



HAL
open science

Structural interactions define assembly adapter function of a type II secretion system pseudopilin

Cristian Escobar, Badreddine Douzi, Geneviève Ball, Brice Barbat, Sebastien Alphonse, Loïc Quinton, Romé Voulhoux, Katrina Forest

► **To cite this version:**

Cristian Escobar, Badreddine Douzi, Geneviève Ball, Brice Barbat, Sebastien Alphonse, et al.. Structural interactions define assembly adapter function of a type II secretion system pseudopilin. *Structure*, 2021, 29 (10), pp.1116-1127.e8. 10.1016/j.str.2021.05.015 . hal-03367306

HAL Id: hal-03367306

<https://hal.science/hal-03367306>

Submitted on 1 Apr 2022

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.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27

Multidisciplinary Interrogation of a Crucial Protein Interface in the Type II Secretion System

Cristian A. Escobar^{*,1,#}, Badreddine Douzi^{*,2,#}, Geneviève Ball², Brice Barbat², Sebastien Alphonse^{2,#}, Loïc Quinton^{3,†}, Romé Voulhoux^{2,†}, Katrina T. Forest^{1,†}

¹ Department of Bacteriology, University of Wisconsin-Madison, Madison, WI

² Laboratoire de Chimie Bactérienne, Institut de Microbiologie de la Méditerranée, CNRS-Aix Marseille Université (UMR7283), Marseille, France.

³ Laboratory of Mass Spectrometry, MolSys Research Unit, University of Liège, Liège, Belgium

*CAE and BD contributed equally to this work.

† To whom correspondence should be addressed:

LQ: Laboratory of Mass Spectrometry, MolSys Research Unit, University of Liège, Liège, Belgium loic.quinton@uliege.be +32 4 3663679 ORCID 0000-0001-8153-9590

RV: LCB IMM CNRS AMU, 31 Chemin Joseph Aiguier, 13402 Marseille Cedex 20, France voulhoux@imm.cnrs.fr +33 4 91 16 46 67 ORCID 0000-0002-5258-5958

KTF: 1550 Linden Dr., Madison WI 53726 forest@bact.wisc.edu 608 265-3566 ORCID [0000-0003-4197-7662](https://orcid.org/0000-0003-4197-7662)

Current address:

CAE Department of Biochemistry, University of Wisconsin-Madison, Madison, WI

BD Université de Lorraine, INRAE, DynAMic, F-54000 Nancy, France

SA Department of Chemistry, CCNY, New York, NY

28 **ABSTRACT** The type IV filament superfamily comprises widespread membrane-associated
29 polymers in prokaryotes. The Type II secretion system (T2SS), a significant virulence pathway in
30 many pathogens, belongs to this superfamily. We have deciphered the molecular position of a
31 missing component of the Xcp T2SS filament using NMR, novel acidic side-chain crosslinking,
32 known 3D structures, and molecular modeling. We demonstrate that the low abundance XcpH is
33 the structural adapter between the trimeric tip complex XcpIJK and the pseudopilus filament of
34 XcpG subunits. The T2SS pseudopilus model reveals that each XcpH, I, J, and K protein caps an
35 XcpG protofilament in a structure compatible with dimensions of the periplasm and the outer
36 membrane-spanning secretin. Unexpectedly, to fulfill its adapter function, the XcpH N-terminal
37 helix is unwound in the filament, akin to the primary XcpG. We provide the first complete model
38 of a type IV filament, a result immediately transferable to understanding of other T2SS and the
39 type IV pili.

40 **INTRODUCTION**

41

42 Bacteria have sophisticated secretory nanomachines that evolved to deliver exoproteins to the
43 bacterial cell surface, into the surrounding medium or directly into host cells. Among these, the
44 type 2 secretion system (T2SS) is a trans-envelope apparatus specialized for secretion of folded
45 proteins from Gram negative bacteria¹. In many cases, these secreted substrates are virulence
46 factors: examples include the cholera toxin of *Vibrio cholera*, the exotoxin A of *Pseudomonas*
47 *aeruginosa*, and the heat labile toxin of enterotoxigenic *Escherichia coli*². The conformational
48 constraint of transporting folded substrates is shared by the Type VII, or ESX, system³ and the
49 Type IX secretion system⁴, but is in clear contrast to the linear export of unfolded polypeptides by
50 the T3SS⁵ or the T4SS⁶. The substrates of the T2SS are not known to share any sequence or 3D
51 structural motifs⁷, and thus an open question is how substrates are individually recognized and
52 ushered out. Dynamic and changing interactions, potentially involving nucleation of structure in
53 intrinsically disordered regions, must be at the heart of this selection process⁸.

54

55 The T2SS accomplishes its task with a distinctive mode of transport involving a pilus-like structure
56 (the pseudopilus) in the periplasmic space, at the interface between an inner membrane assembly
57 platform and a large outer membrane pore, the secretin⁹. Structural, functional and evolutionary
58 commonalities to the extracellular Type IV pili critically inform the T2SS field and *vice versa*¹⁰.
59 Thanks to recent spectacular developments in cryo-electron microscopy, the 3D structures of T2SS
60 secretins¹¹⁻¹⁴ and the filament formed upon overexpression of the pseudopilin subunit PulG¹⁵ have
61 been solved, and reveal important structural features. While secretins form a tightly gated giant
62 double β -barreled pore of 80 Å diameter in the outer membrane, PulG adopts a right-handed
63 homopolymeric helix with a diameter of 70 Å, compatible with observations that a filament
64 extends through the secretin when the pseudopilin is overexpressed^{16,17}. In this pseudopilus, inter-
65 subunit contacts occur along hydrophobic N-terminal α helices of the pseudopilin, with a special
66 role for the negatively charged glutamic acid at position 5 in both recruiting and stabilizing
67 interactions within the filament¹⁸⁻²¹. It has long been assumed that under native expression levels
68 the pseudopilin forms a short and transient thread entirely within the periplasm during secretion²²,
69 a model which has been difficult to verify experimentally until recently, when a cryo-electron

70 tomography reconstruction of the *Legionella pneumophila* T2SS revealed pseudopilus density in
71 ~20% of the complexes²³.

