



HAL
open science

The evolution of protein secretion systems by co-option and tinkering of cellular machineries

Rémi Denise, Sophie S. Abby, Eduardo Rocha

► To cite this version:

Rémi Denise, Sophie S. Abby, Eduardo Rocha. The evolution of protein secretion systems by co-option and tinkering of cellular machineries. *Trends in Microbiology*, 2020, 28 (5), pp.372-386. 10.1016/j.tim.2020.01.005 . pasteur-02626815

HAL Id: pasteur-02626815

<https://pasteur.hal.science/pasteur-02626815v1>

Submitted on 26 May 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0 International License

1 The evolution of protein secretion systems by co-option and tinkering
2 of cellular machineries

3 Rémi Denise^{1,2}, Sophie S Abby^{3,*}, Eduardo P C Rocha^{1,*}

4 ¹ Microbial Evolutionary Genomics, Institut Pasteur, CNRS, UMR3525, Paris, 75015, France.

5 ² Sorbonne Université, Collège doctoral, F-75005 Paris, France

6 ³ Univ. Grenoble Alpes, CNRS, Grenoble INP, TIMC-IMAG, 38000 Grenoble, France

7 * corresponding authors

8

9 **Abstract**

10 The evolution of protein secretion systems of Bacteria, and related nanomachines, remains
11 enigmatic. Secretion is important for biotic and abiotic interactions, and secretion systems
12 evolved by co-option of machinery for motility, conjugation, injection, or adhesion. Some
13 secretion systems emerged many times, whereas others are unique. Their evolution occurred
14 by successive rounds of gene accretion, deletion, and horizontal transfer, resulting in
15 machines that can be very different from the original ones. The frequency of co-option
16 depends on the complexity of the systems, their differences to the ancestral machines, the
17 availability of genetic material to tinker with, and possibly on the mechanisms of effector
18 recognition. Understanding the evolution of secretion systems illuminates their functional
19 diversification and could drive the discovery of novel systems.

20

21 Keywords: Molecular evolution; horizontal gene transfer; exaptation; secretion; functional
22 innovation.

23

24 **Glossary**

25 Co-option/Exaptation: use of an existing biological function for a novel adaptive purpose.

26 Diversifying selection: process where natural selection favors the rapid diversification of a
27 gene; frequent in proteins targeted by immune defenses in bacterial pathogens.

28 Effector protein: a secreted protein that has an effect on another cell, *e.g.* a virulence factor
29 that subverts the function of an eukaryotic cell.

30 Functional innovation: change in a function to provide an adaptive response.

31 Monoderms/Diderms: Bacteria lacking/having an outer membrane. Diderms tend to be gram
32 negative and monoderms gram positive, but several exceptions are known.

33 Neo-functionalization: a process where a gene acquires a new function.

34 Protein secretion systems: systems allowing the transfer of proteins across the outer
35 membrane of diderm Bacteria. Systems with homologous components may exist in
36 monoderms, and are sometimes also called protein secretion systems.

37 Sub-functionalization: a process where a gene with several functions specializes in a subset of
38 them.

39 Tinkering: recruitment of a component into a new biological system (pathway, protein
40 complex, regulon).

41

42 Main text

43 The mechanisms and the plausibility of the evolution of complex functions have been one of
44 the most intriguing and contentious points in evolutionary biology [1]. The increasing
45 mechanistic understanding of some of these functions opens the possibility of integrating
46 molecular and evolutionary biology to detail their evolution. Microbes have an increasingly
47 important role in these studies because they can be easily manipulated, evolved, and analyzed
48 in the laboratory. In nature, their large effective population sizes and their ability to exchange
49 genes horizontally render selection processes efficient [2-4]. Furthermore, Bacteria and
50 Archaea emerged in the planet billions of years before Eukaryotes, and were responsible for
51 many structural, biochemical and genetic innovations.

52 The emergence of novel functions requires the modification of existing genes, the acquisition
53 of genes from other genomes or the establishment of novel associations between gene
54 products. This tinkering of genetic structures by evolutionary processes [5] results from
55 mutations, deletions or accretions of genetic material. These may eventually recombine. The
56 resulting variants — in terms of biochemistry, cell localization, genetic regulation — are
57 submitted to natural selection and eventually purged or amplified in the population. In
58 addition, horizontal gene transfer in Bacteria and Archaea has the potential to spread novel
59 functions to distinct species, as currently observed in the evolution of antibiotic-resistant
60 Bacteria [6, 7]. Genes newly arriving at a genome by transfer are sometimes under weak or
61 no selection, many are probably never expressed, providing genetic material that may become
62 the substrate for further functional innovation. For example, genomes contain many mobile
63 genetic elements that can evolve novel functions for the host advantage [8]. In this process, a
64 gene or system that originally evolved to respond to a given adaptive need is co-opted (or
65 exapted) to provide a function that tackles a different need [9]. This article details how the
66 tinkering of existing cellular machineries was at the origin of most, if not all, protein secretion
67 systems of Bacteria.

68

69 Protein secretion systems and related nanomachines

70 Bacteria and Archaea use secreted proteins (effectors and auxiliary proteins) to protect
71 themselves, manipulate their environment, and interact with other individuals [10, 11].
72 Protein secretion systems are defined as machines that transfer proteins across the outer

73 membrane in diderm Bacteria [12], and are called “TXSS” for type “X” secretion system, where
74 X is a number from 1 to 9 (for broad reviews see [13, 14]). Protein secretion systems transfer
75 proteins directly from the cytoplasm (T1SS, T3SS, T4SS, T6SS, T7SS) or from the periplasm
76 (T2SS, T5SS, T8SS, T9SS). Some systems deliver proteins into other cells using specialized pili
77 (T3SS, T4SS, T6SS), whereas the others deliver them in the extracellular space. The complexity
78 of protein secretion systems is variable. They all need to assemble at the cell envelope and
79 transport at least one protein across the outer membrane. But in some cases, they also
80 transport proteins across the cytoplasmic membrane, deliver proteins directly into other cells,
81 and have complex mechanisms to recognize effectors. As a consequence, some systems are
82 very simple and made of one single protein (T5SS autotransporters), whereas others have
83 more than a dozen different components. While we adhere here to the convention that
84 secretion systems only concern diderms [12], it is important to note that some systems, *e.g.*
85 T4SS, have homologous machineries in monoderms (bacteria lacking an outer membrane).
86 Such systems have usually specific adaptations to the differences in the cell envelope and may
87 also secrete proteins (*e.g.* pilins).

88 The comparison of the key components of protein secretion systems, and related machineries,
89 shows an extensive network of homology between components (Figure 1). Some of these
90 components have homologs across several secretion systems. This is most remarkably the
91 case of the large family of AAA+ ATPases present in T1SS, T2SS, T4SS, and T6SS. Members of
92 this family also perform other types of cellular functions, such as chromosome segregation,
93 and they are probably very ancient [15, 16]. Yet, the majority of components has recognizable
94 homologs in only a few other nanomachines. These evolutionary associations result in
95 numerous structural and sequence similarities. For example, the T2SS is evolutionarily related
96 to the super-family of type IV filaments (TFF) [17, 18], the T3SS is homologous to the secretion
97 system of the bacterial flagellum [19, 20], the T4SS is closely related to the conjugative pilus
98 [21, 22], and a key part of the T6SS is strikingly similar to bacteriophage tails [23, 24]. The
99 study of the evolutionary processes underlying these similarities has the power to illuminate
100 many aspects of the structure, assembly, genetics, and taxonomic distribution of secretion
101 systems. It may also provide means of finding novel systems and ways to manipulate secretion
102 [25]. In the following sections, we describe the evolution of four complex protein secretion
103 systems to illustrate the avenues/paths underlying their natural history and thereby reveal
104 some common evolutionary principles.

105

106 The T4SS evolved novel functions by minor tweaks

107 Many biological systems have multiple functions that can evolve in a few steps. An interesting
108 example is provided by the T4SS, a system that originated for DNA exchange by conjugation
109 and gave rise to a protein secretion system that is used by many pathogenic bacteria to secrete
110 virulence factors into Eukaryotic cells [26] (Figure 2b). In conjugation, a nucleoprotein complex
111 composed of a single-stranded DNA molecule and the relaxase is transferred into another cell
112 by the mating pair formation (MPF) apparatus (reviewed in [26-28], see Figure 2b). The MPF
113 of diderms contains a T4SS whose phylogeny can be divided into eight major clades with some
114 different accessory components [22]. Phylogenetic analyses indicate that the T4SS was
115 originally involved in conjugation in diderm Bacteria and only later was transferred to
116 monoderms (Bacteria and Archaea) where two sub-types adapted to the absence of an outer
117 membrane. Two AAA+ ATPases of the T4SS — VirB4 and the coupling protein (T4CP) — arose
118 from an event of gene duplication preceding the emergence of the modern systems. The T4CP
119 specialized in mediating the interaction between the relaxase and the T4SS, whereas VirB4
120 became tightly involved in the function and/or the assembly of T4SS. Other homologous
121 ATPases are found in some sub-types of T4SS, *e.g.* the VirB11 ATPase that was recruited from
122 a TFF [29]. There are alternative mechanisms of DNA transfer by conjugation that are driven
123 by AAA+ ATPases. The protein TraB transports double stranded DNA in Actinobacteria forming
124 mycelia [30] and the protein TdtA has a key role in a recently discovered transformation-
125 dependent conjugation process (transjugation) [31]. Hence, one family of ATPases has
126 diversified early in natural history to provide a very large panel of functions, many of which
127 implicate interaction with membrane components and transport of molecules across the cell
128 envelope.

