On the path to uncover the bacterial type II secretion system

Badreddine Douzi1,2, Alain Filloux3 and Romé Voulhoux1,*

1Laboratoire d’Ingénierie des Systèmes Macromoléculaires (CNRS-LISM-UPR 9027), Aix-Marseille Universités, Institut de Microbiologie de la Méditerranée, 31 Chemin Joseph Aiguier, 13402 Marseille Cedex 20, France
2Laboratoire Architecture et Fonctions des Macromolécules Biologiques (CNRS-AFMB-UMR 6098), Aix-Marseille Universités, Case 932, 163 avenue de Luminy, 13288 Marseille Cedex 09, France
3Centre for Molecular Microbiology and Infection (CMMI), Division of Cell and Molecular Biology, Imperial College London, London SW7 2AZ, UK

Gram-negative bacteria have evolved several secretory pathways to release enzymes or toxins into the surrounding environment or into the target cells. The type II secretion system (T2SS) is conserved in Gram-negative bacteria and involves a set of 12 to 16 different proteins. Components of the T2SS are located in both the inner and outer membranes where they assemble into a supramolecular complex spanning the bacterial envelope, also called the secreton. The T2SS substrates transiently go through the periplasm before they are translocated across the outer membrane and exposed to the extracellular milieu. The T2SS is unique in its ability to promote secretion of large and sometimes multimeric proteins that are folded in the periplasm. The present review describes recently identified protein–protein interactions together with structural and functional advances in the field that have contributed to improve our understanding on how the type II secretion apparatus assembles and on the role played by individual proteins of this highly sophisticated system.

Keywords: type II secretion system (T2SS); secreton; pseudopilus; secretin; protein–protein interaction

1. INTRODUCTION

Gram-negative bacteria are surrounded by a dual membrane structure establishing an interface between the environment and the interior of the cells. The two membranes are separated by an aequous periplasmic space containing a rigid peptidoglycan layer. This cell envelope constitutes a highly selective barrier for uptake and release of various compounds. Gram-negative bacteria have evolved several highly specialized secretory pathways to release proteins into their surrounding environment. Among them, the type II secretion pathway is a two-step process dedicated to the secretion of folded and/or oligomeric exoproteins. This ability to secrete large molecules is extremely valuable and is achieved by a sophisticated molecular nano-machine embedded in the bacterial envelope called the secreton.

The type II secretion pathway is conserved in Gram-negative bacteria [1] with prevalence in bacterial pathogens of plants (Pseudomonas fluorescens, Erwinia or Xanthomonas species), animals (Aeromonas hydrophila) and humans (Klebsiella oxytoca, Pseudomonas aeruginosa, Vibrio cholerae or Legionella pneumophila) [2–5]. The number of proteins secreted via the T2SS by any given organism is variable and ranges from one, in the case of K. oxytoca [6], to more than ten in P. aeruginosa [7], V. cholerae [8] or L. pneumophila [9]. The functions of these proteins are extremely diverse and include toxins [10,11], surface-associated virulence factors [12,13], cytochromes [14] and a broad range of enzymes that hydrolyse macromolecules such as lipids, polysaccharides and proteins [15].

(a) Genetic organization

Typical type II secretion systems (T2SSs) are encoded by a set of 12 to 16 gsp (general secretion pathway) genes organized into large operons including the conserved ‘core’ genes denoted gspC to OA, or gspP to gspS (figure 1). Because a different nomenclature is used for Pseudomonas and non-Pseudomonas T2SSs, the alternative gene or protein nomenclature is indicated throughout the review. For example, in GspER the ‘R’ refers to the Pseudomonas XcpR T2SS component, which is reciprocally called XcpER. Apart from rare exceptions, mutation in any gsp gene prevents secretion and causes accumulation of the exoproteins in the periplasm. The genetic organization of the T2SS clusters is remarkably conserved. However, in some species, the position of the gspCPDQ genes is peculiar (figure 1). In the P. aeruginosa xcp cluster, these two genes form an operon divergent from the operon containing the gspER-MZ genes. In Xanthomonas

* Author for correspondence (voulhoux@imm.cnrs.fr).

One contribution of 11 to a Theme issue ‘Bacterial protein secretion comes of age’.

DOI: 10.1098/rstb.2011.0204

© 2012 The Royal Society
campestris, the gspCPDQ genes are found after gspM2 at the end of the gsp operon (figure 1). Exceptions are with the T2SS genes in Burkholderia pseudomallei and Legionella pneumophila, where gspCP and gspDQ are not next to each other. Finally, it should be noted that the whole P. aeruginosa hxc (for homologous to xcp) cluster [16] has a radically different organization of its genes. This T2SS is used by P. aeruginosa for the secretion of a single exoprotein, the alkaline phosphatase LapA, and the hxc genes are expressed in phosphate-limited growth conditions [16]. The specificity of the P. aeruginosa Hxc system versus the more general Xcp machinery is probably not linked with growth in phosphate starvation conditions since phosphate-regulated phospholipases (PlcH, –N and –B) [17,18], or alkaline phosphatase, PhoA [19], are all secreted via the Xcp machinery. Interestingly, Durand et al. [20] recently identified specific Hxc phenotypes suggesting the existence of two T2SS subtypes called T2aSS and T2bSS to which, Xcp and Hxc, respectively, belong. Indeed, the authors propose that the secretion process of the Hxc T2bSS of P. aeruginosa involves a pseudopilus whose structure and stability may differ from the one commonly found in Xcp and other known T2aSSs. A
supramolecular complex, called the secreton, made of the different Gsp proteins.