72
73 In addition to the major pseudopilin, additional low abundance (or *minor*) pseudopilins have been
74 identified in the T2SS, based on being encoded within the same operon and sharing amino acid
75 sequence with the pseudopilin, plus undergoing maturation by the same prepilin peptidase²⁴⁻²⁶.
76 The four minor pseudopilins are generally designated as the H, I, J, and K proteins (variously
77 prefixed with Gsp-, Pul-, Xcp-, depending on the particular T2SS, please see on line Materials and
78 Methods for nomenclature notes). While none of these has been directly observed in a T2SS
79 pseudopilus, the periplasmic domains of minor pseudopilins GspI, GspJ, and GspK have the ability
80 to assemble into a ternary complex whose crystal structure has been solved. The four periplasmic
81 domains of GspH, GspI, GspJ and GspK, form a quaternary complex^{27,28}, but its structure is
82 unknown. As a group, the minor pseudopilins are known to play a role in initiation and control of
83 pseudopilus assembly and potentially retraction. The complex furthermore interacts with secretion
84 substrates³² and with the periplasmic domains of the secretin³³. Although there is strong evidence
85 that the quaternary complex is found, at least transiently, at the tip of the pseudopilus, its structure
86 has been elusive. Minor pseudopilin subunits could not be resolved in the *L. pneumophila* T2SS
87 cryo electron tomograms²³. In a recent T4P cryo electron tomography study, minor pilins were
88 attributed to density blobs but no conclusions could be drawn about which protein belonged to
89 which density nor about the molecular nature of their interactions³⁴

90
91 While many reports have established the importance of the minor pseudopilins and pilins as a
92 group for initiation and/or assembly of (pseudo)pili, fewer have analyzed their individual
93 contributions to the secretion process. Nonetheless, specific roles have been assigned or proposed
94 for all four minor pseudopilins: GspJ primes filament assembly through the recruitment of other
95 minor pseudopilins³⁷, GspI serves as a protein-protein interaction hub²⁷, and GspK controls
96 pseudopilus length and/or number^{27,37}. GspH interacts with GspJ through its globular domain and
97 with the structurally similar major pseudopilin via its N-terminal hydrophobic α -helices^{18,27,39,40},
98 thus it is tempting to suggest GspH plays an adapter function between the tip and the body of the
99 pseudopilus. However, the essentiality of GspH for secretion has been called into question because

100 its absence can be overcome by overproduction of other minor pseudopilins³⁷. There are no
101 molecular data on how GspH is embedded in the quaternary complex.

102

103 Based on x-ray crystallographic structures of constructs representing periplasmic domains of
104 GspH, GspI, GspJ, and GspK from several microbial species solved alone and/or in dimeric or
105 trimeric subcomplexes^{28,39-43}, each pseudopilin fold resembles major T2SS pseudopilins and Type
106 IV pilus (T4P) major pilins characterized by an extended N-terminal α -helix, the first half of which
107 (α 1N) is not included in the soluble crystallography targets but is expected to be exposed to the
108 membrane environment, and the second half of which (α 1C) is partially buried in a globular
109 α/β domain. Despite this common overall topology, it is important to bear in mind that the
110 periplasmic domains of the minor pseudopilins interact to form the soluble quaternary complex in
111 the absence of α 1N, and since these minor subunits are at the tip of the pseudopilus they cannot
112 depend on vertical interactions with major subunits for upward assembly interactions. Within this
113 general framework, each minor pseudopilin has distinctive sequence and structural hallmarks.
114 GspH is most similar in size, sequence and structure to the major pseudopilin. GspI is the smallest
115 of the four. GspJ is larger, and displays a pronounced groove on its surface. Finally, GspK carries
116 an additional structural domain of over 100 amino acids. GspH, GspI, and GspJ share the glutamate
117 at position 5 with GspG, whereas GspK does not.

118

119 Based on two crystal structures of ternary complexes of minor pseudopilin periplasmic domains
120 (3CI0 from the enterotoxigenic *E. coli* Gsp T2SS and 5VTM from the *P. aeruginosa* Xcp T2SS
121 equivalent minor pilins XcpIJK^{28,40}), we know the stabilizing interactions among the periplasmic
122 domains (designated here with a subscript “p”) GspI_p, GspJ_p, and GspK_p include buried salt bridges
123 between α 1C residues for all three pairwise interactions: XcpI_p Asp51 to XcpK_p Arg45, XcpJ_p
124 Glu58 to XcpK_p Arg45, and XcpI_p Asp51 to XcpJ_p Glu58. The interaction is also stabilized by salt
125 bridges involving amino acids of XcpI_p α 1 and XcpJ_p β 11, including XcpI_p Asp51 and XcpK_p
126 Arg195. Contacts at the bottom of the bundle of helices further lock the ternary complex in place.
127 Double involvement of XcpI_{D51} in XcpI_p:J_p & XcpI_p:XcpK_p interfaces is evidence to bolster the
128 original finding that XcpI_p supports the linear organization of XcpJ_p:I_p:K_p interactions in the
129 complex²⁷. Somewhat perplexingly, the largest buried surface area between any two of the soluble
130 proteins in both ternary complexes is between XcpJ_p (GspJ_p) and XcpK_p (GspK_p), with a total

131 buried surface area of $\sim 3600 \text{ \AA}^2$ in 5VTM. And yet, these two subunits do not interact with each
132 other in solution in the absence of XcpI_p²⁷.

133

134 Several lines of evidence support downward addition of XcpH to the base of the XcpJ:I:K complex
135 through an interaction between the globular domains of XcpH and XcpJ. XcpH_p completes the
136 quaternary complex by interacting with XcpJ_p²⁷. Additionally, $\alpha 1$ of XcpH is essential in this
137 interaction and XcpH_p can be positioned most convincingly in a small angle x-ray scattering
138 envelope of the quaternary complex below XcpJ_p²⁸. The structural compatibility of the ternary tip
139 complex GspI_pJ_pK_p with downward addition of GspH below GspJ was proposed by Korotkov *et*
140 *al.* in agreement with their previous suggestion that GspH could be positioned at the top of a GspG
141 fiber but not at the bottom^{39,40}. The lack of Glu5 in GspK can also be taken as evidence that GspK
142 does not require helix:helix packing stabilization via a salt bridge with a higher subunit in the
143 filament and thus is located at the top of the pseudopilus³¹.

144

145 In this study we set out to validate the biological importance of XcpH and to decipher the molecular
146 details of the XcpH:I:J:K interaction. Because all available data to date suggest an unstable,
147 dynamic and/or small interaction interface with XcpH, we turned to a collection of compatible
148 techniques suited to interrogation of such interactions. We show that XcpH is indeed critical for
149 efficient Type II secretion in *P. aeruginosa*. We demonstrate it interacts specifically with the $\alpha 1$
150 helix of XcpJ, and we narrow the interaction residues to a well-defined interface that is highly
151 conserved among T2SSs. Our experimentally determined constraints allow us to synthesize a
152 molecular model of the XcpHIJK quaternary structure, and to place this in the context of the
153 complete pseudopilus filament, with biological implications for the role of the GspH/XcpH low
154 abundance pseudopilin. Moreover, the molecular organization involving an inner membrane
155 platform, priming by minor pilins, and a helical subunit assembly that is remarkably well
156 conserved both evolutionarily and structurally in other type IV filamentous nanomachines^{10,44}
157 including the canonical type IV pilus (T4P)^{19,31,38}, competence pilus^{45,46}, and to some extent the
158 archaellum⁴⁷⁻⁴⁹. Thus our work has wide-ranging implications for understanding assembly of
159 several fundamental microbial organelles.

