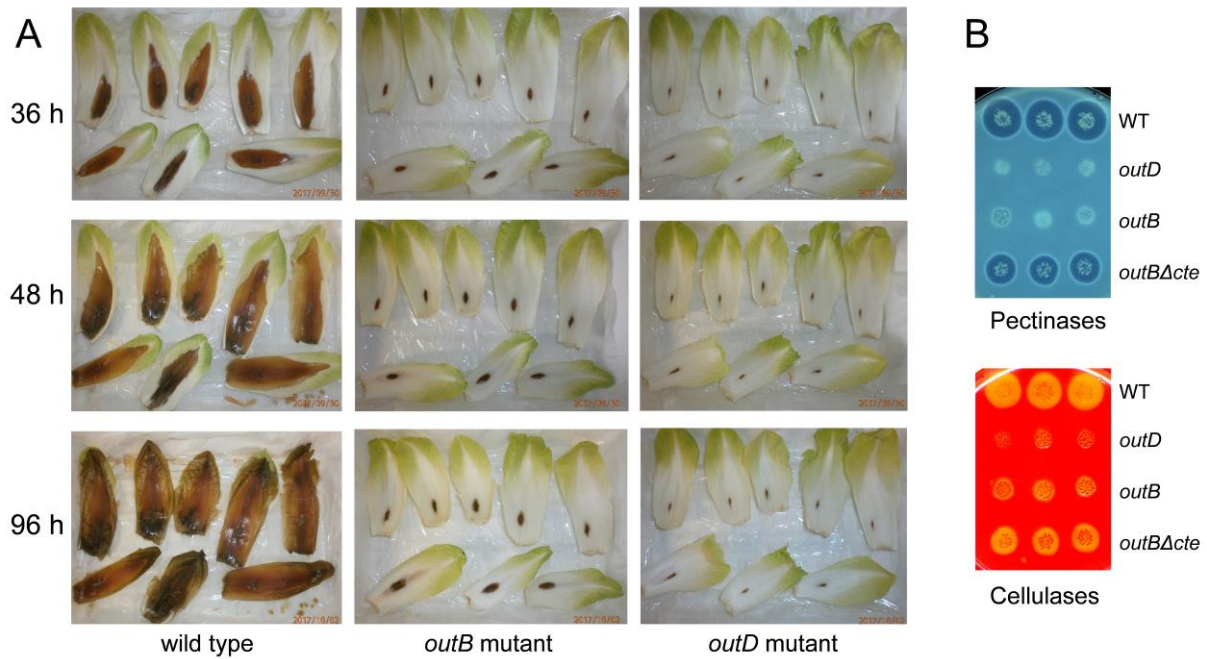
Figure 1. OutB is required for the type 2 secretion and full virulence of *D. dadantii*.

A. Pathogenicity tests with chicory leaves. Chicory leaves were inoculated with 10^6 cells of *D. dadantii*

wt, *outB* or *outD* strains and incubated at 28°C for the indicated time. **B.** Pectinase and cellulase

secretion assays. *D. dadantii* strains carrying the indicated mutations were grown for 14 h at 30°C on

plates containing polygalacturonate or carboxymethylcellulose and flooded with cupper acetate or

Congo Red (upper and lower panels, respectively). Size of halo reflects the level of pectinase and