(c) Structural organization of the secreton
Based on data obtained by many different experimental approaches, including subcellular localization, protein–protein interactions between individual components of the T2SS and resolution of protein structure, the current model for the secreton is represented by three functional sub-complexes (figure 2b). An inner membrane platform (IMP) (figure 2b, green) is composed of the GspDQ$_F$, LS, LY and MZ IM proteins; the cytoplasmic traffic ATPase GspE$_K$ is associated with this through an interaction with the bitopic protein GspLY [22–25]. The secreton also contains five proteins that display homologies with the type IV pilin PilA and are designated pseudopilins [26–28]. These proteins have been proposed to be involved in the formation of a fibrillar piston-like structure, the pseudopilus (figure 2b, orange/red) [29–33]. Finally, GspD$_Q$, the OM component of the system, belongs to the secretin family and likely constitutes the channel giving T2SS substrates access to the extracellular medium (figure 2b, blue) [34,35]. Whereas the proton motive force has been shown to be involved in the translocation of T2SS substrates across the OM [36,37], GspE$_K$, which contains motifs characteristic of traffic ATPases, also contributes to energize the T2SS-dependent process [38,39] and could drive the pseudopilus through the GspD$_Q$ channel, pushing out exoproteins to the external medium [15,40,41].

(d) Cellular localization of the secreton
In P. aeruginosa the number of assembled secreton machines is thought to be relatively low and has been estimated at 50–100 secretons per cell [34]. Moreover, while results obtained in K. oxytoca and V. cholerae [42–44] with GFP-fused Gsp proteins indicate a circumferential distribution of the machinery into foci, the P. aeruginosa Xcp secreton was proposed to be polar. This was shown by adding a Lumino tag onto XcpR or XcpS or by the visualization of protease secretion with an intramolecularly quenched casein conjugate [45]. Such discrepant results could be due to artefacts related to the artificial production of the reporters used, as clearly demonstrated by Lybarger et al. [42]. Alternatively, it cannot be ruled out that cellular localization of T2SSs might vary from one species to another. Interestingly, further localization experiments of the secreton, which were performed in various gsp backgrounds, indicate that in contrast to other Gsp proteins, secretin does not need other secreton components for correct localization in the bacterial envelope, thus suggesting an assembly of the secreton from the OM [42]. This relatively new concept of molecular machines assembly from their OM secretins was also recently proposed for the type III secretion machinery [46].

In this review, we will summarize what is currently known about the individual organization of the three secreton sub-complexes briefly outlined in this introduction. We will particularly highlight new findings on solved protein structures and protein–protein interactions among and between the three sub-
secretin PulD mislocalizes to the IM [53], indicating secretin transport since in its absence, from proteolytic degradation. PulS is also involved in ties since it is involved in the protection of the secretin that a lipid-anchored chaperone is required for efficient and correct insertion of the secretin into the OM. To date, genes encoding GspS members have not been found in all T2SSs (figure 1). Therefore, it cannot be ruled out that genes with low homologies, which are not associated with the gsp cluster, could encode proteins playing the same function as GspS [54]. Alternatively, secretin transport is assisted either directly for liposecretin [56] or via their pilotin [58]. Interestingly, it has been shown that transport to the OM of the T2SS secretin PulD is not dependent on the general Bam OM protein transport pathway [57]. It is therefore possible that secretins use the Lol lipoprotein transport route for their transport to the OM either directly for liposecretin [56] or via their pilotin [58]. Alternatively, it cannot be excluded that some secretins use the Bam pathway as was demonstrated for the Neisseria type IV pilus secretin PiQ [59].

Secretin monomers are bipartite proteins ranging from 50 to 70 kDa in size. Homology among members of the secretin family resides in the C-terminal half of the protein that is important for oligomerization, whereas the N terminus is conserved only within subgroups from related transport pathways and thought to be involved in system-specific interactions [60]. Recently, Reichow et al. [61] have solved the low-resolution structure of the V. cholerae full-length GspDQ (EspDQ) secretin using cryo-electron microscopy (figure 3). The cryo-EM reconstitution of the V. cholerae secretin at 19 Å resolution suggests a dodecameric structure reminiscent of a barrel with a large internal channel containing two compartments separated by a closed periplasmic gate, a periplasmic vestibule and an extracellular chamber located in the OM (figure 3).