160 **RESULTS**

161

162 **The four minor pilins are essential for secretion.**

163 In our quest to understand the structure-function relationship of the minor pseudopilin complex,
164 we endeavored to systematically assess the independent requirement of each of the four minor
165 pseudopilins of the *P. aeruginosa* Xcp T2SS during the secretion process. Individual in-frame
166 deletion mutants were constructed in the PA01 reference strain for *xcpH*, *xcpI*, *xcpJ* and *xcpK* (see
167 online Materials and Methods for nomenclature). Each mutant was assessed for its ability to secrete
168 the major Xcp T2SS effector LasB (**Figure 1a**) and for the ability of secreted LasB to degrade
169 casein on milk plates (**Figure 1b**). All four individual deletion strains were similar in their inability
170 to secrete LasB and lacked protease activity, like the negative control strain in which the entire
171 T2SS operon is missing, thus demonstrating the absolute requirement of each minor pseudopilin
172 in the Xcp T2SS secretion process. This finding is supported by complementation of each secretion
173 deficient phenotype by introduction of the corresponding wild type gene on a plasmid. Our data
174 are apparently in conflict with a 2018 publication suggesting that XcpH and XcpK are not required
175 for T2S in *P. aeruginosa*²⁸. Our finding that XcpH is indispensable for secretion prompted us to
176 further explore the 3D integration of XcpH into the XcpHIJK quaternary complex.

177

178 **Chemical shift perturbation studies of labeled XcpH_p in the presence of XcpJ_p.**179 *XcpH_p solution NMR assignments*

180 To discover sites of interaction between XcpH and XcpJ, we performed solution NMR
181 experiments. We initially isotopically labeled the ~16 kDa XcpH_p, a feasible NMR target, with
182 ¹⁵N and ¹³C, and carried out amino acid assignments. Regions heavily crowded in the HSQC
183 spectrum were excluded, leading to 85% coverage of the XcpH_p amino acid sequence (**Figure 2a**).
184 The high signal dispersion of the HSQC spectrum indicates an overall well-folded protein, while
185 the area in the spectrum with high resonance overlap and strong signal (around 8.2 ppm in the ¹H
186 dimension and 121 ppm in the ¹⁵N dimension) indicates a highly dynamic region of the protein.

187

188 We employed these assignments to experimentally determine the secondary structure elements
189 within XcpH_p (**Figure 2b**). Furthermore, we generated two 3D homology models of XcpH_p using

190 PDB 2KNQ and 2QV8³⁹. Each 3D model matched our experimental secondary structure
191 assignments well (not shown).

192

193 A subset of XcpH_p amino acids, in particular lysines, could not be assigned in our NMR spectra.
194 The inability to assign these lysines implies a flexible region of the protein backbone. We tested
195 the role of these residues in secretion by creating an in-frame deletion of amino acids 108-117,
196 which form a long excursion between β 3 and β 4, resembling an old-fashioned mouse trap spring
197 in one homology model (**Supplemental Figure 1**). We found that trans complementation with the
198 *xcpH_{Δ34}* deletion restores LasB secretion and activity to the *ΔxcpH* strain (**Figure 1b, c**). Thus
199 residues 108-117 are not required for LasB secretion, and therefore an disorder to order transition
200 of this region cannot be required for secretion of LasB either.

201

202 *Specific sites of spectral changes along the XcpH_p helix suggest binding interface: Chemical Shift*
203 *Perturbation*

204 With XcpH_p chemical shifts in hand, we performed chemical shift perturbation (CSP) analysis to
205 identify structural regions impacted in XcpH_p upon binding of its partner XcpJ_p. It is expected that
206 resonances from residues close to the protein-protein interface will be shifted by the presence of
207 the binding partner. Thus, a sample of 100 μ M ¹⁵N uniformly labeled XcpH_p was titrated with
208 increasing amounts of XcpJ_p and the 2D ¹H-¹⁵N HSQC spectrum was collected. XcpJ_p
209 concentrations above 80 μ M could not be used since at that concentration there was a significant
210 signal intensity loss due to the large size of the complex. Significant CSPs were observed in
211 multiple places in the amino acid sequence (**Figure 3a**). By mapping these CSPs on a structural
212 model of XcpH_p, we observed one subset of amino acid residues located close to the tip of the
213 alpha helical spine of XcpH_p (residues G42 to G59 in the helix, and G130 and G131 on the nearby
214 β 4- β 5 loop) that are perturbed by the presence of XcpJ_p (**Figure 3b**). In addition, there is a second
215 set of residues with significant CSP located on the C-terminal β sheet of XcpH_p (residues F135 to
216 R138 in β 5, L149 and S153 in β 6 and E162 and A164 in β 7). These additional perturbed sites may
217 indicate a secondary point of interaction with XcpJ_p, or changes in XcpH_p propagated through its
218 structure after binding to XcpJ_p. Global smaller amplitude changes across much of XcpH_p suggest
219 that XcpH_p:J_p interaction subtly affects dynamics or the structure of the entire XcpH_p molecule.

220

221 Site specific cysteine incorporation in XcpJ_p provides XcpH_p:J_p interaction sites.*222 Cysteine cross-linking*

223 The knowledge that a significant structural interaction interface for the major pseudopilins is the
224 packing of their $\alpha 1$ helices¹⁵, combined with our CSP results implicating the XcpH $\alpha 1C$ in the
225 XcpH_p:J_p interaction, led us to hypothesize that $\alpha 1C:\alpha 1C$ packing forms a major interaction
226 between these two minor pseudopilins. We introduced cysteines at surface-exposed $\alpha 1C$ positions
227 in XcpH_p (V46, D49, L53) and XcpJ_p (R46, R53) and mixed purified proteins under oxidizing
228 conditions. In every instance, homodimers are formed (**Figure 4**). Heterodimers between XcpH_p
229 and XcpJ_p were also observed in some prominent cases, most notably for XcpJ_{R46C}:H_{L53C} (**Figure**
230 **4**). Heterodimers were weak or absent when XcpH_{pV46C} was one of the partners. Although these
231 results provide no quantitative measure of affinity or specificity, they served as motivation to use
232 XcpJ_p $\alpha 1C$ cysteine variants as probes of specific sites of interaction in our NMR platform.