cellulase secretion, respectively.
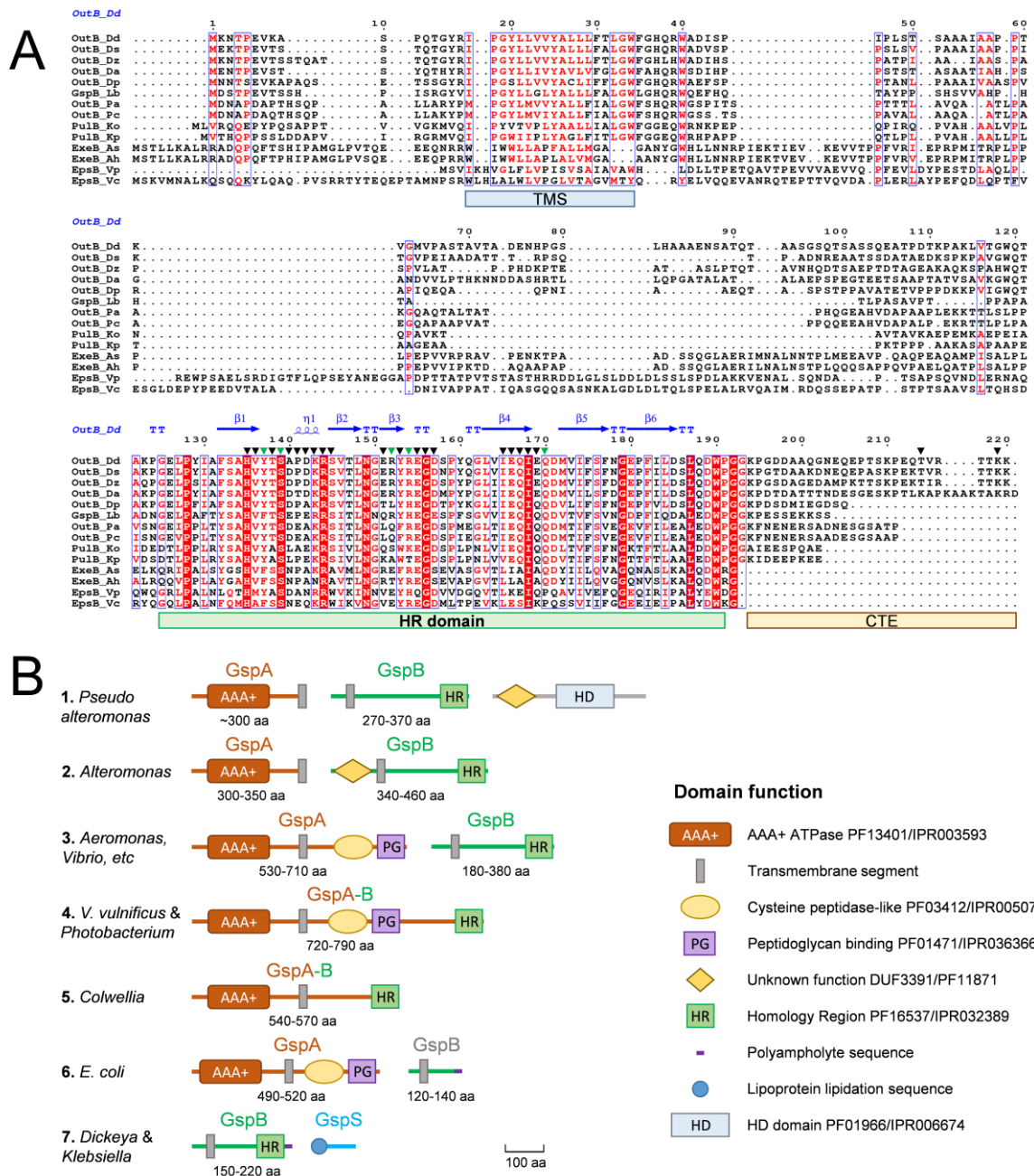
Figure 2. GspB organization. A. Alignment of representative GspB sequences. Multiple sequence alignment was performed with Clustal Omega (ebi.ac.uk) and ESPript server (107). The secondary structure elements are shown for the OutB HR domain (PDB entry 4WFW). The residue numbering is this for *D. dadantii* OutB. The groups of identical and similar residues are highlighted in red or are in red, respectively. The residues substituted with *p*BPA and used in photo-cross-linking are shown with triangles; with green triangles are the residues generating an abundant complex with OutD. Positions of the transmembrane segment (TMS), the homology region (HR) and the C-terminal extension (CTE) are indicated with coloured bars. Accession numbers of the used protein sequences are listed in Table S6. **B.