The N terminus of T2SS secretins comprises four structurally independent domains, N0–N3 (figure 4). The three-dimensional structure of the N0–N1–N2 region of the enterotoxigenic Escherichia coli (ETEC) GspDQ secretin has been solved at 2.8 Å by X-ray crystallography [62]. This structure has been used to generate a 12-fold symmetrical ring, which was fitted into the density map of the full length V. cholerae EpsDQ obtained by cryo-EM [61] (figure 3). The reconstruction confirmed that the C-terminal and N-terminal regions are two structurally separate domains located in the OM and the periplasm, respectively. Since only low-resolution structures have been obtained for the

Figure 3. Electron microscopy structure of T2SS secretin. Cryo-EM reconstitution of V. cholerae T2SS secretin GspDQ at 19 Å resolution (EMDB1763 and adapted from [61] by permission from Nature Publishing Group). The GspDQ cryo-EM density reveals a cylindrical channel assembly 155 Å in diameter and 200 Å in length. In side view, three domains are identified from top to bottom: the extracellular cap, the outer-membrane domain and the periplasmic vestibule domain. In a cutout view, secretin contains an extracellular chamber limited by an extracellular gate and a periplasmic gate. The vestibule domain shows a constriction which results in a narrowing of the channel diameter from 75 to 55 Å. The crystal structure of the N-terminal periplasmic subdomains N0–N1–N2 from ETEC [62] is fitted into the GspDQ periplasmic vestibule (adapted from [47] by permission from Elsevier).
pore-forming part of the secretin, the precise structural folding of this domain remains unknown. Like the majority of the bacterial OM proteins, secretins are predicted to adopt a β-barrel structure [63]. For example, the topology of the XcpQD secretin together with predictions for β-strands in the primary amino acid sequence has previously been assessed [64]. If this is the case, whether this domain is formed by one large single homomultimeric β-barrel or by the assembly of 12 individual β-barrels remains an open question. Alternatively, it is still a likely possibility that secretin pores do not form β-barrels but adopt alternative α-helical folds. Such an original fold was first identified in the *E. coli* capsular polysaccharides OM pore Wza [65], but is also required for insertion of the type IV secretion OM protein VirB10 in *Agrobacterium tumefaciens* [66] or the *P. aeruginosa* PelC protein [67]. The hypothesis of an α-helical fold in T2SS secretins is supported by the observation that transport of this group of proteins does not involve the Bam complex [57] which is required for insertion of OM proteins forming β-barrels [59,68].

A recent cryo-EM study showed that the T2SS-dependent cholera toxin binds in the lower part of the periplasmic vestibule of the *V. cholerae* EpsDQ secretin [69]. This observation is in agreement with the previous interaction found between the T2SS secretin OutDQ and its cognate substrate PelB [41] and confirms the role of the secretin in substrate binding. How the substrate further travels through the pore is still unknown but the closed state of the channel suggests that several conformational changes might occur in the secretin core to accommodate the
substrate and to trigger communication between the different chambers. In addition, and as we shall see in §3–7, other components of the secreton such as the transperiplasmic protein GspCp and the pseudopili are also involved in this process since they also interact with the substrate [70].

3. THE TRANS-PERIPLASMIC PROTEIN GSPCP

GspCp proteins are bitopic IM proteins that are the least conserved of the T2SS components. Proteins of this family are organized into several domains including an N-terminal cytoplasmic region, a transmembrane (TM) domain, a highly conserved central periplasmic domain (homology region, HR) and a C-terminal part containing specific secondary structures such as coiled-coil domains (in P. aeruginosa and Pseudomonas alcaligenes, for instance) or PDZ domains (in K. oxytoca and E. chrysanthemi) (figure 4) [71]. GspCp was shown to be active as a dimer and self-associates by its TM domain, which is not a simple membrane anchor but plays an active role in the function of the protein [72]. The HR domain of V. cholerae GspCp (EpsCp) was shown to interact directly with the periplasmic N0 domain of the secretin EpsDQ [44,73]. This interaction was also seen in E. chrysanthemi where the interaction site of OutCp on OutDQ was localized between residues 139 and 158 of the HR domain. This peptide called OutCp-sip ‘secretin interacting peptide’ recognizes two different sites on the OutDQ secretin. The first one is localized on domain N0, whilst the second one sits astride domains N2 and N3 [74]. Furthermore, mapping of the interaction between XcpFC and XcpQD in P. aeruginosa showed that the N3 domain is essential for this interaction [70]. These different interaction sites between secretins and GspCs suggest that GspCp/DQ partners have evolved various strategies to interact with each other.

Several genetic data indicate that GspCp and GspDQ form a functional couple determining the specificity of the machinery. For example, in the very closely related E. chrysanthemi and E. carotovora Out systems, all genes are individually exchangeable except for outCp and outDQ [75]. Similar investigations performed comparing P. aeruginosa and P. alcaligenes also indicate that XcpFC and XcpQD are determinants of substrate specificity [76]. In order to localize the GspCp domain directly or indirectly involved in substrate specificity, chimeras between E. chrysanthemi OutCp and P. aeruginosa XcpFC domains have been generated and their ability to support secretion in P. aeruginosa tested [71]. Interestingly, XcpFC chimeras containing the TM, HR or C-terminal domain of OutCp remain functional, indicating that none of these domains play a role in substrate specificity. However, the replacement of the intermediary domain between TM and HR called TM/HR (figure 4) leads to secretion defect, suggesting that this domain plays an essential role in specificity and potentially in substrate recognition. Recently, a set of in vitro experiments has revealed direct interactions between purified XcpFC and two substrates of the Xcp T2SS, the elastase (LasB) and the lipase (LipA) [70]. Importantly, no interaction was detected using the substrate of the second P. aeruginosa T2SS (Hxc), i.e. the alkaline phosphatase (LapA). These observations revealed that the species-specificity of the T2SS mechanism is largely contributed by the exoproteins and involves GspCp and GspDQ, which directly interact with cognate substrates.