233

234 Paramagnetic relaxation enhancement

235 To identify local interactions and specific amino acid involvements in the XcpH:XcpJ interface,
236 we performed paramagnetic relaxation enhancement (PRE) experiments guided by cysteine cross-
237 linking results. In this approach, the effect of a specifically incorporated spin-label (1-Oxyl-
238 2,2,5,5-tetramethylpyrroline-3-methyl) methanethiosulfonate (MTSL) within XcpJ_p on the ¹⁵N
239 XcpH_p spectrum is observed; in particular, XcpH_p residues in proximity to the spin label on XcpJ_p
240 will experience a decrease in signal intensity. XcpJ_p cysteine variants R46C and R53C were spin-
241 labeled, as these positions in XcpJ_p were implicated in cysteine cross-linking experiments
242 described above to participate in an XcpH $\alpha 1C$ - XcpJ $\alpha 1C$ packing interaction (**Figure 4**). In
243 addition, XcpH_{pT178C} and XcpH_{pE180C} were created and used for spin-labeling; these amino acids
244 are on the face of XcpJ opposite $\alpha 1C$, where we did not expect any change in signal from spin
245 label incorporation. We chose to perform PRE experiments using a 10:3 molar ratio of ¹⁵N-U
246 XcpH_p to spin-labeled XcpJ_p variants. Although a 1:1 ratio would in theory have provided more
247 quantitative results, for this protein pair, it would have led to signal loss as seen in the previously
248 described titration.

249

250 The spin label MTSL at positions 46 or 53 within XcpJ_p has a marked effect on XcpH_p signal
251 intensities (**Figure 5a**). XcpJ_{pR46C}-MTSL induced a signal reduction on XcpH_p $\alpha 1C$ (residues 42

252 to 54), with a maximum effect on D49. In addition, residues located in the loop between XcpH_p
253 β 4 and β 5 are also affected (G130 and E131) (**Figure 5**), in agreement with CSP results. XcpJ_pR53C-
254 MTSL had an effect on similar amino acid residues. Within XcpH_p α 1C the overall signal intensity
255 decrease was slightly less compared to the C46 label, but interestingly the effect was shifted up
256 the helix with the residue most greatly affected by the paramagnetic label, XcpH_p-N55, located at
257 the tip of XcpH_p (**Figure 5**). No significant effect was observed on XcpH_p when XcpJ_p was labeled
258 at positions T178C or E180C (**Supplemental Figure 2**). Thus, the PRE experiment clearly
259 supports that the interface between XcpH_p and XcpJ_p is mediated by α 1C interactions with
260 involvement of residues in the β 4- β 5 loop of XcpH. In addition, the difference in signal intensity
261 decrease and the shift on the most affected residue towards the tip of XcpH_p when the MTSL label
262 is at XcpJ position 53 as opposed to 46 suggest that the helical register is organized with XcpJ_p
263 above XcpH_p in the heterodimer.

264

265 **Cross-linking Mass Spectrometry using acidic cross links *via* the DMTMM activating** 266 **reagent**

267 In order to extend our site-specific information from engineered disulfide bonds and PRE
268 measurements, we applied a recently developed methodology for crosslinking of acidic side chains
269 within a defined distance, followed by mass spectrometry identification of cross-linked peptides⁵⁰.
270 We established our pipeline using a 1:1 molar ratio of XcpH_p and XcpJ_p (Supplemental Methods
271 and **Supplemental Figure 3**). Previous work has established that XcpJ_p is indispensable for XcpH_p
272 involvement in the quaternary complex, which was the motivation for using this pair in our NMR
273 experiments also^{27,28}. Subsequently, we interrogated the complete quaternary complex of
274 XcpH_pI_pJ_pK_p, known already to form *in vitro* a quaternary complex with a molar ratio of 1:1:1:1²⁸.
275 These latter experiments confirmed the set of acidic crosslinks between XcpH_p and XcpJ_p are the
276 same in the heterodimer as in the heterotetramer, and thus we focus here on the quaternary complex
277 results.

278

279 The four minor pseudopilin soluble constructs XcpH_p, XcpI_p, XcpJ_p and XcpK_p contain 25, 14, 36
280 and 42 acidic residues, respectively, which are distributed along their amino acid primary
281 sequences. Thus the use of adipic acid dihydrazide (ADH) coupled to 4-(4,6-dimethoxy-1,3,5-
282 triazin-2-yl)-4-methyl-morpholinium chloride (DMTMM) as a coupling reagent appeared

283 promising. DMTMM activates carboxylic acid functionalities, usually poorly reactive, allowing
284 coupling of the crosslinking reagent ADH to the extended amino acid side chains with high
285 efficiency⁵⁰. When added in large excess to a freshly prepared aqueous solution containing a
286 1:1:1:1 molar ratio of XcpH_p, XcpI_p, XcpJ_p and XcpK_p, DMTMM/ADH reagents formed covalent
287 bonds (**Figure 6a**), which connect proximal acidic functionalities with a maximum linking
288 distance of 20-25 Å⁵¹. The most likely tetramer band observed on SDS-PAGE of this mixture was
289 excised and analyzed by a classical proteomic bottom-up approach to validate the presence of all
290 four proteins. Each protein was identified unambiguously, with a sequence coverage of 92% for
291 XcpH_p (61 peptides), 78% for XcpI_p (31 peptides), 86% for XcpJ_p (71 peptides) and 91% for
292 XcpK_p (146 peptides).

293
294 Due to the long length of the linkers, these cross links provide complex and valuable information
295 to constrain a model of the quaternary structure of the assembled XcpH_pI_pJ_pK_p complex. Most
296 revealing for our goals of understanding the placement of XcpH within the pseudopilus quaternary
297 complex were crosslinks between glutamate or aspartate side chains of XcpH_p with any of the other
298 three soluble pseudopilin subunits in the complex. Many inter-subunit crosslinks between XcpH_p
299 and XcpJ_p were identified between α1C of XcpH_p, in particular within the peptide
300 LAGLIGVLTDEAVLDNR, and the several peptides that together span the analogous α1C of
301 XcpJ_p and the β hairpin that follows it (**Figures 6b, 6c**). An example MS/MS spectrum reveals
302 intense peaks defining the connection between the two peptides (**Figure 6b**), with both chains
303 unambiguously characterized by many fragment ions distributed along the peptide sequences.
304 Moreover, the large mass difference observed between y₆α⁺ and y₇α⁺ characterizes precisely the
305 position of crosslinking. The same α1C region of XcpH_p also cross links with two regions of XcpK_p
306 and one region of XcpI_p therefore suggesting they are in physical proximity (**Figure 6c, Table**
307 **S3**). Interestingly, an acidic residue located in the XcpH_p β₄β₅ loop shown by PRE analysis to be
308 in close proximity with the α1C region of XcpJ_p (**Figure 5b**) is also involved in cross links with
309 XcpJ_p, XcpI_p and XcpK_p (**Figure 6c, Table S3**).