** Typical gene and domain organizations of GspB and related GspA and GspS proteins. For each GspB representative shown in Fig. S14, gene synteny and domain organization were analyzed with NCBI (www.ncbi.nlm.nih.gov), InterPro (www.ebi.ac.uk/interpro) and Pfam (pfam.xfam.org) data bases and summarized into seven architypes, named according to the most abundant or most studied representative bacteria. The protein size range (in amino acids) is indicated for each group. Of note is that more than 80% of analyzed GspB proteins belong to the archetype #3.
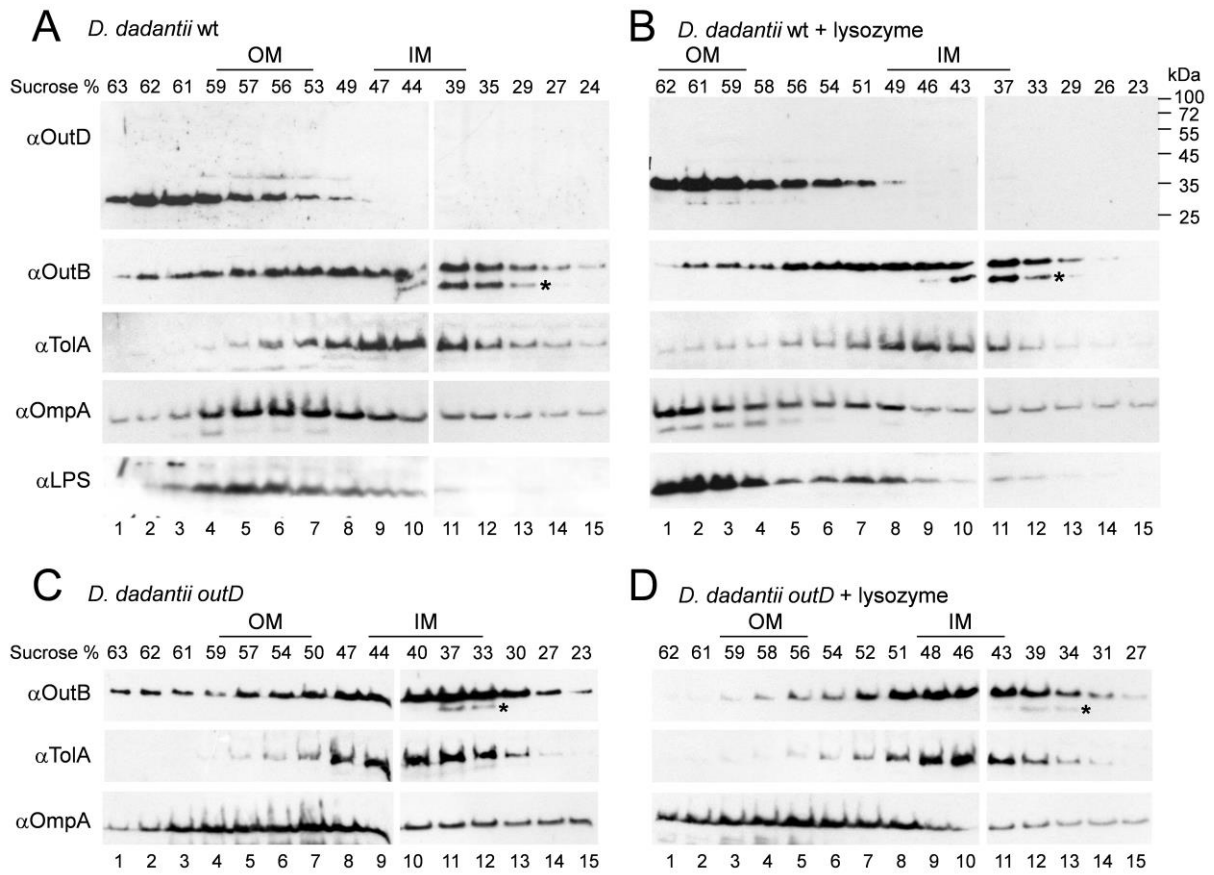
Figure 3. OutB is associated with both the inner and the outer membranes. *D. dadantii* A5652 wt (**A** and **B**) and A5653 *outD* (**C** and **D**) cells were grown on pectin containing plates, broken in French press and membranes