4. THE ATPASE OF THE SYSTEM: GSPER

A functional T2SS requires the presence of a traffic ATPase, GspEr. Traffic ATPases are involved in several other transport machines such as type IV secretion, conjugation and type IV piliation systems. Structural analysis on the type IV secretion and type IV piliation ATPases indicated that they may function as dynamic hexamers [77,78]. The members of the traffic ATPase superfamily are characterized by two nucleotide-binding motifs designated Walker A and B boxes and also His and Asp boxes (figure 4) [79]. GspEr proteins have a characteristic Walker box A containing the P-loop of an NTP-binding motif, and a less well-defined Walker B box in which the second conserved aspartate residue is replaced by either a glycine or an alanine. Mutation of a conserved glycine residue within the Walker A motif of GspEr from P. aeruginosa, K. oxytoca, E. chrysanthemi or V. cholerae causes the bacteria to be secretion-defective, showing the important role played by this protein in the secretory process [80–83]. Mutations in the less conserved Walker B box have little or no effect on the secretion process [83]. The T2SS traffic ATPase family is distinct from other ATPases in three additional conserved regions: first the aspartate box ‘Asp Box’ between the Walker A and B boxes consisting of two short aspartate-rich motifs (figure 4), which is required for the function of GspEr in the secretion process and may be involved in the formation and stabilization of the nucleotide-binding fold by interacting with Mg$^{2+}$ [83]; second, the His box, including two histidine residues, which is located downstream of the Walker B box, although the role of the His box in GspEr function is still unknown; last, a tetrapeptide (Cys4) motif that appears to be essential for function, since replacement of any of the cysteine residues by a serine within the K. oxytoca GspEr leads to a large decrease in pullulanase secretion (figure 4) [15].

GspEr traffic ATPases lack hydrophobic domains and exhibit the general characteristics of a cytoplasmic protein (figure 4). However, they were found to be associated with the IM through an interaction with the bitopic protein GspLr [80,84]. Results obtained from V. cholerae indicate that ATP hydrolysis by the EpsEr/EpsLr complex is stimulated by acidic phospholipids, whereas the activity of EpsEr alone is unaffected [85]. Further mutagenesis revealed that the membrane-proximal region of the cytoplasmic domain of EpsLr subtly controls the interaction of EpsEr with the cytoplasmic membrane and influences its oligomerization, thereby stimulating its ATPase activity [85]. Other results from X. campesstris have shown that XpsEr oligomerization, as well as its association with XpsLr, requires ATP binding but not ATP hydrolysis, thus indicating that association between XpsEr and XpsLr is needed for ATPase activity [86].

The crystal structure of a truncated V. cholerae EpsEr protein lacking the N-terminal 90 residues was determined with or without the nucleotide bound [87].
These structures reveal a two-domain architecture with the five characteristic motifs of the GspER subfamily clustering around the nucleotide-binding site in the C-terminal domain. The EpsER subunits form a right-handed helical arrangement in the crystal with extensive and conserved contacts between the C and N domains of neighbouring subunits, thus suggesting that EpsER is organized as a hexameric structure. The hexameric state of GspER is confirmed by results obtained in *V. cholerae* and *X. campestris* showing that optimal ATPase activity is obtained with hexameric GspER [38,86]. The crystal structure of the N-terminal part of *V. cholerae* EpsER in complex with the cytoplasmic domain of *V. cholerae* EpsL shows that these two proteins form a heterotetramer in which EpsL forms a central dimer and EpsER binds at the periphery [88].

Amino acid sequence alignments have shown that XpsER of *X. campestris* contains an additional N-terminal extension not found in most other GspERs. This additional domain appears to be essential for XpsL binding, therefore indicating that a more sophisticated interaction process between GspER and GspL might occur within the Xps secretory system of *X. campestris* [89]. To date, the structure of a full-length GspER has not been reported and this would provide key and definite structural information about the architecture of GspER monomer and multimer.

5. **GspF\(\text{S, L}^\text{Y}\) AND M\(\text{Z}\): THE IMP STABILIZERS**

GspF\(\text{S}\) is a polytopic integral membrane protein with a small periplasmic loop and two large cytoplasmic domains connected by three TM regions (figure 4) [90,91]. Two-hybrid studies have shown that the N-terminal domain of the *E. chrysanthemi* OutF\(\text{S}\) protein interacts both with OutER and OutL\(\text{Y}\) [24] suggesting that OutF\(\text{S}\) could participate in the stability of the IMP [92]. Construction of a chimera between *P. aeruginosa* and *P. putida* XcpSF has shown that interaction with other T2SS components is mediated by the cytoplasmic domains [25].