129 Conjugation is intrinsically a mechanism of protein secretion because the relaxase is
130 recognized by the T4CP and then secreted by the T4SS with the covalently linked DNA [32, 33].
131 This facilitated its co-option into a machine that is specialized in the secretion of proteins. This
132 process was followed by the evolution of the ability to secrete a broad range of effectors (Box
133 1). Recently, a T4SS was shown to secrete toxins into other bacteria broadening the type of
134 hosts into which T4SS can deliver proteins [34]. T4SS thus evolved to transfer mobile genetic
135 elements or protein effectors to both Eukaryotes and Bacteria, and phylogenetic analyses
136 show many independent co-options of T4SS into specialized protein secretion systems [22].

137 The experimentally verified ones have all occurred in just two of the eight T4SS clades (named
138 T and I). We don't know the reasons for this apparent concentration of co-option events in
139 two clades. One possibility is that some conjugative pili are better at interacting with different
140 cell types, as suggested by the observation that broad host range conjugative systems tend to
141 be of type T [35]. Alternatively, the evolution of the ability to recognize different effectors may
142 be simpler in certain types of conjugative systems. While conjugative pilus typically recognize
143 one or a few different relaxases, the T4SS of *Legionella* spp. can deliver numerous different
144 proteins because its T4CP binds adaptors that recruit distinct subsets of effectors [36, 37]. The
145 ability to recognize different effectors, and how these effectors arise, remains one of the least
146 understood aspects of the evolution of protein secretion systems (Box 1).

147 The functional promiscuity of the T4SS, *i.e.* its ability to intrinsically secrete proteins,
148 implicates that the evolution of a protein secretion system by neo-functionalization of an
149 adequate conjugation pilus can take place in a small number of steps: acquisition of a T4CP-
150 interacting domain by the effector (or horizontal transfer of the effector) and loss of the
151 relaxase. Accordingly, T4SS with intermediate properties have been observed, including some
152 systems that are able to secrete the relaxase without attached DNA [38], and systems that are
153 able of both DNA conjugation and secretion of virulence factors [39]. In contrast, the evolution
154 of a T4SS of Campylobacterales into a competence pilus for DNA uptake, first discovered in
155 *Helicobacter pylori*, required many more changes and seems to have taken place only once
156 [40]. In summary, the evolution of T4SS illustrates how a few changes can dramatically affect
157 the function and role of a molecular system.

158

159 Initial specialization of flagellar secretion led to the injectisome

160 The bacterial flagellum includes a secretion system — the flagellar T3SS (F-T3SS) — to export
161 its filament to the outside of the cell. Flagella are complex machines containing the F-T3SS,
162 the motor, the hook and the filament (Figure 2a). While the key function of the flagellum is to
163 allow motility, it can provide many other functions including adhesion, biofilm formation, and
164 interaction with immune systems [41]. Several bacteria also use the flagellum to secrete toxins
165 and other proteins to the extracellular milieu [42-44], suggesting that this trait evolved
166 multiple times in the F-T3SS. *Buchnera* spp. are endo-mutualistic bacteria with very small
167 genomes containing a few flagellar genes. These genes are insufficient to produce a complete
168 flagellum. Instead, they code for a flagellar basal body devoid of most extracellular structures

169 at the cell surface that could function as a secretion system [45]. The extracellular components
170 of the flagellum are not required for secretion and flagellar motility is useless in this obligatory
171 intracellular bacteria, suggesting that *Buchnera* co-opted the F-T3SS for protein secretion. This
172 reduction of the flagellar structure is reminiscent of the mechanism of ejection of the
173 flagellum hook and filament observed in diverse γ -Proteobacteria bacteria to cut the costs on
174 cell motility under nutrient depletion [46]. Hence, simple gene losses could have driven the
175 specialization of the F-T3SS into systems that cannot provide motility to the cell, but are able
176 to secrete proteins in the extracellular space.

177 One particular case of flagellar reduction led to the subsequent evolution of the non-flagellar
178 T3SS or injectisome (NF-T3SS, here shortened to T3SS). The T3SS has many genes homologous
179 to the flagellum, even if it has fewer components (flagella are encoded by ~50 genes, T3SS by
180 less than half) [20, 47, 48] (Figure 2a). The analysis of the components of the T3SS, and of their
181 phylogeny, shows that most core components of the T3SS were directly derived from the
182 ancestral flagellum, including the filament and the F-T3SS. This process was accompanied by
183 the loss of many flagellar genes, which may have led to intermediate systems involved in
184 protein transport to the periplasm or to the extracellular space [19]. A few systems of
185 unknown function in *Myxococcales* spp. may be representative of such intermediate systems.
186 Some crucial gene gains then led to the T3SS, a machine that secretes a plethora of effectors
187 directly into eukaryotic cells.

188 The only recognizably ubiquitous component of the T3SS that lacks a homolog in the flagellum
189 is the pore-forming secretin, which was shown by phylogenetic inference to have been
190 acquired by T3SS at least three times from different systems (see Box 2 and Figure 3). Other
191 key components found in the T3SS and lacking in the flagellum are the devices puncturing the
192 eukaryotic cells (translocon and tip of the needle, Figure 2a). They provide functions relevant
193 for the T3SS, but not for the F-T3SS. The relations of homology between these two
194 components across sub-types of T3SS cannot always be ascertained, because they evolve very
195 fast and the analogous components often lack significant sequence similarity. For the same
196 reason, it's not known if they were co-opted from other cellular machineries. The rapid
197 evolution of these components may result from diversifying selection, since they are in direct
198 contact with the host and are targeted by immune responses [49]. Following the initial crucial
199 steps of gene losses, the gains of secretins and needles led to the modern T3SS. Specific T3SS
200 sub-types are specialized in puncturing either animal, plant or fungal cells [50] (Figure 2a).

201 Intriguingly, and contrary to T4SS and T6SS, we have no knowledge of reports showing the use
202 of T3SS to secrete effectors into other bacterial cells (see Outstanding Questions). T3SS have
203 spread by horizontal gene transfer to many diderm Bacteria. This is a mechanism that affects
204 much less frequently the evolution of flagellar systems [51, 52]. As a result, some Bacteria
205 (*e.g. Burkholderia*) encode multiple T3SS to interact with multiple types of eukaryotic hosts
206 [53]. In summary, the natural history of T3SS shows that nanomachines enduring processes of
207 functional simplification can subsequently acquire novel complex functions.

208

209 TFF evolved different functions from a common set of components

210 The T2SS uses a complex machinery to secrete proteins from the cell periplasm to the
211 extracellular space. It initially evolved from a type IVa pilus (T4aP) presumably involved in
212 twitching motility. Together with other systems of Archaea and Bacteria (both monoderm and
213 diderm), T2SS and T4aP are part of the super-family of type IV filaments (TFF) [18, 54]. TFFs
214 typically include five to six core components: AAA+ ATPase(s), major and minor pilins,
215 cytoplasmic membrane platform(s), a prepilin peptidase and sometimes a secretin (Figure 4).
216 The phylogeny of the ATPase family places the root of the tree between the bacterial and the
217 archaeal groups of TFF, suggesting that the system pre-dated the last common ancestor of all
218 cellular life forms [18]. Most other key components have recognizable homologs across TFF,
219 suggesting that they were already present in the ancestor of all TFF. Unfortunately, they rarely
220 have homologs in non-TFF systems, implicating that only the ATPases can be used to trace the
221 very initial stages of TFF evolution. The ancestral system diversified by a succession of
222 mutations, gene deletions, duplications, fissions and fusions that led to TFFs with very
223 different functions. They include adhesion (many TFFs), protein secretion (T2SS), twitching
224 motility (T4aP and T4bP), flagellar motility (Archaeellum, unrelated with the F-T3SS) and DNA
225 uptake (Com, T4aP) [40, 55-58] (Figure 4). For example, T2SS have evolved pseudo-pili
226 producing shorter structures than the pili of the ancestral T4aP from which they were derived.
227 These specialized components drive the transport of the effectors across the outer membrane
228 [59].