were separated on sucrose gradient and analyzed by immunoblotting with the indicated antibodies. In **B** and **D**, the broken cell extracts were treated with lysozyme prior loading onto the gradient. The position of the inner and the outer membrane fractions are indicated according to the positions of LPS and OmpA (OM) as well as TolA (IM). A degradative product of OutB is indicated with an asterisk.
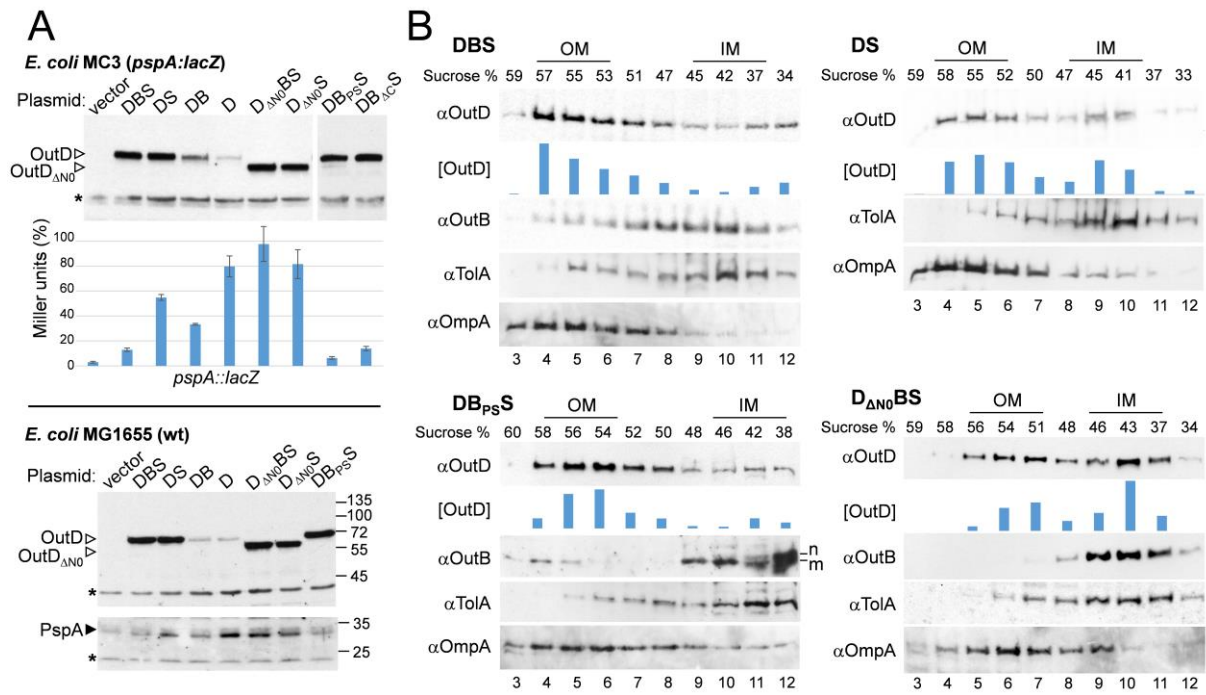
Figure 4. OutB suppresses PspA response caused by OutD and assists OutD targeting in the OM. A. *E. coli* MC3 (upper panels) and MG1655 (lower panels) carrying pREP4 (*lacI^Q*) and pGEMT plasmid with the indicated combinations of *outD, outB* and *outS* (D, B and S, respectively) were grown at 28°C for 14 h. The level of OutD and PspA was probed by immunoblotting or by β-galactosidase activity measurement, for *pspA::lacZ*. The β-galactosidase values are the mean of three cultures; one of which was also probed by immunoblotting with anti-OutD. Asterisk indicates a nonspecific cross-reacting protein, used as a loading control. **B.** *E. coli* MC3 cells grown as in **A** were broken in French press and membranes were separated on sucrose gradient and analyzed by immunoblotting with the indicated antibodies. The positions of IM and OM are indicated according to the location of TolA and OmpA, respectively. The quantity of OutD in the fractions, [D], was estimated with Image Lab software (Bio-Rad). The exposition time of DS panel with OutD antibodies was 10 min versus 3 min with DBS panel, consistent with a lower OutD content. Non-matured and matured versions of OutB$_{PS}$ are indicated with 'n' and 'm'.
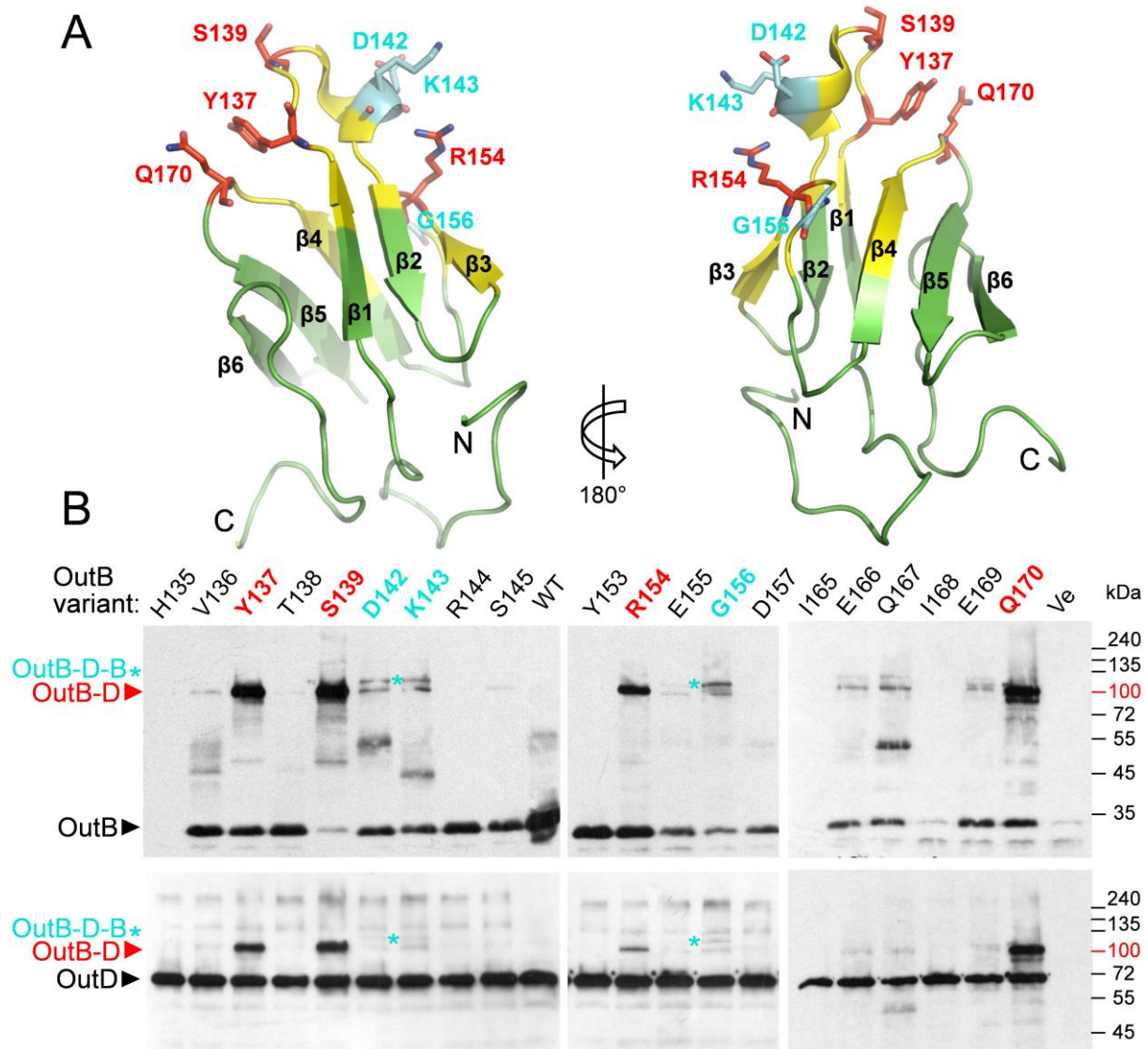
1