GspL\(\text{Y}\) is a bitopic IM protein organized in three domains, the C-terminal domain localized in the cytoplasmic compartment, the TM domain, and the periplasmic domain (figure 4) [22]. The structures of both cytoplasmic and periplasmic domains of EpsL in *V. cholerae* have been solved at 2.7 and 2.3 A, respectively [93,94]. The cytoplasmic part is composed of subdomains I, II and III and was shown to interact with the N-terminal part of GspER through subdomains II and III [88].

GspM\(\text{Z}\) is a bitopic protein [22] with a short cytoplasmic domain, a TM domain and a periplasmic domain (figure 4) involved in homo-dimerization [24,95,96]. GspM\(\text{Z}\) was shown to be required for GspL\(\text{Y}\) stability since the amount of the former is greatly dependent on the presence of the latter [23]. Studies on GspM\(\text{Z}\) variants in *P. aeruginosa* revealed that three periplasmic domains of the protein were found to be important for interaction with GspL\(\text{Y}\). Two distinct stabilizing domains were localized, respectively, at the beginning and at the end of the periplasmic part of the protein whereas the third one, localized next to the TM domain, also required the presence of the transperiplasmic protein GspC\(\text{p}\) to promote GspL\(\text{Y}\) stabilization [97]. The influence of GspC\(\text{p}\) on the stability of the GspL\(\text{Y}\)/GspM\(\text{Z}\) complex was also observed in *X. campestris* since GspL\(\text{Y}\) dissociates faster from the GspL\(\text{Y}\)/GspM\(\text{Z}\) complex than from the GspC\(\text{p}\)/L\(\text{Y}\)/M\(\text{Z}\) one [98]. In addition, antibodies against GspM\(\text{Z}\) co-immunoprecipitated GspL\(\text{Y}\), GspC\(\text{p}\) and GspER from detergent-solubilized cell extracts, confirming the existence of a complex containing these four proteins [99].

6. **THE PSEUDOPILUS: A CENTRAL STRUCTURE OF THE T2SS MACHINE**

Six of the 12 conserved gsp genes are dedicated to the formation of a periplasmic pilus-like structure called the pseudopilus (figure 1). Five of those genes, *gspF\(\text{T}\)K\(\text{X}\), encode the pseudopilins which are the constitutive elements of the pseudopilus, whereas *gspO\(\text{A}\)* encodes the prepilin peptidase involved in their maturation [26–28,100,101]. Like the closely related type IV pilins involved in type IV pilus formation, the five pseudopilins are synthesized as precursors with a short leader peptide of 6–7 mostly charged residues that is cleaved off by the prepilin peptidase PilD/GspO\(\text{A}\). Mature pilins and pseudopilins are characterized by a highly conserved N-terminal hydrophobic domain of about 20 residues followed by a C-terminal extension specific for each pilin and pseudopilin [15]. Topology studies have shown that pilins and pseudopilins are bitopic IM proteins with a single N-terminal trans-membrane domain segment and a periplasmic C-terminal globular domain (figure 4) [27,101]. Similar to typical IM proteins, pseudopilins use the Sec/SRP pathway for their membrane targeting and insertion [90,102]. While it is an essential step, the significance and role of pseudopilin maturation by the peptidase is unknown. Nevertheless, the removal of cytoplasmic positive charges may facilitate extraction of the protein from the membrane. Interestingly, it was shown that pseudopilins co-fractionate with both IM and OM fractions [28,103], suggesting either a re-localization of these proteins to the OM after processing or more likely the formation of a supramolecular complex.

Among the five pseudopilins, GspGT is the most abundant and is therefore called the major pseudopilin [27], in contrast to GspH\(\text{D}\), *I*, *J*\(\text{B}\) and *K*\(\text{X}\), which are named minor pseudopilins. Biochemical data obtained in *X. campestris* revealed the presence of the major pseudopilin, XpsG\(\text{T}\), within a large complex of about 440 kDa [103]. This observation clearly favoured the formation of a pilus-like structure spanning the bacterial envelope. In agreement with this hypothesis, it was shown that when GspGT is overproduced it is able to assemble into an unusually long fibrillar structure protruding out of the cell, which closely resembles the type IV pilus [32,33,104,105]. Such a structure, also called a hyper pseudopilus (HPP), is only obtained upon overproduction of GspGT pseudopilins and probably represents an uncontrolled elongation of what could be a physiologically relevant pseudopilus. Based on crystallographic and electron microscopy data, an assembly model of GspGT into HPP has been generated by a molecular modelling approach. This pseudo-atomic model was
Three of the four minor pseudopilins (GspIV, JW, inv) involved in their helical oligomerization [29,107,109–111], but their role in pseudopilus formation is undeniable. Results obtained with K. oxytoca or P. aeruginosa pseudopilins have shown that GspIV is essential for HPP formation and could play a major role during the initiation step of fibre formation [33,105]. It was also proposed that HPP elongation is controlled by GspKX [101] since its length varied depending on the number of GspKX subunits produced [104,105]. While none of the four minor pseudopilins have ever been detected in HPP structures [32,33,104], their presence in a native pseudopilus became clear in light of identified protein–protein interactions and resolved three-dimen- sional structures. The four minor pseudopilins share the typical αβ-fold commonly found in major pilins and pseudopilins with a long N-terminal α-helix involved in their helical oligomerization [29,107,109–111]. Three of the four minor pseudopilins (GspJW, KX and KX) have been co-crystallized and shown to form a ternary complex through their globular domains [29]. Recently, a quaternary complex containing the globular domains of the four minor pseudopilins was identified, indicating that the fourth minor pseudopilin, GspHU, also belongs to the minor pseudoplin complex [30]. Structural prediction using the entire amino acid sequence of the minor pseudopilins, i.e. including also the N-terminal hydrophobic domain, suggests a similar helical assembly as the one observed for GspGT. Since extra major pseudopilins could only be added beneath the minor pseudopilins complex [29], it was proposed that the tetrameric complex was localized at the tip of the GspG7 pseudopilus. Given that GspH72 was previously shown to interact with the major pseudopilin GspGT [103,119], GspHU could constitute the hinge between the tip and the core of the pseudopilus [30].