310
311 Crosslinks between XcpI_p and XcpJ_p as well as a single crosslink identified between XcpI_p and
312 XcpK_p were all in agreement with the crystallographically determined structures of GspI_pJ_pK_p and
313 XcpI_pJ_pK_p. Moreover, intrasubunit crosslinks agreed with known protein 3D structures (data not

314 shown). This rich library of crosslinks between amino acid side chains within the quaternary
315 complex provides strong constraints for modeling of a three-dimensional structure of the
316 XcpH_pI_pJ_pK_p complex.

317

318 **Integrating experimental data to build quaternary complex model**

319 *XcpH_pI_pJ_pK_p complex*

320 Data from spin labelling studies and acidic cross linking mass spectrometry were used as restraints
321 on the interaction of XcpH_p with XcpI_pJ_pK_p, to model the quaternary complex using HADDOCK
322 software for protein-protein docking. Docking used the structure of the XcpI_pJ_pK_p ternary complex
323 (PDB: 5VTM), the XcpH_p homology model, PRE experimental results and acidic cross-linking
324 interactions as input (**Supplemental Tables 3 & 4**). The best model obtained is in good agreement
325 with these experimental restraints (**Figure 7**). In this model, the α 1C of XcpH_p is in direct contact
326 with that of XcpJ_p, and runs approximately parallel to the helical bundle in the XcpI_pJ_pK_p complex
327 (**Figure 7**). XcpH_p interaction with the ternary complex is maintained by a large network of
328 electrostatic interactions and hydrogen bonds (**Figure 7**). One group of contacts corresponds to an
329 electrostatic interaction between helices, in particular including a salt bridge between side chain
330 XcpH_p D49 and XcpJ_p R42 (possibly also to XcpJ_p R46, given the rotamers available to Arg side
331 chains and the limitations of resolution of this model) as well as numerous main chain and side
332 chain H-bonds. A second group comprises β -strand to β -strand contacts. In this interaction, XcpH_p
333 loop β 6- β 7 G155 and F156 backbone carboxyl groups form hydrogen bonds with XcpJ_p β 7 R131
334 side chain and Q133 main chain, respectively. Other contacts observed between XcpH_p and XcpJ_p
335 correspond to: 1) XcpH_p loop α 1- β 1 and XcpJ_p loop β 4- β 5 (N55-R92), 2) XcpH_p loop β 4- β 5 and
336 XcpJ_p loop β 4- β 5 (S129-R92) (**Figure 7**). Remarkably, although no data from the chemical shift
337 perturbation experiments were explicitly included in the HADDOCK restraints, many of the
338 interactions in the model are recapitulated as chemical shift perturbations, for example the very
339 high signal interaction within XcpH_p β 7 is explained by its close approach to XcpJ_p.

340

341 The use of acidic cross-linking restraints for XcpH_p with the two other subunits from the
342 quaternary complex disambiguates the XcpH_p position *vis a vis* these proteins, for example the
343 crosslinks to XcpK_p help to position the tip of XcpH_p α 1C close to the XcpK_p β -sheet, allowing
344 additional contacts that stabilize the complex. These include XcpH_p-XcpK_p residues E50-R283,

345 D54-R283, R56-D282, R56-R284 and N55-G309 (**Figure 7**). We were able to further support this
346 quaternary structure by analyzing the co-evolution of residues in XcpH_p and XcpJ_p.
347 (**Supplemental Figure 4**). The analysis similarly points to XcpH polar amino acids at the tip of
348 α 1C (D49, E57) and in the β 4- β 5 loop (S128, S129, E131) and pairs them with two sets of XcpJ
349 co-evolved amino acids located within a large flap at the tip of XcpJ (R92 and R99, forming
350 electrostatic interactions), or in the vicinity of α 1C (V45, R189 and W107, potentially contributing
351 to packing geometry) (**Supplemental Figure 4**). This co-evolution result adds credence to the
352 robustness and physiological relevance of our XcpH:XcpJ interaction model. Altogether, we
353 present a model of the XcpH_pI_pJ_pK_p quaternary complex supported by experimental evidence,
354 which shows the interaction of XcpH with XcpJ and XcpK is maintained by several electrostatic
355 interactions.

356

357 *Filament model*

358 The quaternary complex presented here reveals the interaction of the minor pseudopilin soluble
359 domains, and in particular places XcpH_p in this ensemble. However, these proteins also interact
360 via their missing hydrophobic α helices (α 1N) in the biological context of the pseudopilus. Thus,
361 we used available experimental data and chemically reasonable restraints to model the complete
362 Type II secretion system pseudopilus. For that purpose, hydrophobic α 1N were added to the
363 XcpH_pI_pJ_pK_p proteins in the soluble quaternary complex model. In addition, the major pseudopilin
364 XcpG filament was modeled using as template the structure of PulG filament (PDB: 5WDA)¹⁵.
365 Then, the quaternary complex was added to the tip of the XcpG filament by aligning XcpH to the
366 first XcpG unit. Finally, the whole model was minimized using PyRosetta⁵². The best model was
367 selected based on the lowest RMSD with respect to the HADDOCK model (**Figure 8**). An
368 important feature included as a restraint in calculating this model was the salt bridge between N-
369 terminal amino groups and the carboxylic acid of E5 of each preceding unit in the filament,
370 including XcpHIJK at the tip of the pseudopilus. This contact neutralizes these charges in the
371 transmembrane helices as seen in other filament structures such as *P. aeruginosa* PAK pilus (PDB
372 5VXY). As expected, E5-F1 contacts continue into the minor pseudopilin complex in the following
373 sequence: XcpG-XcpH-XcpJ-XcpI-XcpK. This arrangement maintains the right-handed nature of
374 the major pseudopilin filament structure. In the filament model, the average distance between E5
375 carboxyl group and F1 amino group is 3.58 ± 0.34 Å between adjacent subunits. In addition,

376 Cisneros *et al.*³⁷ showed the close contact between the residues 16 and 10 of neighboring subunits
377 (PulJ-PulI and PulI-PulK). Even though these contacts were not used as restraints during
378 modelling, the current model reproduces them between the minor and major pseudopilins
379 throughout the filament (**Figure 8** and data not shown). Thus, the C α -C α distance from residue
380 10 to 16 of the neighbouring subunits in the PulG filament is 8.5 Å, in the model that distance is
381 on average 8.9 ± 1.4 Å. Strikingly, to maintain the restraint contacts, in particular the E5 to F1 salt
382 bridge between XcpH and XcpJ, it was necessary to allow XcpH α 1N between residues 20 and 26
383 to unravel. Although there is no other experimental evidence that the XcpH helix adopts this melted
384 secondary structure, the equivalent amino acids in the major T2SS pseudopilin and the major T4P
385 pilin are strikingly extended in high resolution filament models^{15,53}. Additionally, the presence of
386 a glycine at XcpH position 25 supports this conclusion, since glycine and proline residues
387 destabilize α helices.