229 The phylogeny of some TFFs suggests complex evolutionary scenarios. Soon after the
230 divergence of the ancestral TFF into an archaeal and a bacterial branch, the ATPase of the
231 bacterial systems was duplicated into a pair of proteins specialized in pilus extension (PilB)
232 and retraction (PilT). This event was thought to have endowed the T4aP with the ability to

233 provide twitching motility by rounds of pilus extension and forceful contraction. Interestingly,
234 the single homologous ATPase of the tight adherence (Tad) pilus systems can extend and
235 contract the pilus [29], and PilT mutants of T4aP can weakly retract the pilus [60]. Hence, the
236 PilB/PilT duplication seems to have allowed a specialization of two ancestral functions (sub-
237 functionalization) where PilT evolved to provide the high-force retraction required for
238 twitching motility. In some TFF there was yet another amplification the ATPase family giving
239 rise to PilU (Figure 1). Its function is to coordinate the activity of retraction with PilT [61 , 62].
240 When the T2SS evolved from a T4aP, the retraction function was unnecessary and the
241 corresponding specialized ATPase was lost.

242 The majority of TFFs specialized in the secretion of protein effectors (T2SS-like) are
243 monophyletic and present across Proteobacteria [63]. But several recently uncovered T2SS-
244 like systems may have evolved in different ways. The phylogenies of the components of the
245 Chlamydiales T2SS-like systems show that they were recruited from diverse TFFs [18]. The
246 T2SS-like systems from Bacteroidetes, *e.g.* those of *Cytophaga*, constitute completely
247 independent co-options of (different) T4aP [18]. This suggests that a specialized protein
248 secretion system may require relatively few steps to evolve from the T4aP, possibly because
249 the latter already secretes a specific family of proteins (pilins). Once the different T2SS-like
250 systems evolved into specialized secretion systems, they became able to secrete many
251 different types of proteins [64, 65] (Box 1). This versatility may have contributed to the
252 evolutionary success of T2SS.

253 The spread of TFFs among Bacteria and Archaea was promoted by horizontal transfers. For
254 example, the Tad pilus originated from an archaeal pilus that was transferred to diderm
255 Bacteria, where it recruited a secretin, and was then transferred to other phyla [18] (see Box
256 2). There is a close interplay between horizontal transfer and the genetic organization of
257 systems because transfer is facilitated when the systems are encoded in a single locus. Such
258 systems, which among TFF include most T2SS and Tad, are more frequently transferred
259 horizontally than those that tend to be encoded in multiple loci [18]. This is presumably
260 because one single event of transfer is enough to provide a novel function to the recipient
261 when all the system is encoded in a single locus [66]. Horizontal transfer spreads these novel
262 functions to other cells where they are eventually fixed when they provide a competitive
263 advantage. Importantly, by transferring systems in new genetic backgrounds this process may

264 drive further innovation to allow the accommodation of the function to its novel cell. In this
265 way, a relatively small set of components can diverge to produce very diverse functions.

266

267 Co-option of phages led to the radical invention of T6SS

268 The T6SS delivers effectors directly into eukaryotic or bacterial cells and is implicated in inter-
269 specific competition, virulence, resource scavenging, and genetic exchanges [67-69]. It is
270 composed of a baseplate-like platform bound to the membrane, which anchors a contractile
271 sheathed tube decorated by a puncturing device that allows toxins to penetrate the target
272 cells (Figure 2c). The structure of many of these components resemble bacteriophage proteins
273 and the overall T6SS resembles an inverted phage tail [23, 24, 70-72], which suggests that a
274 large part of the T6SS derived from a phage. Unfortunately, sequence similarities between
275 phage and T6SS homologs are very low and preclude the study of the initial processes of co-
276 option using standard phylogenetic approaches. The similarities between T6SS and phages
277 extend to the assembly/disassembly dynamics of the respective structures. In particular, T6SS
278 sheaths are contractile in order to project the tube towards target cells, in a similar fashion
279 that phage tails contract to deliver DNA into target cells [73]. This activity is dependent on the
280 ClpV ATPase [73], which is an AAA+ ATPase homologous to those of T2SS and T4SS, but lacking
281 close homologs in phages (see Figure 1). The T6SS also has two components homologous to
282 the type IVb pilus (T4bP) of the TFF super-family, which led to its initial misnaming as a T4bP
283 [74]. This suggests that the co-option of a phage tail was subsequently integrated with
284 components from other cellular machines to produce extant T6SS.

285 Some genomes encode many different T6SS. For example, *Burkholderia thailandensis* have
286 five different systems, one of which is specialized in the interaction with the host and another
287 with competing Bacteria [75]. As some structural components of the T6SS are also effectors
288 (Box 1), different systems may be associated with different effectors. In spite of the
289 multiplicity of T6SS in certain genomes, the four known major variants of T6SS, called T6SSⁱ to
290 T6SS^{iv}, tend to be present in different types of Bacteria. Variants of the most widespread T6SS
291 (T6SSⁱ) were initially found in *Francisella* (T6SSⁱⁱ) and in Bacteroidetes (T6SSⁱⁱⁱ) [76-78]. In spite
292 of having specific components, these systems include most of the T6SSⁱ core components and
293 are presumed to have derived from the same ancestor. More remarkably, the T6SS^{iv}
294 resembles the other variants in terms of the structure of the puncturing device, but seems to
295 lack components not associated to phages, such as the ATPase or the trans-envelope complex

296 [79]. The phylogenetic analyses of sheath protein sequences suggest that it could have
297 emerged from phages independently from other T6SS (Figure 2c) [79]. These results suggest
298 that even such complex co-option events may occur several times in parallel. This might seem
299 surprising, but one should take into consideration that the accumulation of machineries in the
300 genome that can be used for such complex evolutionary processes may increase the frequency
301 at which they occur. Half of the bacteria are lysogens and many of these encode several
302 prophages [80], and these may provide abundant genetic material for evolutionary tinkering
303 leading to the evolution of puncturing devices.

304

305 **Concluding Remarks and Perspectives**

306 The evolution of secretion systems illuminates some general properties of the processes of
307 co-option. The evolution of T4SS exemplifies how a small number of changes suffices to evolve
308 a protein secretion system when there is functional promiscuity in the original system (which
309 is inherently capable of secreting a protein attached to the DNA). The simplicity of this process
310 explains why it evolved many times independently in natural history [22]. Analogously, the
311 evolution of the T2SS was facilitated by the fact that most TFF secrete to the extracellular
312 space the proteins composing their filaments. In both cases, the evolutionary success of the
313 innovations may be due to the ability of the derived secretion systems to secrete a large panel
314 of effectors. Unfortunately, the evolution of mechanisms of effector recognition remain
315 poorly known (Box 1) and this hampers our ability to fully understand these processes.

316

317 The evolution of the two other systems was much more complex. The loss of genes of the F-
318 T3SS results in systems that may be involved in protein transport. This initial sub-
319 functionalization was followed by the accretion of novel components that led to the T3SS.
320 Some steps of this process occurred multiple times: the loss of the motility component of
321 flagella, the subsequent acquisition of secretins (Box 2), and possibly the acquisition of
322 translocons [19]. Since all known T3SS are monophyletic, the complete path from the
323 flagellum to the T3SS occurred only once. This suggests that the frequency of the processes of
324 co-option depends on the complexity of the evolutionary path. Similarly, the T6SS endured a
325 complex evolutionary process that one might think unique in natural history. The T6SS^{iv}, which
326 may have emerged independently from the other T6SS, seems to lack some of the functions
327 of the other systems [79], and could result from the initial steps of phage tail co-option that

328 were extended in the case of the other T6SS. If correct, both T3SS and T6SS evolved by an
329 initial process of sub-functionalization (co-opting basal flagellar basal bodies in one case and
330 phage tails in the other) that produced intermediary structures. These were later integrated
331 with other cellular components to produce a complex novel secretion system. One could
332 regard such processes as rare, but the frequency of co-option events is certainly also shaped
333 by the availability of machineries that can be tinkered by natural selection for functional
334 innovation. Both flagellar loci and prophages are frequent in bacterial genomes.

335 The existence of analogous functions across secretion systems and other cellular machineries,
336 sometimes performed by homologous proteins (Figure 1), may facilitate the combination of
337 components of different origins into novel systems. For example, the T4SS of type I, also called
338 T4SSb, includes a T4bP that is essential for conjugative transfer of plasmid R64 in liquid media
339 and for adherence to host cells by the T4SS of *Legionella* [81]. A combination of a T4SS and a
340 TFF secretin is used by some *Haemophilus* spp. to export DNA to the environment [82]. Finally,
341 the T1SS is a combination of two systems that pre-existed autonomously: an outer membrane
342 porin and an ABC transporter [83]. The frequency with which components were combined to
343 produce novel systems during natural history suggests that novel systems may be uncovered
344 by looking at atypical combinations of known components. Furthermore, it opens the
345 possibility that humans could engineer secretion systems with novel features by recruiting
346 components of diverse membrane-associated machines.