Taking into account the resemblance to the type IV piliation system, the phenotypic, structural and interaction network data lead to a reasonable model for pseudopilus architecture and biogenesis which consists of the prior assembly of the quaternary tip complex in the IM. The tip complex could then be driven to the secretin by the addition of major pseudopilin subunits underneath. Given that GspKX has a large globular domain and is located at the very end of the pseudopilus, its involvement in pseudopilus length control could consist of stopping pseudopilus growth when contacting the secretin. Therefore, when the physiological stoichiometry between GspKX and GspGT is respected, the length of the pseudopilus is likely to be restricted to the periplasmic space. Only when a large excess of GspGT is produced, may the pseudopilus grow without a tip and therefore pass through the secretin channel. This model, which requires experimental validation, is nevertheless supported by the direct interaction observed between minor pseudopilins and secretin [61,96].

The pseudopilus assembly likely occurs as in the type IV piliation process with the energetic assistance of the GspER and PilB traffic ATPases, respectively. If the retraction process exists in type II secretion, it could be associated with a piston-like mechanism of the pseudopilus, but probably involves a different mechanism since there is no counterpart in the T2SS for the PilT ATPase that disassembles type IV pili [120,121]. Interestingly, the minor pseudopilin GspKX was found to interact with GspGT and this interaction triggers a destabilization of GspGT [105]. An alternative retraction process can therefore be proposed for the T2SS pseudopilus, i.e. upon contact with the secretin pore, GspKX acquires, possibly upon conformational changes, the capacity to interact with GspGT thus leading to pseudopilus collapse. An ATPase-free retraction event might thus be sufficient to support the disassembly of a short trans-periplasmic pseudopilus.

7. SUBSTRATE RECOGNITION AND TRANSPORT BY THE T2SS MACHINE

The type II secretion apparatus is widespread in Gram-negative bacteria and a wide variety of enzymes and toxins use this pathway. We have already alluded to the species-specificity of this system; i.e. cognate exo- proteins from one T2SS are not recognized by another machinery [75]. T2SS substrates are loaded on the nanomachine in the periplasm and translocated across the OM in a folded conformation [122–124]. Moreover, studies in E. chrysanthemi, K. oxytoca and P. aeruginosa have demonstrated that disulphide bridges are formed within exoproteins before secretion [125–128]. The high specificity demonstrated for T2SSs and their substrates, as well as their specific recognition in the periplasm among all other resident proteins, suggests the existence on folded substrates of a secretion motif that is required for T2SS recognition. Many studies have been carried out on T2SS-dependent exoenzymes in order to define this secretion motif, which is still a biological puzzle [129]. They all converge to the idea of a conformational signal gathering several motifs spread along the primary amino acid sequence of the protein. With the K. oxytoca pullulanase PulA, it was shown that two non-adjacent regions were together necessary to promote translocation of PulA-β-lactamase hybrid proteins across the OM [130]. Another study suggested that at least three regions of PulA might contain information that influences its secretion [131]. It was also suggested that P. aeruginosa exotoxin A (ToxA) contains two separate secretion signals [132], while alteration of another region also affects secretion efficiency [133]. Finally, the polygalac-turonase PehA of E. carotovora was found to contain three separate domains involved in T2SS targeting [134,135]. As reported above, the secretion signal may be composed of residues from different locations in the linear polypeptide chain, which are brought together into a conformational patch during protein folding [136]. One alternative to a single structural motif is that successive specific interactions lead to the secretion of exoproteins. These interactions may involve different secretion signals that are not essential individually but are required simultaneously, or sequentially, for optimal secretion. Interestingly, secretion of the E. chrysanthemi cellulase Cel5 involves a transitory intramolecular interaction between the cellulase binding domain and