2 **Figure 5. Photo-crosslinking maps OutB sites interacting with OutD. A.** OutB HR structure is formed

3 by two three-stranded anti-parallel β-sheets coming together to form a β-sandwich. The residues

4 generating abundant OutB-D complexes are in red, those generating putative OutB-D-B complex are

5 in cyan and other residues probed by photo-crosslinking are in yellow. **B**. *In vivo* site specific photo-

6 crosslinking. OutB *p*BPA substitutions (indicated at the top) were expressed from DBS plasmid in *E.*

7 *coli* MG1655/pREP4/pSupBPA. Cells were irradiated by UV (365 nm) for 3 min and analyzed by

8 immunoblotting with anti-OutB and OutD antibodies, upper and lower panel, respectively. Only UV-

9 irradiated cells extracts are shown. An equivalent amount of cells was loaded into each well. OutB,

10 OutD and their complexes are indicated by an arrow or by an asterisk.

11
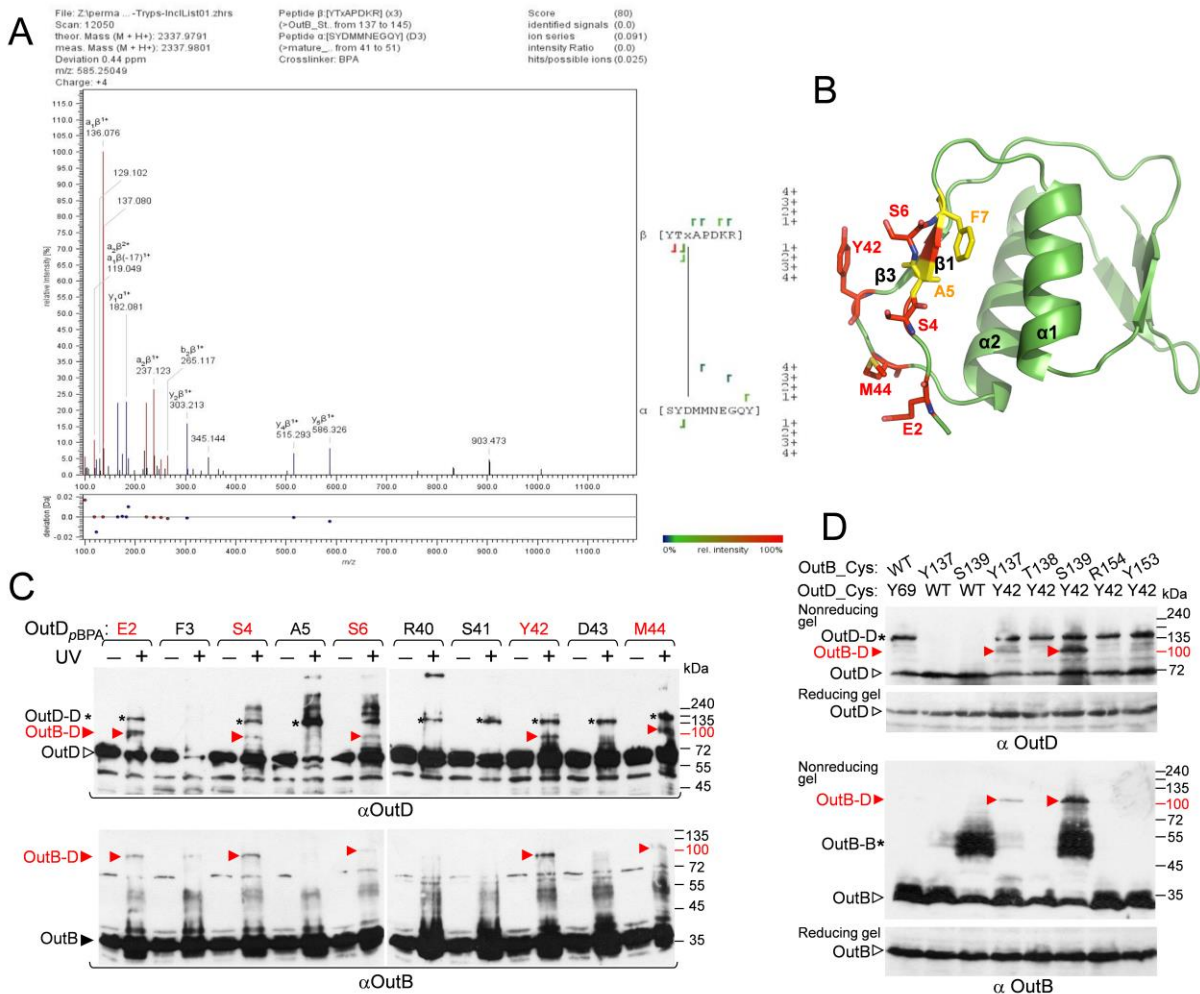
1
2 **Figure 6. Mapping of the OutD site interacting with OutB. A.** LC MS/MS analysis shows that
3 OutB_S139*p*BPA interacts with SYDMMNEGQY peptide of OutD. OutB_S139*p*BPA-OutD complex was
4 purified by Strep-Tactin chromatography followed by SDS-PAGE and in-gel digestion (Fig S9). The
5 peptides were subjected to a high-resolution LC MS/MS analysis and MS spectra were assigned with
6 StavroX. Shown is a representative MS/MS spectrum with the assigned OutD (α) and OutB (β) peptides
7 and crosslinking site, where 'x' indicates the *p*BPA residue. The StavroX validation Score of 80 is above
8 the score cut-off of 48 (Fig S10). **B.** OutD N0 residues probed by photo cross-linking: these generating
9 OutB-D and OutD-D complexes are in red and in yellow, respectively. **C**. *In vivo* photo cross-liking maps
10 the OutD residues interacting with OutB. OutD_*p*BPA substitutions (indicated at the top) were
11 expressed from a DBS plasmid in *E. coli* MG1655/pREP4/pSupBPA. Cells were irradiated by UV and
12 analysed by immunoblotting with OutB and OutD antibodies. **D.** Disulfide bonding analysis maps an
13 OutB-OutD interacting site. *E. coli* MG1655/pREP4 carrying a DBS plasmid with the indicated cysteine
14 substitutions in OutD and/or OutB were grown aerobically to allow formation of disulfide bonds in the
15 periplasm. The remaining free thiol groups were then blocked with iodoacetamide and the extent of
16 disulfide bonding was assessed in nonreducing gel, followed by immunoblotting with OutD and OutB
17 antibodies. The same samples were analyzed in reducing gel with 2-mercaptoethanol to estimate the
18 quantities of OutB and OutD. An equivalent amount of cells was loaded into each well. OutB, OutD and
19 their complexes are indicated by an arrow or by an asterisk.