388

389 Another feature observed in the filament model, specifically related to the minor pseudopilin tip,
390 is the capping of each XcpG protofilament by a minor pseudopilin and the off-center placement of
391 XcpK. If the major pseudopilin filament is separated into 4 protofilaments, each one is capped by
392 a minor pseudopilin (**Figure 8**). The exception to this is XcpJ; although the XcpJ globular domain
393 is directly above one protofilament, a small gap exists between XcpJ and the XcpG unit directly
394 below it. An additional consequence of capping one XcpG protofilament with XcpH is that the
395 XcpK globular domain protrudes beyond the major pseudopilin filament diameter, giving the
396 appearance of a hook-like structure to the filament tip (**Figure 8**).

397

398 While our model cannot be considered an atomic structure of the Type II secretion pseudopilus, it
399 provides a valuable path to testable predictions such as the importance of XcpH helix unravelling,
400 possible substrate binding sites available on the filament tip, the presence of dynamic contacts
401 between XcpG and minor pseudopilins, and the importance of the XcpK globular domain for
402 secretion.

403

404 **DISCUSSION**

405 Study of dynamic interactions prioritized by binding affinity is challenging but we have used
406 chemical, structural, microbiological, and computational approaches to describe the complete

407 ultrastructure of the T2SS pseudopilus and in particular to propose a specific and testable model
408 for the location of the adapter unit XcpH (GspH) between XcpIJK and the filament formed by
409 XcpG. Addition of XcpH would be the next step in a dynamic chronological pathway following
410 XcpIJK association at the periplasmic membrane. Subsequently, XcpG addition absolutely
411 requires the α 1N helices, as there is no interaction between the soluble domains of XcpH and
412 XcpG (**Supplemental Figure 5**), nor indeed between the soluble domain of XcpG and the soluble
413 domain of any of the quaternary complex soluble domains individually²⁷. We have amassed
414 independent and complementary data from several techniques in this integrative study. In
415 particular, the pioneering methodologies in acidic crosslinking represent an important technical
416 breakthrough in the experimental approach to studying protein:protein interactions in multipartner
417 complexes.

418

419 This study provides a rationale for the observation that at least in the *P. aeruginosa* Xcp T2SS,
420 XcpH is required for secretion. It serves as the adapter that connects the regular helical assembly
421 of XcpG subunits to the asymmetric complex that catalyzes initiation of the filament and interacts
422 with substrates and other components of the secretion system. It may be possible that XcpH serves
423 other as yet unverified roles in the system as well. For example, given its very strong structural
424 homology to the major XcpG subunit, XcpH might occasionally integrate into the filament and
425 thereby destabilize it, possibly initiating disassembly of the pseudopilus³¹.

426

427 Previous interaction studies using surface plasmon resonance²⁷ testing pairwise interactions of the
428 soluble domains of XcpG, H, I, J and K as well as our current NMR results (**Supplemental Figure**
429 **5**) reveal that XcpH interacts exclusively with XcpJ. In addition, the binding of XcpH to XcpJ is
430 more efficient when XcpJ is engaged in the XcpIJK ternary complex (Figures 2 & 3 in ²⁷). These
431 data suggested that conformational changes in XcpJ upon binding to XcpI and XcpK create a more
432 favorable docking platform for XcpH and/or favor the downward addition of XcpH to the ternary
433 complex over that of XcpG when both are available within the inner membrane. Our current data
434 supporting many pairwise interactions among the quaternary complex proteins, including between
435 XcpH and each of the other three, fit well with the hypothesis that alone the soluble domains have
436 weak and transient interactions with each other which in some cases are below detection but that
437 within the full ensemble, these interactions are revealed. Our model furthermore accentuates that

438 the central α helices are, not surprisingly, fundamentally important for more stable interactions.
439 Future experiments will rely on the α 1N helical tails being part of the interactions.

440

441 Based on the near atomic resolution of the XcpQ secretin¹² and taking into account that the
442 pseudopilus tip physically interacts with the secretin³³, a satisfying outcome of our data-driven
443 modeling of the complete pseudopilus is the observation that the filament (~6 nm) docks readily
444 into the periplasm-facing vestibule (~8 nm)¹² of the XcpD secretin (**Supplemental Figure 6**).
445 Moreover, the positioning of the Xcp pseudopilus plus secretin complex into the cryo-tomographic
446 map of the complete T2SS in its natural context, the bacterial envelope²³, indicates that 8 to 16
447 pseudopilin subunits in addition to the tip complex are required to span the periplasm between the
448 IM and the entry into, or the internal periplasmic gate of, the secretin interior, respectively
449 (**Supplemental Figure 6**).

450

451 The results presented here provide insights into not only the *P. aeruginosa* T2SS but indeed into
452 the dozens of T2SS recovered in proteobacteria with similar genome content of five pre-
453 pseudopilins equivalent to XcpG, H, I, J, and K². The conservation of amino acid sequences of
454 these proteins across species implies their protein:protein interaction interfaces are also conserved,
455 and coevolved as shown here for the XcpH_p:XcpJ_p interface. By extension, both the minor
456 pseudopilins of the T2SS and the minor pilins of the T4P can be expected to form a similar
457 quaternary complex in which an adapter minor (pseudo)pilin subunit sets the stage for downward
458 addition of the major (pseudo)pilin subunit to form the extended filament by interacting via its
459 periplasmic domain with the three upper minor subunits and via its unraveled α 1C with the first
460 major subunit. Correspondingly, the set of *E. coli* T4P minor pilins restore assembly of the major
461 T2SS pseudopilin PulG in absence of minor pseudopilins PulHJK³⁷. An open question is whether
462 the secretion ATPase motor protein, GspE, is required for insertion and/or α 1C unraveling in this
463 adapter minor pilin at the junction between the tip complex and the filament itself.

464 **ONLINE MATERIALS AND METHODS**465 ***Gene and protein nomenclature***

466 Here we have extended the general T2SS nomenclature to the *P. aeruginosa* Xcp T2SS
467 components. Thus, former XcpT, U, V, W and X proteins are now labelled XcpG, H, I, J and K in
468 agreement with their homologs in other T2SSs. All pseudopilin soluble domains (deleted for their
469 N-terminal hydrophobic domains (α 1N)) used in this study are designated with a subscript p
470 indicating periplasmic. According to the standard in the field, polypeptides are numbered with
471 amino acid 1 as the first residue in the mature pseudopilin following prepilin peptidase cleavage.