Phil. Trans. R. Soc. B (2012)
a region close to the active site \[137\]. In this case, the exoprotein could adopt a secretion-competent conformation prior to secretion and another conformation once released in the extracellular medium \[138\]. This sequence of events is probably not the case with \textit{P. aeruginosa} and \textit{V. cholerae} T2SSs for which interactions found between substrates and secreton components were detected using purified secreted proteins \[41,61,70\]. Obviously the question of specific substrate recognition is far from being resolved. One may consider that co-evolution of T2SS machines together with bacteria. Based on structural data, protein–protein interactions, and phenotypic observations described in this review, it is possible to propose an innovative model for secreton biogenesis and functioning (figure 5). In this model, secreton biogenesis starts by the insertion of secretin in the OM, thus defining the secretion site \[42\]. In the second step, the transperiplasmic element GspCp binds the secretin, therefore allowing docking of IMP. Indeed, cellular localization experiments performed in \textit{V. cholerae} have shown that, in contrast to secretin, GspCp needs GspDQ but not GspMz which itself needs both GspCp and GspDQ for proper localization \[42\]. We propose that the transitory periplasmic T2SS substrates are first recruited by the peripheral element GspCp and then transferred to the secretin vestibule. The substrate could then be contacted by the pseudopilus tip, and could be pushed and expelled from the cell through the secretin by a growing pseudopilus. This tentative model is in good agreement with the majority of the data collected so far about the T2SSs, but as any working model, it should be challenged and used to design experimental approaches that will confirm or disprove the views presented in this review. For example: (i) Is there a dedicated location for the T2SS in the cell envelope? If yes, how is the first element of the system, say the secretin, targeted there? (ii) Even if we are getting closer to understand substrate recognition by the machinery, the identity of the secretion signal remains enigmatic. (iii) Does the binding of the substrate to XcpPc trigger the assembly of the whole system as previously shown for the type I secretion system \[139\] or does it only trigger pseudopilus elongation? (iv) Does the pseudopilus elongate upon contact of the substrate on GspCp or later when it is positioned in the secretin vestibule? (v) How is the substrate released from the secretin? Is it through a mechanical movement operated by the pseudopilus or following conformational changes within the secretin or both? (vi) Does the pseudopilus effectively

8. A MODEL FOR SECRETON ASSEMBLY AND MODE OF ACTION

Type II protein secretion occurs in two steps. Secreted proteins are first exported across the IM and then released in the extracellular medium thanks to a sophisticated machine, the secreton. The secreton is composed of at least 12 different proteins embedded in the bacterial envelope and organized in a large multi-protein complex capable of secreting a wide range of folded exoproteins across the OM of Gram-negative bacteria. Based on structural data, protein–protein interactions, and phenotypic observations described in this review, it is possible to propose an innovative model for secreton biogenesis and functioning (figure 5). In this model, secretin is the first element to be inserted in the OM, thus defining the secretion site \[42\]. In the second step, the transitory periplasmic element GspCp binds the secretin, therefore allowing docking of IMP. Indeed, cellular localization experiments performed in \textit{V. cholerae} have shown that, in contrast to secretin, GspCp needs GspDQ but not GspMz which itself needs both GspCp and GspDQ for proper localization \[42\]. We propose that the transitory periplasmic T2SS substrates are first recruited by the peripheral element GspCp and then transferred to the secretin vestibule. The substrate could then be contacted by the pseudopilus tip, and could be pushed and expelled from the cell through the secretin by a growing pseudopilus. This tentative model is in good agreement with the majority of the data collected so far about the T2SSs, but as any working model, it should be challenged and used to design experimental approaches that will confirm or disprove the views presented in this review. For example: (i) Is there a dedicated location for the T2SS in the cell envelope? If yes, how is the first element of the system, say the secretin, targeted there? (ii) Even if we are getting closer to understand substrate recognition by the machinery, the identity of the secretion signal remains enigmatic. (iii) Does the binding of the substrate to XcpPc trigger the assembly of the whole system as previously shown for the type I secretion system \[139\] or does it only trigger pseudopilus elongation? (iv) Does the pseudopilus elongate upon contact of the substrate on GspCp or later when it is positioned in the secretin vestibule? (v) How is the substrate released from the secretin? Is it through a mechanical movement operated by the pseudopilus or following conformational changes within the secretin or both? (vi) Does the pseudopilus effectively

---

**Figure 5.** Model of secreton assembly and operation. Schematic of secreton biogenesis as discussed in this review. The first part of the secreton to be assembled is proposed to be the secretin \(1\). The next step could be the successive recruitment of the trans-periplasmic protein GspCp \(2\) and of the inner membrane surface \(3\). The recognition of the substrate by the T2SS takes place in the periplasm and may involve a peripheral element of the secreton, GspCp \(4\). The substrate is then transferred to the secretin vestibule \(5\) in which it could contact the pseudopilus tip complex that is emerging from the inner membrane surface \(6\). The exoprotein could then be released in the extracellular medium through the secretin pore \(7\). The secretin and periplasmic domain of GspCp are shown in blue, the components of the inner membrane surface are shown in green, the pseudopilus and the secreted proteins are shown in orange/red and yellow, respectively.
retract? If yes, what are the molecular mechanism and energy source associated with this event?