347 This review focused on the most-studied secretion systems, but many relations of homology
348 between protein secretion systems and other machines have been reported. It was recently
349 proposed that some core components of the flagellum and T3SS are a functional unit capable
350 of forming nanotubes between cells for the transport of nutrients [84, 85]. The ESX inner
351 membrane transport system in monoderms has components homologous to the T4SS and to
352 the T7SS of *Mycobacteria* [86, 87], the autotransporters (T5SS) are structurally homologous
353 to the porins of the novel FAP system that exports amyloid subunits in *Pseudomonas* [88], and
354 there are intriguing homologies between components of several secretion systems and a
355 system putatively involved in the transport of proteins between mother and forespore cells in
356 Firmicutes [89]. Finally, a TonB-dependent transporter (TBDT) of *Myxococcus xanthus*, a
357 widespread family of systems usually involved in protein and nutrient import, was recently
358 shown to be involved in the two-step protein secretion of a protein [90]. Further work is
359 needed to understand the evolution of some other secretion systems. Notably, the T9SS is

360 present in many bacteria of the Fibrobacteres-Chlorobi-Bacteroidetes super-phylum where it
361 is involved in both gliding motility and protein secretion [91, 92], but its evolutionary
362 associations with other cellular components are unknown. The T7SS is present in
363 Mycobacteria and is regarded as a *bona fide* secretion system because these bacteria have an
364 outer membrane (that differs in composition from those of Proteobacteria) [87]. The T7SS is
365 partly homologous to systems involved in protein transport across the cytoplasmic membrane
366 of Firmicutes [86, 87]. Interestingly, these two systems contain AAA+ ATPases like those of
367 T4SS and T2SS, suggesting ancient evolutionary associations that need to be unravelled.
368 Based on past experience, there are probably protein secretion systems yet to be discovered,
369 especially in the numerous clades that are being uncovered by environmental metagenomics.
370 The evolutionary principles described in this article suggest that many of these novel systems
371 will include components homologous to those of known membrane-associated machines. This
372 informed hypothesis can be leveraged to identify the novel systems. Such endeavour can now
373 also rely on abundant genomic data to identify components homologous to those of
374 membrane-associated machineries, metagenomics data to investigate conditions and
375 environments where they are expressed, structural and microbial cell biology to understand
376 their function, and evolutionary biology to integrate this information. Altogether, these
377 elements will likely accelerate the pace of discovery of novel protein secretion systems (see
378 Outstanding Questions). In turn, these may illuminate some yet obscure aspects of the very
379 ancient history of complex machineries at the origin of known secretion systems, such as the
380 flagellum or the conjugative system.

381

382 **Box 1: Where do all effectors come from?**

383 At a given moment in evolution, protein secretion systems acquired the key ability to
384 recognize novel effectors. This is particularly remarkable for systems able to secrete many
385 different effectors, such as T2SS, T3SS, T4SS, T6SS, or T9SS, many of which evolved from
386 machines secreting only a few different proteins (pilins, flagellins, relaxases, etc). Systems
387 capable of secreting many different effectors may be more economical than systems secreting
388 a single one, because the production of one single machine is sufficient to secrete a large panel
389 of proteins. Furthermore, some sets of effectors must interact to be efficient, which means
390 they must be secreted at the same time. For example, T3SS effectors require co-secretion of
391 specific protein translocases to traverse the membrane of eukaryotic cells [93]. Selection for
392 the ability to secrete many different effectors in a single cell may have been a driver of the
393 evolution of the complex machineries of T2SS, T3SS, T4SS, and T6SS [48]. On the other hand,
394 systems with fewer components, but usually associated with one or a few effectors like the
395 T1SS and the T5SS, can be horizontally transferred with their effectors across bacteria.

396 Secretion systems discriminate effectors using mechanisms that provide some clues on how
397 effector recognition evolved. Some T6SS effectors are structural components of the system
398 [94], and some T4SS effectors have domains that are structurally similar to the recognition
399 domain of the relaxase [95]. In both cases, genetic fusions of proteins with domains
400 recognized by the secretion system can result in novel effectors. Such gene fusions are
401 frequent in secreted proteins, as revealed by the combinatorial variation of domains of
402 secreted polymorphic toxins [96].

403 The evolution of other mechanisms of effector recognition by secretion systems is unclear,
404 since the mechanisms of effector recognition are themselves poorly understood. Interestingly,
405 some effectors can evolve to be recognized by multiple systems. For example, a toxin was
406 recently shown to be secreted by both T2SS and T3SS in *Vibrio* [97]. It is unclear if this involved
407 adaptation of the secretion system to the effector or the other way around. This raises the
408 complementary question of how effectors evolve to be recognized by the secretion systems.
409 Some effectors might evolve by co-option, just like their cognate secretion systems. For
410 example, a large fraction of the repertoire of *Legionella* T4SS effectors may have been co-
411 opted from proteins of Eukaryotes [36], and *Burkholderia* deploy a T3SS anti-fungal protein
412 that may have been co-opted from a prophage tail-like protein [98]. Proteins that evolved to
413 become secreted by a specialized system can subsequently endure processes of gene

414 duplication and transfer to diversify into novel functions. Actually, genes encoding secreted
415 proteins are very often on mobile genetic elements, such as plasmids [10] and temperate
416 phages [99].
417

418 **BOX 2. Sharing and recruiting components**

419 The transfer of a system to another bacterium may require the accretion of novel functions to
420 fit in the novel genetic background. When the pertinent functions are encoded in the new
421 host genome, they may be initially shared with the other systems. An interesting example is
422 provided by the T1SS, a system composed of an ABC transporter, a porin, and a fusion protein
423 connecting the two. Several T1SS of *E. coli* share the porin function encoded by a single gene
424 (*tolC*), which is also involved in the transport of small molecules [100]. Hence, a single gene,
425 distant from the other genes of the T1SS, provides multiple functions to the cell. Similar gene
426 sharing is observed in TFF, where some prepilin peptidases contribute to the assembly of both
427 T4P and T2SS [101]. However, a component that is shared by several systems accumulates
428 functional, structural and regulatory constraints [102]. This may lead to subsequent gene
429 duplication and sub-functionalization events that will eventually result in the presence of
430 multiple homologous genes in the genome (one per system).

431 The secretin has been recruited by many systems independently. At the initial stages of this
432 process there was probably a single gene, which implicates that the protein may have been
433 shared between systems. The secretin was co-opted by the ancestor of the Tad where it
434 allowed the pilus to cross the outer membrane after its transfer from Archaea to diderm
435 Bacteria. Even more strikingly, the co-option of the secretin from the TFF took place three
436 times independently during the evolution of the T3SS, one of which from a Tad pilus that had
437 obtained it previously from a T4aP (Figure 3). The secretin was also co-opted from the T2SS
438 by some filamentous phages that use this porin to secrete virions from living cells (Figure 3).
439 The grafting of secretins into secretion systems requires a remarkable structural flexibility,
440 especially if this takes place while the same protein remains a component of another
441 machinery. Secretins of the T2SS assemble independently of the rest of the system, and the
442 substitution of a single amino-acid in some T4P secretins was shown to make them capable of
443 self-assembly [103]. Yet this structure is not stable in the absence of the inner-membrane
444 components with which its N-terminus domains interact [104, 105]. It is tempting to speculate
445 that the ability of secretins to assemble autonomously from the rest of the system for a short
446 period of time has facilitated the initial recruitment of the secretin by many different systems.

447

448

449 [Acknowledgements](#)

450 We thank the many people with whom we have collaborated and discussed on the topic of
451 protein secretion systems and related machineries along the years, notably Fernando de la
452 Cruz, Julien Guglielmini, Bertrand Néron, Olivera Francetic and Elie Dassa. We are grateful to
453 Claude Parsot, Laura Gomez-Valero, and two anonymous reviewers for comments and
454 suggestions on an earlier version of this manuscript.

455

456

457 **Figure legends**

458

459 **Figure 1 – Pairwise HMM profile alignments between all the proteins of the TXSS-related**
460 **systems.** The HMM profiles were obtained from TXSScan [18, 63]. The color of nodes
461 represents systems in which the proteins were found. To establish relationships of homology
462 between the components of the different systems, *i.e.* to draw edges between nodes, we
463 made pairwise alignments of their HMM protein profiles using HHSearch v3.0.3 (p-value
464 threshold of 0.001). Groups of proteins that gathered more than two components from at
465 least two systems are displayed. The function attributed to each group is written in black in
466 its background. Given the current difficulty in precisely delineating the functions of the TFF of
467 Archaea, they were grouped under “Archaeal-T4P”.