1
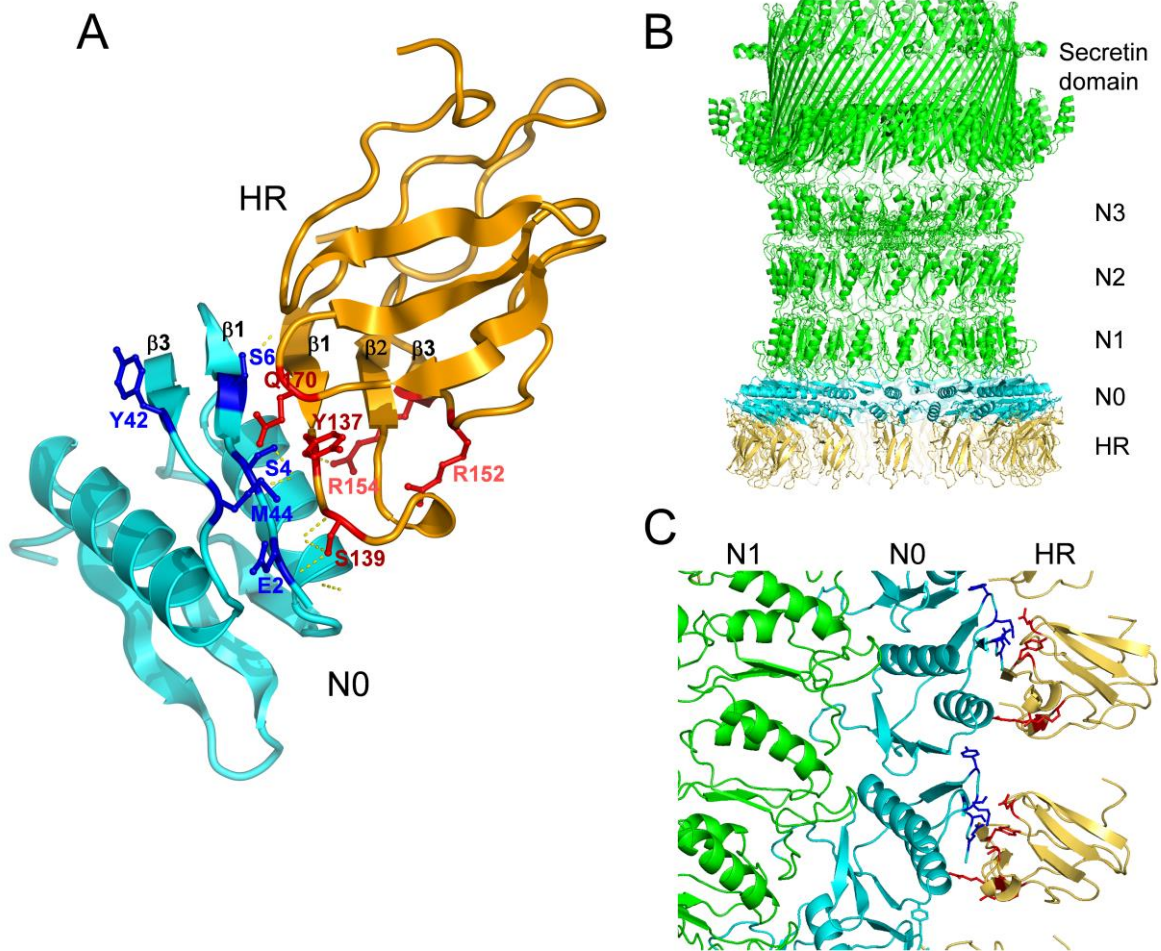
Figure 7. Model of the OutB HR/OutD N0 complex of *D. dadantii*. A. OutB HR/OutD N0 interdomain

interface modelled according to the *in vivo* cross-linking data. **B.** Model of the 15-meric secretin

channel in complex with OutB HR domains, assuming all N0 domains are occupied, front view. **C.**

Close up view of the interacting regions between OutB HR and OutD N0; two adjacent N0/HR

complexes are shown. OutB HR domains are in gold, OutD is in green with N0 domains in cyan.
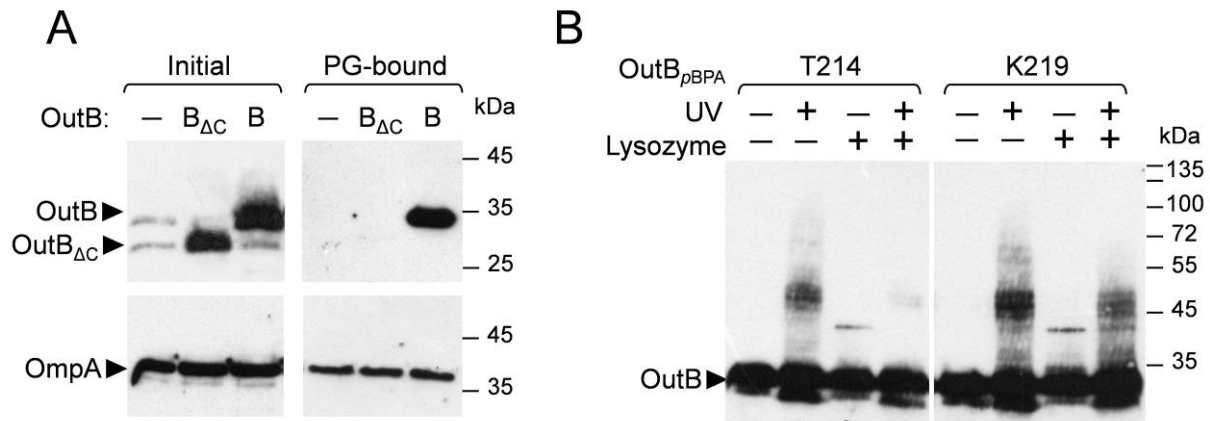
1