Whereas all these questions remained to be addressed, it is remarkable to see the improvement of our understanding of the T2SS over the past few years. This is largely due to the ever increasing performance of structural and biochemical techniques that have generated a lot of new data and come to complement and back up all the original genetic data. Ideally, one would like to see the three-dimensional reconstruction of the whole T2SS machine and even better have this megastucture in motion while transporting the exoprotein. Such level of achievement might not be so far away considering the advances that are currently being made in understanding motion in proteins using molecular dynamics computer simulations [31,140].

We should like to thank M. Tegoni, C. Cambillau, C. Bernard, S. Alphonce, K. T. Forest, L. Franz and D. Dyer for fruitful collaborations, as well as E. Durand, G. Ball, B. Lee, S. Bleves and G. Michel for their contribution to the research undertaken in our laboratory on type II secretion. We also should like to thank M. Tegoni, C. Cambillau, C. Bernard, B. Douzi and E. Durand, G. Ball, B. Lee, S. Bleves and G. Michel for their contribution to the research undertaken in our laboratory on type II secretion. Research on the bacterial type II secretion system in the laboratory of Rémé Voulhoux is supported by the Centre National de la Recherche Scientifique (CNRS) and the ‘3D-Pilus’ young researcher ANR grant (ANR-JC 07-183230). Badreddine Douzi is supported by a BDI-PED PhD grant. Alain Filloux is supported by the Royal Society.

REFERENCES

1 Filloux, A., Bally, M., Ball, G., Aktrim, M., Tommassen, J. & Lazdunski, A. 1990 Protein secretion in Gram-negative bacteria: transport across the outer membrane involves common mechanisms in different bacteria. EMBO J. 9, 4323–4329.


Phil. Trans. R. Soc. B 2012


55 Stroozen, T. G., Stanley, H., Gu, Y., Boyd, J., Bagdasarian, M., Sandkvist, M. & Howard, S. P. 2011 Involvement of the GspAB complex in assembly of the type II secretion system secretin of 

56 Aeromonas and Vibrio species. 


59 J. Biol. Chem. 284, 33 815–33 823. 

60 Collin, S., Guivout, I., Chami, M. & Pugsley, A. P. 2007 YaeT-independent multimermization and outer membrane association of secretin PulD. 


63 J. Biol. Chem. 286, 38 833–38 843. (doi:10.1074/jbc.M111.279851) 

64 Voulhoux, R. & Filloux, A. 2010 The C-terminal amphipathic alpha-helix of 

65 Pseudomonas aeruginosa 

66 secretion machineries of 

67 Erwinia chrysanthemi type II secretion system in 

68 Erwinia chrysanthemi. 

69 J. Biol. Chem. 281, 33 152–33 162. (doi:10.1074/jbc.M060245200) 

70 Korotkov, K. V., Krumm, B., Bagdasarian, M. & Hol, W. G. 2006 Structural and functional studies of EpsC, a crucial component of the type 2 secretion system from 

71 Vibrio cholerae. 


73 Loghin, F. H. & Shevchik, V. E. 2006 The single transmembrane segment drives self-assembly of OutC and the formation of a functional type II secretion system in 

74 Erwinia chrysanthemi. 

75 J. Biol. Chem. 281, 33 152–33 162. (doi:10.1074/jbc.M060245200) 

76 Lindeberg, M., Salmond, G. P. & Collmer, A. 1996 Complementation of deletion mutations in a cloned functional cluster of 

77 Erwinia chrysanthemi out genes with 

78 Erwinia carotovora out homologues reveals OutC and OutD as candidate gatekeepers of species-specific secretion of proteins via the type II pathway. 


80 de Groot, A., Koster, M., Gerard-Vincent, M., Gert- 

81 ritse, G., Ladzunski, A., Tommassen, J. & Filloux, A. 2001 Exchange of Xcp (Gsp) secretion machineries between 

82 Pseudomonas aeruginosa and 

83 Pseudomonas alca- 

84 ligenes: species specificity unrelated to substrate recognition. 


86 Savvides, S. N. et al. 2003 VirB11 ATPases are dynamic hexameric assemblies: new insights into bacterial type IV secretion. 


92 Sandkvist, M., Bagdasarian, M., Howard, S. P. & Dirla, V. J. 1995 Interaction between the autokinase EpsE and EpsL in the cytoplasmic membrane is required for extracellular secretion in 

93 Vibrio cholerae. 

94 EMBO J. 14, 1664–1673. 

95 Turner, L. R., Lara, J. C., Nunn, D. N. & Lory, S. 1993 Mutations in the consensus ATP-binding sites of XcpR and PilO eliminate extracellular protein secretion and pilus biogenesis in 

96 Pseudomonas aeruginosa. 

97 J. Bacteriol. 175, 4962–4969. 

98 Py, B., Loiseau, L. & Barras, F. 1999 Assembly of the type II secretion machinery of 

99 Erwinia chrysanthemi: direct interaction and associated conformational change between OutE, the putative ATP-binding


100 Forest, K. T. & Tainer, J. A. 1997 Type 4 pilus structure: outside to inside and top to bottom—a