468

469 **Figure 2 – The evolution of protein secretion systems delivering effectors directly into other**
470 **cells.** (a) The diversification of the T3SS from the bacterial flagellum involved initially the loss
471 of flagellum-specific genes and the motility function. The subsequent multiple acquisitions of
472 pore-forming secretins, translocons, and a few other genes led to the extant T3SS. On the
473 right, the rooted cladogram represents the history of the T3SS [19]. The *Myxococcales* system
474 is not a genuine T3SS since, to the best of our knowledge, it lacks an outer membrane porin.
475 (b) Conjugative apparatuses, involved in ssDNA conjugation, were co-opted multiple times
476 independently into T4SS secreting proteins into other cells. On the right, the rooted cladogram
477 represents the evolution of T4SS [22]. MPF stands for mating pair formation and includes the
478 T4SS. (c) The T6SS resulted from the co-option of contractile tail phage genes (and their
479 integration with other genes). This resulted in a contractile structure able to puncture
480 eukaryotic or bacterial cells and deliver effectors, often toxins. On the right is presented a non-
481 rooted cladogram of the history of T6SS [79]. MAC stands for metamorphosis- inducing
482 structures. Afps stands for insecticidal anti-feeding prophages. The drawings of the systems
483 are based on [48]. OM stands for outer-membrane, IM for inner (cytoplasmic) membrane; the
484 periplasm is shown in brown between the IM and OM; “Enterobacteria” stands for enterobacteria.

485

486

487 **Figure 3 – Phylogeny of secretin proteins.** The tree was built using the secretin domain of the
488 protein sequences from [18], with the addition of T3SS and phage sequences from [19]. We
489 aligned the sequences using MAFFT v7.273 (einsi algorithm) [106], selected informative sites
490 in the multiple alignment using Noisy v1.5.12 [107] (default parameters), and inferred the
491 maximum-likelihood tree from these alignments with IQ-TREE v1.6.7.2 [108] (using the best
492 evolutionary model, options -MF, BIC criterion, -allnri, -ntop 1000, -nm 10000). Node supports
493 displayed at nodes were estimated using the option -bb 1000 for ultrafast bootstraps [109].
494 The tree is consistent with the results of several previous studies. The root was positioned
495 between T4bP and the remaining clades, as suggested elsewhere [18, 19, 103]. If correct,
496 secretins were first components of the TFF superfamily and subsequently co-opted by phages
497 and T3SS (three times independently from Tad for Rhizobiales, from T2SS for Chlamydiae, and
498 from T4aP for other proteobacterial T3SS). They were also recruited by the Tad upon the
499 transfer of the ancestor of this system from Archaea. The color of circles at the tip of the tree
500 corresponds to different bacterial phyla. The colors of groups and drawings on the right depict
501 the different systems where secretins have been identified.

502

503 **Figure 4 & KEY FIGURE – Diversification of the type IV filament (TFF) superfamily around a**
504 **common set of components.** The TFF superfamily diversified into many different systems
505 using a few homologous core components (in the middle) and integrating some new ones (in
506 circles, described in the bottom legend). One should note that some of the acquired
507 components may have been already present in the ancestral system but evolved so fast that
508 one can't trace homology. The different machines were able to diversify into functions as
509 different as secretion of toxins, DNA uptake, motility, adhesion to surface. Components
510 colored in the same way correspond to homologs. The drawings of the systems are based on
511 [110].

512 **References**

- 513 1. Pál, C. and Papp, B. (2017) Evolution of complex adaptations in molecular systems. *Nature*
514 *Ecol Evol* 1, 1084.
- 515 2. Ochman, H. et al. (2000) Lateral gene transfer and the nature of bacterial innovation.
516 *Nature* 405, 299-304.
- 517 3. Lenski, R.E. (2017) Experimental evolution and the dynamics of adaptation and genome
518 evolution in microbial populations. *ISME J* 11, 2181-2194.
- 519 4. Bobay, L.-M. and Ochman, H. (2018) Factors driving effective population size and pan-
520 genome evolution in bacteria. *BMC evolutionary biology* 18, 153.
- 521 5. Jacob, F. (1977) Evolution and tinkering. *Science* 196, 1161-1166.
- 522 6. Martínez, J.L. (2008) Antibiotics and antibiotic resistance genes in natural environments.
523 *Science* 321, 365-367.
- 524 7. Wiedenbeck, J. and Cohan, F.M. (2011) Origins of bacterial diversity through horizontal
525 genetic transfer and adaptation to new ecological niches. *FEMS Microbiol Rev* 35, 957-976.
- 526 8. Touchon, M. et al. (2014) The chromosomal accommodation and domestication of mobile
527 genetic elements. *Curr Opin Microbiol* 22, 22-29.
- 528 9. Gould, S.J. and Vrba, E.S. (1982) Exaptation-A Missing Term in the Science of Form.
529 *Paleobiology* 8, 4-15.
- 530 10. Nogueira, T. et al. (2009) Horizontal Gene Transfer of the Secretome Drives the Evolution
531 of Bacterial Cooperation and Virulence. *Curr Biol* 19, 1683-91.
- 532 11. Granato, E.T. et al. (2019) The evolution and ecology of bacterial warfare. *Curr Biol* 29,
533 R521-R537.
- 534 12. Desvaux, M. et al. (2009) Secretion and subcellular localizations of bacterial proteins: a
535 semantic awareness issue. *Trends Microbiol* 17, 139-45.
- 536 13. Dalbey, R.E. and Kuhn, A. (2012) Protein traffic in Gram-negative bacteria--how exported
537 and secreted proteins find their way. *FEMS Microbiol Rev* 36, 1023-45.
- 538 14. Costa, T.R. et al. (2015) Secretion systems in Gram-negative bacteria: structural and
539 mechanistic insights. *Nature Rev Microbiol* 13, 343-359.
- 540 15. Planet, P.J. et al. (2001) Phylogeny of genes for secretion NTPases: identification of the
541 widespread *tadA* subfamily and development of a diagnostic key for gene classification. *Proc*
542 *Natl Acad Sci U S A* 98, 2503-8.
- 543 16. Iyer, L.M. et al. (2004) Comparative genomics of the FtsK-HerA superfamily of pumping
544 ATPases: implications for the origins of chromosome segregation, cell division and viral
545 capsid packaging. *Nucleic Acids Res* 32, 5260-79.
- 546 17. Whitchurch, C.B. et al. (1991) Characterisation of a *Pseudomonas aeruginosa* twitching
547 motility gene and evidence for a specialised protein export system widespread in eubacteria.
548 *Gene* 101, 33-44.
- 549 18. Denise, R. et al. (2019) Diversification of the type IV filament superfamily into machines
550 for adhesion, protein secretion, DNA uptake, and motility. *PLoS Biol* 17, e3000390.
- 551 19. Abby, S.S. and Rocha, E.P. (2012) The non-flagellar type III secretion system evolved from
552 the bacterial flagellum and diversified into host-cell adapted systems. *PLoS Genet* 8,
553 e1002983.
- 554 20. Ginocchio, C.C. et al. (1994) Contact with epithelial cells induces the formation of surface
555 appendages on *Salmonella typhimurium*. *Cell* 76, 717-24.
- 556 21. Weiss, A.A. et al. (1993) Molecular characterization of an operon required for pertussis
557 toxin secretion. *Proc Natl Acad Sci U S A* 90, 2970-4.