472

473 ***Bacterial strains and plasmids***

474 *Escherichia coli* K-12 DH5 α (laboratory collection) and BL21(DE3) pLysS (laboratory collection)
475 were used for cloning procedures and soluble protein production, respectively. *Pseudomonas*
476 *aeruginosa* PAO1 wild type (laboratory collection), PAO1 $\Delta xcpH$ (this study), $\Delta xcpI$ (this study),
477 $\Delta xcpJ^{43}$, and $\Delta xcpK^{26}$ or PAO1 Δxcp (also called DZQ40)⁵⁴ strains were used for *in vivo*
478 complementation assay. Construction of the *xcpH* and *xcpI* *P. aeruginosa* PAO1 deletion strains
479 was performed as described previously¹⁶ using the pKN- Δ H and pKN- Δ I mutator plasmids (**Table**
480 **S1**).

481

482 Plasmids and oligonucleotides used in this study are listed in **Table S1** and **Table S2**, respectively.
483 Site directed mutagenesis was performed using quick change technology (Stratagen) or inverse
484 PCR. The gene encoding the XcpH_p protein deleted for its mouse trap domain was cloned into the
485 pJN105 vector using SLIC technology⁵⁵. Plasmid pET-XcpH_p was constructed following the
486 strategy used by Durand et al.¹⁶ to construct pET-XcpG_p. All constructs have been verified by
487 DNA sequencing.

488

489 ***LasB secretion and protease activity on plates.***

490 Preparation of culture supernatants from *P. aeruginosa* for analysis of secreted proteins has been
491 described⁵⁶. Gel analysis was standardized so that the volume of supernatant equivalent to 2 OD₆₀₀
492 units of culture was used for each sample (**Figures 1a, 1c**). Protease activity was tested by spotting
493 5 μ l of bacterial culture grown in LB to early stationary phase on TSA plates containing 1.5%
494 lyophilized milk and appropriate antibiotics, followed by 14 h incubation at 30°C.

495

496 ***Protein production and purification***

497 Production and purification of XcpH_p and XcpJ_p with or without cysteine substitutions for NMR
498 purposes (except XcpG_p/H_p NMR CSP, see Supplemental Material) was carried out with minor
499 modifications from published procedures³². Expression was performed in BL21(DE3) pLysS *E.*
500 *coli* carrying the plasmids pETG-20A-XcpH_p or pETG-20A-XcpJ_p. For ¹³C and ¹⁵N uniform
501 protein labeling, cells were grown in 3 L of LB at 37°C to OD₆₀₀ ~0.6. Cells were collected and
502 washed with M9 media without glucose or ammonium chloride to remove remaining LB media.
503 After washing, cells were transferred to 1 L M9 media supplemented with 1 g of ¹⁵N-U ammonium
504 chloride and 2 g of ¹³C-U glucose (Cambridge Isotope Laboratories) and incubated at 25°C for 2
505 hours before inducing with 1 mM IPTG overnight. Purification of XcpH_p proceeded using a
506 combination of nickel affinity and size exclusion chromatography. The cell pellet was resuspended
507 in Buffer A (50 mM tris-HCl buffer pH 8.0 containing 300 mM NaCl) and lysed using a French
508 press. Supernatant from lysate clarified by centrifugation at 25,000 x G for 30 min. was loaded
509 into a 5 mL Ni-NTA column (Quiagen) and washed with 100 mL of Buffer A containing 50 mM
510 imidazole. Protein was eluted from the column using Buffer A with 500 mM imidazole. Peak
511 fractions containing target protein were pooled and treated with TEV protease (purified in house)
512 added to a final concentration of ~40 µg/mL and the sample was dialyzed overnight at 4°C against
513 Buffer A to remove imidazole. Removal of TEV protease proceeded by loading samples into a 1
514 mL Ni-NTA column which was washed with the same buffer. Fractions containing XcpH_p were
515 collected and concentrated for size exclusion chromatography in a Sephacryl S-100 HiPrep 16/60
516 column (GE). During this step, buffer was exchanged to NMR buffer (25 mM sodium phosphate
517 buffer pH 6.5 containing 25 mM NaCl). XcpJ_p was expressed and purified using the identical
518 protocol but without ¹³C and ¹⁵N labeling.

519

520 Production and purification of XcpH_p, XcpI_p, XcpJ_p, XcpK_p, XcpH_pV46C, XcpH_pD49C, XcpH_pL53C,
521 XcpJ_pR46C and XcpJ_pR53C for cysteine cross-linking (and cross-linking MS experiments was
522 performed in BL21(DE3) pLysS *E. coli* carrying the corresponding pETG-20A plasmid (**Table**
523 **S1**) as published.²⁷

524

525 ***In vitro cysteine cross-linking***

526 40 μ M of a single cysteine variant or two variant proteins in 1:1 molar ratio were incubated in
527 Tris-HCl 50 mM, NaCl 100mM, pH 8 in a total volume of 500 μ L. The mixture was supplemented
528 with DTT (20 mM) to allow reduction of intra-chain disulfide bands formed during purification
529 steps. The mixture was dialyzed against 300 mL of Tris-HCl 50 mM, NaCl 100 mM, pH 8 buffer
530 for 2 h at RT to allow DTT removal and cysteine oxidation. To analyze disulfide bond formation,
531 75 μ L of each reaction was mixed with 25 μ L of Leammli loading buffer with or without reductant
532 as appropriate. The samples were heated for 5 min at 95°C and then visualized by 12% Coomassie
533 blue stained SDS-PAGE.

534

535 ***Acidic Crosslinking of XcpH_p + XcpJ_p and XcpH_p + XcpI_p + XcpJ_p + XcpK_p***

536 *Preparation of multimers* For assembling the dimer, 17 μ L of XcpH_p at 4 mg/mL were mixed to
537 25 μ L of XcpJ_p at 4 mg/ml, leading to a mixture of 4 nmol of each protein (molar ratio 1:1). For
538 the tetramer, 17 μ L of XcpH_p (4 mg/mL), 12 μ L of XcpI_p (4 mg/mL), 25 μ L of XcpJ_p (4 mg/mL)
539 and 34 μ L of XcpK_p (4 mg/mL) were mixed to yield a 4 nmol solution of each protein (molar ratio
540 1:1:1:1). Samples were dried under vacuum, and dissolved in 100 μ L of PBS buffer pH 7.4 to
541 achieve a final concentration of each protein of 40 μ M. Proteins were kept at 25°C for 1 h to let the
542 association of the dimer or tetramer occur.

543

544 *Cross-linking reaction* To crosslink the dimer, 10 μ L of ADH (adipic acid dihydrazide) and 16 μ L
545 of DMTMM ((4-(4,6-dimethoxy-1,3,5-triazin-2-yl)-4-methyl-morpholinium chloride) were added
546 to the solution of XcpH_p:J_p complex, reaching final concentrations of ADH and DMTMM equal
547 to 46 mM (8 and 12.7 mg/mL, respectively). For the tetramer, 23 μ L of ADH and 36 μ L of
548 DMTMM were added to the solution of XcpH_p:I_p:J_p:K_p complex, reaching final concentrations of
549 ADH and DMTMM equal to 81 mM (14.5 and 22.6 mg/mL, respectively). These concentrations
550 were chosen to impose a molar excess of crosslinking reagent of more than 1,000. The mixture
551 was incubated 2 h at 37°C under agitation (750 rpm) to allow intensive crosslinking reaction. The
552 reaction was quenched by reagent removal using a ZebaSpinDesalting Column (0.5mL, 7k,
553 Pierce), employed according to the manufacturer recommendations. The resulting sample was
554 dried using a Speedvac system.