2 **Figure 8. OutB interacts with peptidoglycan. A.** *In vivo* chemical crosslinking with DTSSP. *D. dadantii*

3 *outB* A5719 cells carrying either an empty vector, or a plasmid with *outB*, or *outBΔcte* (-, B and B$_{ΔC}$,

4 respectively, in 'Initial') were treated with DTSSP. The PG was next extracted by SDS-boiling

5 procedure and PG-bound proteins were eluted with 2-mercaptoethanol and analyzed by

6 immunoblotting with anti-OutB or anti-OmpA ('PG-bound'). 25-fold less material, in cell equivalent,

7 was loaded in 'Initial' in comparison with 'PG-bound'. **B.** *In vivo* site specific photo-crosslinking. OutB

8 *p*BPA substitutions (T214 and K219) were expressed in *D. dadantii outB* A5719. Cells were irradiated

9 by UV, treated or not with lysozyme and analyzed by immunoblotting with anti-OutB.

10
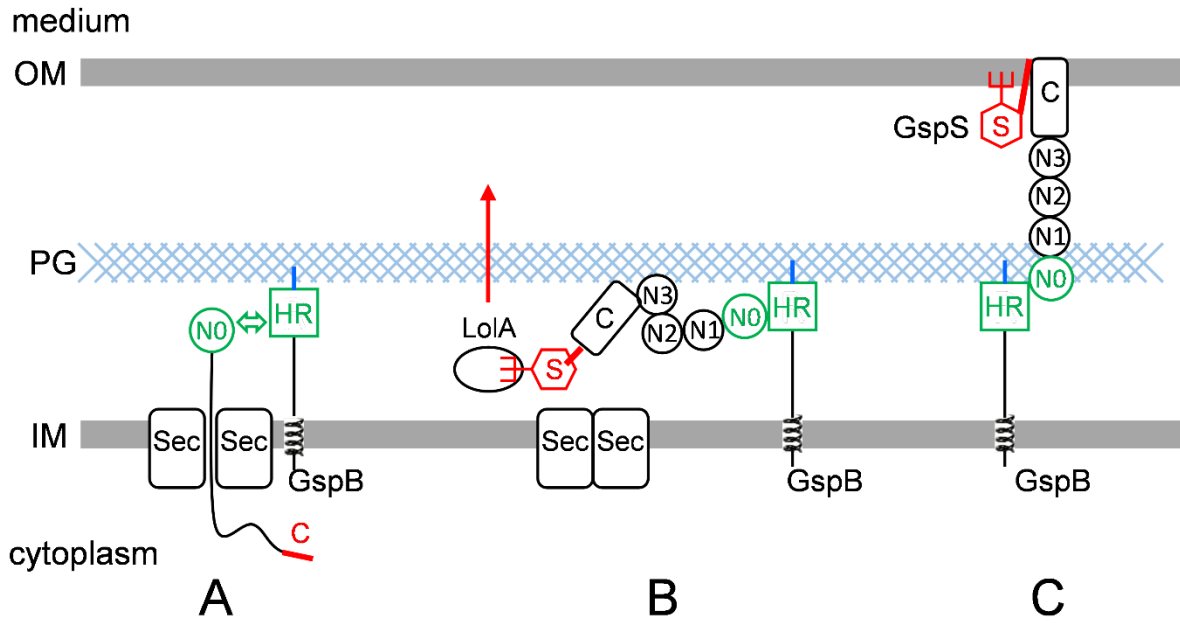
Figure 9. OutB and OutS play complementary functions in targeting, assembly and scaffolding of secretin. **A.** OutB acts early during OutD export by binding the N0 domain, the first part of the secretin releasing into the periplasm. In this way, OutB could orient the secretin and prevent its mis-insertion into the IM. **B.** In the later step, pilotin OutS (red hexagon) binds the C-terminal S-domain (red segment) and, with the assistance of the Lol machinery, targets the C-terminal part of the secretin into the OM where several secretin subunits form a channel. **C.** OutB binds to and stabilizes the N-terminal entry of the secretin channel providing scaffolding with the IM and the PG layer. For simplicity, only one subunit of each protein is shown.