- 558 22. Guglielmini, J. et al. (2013) Evolution of Conjugation and Type IV Secretion Systems. *Mol*
559 *Biol Evol* 30, 315-331.
- 560 23. Pukatzki, S. et al. (2007) Type VI secretion system translocates a phage tail spike-like
561 protein into target cells where it cross-links actin. *Proc Natl Acad Sci U S A* 104, 15508-15513.
- 562 24. Logger, L. et al. (2017) Type VI secretion TssK baseplate protein exhibits structural
563 similarity with phage receptor-binding proteins and evolved to bind the membrane complex.
564 *Nature Microbiol* 2, 17103.
- 565 25. Xu, Q. et al. (2016) A distinct type of pilus from the human microbiome. *Cell* 165, 690-
566 703.
- 567 26. Alvarez-Martinez, C.E. and Christie, P.J. (2009) Biological diversity of prokaryotic type IV
568 secretion systems. *Microbiol Mol Biol Rev* 73, 775-808.
- 569 27. de la Cruz, F. et al. (2010) Conjugative DNA Metabolism in Gram-negative Bacteria. *FEMS*
570 *Microbiol Rev* 34, 18-40.
- 571 28. Grohmann, E. et al. (2018) Type IV secretion in Gram-negative and Gram-positive
572 bacteria. *Mol Microbiol* 107, 455-471.
- 573 29. Ellison, C.K. et al. (2017) Obstruction of pilus retraction stimulates bacterial surface
574 sensing. *Science* 358, 535-538.
- 575 30. Ghinet, M.G. et al. (2011) Uncovering the Prevalence and Diversity of Integrating
576 Conjugative Elements in Actinobacteria. *PLoS ONE* 6, e27846.
- 577 31. Blesa, A. et al. (2017) The transjugation machinery of *Thermus thermophilus*:
578 Identification of TdtA, an ATPase involved in DNA donation. *PLoS Genet* 13, e1006669.
- 579 32. Llosa, M. et al. (2002) Bacterial conjugation: a two-step mechanism for DNA transport.
580 *Mol Microbiol* 45, 1-8.
- 581 33. Trokter, M. and Waksman, G. (2018) Translocation through the Conjugative Type IV
582 Secretion System Requires Unfolding of Its Protein Substrate. *J Bacteriol* 200, e00615-17.
- 583 34. Souza, D.P. et al. (2015) Bacterial killing via a type IV secretion system. *Nat Commun* 6,
584 6453.
- 585 35. Suzuki, H. et al. (2010) Predicting plasmid promiscuity based on genomic signature. *J*
586 *Bacteriol* 192, 6045-55.
- 587 36. Gomez-Valero, L. et al. (2019) More than 18,000 effectors in the *Legionella* genus
588 genome provide multiple, independent combinations for replication in human cells. *Proc*
589 *Natl Acad Sci* 116, 2265-2273.
- 590 37. Christie, P.J. et al. (2017) Biological diversity and evolution of type IV secretion systems.
591 In *Type IV Secretion in Gram-Negative and Gram-Positive Bacteria*, pp. 1-30, Springer.
- 592 38. Draper, O. et al. (2005) Site-specific recombinase and integrase activities of a conjugative
593 relaxase in recipient cells. *Proc Natl Acad Sci* 102, 16385-16390.
- 594 39. Vogel, J. et al. (1998) Conjugative transfer by the virulence system of *Legionella*
595 *pneumophila*. *Science* 279, 873-6.
- 596 40. Johnston, C. et al. (2014) Bacterial transformation: distribution, shared mechanisms and
597 divergent control. *Nat Rev Microbiol* 12, 181-96.
- 598 41. Chaban, B. et al. (2015) The flagellum in bacterial pathogens: for motility and a whole lot
599 more. In *Seminars in cell & developmental biology*, pp. 91-103, Elsevier.
- 600 42. Young, G.M. et al. (1999) A new pathway for the secretion of virulence factors by
601 bacteria: the flagellar export apparatus functions as a protein-secretion system. *Proc Natl*
602 *Acad Sci U S A* 96, 6456-61.
- 603 43. Konkel, M.E. et al. (2004) Secretion of virulence proteins from *Campylobacter jejuni* is
604 dependent on a functional flagellar export apparatus. *J Bacteriol* 186, 3296-303.

605 44. Scanlan, E. et al. (2017) A quantitative proteomic screen of the *Campylobacter jejuni*
606 flagellar-dependent secretome. *J Proteomics* 152, 181-187.

607 45. Maezawa, K. et al. (2006) Hundreds of flagellar basal bodies cover the cell surface of the
608 endosymbiotic bacterium *Buchnera aphidicola* sp. strain APS. *J Bacteriol* 188, 6539-43.

609 46. Ferreira, J.L. et al. (2019) γ -proteobacteria eject their polar flagella under nutrient
610 depletion, retaining flagellar motor relic structures. *PLoS Biol* 17, e3000165.

611 47. Pallen, M.J. and Matzke, N.J. (2006) From The Origin of Species to the origin of bacterial
612 flagella. *Nature Rev Microbiol* 4, 784-790.

613 48. Galán, J.E. and Waksman, G. (2018) Protein-injection machines in bacteria. *Cell* 172,
614 1306-1318.

615 49. Guttman, D.S. et al. (2006) Diversifying selection drives the evolution of the type III
616 secretion system pilus of *Pseudomonas syringae*. *Mol Biol Evol* 23, 2342-2354.

617 50. Troisfontaines, P. and Cornelis, G.R. (2005) Type III secretion: more systems than you
618 think. *Physiology* 20, 326-39.

619 51. Nguyen, L. et al. (2000) Phylogenetic analyses of the constituents of Type III protein
620 secretion systems. *J Mol Microbiol Biotechnol* 2, 125-44.

621 52. Gophna, U. et al. (2003) Bacterial type III secretion systems are ancient and evolved by
622 multiple horizontal-transfer events. *Gene* 312, 151-163.

623 53. Sun, G.W. and Gan, Y.H. (2010) Unraveling type III secretion systems in the highly
624 versatile *Burkholderia pseudomallei*. *Trends Microbiol* 18, 561-8.

625 54. Berry, J.-L. and Pelicic, V. (2014) Exceptionally widespread nanomachines composed of
626 type IV pilins: the prokaryotic Swiss Army knives. *FEMS Microbiol Rev* 39, 134-154.

627 55. Tomich, M. et al. (2007) The tad locus: postcards from the widespread colonization
628 island. *Nature Reviews. Microbiology* 5, 363-375.

629 56. Cianciotto, N.P. and White, R.C. (2017) Expanding role of type II secretion in bacterial
630 pathogenesis and beyond. *Infection and immunity* 85, e00014-17.

631 57. Makarova, K.S. et al. (2016) Diversity and evolution of type IV pili systems in archaea.
632 *Frontiers Microbiol* 7, 667.

633 58. Roux, N. et al. (2012) Neglected but amazingly diverse type IVb pili. *Res Microbiol* 163,
634 659-73.

635 59. Lopez-Castilla, A. et al. (2017) Structure of the calcium-dependent type 2 secretion
636 pseudopilus. *Nature Microbiol* 2, 1686.

637 60. Zöllner, R. et al. (2019) Motor Properties of PilT-Independent Type 4 Pilus Retraction in
638 *Gonococci*. *J Bacteriol* 201, e00778-18.

639 61. Adams, D.W. et al. (2019) The type IV pilus protein PilU functions as a PilT-dependent
640 retraction ATPase. *PLoS Genet* 15, e1008393.

641 62. Chlebek, J.L. et al. (2019) PilT and PilU are homohexameric ATPases that coordinate to
642 retract type IVa pili. *PLoS Genet* 15, e1008448.

643 63. Abby, S.S. et al. (2016) Identification of protein secretion systems in bacterial genomes.
644 *Sci Rep* 6, 23080.

645 64. DebRoy, S. et al. (2006) *Legionella pneumophila* type II secretome reveals unique
646 exoproteins and a chitinase that promotes bacterial persistence in the lung. *Proc Natl Acad*
647 *Sci U S A* 103, 19146-51.

648 65. Korotkov, K.V. and Sandkvist, M. (2019) Architecture, function, and substrates of the
649 type II secretion system. *EcoSal Plus* 8, 10.1128/ecosalplus.ESP-0034-2018.

650 66. Lawrence, J.G. and Roth, J.R. (1996) Selfish operons: horizontal transfer may drive the
651 evolution of gene clusters. *Genetics* 143, 1843-1860.

652 67. Mougous, J.D. et al. (2006) A virulence locus of *Pseudomonas aeruginosa* encodes a
653 protein secretion apparatus. *Science* 312, 1526-30.

654 68. Hood, R.D. et al. (2010) A type VI secretion system of *Pseudomonas aeruginosa* targets a
655 toxin to bacteria. *Cell host & microbe* 7, 25-37.

656 69. Borgeaud, S. et al. (2015) The type VI secretion system of *Vibrio cholerae* fosters
657 horizontal gene transfer. *Science* 347, 63-67.

658 70. Leiman, P.G. et al. (2009) Type VI secretion apparatus and phage tail-associated protein
659 complexes share a common evolutionary origin. *Proc Natl Acad Sci U S A* 106, 4154-9.

660 71. Pell, L.G. et al. (2009) The phage lambda major tail protein structure reveals a common
661 evolution for long-tailed phages and the type VI bacterial secretion system. *Proc Natl Acad Sci U S A* 106, 4160-5.

662 72. Lossi, N.S. et al. (2013) The HsiB1C1 (TssB-TssC) complex of the *Pseudomonas aeruginosa*
663 type VI secretion system forms a bacteriophage tail sheathlike structure. *J Biol Chem* 288,
664 7536-7548.

665 73. Basler, M. et al. (2012) Type VI secretion requires a dynamic contractile phage tail-like
666 structure. *Nature* 483, 182-6.

667 74. Cascales, E. (2008) The type VI secretion toolkit. *EMBO reports* 9, 735-741.

668 75. Schwarz, S. et al. (2010) Burkholderia type VI secretion systems have distinct roles in
669 eukaryotic and bacterial cell interactions. *PLoS Pathogens* 6, e1001068.

670 76. Ludu, J.S. et al. (2008) The *Francisella* pathogenicity island protein PdpD is required for
671 full virulence and associates with homologues of the type VI secretion system. *J Bacteriol*
672 190, 4584-95.

673 77. Russell, A.B. et al. (2014) Type VI secretion system effectors: poisons with a purpose. *Nat*
674 *Rev Microbiol* 12, 137-48.

675 78. Russell, A.B. et al. (2014) A type VI secretion-related pathway in *Bacteroidetes* mediates
676 interbacterial antagonism. *Cell Host Microbe* 16, 227-36.

677 79. Böck, D. et al. (2017) In situ architecture, function, and evolution of a contractile
678 injection system. *Science* 357, 713-717.

679 80. Touchon, M. et al. (2016) Genetic and life-history traits associated with the distribution
680 of prophages in bacteria. *ISME J* 10, 2744-2754.

681 81. Komano, T. et al. (2000) The transfer region of IncI1 plasmid R64: similarities between
682 R64 tra and legionella icm/dot genes. *Mol Microbiol* 35, 1348-59.

683 82. Jurcisek, J.A. et al. (2017) Nontypeable *Haemophilus influenzae* releases DNA and DNABII
684 proteins via a T4SS-like complex and ComE of the type IV pilus machinery. *Proc Natl Acad Sci*
685 114, E6632-E6641.

686 83. Spitz, O. et al. (2019) Type I Secretion Systems—One Mechanism for All? *Microbiology*
687 *spectrum* 7, 10.1128/microbiolspec.PSIB-0003-2018.

688 84. Bhattacharya, S. et al. (2019) A ubiquitous platform for bacterial nanotube biogenesis.
689 *Cell reports* 27, 334-342. e10.

690 85. Pal, R.R. et al. (2019) Pathogenic *E. coli* extracts nutrients from infected host cells
691 utilizing injectisome components. *Cell* 177, 683-696. e18.

692 86. Pallen, M.J. (2002) The ESAT-6/WXG100 superfamily—and a new Gram-positive secretion
693 system? *Trends in microbiology* 10, 209-212.

694 87. Gröschel, M.I. et al. (2016) ESX secretion systems: mycobacterial evolution to counter
695 host immunity. *Nature Rev Microbiol* 14, 677.

696 88. Rouse, S.L. et al. (2017) A new class of hybrid secretion system is employed in
697 *Pseudomonas amyloid* biogenesis. *Nature communications* 8, 263.

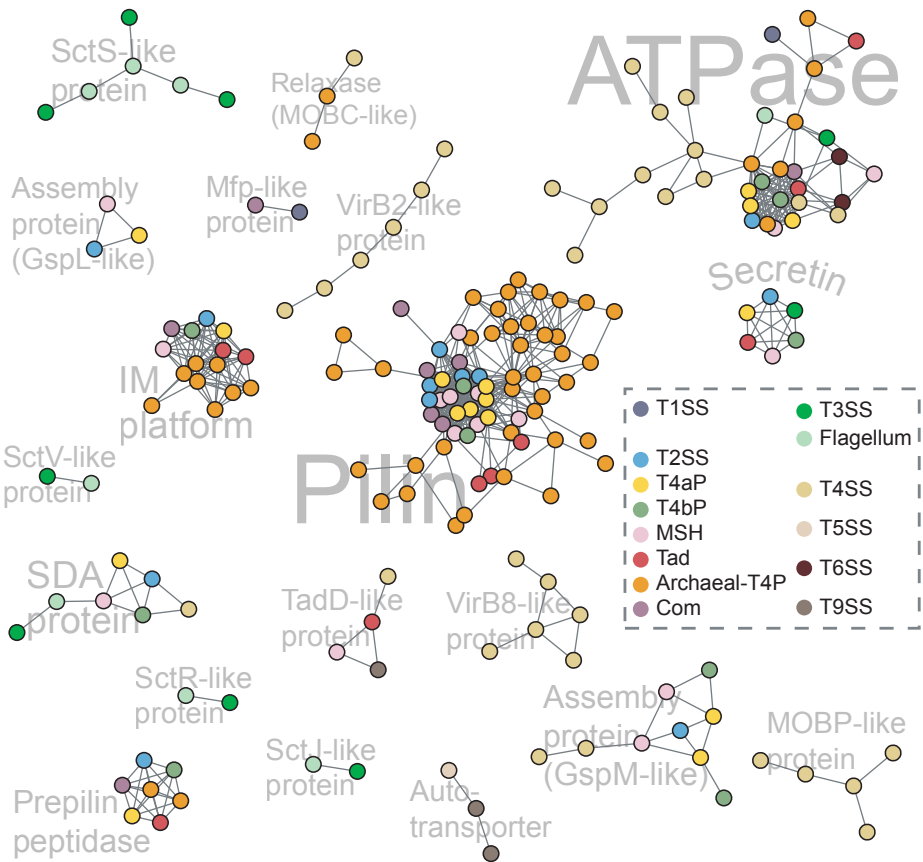
- 699 89. Morlot, C. and Rodrigues, C.D. (2018) The New Kid on the Block: a specialized secretion
700 system during bacterial sporulation. *Trends in microbiology* 26, 663-676.
- 701 90. Gómez-Santos, N. et al. (2019) A TonB-dependent transporter is required for secretion of
702 protease PopC across the bacterial outer membrane. *Nature communications* 10, 1360.
- 703 91. Lauber, F. et al. (2018) Type 9 secretion system structures reveal a new protein transport
704 mechanism. *Nature* 564, 77.
- 705 92. McBride, M.J. (2019) Bacteroidetes Gliding Motility and the Type IX Secretion System.
706 *Microbiology spectrum* 7, 10.1128/microbiolspec.PSIB-0002-2018.
- 707 93. Blocker, A. et al. (1999) The tripartite type III secretin of *Shigella flexneri* inserts IpaB and
708 IpaC into host membranes. *J Cell Biol* 147, 683-693.
- 709 94. Hachani, A. et al. (2016) Type VI secretion and anti-host effectors. *Curr Opin Microbiol*
710 29, 81-93.
- 711 95. Kwak, M.-J. et al. (2017) Architecture of the type IV coupling protein complex of
712 *Legionella pneumophila*. *Nature Microbiol* 2, 17114.
- 713 96. Jamet, A. and Nassif, X. (2015) New players in the toxin field: polymorphic toxin systems
714 in bacteria. *MBio* 6, e00285-15.
- 715 97. Matsuda, S. et al. (2019) Export of a *Vibrio parahaemolyticus* toxin by the Sec and type III
716 secretion machineries in tandem. *Nature Microbiol* 4, 781.
- 717 98. Swain, D.M. et al. (2017) A prophage tail-like protein is deployed by Burkholderia
718 bacteria to feed on fungi. *Nature communications* 8, 404.
- 719 99. Tobe, T. et al. (2006) An extensive repertoire of type III secretion effectors in *Escherichia*
720 *coli* O157 and the role of lambdoid phages in their dissemination. *Proc Natl Acad Sci U S A*
721 103, 14941-6.
- 722 100. Wandersman, C. and Delepelaire, P. (1990) TolC, an *Escherichia coli* outer membrane
723 protein required for hemolysin secretion. *Proc Natl Acad Sci* 87, 4776-4780.
- 724 101. Marsh, J.W. and Taylor, R.K. (1998) Identification of the *Vibrio cholerae* type 4 prepilin
725 peptidase required for cholera toxin secretion and pilus formation. *Mol Microbiol* 29, 1481-
726 92.
- 727 102. Duboule, D. and Wilkins, A.S. (1998) The evolution of "bricolage". *Trends in Genetics* 14,
728 54-59.
- 729 103. Nickerson, N.N. et al. (2012) A Single Amino Acid Substitution Changes the Self-
730 Assembly Status of a Type IV Piliation Secretin. *J Bacteriol* 194, 4951-4958.
- 731 104. Korotkov, K.V. et al. (2011) Secretins: dynamic channels for protein transport across
732 membranes. *Trends Biochem Sci* 36, 433-43.
- 733 105. Crago, A.M. and Koronakis, V. (1998) *Salmonella* InvG forms a ring-like multimer that
734 requires the InvH lipoprotein for outer membrane localization. *Mol Microbiol* 30, 47-56.
- 735 106. Katoh, K. and Standley, D.M. (2013) MAFFT multiple sequence alignment software
736 version 7: improvements in performance and usability. *Mol Biol Evol* 30, 772-80.
- 737 107. Dress, A.W. et al. (2008) Noisy: identification of problematic columns in multiple
738 sequence alignments. *Algorithms Mol Biol* 3, 7.
- 739 108. Nguyen, L.T. et al. (2015) IQ-TREE: a fast and effective stochastic algorithm for
740 estimating maximum-likelihood phylogenies. *Mol Biol Evol* 32, 268-74.
- 741 109. Hoang, D.T. et al. (2017) UFBoot2: improving the ultrafast bootstrap approximation.
742 *Mol Biol Evol* 35, 518-522.
- 743 110. Korotkov, K.V. et al. (2012) The type II secretion system: biogenesis, molecular
744 architecture and mechanism. *Nature Reviews. Microbiology* 10, 336-51.
- 745

Highlights

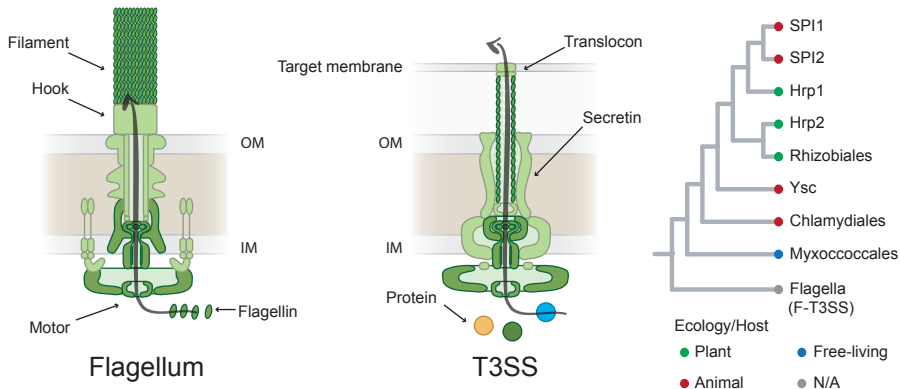
- Recent discoveries in molecular, structural and cell biology revealed interactions between components and unexpected functions for protein secretion systems. They illuminate the evolutionary history of these machines.
- Many protein secretion systems evolved from evolutionary tinkering of other pre-existing cellular machines by co-option of components or entire functional modules.
- The extent of changes involved in these evolutionary processes depends on the complexity of the systems, functional differences to the original co-opted system, abundance of genetic elements to be tinkered with, and, possibly, the evolution of the mechanisms for interacting with effectors and discriminate them.
- Some protein families, like the secretin, were recruited multiple times to different protein secretion systems and may have endured periods where the same gene was shared by several systems.

Outstanding questions

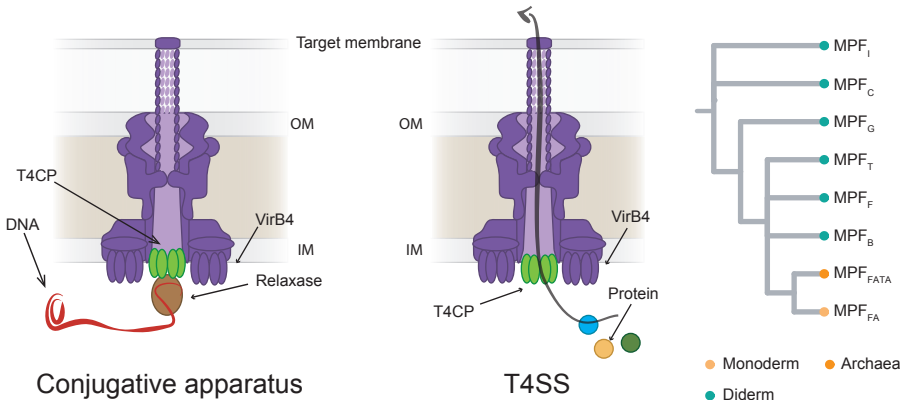
- Are there novel systems yet to be discovered?
- Are systems that deliver proteins inside other cells exclusively present in diderm Bacteria? If so, why?
- Are there specific structural traits facilitating the integration of components in a secretion system?
- Can we combine components from known systems to design novel ones with specific functions?
- How did the recognition of effectors at the onset of secretion systems evolve?
- Can T3SS be used to secrete proteins directly into bacterial cells?



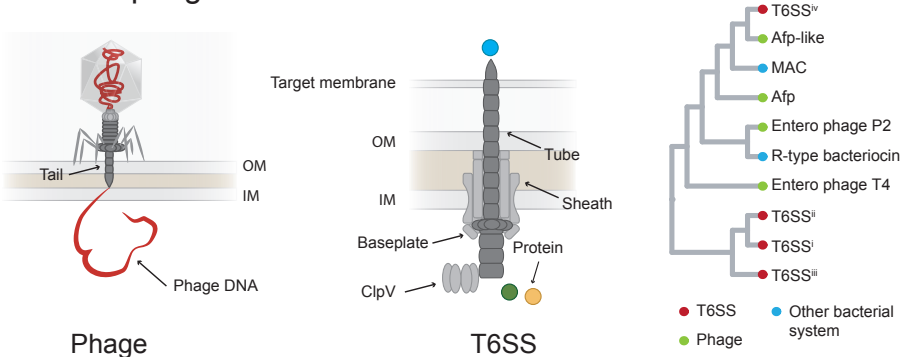
a- From the flagellum to the T3SS

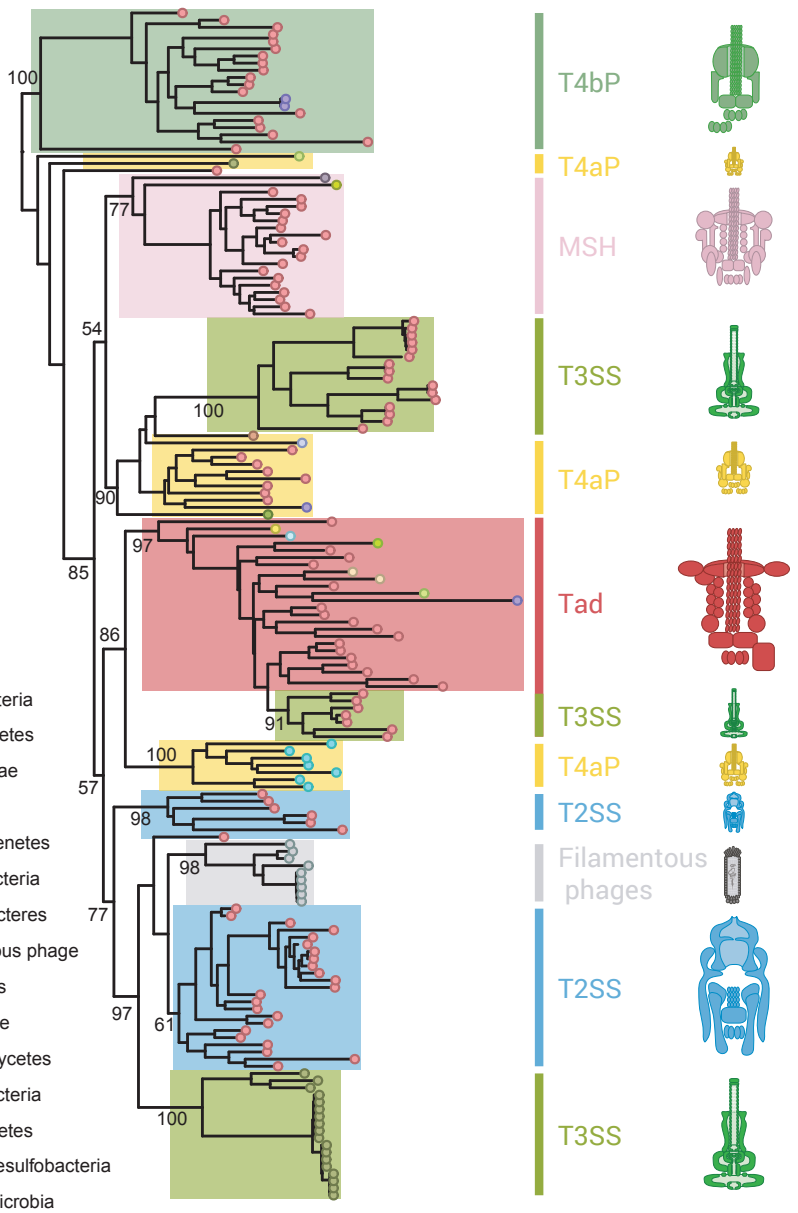


b- The T4SS: from conjugation to protein secretion



c- From a phage to the T6SS





Phylum

- Acidobacteria
- Bacteroidetes
- Chlamydiae
- Chlorobi
- Chrysiogenetes
- Cyanobacteria
- Deferribacteres
- Filamentous phage
- Firmicutes
- Nitrospirae
- Planctomycetes
- Proteobacteria
- Synergistetes
- Thermodesulfobacteria
- Verrucomicrobia

T4bP
 T4aP
 MSH
 T3SS
 T4aP
 Tad
 T3SS
 T4aP
 T2SS
 Filamentous phages
 T2SS
 T3SS