555

556 *SDS-PAGE* Cross-linked species were imaged by classical SDS-PAGE analysis. The protein
557 mixture was suspended in 170 μ L of 8M urea, leading to a global protein concentration of 48 μ M.
558 20 μ g of proteins (=10 μ L) were dried before being solubilized in 20 μ L of Laemmli buffer (Tris
559 HCl pH 6.8 (65mM), SDS 2%, glycerol 20% and DTT 350mM) and a spatula tip of bromophenol
560 blue. The sample was heated at 100°C for 5 min, and 10 μ L (~10 μ g) were loaded on the gel
561 (NuPageTM 4-12% Bis-Tris-Gel, 1.0 mm x 10 wells). The migration starts with a voltage of 200V
562 applied for 40min (stacking), and then 150V for the separation. The electrophoresis system is
563 switched-off when the migration blue line is localized at the extremity of the gel. The gel is fixed
564 3h in a bath composed of 50% ethanol, 47% water and 3% phosphoric acid, washed three times
565 with ultrapure water, and finally incubated in water containing 34% methanol, 17% ammonium
566 sulfate and 3% phosphoric acid. The coloration of the bands was made during 3 days with
567 Coomassie blue G250 added at 360 mg/l in the solution. The gel was finally washed several times
568 in pure water to remove the excess of Coomassie blue.

569

570 *In-gel digestion* The bands corresponding to the XcpH_p:XcpJ_p dimer or
571 XcpH_p:XcpI_p:XcpJ_p:XcpK_p tetramer, respectively, were excised, and cut into small pieces.
572 Resulting pieces of gel were then washed by 50 μ L of 50mM NH₄HCO₃ then centrifuged (5min,
573 600 rpm) to remove the supernatant. The same step was repeated with 50 μ L of 50/50 (v/v) 50mM
574 NH₄HCO₃/acetonitrile. The two previous steps were repeated twice. To ensure a good penetration
575 of the enzymes into the pieces of gel, the latter were dehydrated twice with 50 μ L of pure
576 acetonitrile removed by centrifugation (5 min, 600 rpm). The gel was then rehydrated at 0°C with
577 3 μ L of a solution containing 1/100 of Lys-C and 1/50 of trypsin in 50mM NH₄HCO₃. The digestion
578 was performed at 37°C for 4 hr. Enzymatic activity was quenched by acidifying the medium using
579 25 μ L of 1% trifluoroacetic acid. Resulting peptides were eluted from the gel by incubating the
580 sample overnight at 20°C under 600 rpm.

581

582 *LC-MS/MS* 1 μ g of the digested material was analysed using a UPLC nanoACQUITY (Waters,
583 UK) coupled to a Q-Exactive Plus Hybrid Quadrupole-Orbitrap Mass Spectrometer (Thermo
584 Scientific, USA). The chromatographic system is equipped with two columns. The first one,
585 dedicated to the trapping of peptides, is a Symmetry C18 (5 μ m, 180 μ m x 20 mm, Waters, UK).
586 The second one which performs the analytical separation is a HSST3 C18 (1.8 μ m, 75 μ m x 250

587 mm, Waters). The dimensions given for the columns are in the order: particle diameters, internal
588 diameters and column lengths. Solvent A was water, acidified with 0.1% formic acid and solvent
589 B was acetonitrile, also acidified with 0.1% formic acid. Cross-linked peptides were first trapped
590 for 3min (98/2, v/v, A/B) at a flow rate of 20 μ L/min before being eluted with a gradient of 57 min
591 at a flow rate of 700 nL/min. The elution started with a linear gradient of B from 2% to 7% in
592 5min, followed by an increase from 7% to 40% in 25min, then to 85% in 3min. This composition
593 A/B 15/85 is kept for 5min before changing to 98/2 for reconditioning the analytical column. The
594 Q-Exactive Plus spectrometer was set in a nanoESI positive mode acquisition for 57 minutes. The
595 acquisition was recorded in full scan MS and data dependent MS/MS, in a mass range m/z 400-
596 1,750. For the MS stage, resolving power was set at 70,000 @m/z 200, with an automatic gain
597 control (AGC) target at 1e6 (or 50 ms as a maximum injection time). For MS/MS, a “Top 12”
598 experiments was applied, meaning that the twelve most intense ions of each MS scan were selected
599 for fragmentation. Singly charged ions, ions with undetermined charge (for example, electronic
600 noise) and ions with signal intensities below the AGC threshold set at 1e3 were excluded from this
601 selection. For precursor ions, the selection window was 2.0 m/z, the AGC target was 1e5 (or 50ms
602 as a maximum injection time) and the resolving power of 17,500 @m/z 200. Normalized collision
603 energy was 25. A dynamic exclusion of 10s was also applied to avoid the redundancy of MS/MS
604 spectra of the same ions.

605 - Data Analysis: Spectrum Identification Machine SIM-XL version 1.5.0.14⁵⁷ was used for
606 identification of cross-linked peptides. ADH-DMTMM crosslinker was set up in the software to
607 create a mass shift of 138.0905 Da for each crosslink and 156.1012 for hydrolyzed monolinks.
608 The possible reaction sites were restricted to C-terminal extremities and acidic side chains of Glu
609 and Asp amino acids. Accuracy on mass measurements was 2 ppm for the precursor and 10 ppm
610 for the fragment ions, to reduce to their maximum the false positives and to increase the
611 reliability of the results. Oxidation of methionine, due to experimental conditions, has also been
612 considered as a variable modification. Only the crosslinked peptides characterized by a SIM-XL
613 internal score of 2.5, 3.0, 2.3, 3.0 or 2.5 and higher have been retained for IK, IJ, HJ, HI, or HK
614 pairs, respectively. All the detected crosslinked peptides have been manually validated by
615 analyzing their corresponding MS/MS spectra for exact masses, numbers of identified fragments,
616 and signal intensity (*e.g.* **Figure 6b**, **Supplemental Figure 3b**).

617 *NMR experiments.*

618 All solution NMR experiments (except XcpH_p/J_p chemical shift perturbation) were run at 37°C in
619 a Bruker 900 MHz NMR spectrometer at the National Magnetic Resonance Facility at Madison.
620 Sample condition used was NMR buffer containing 0.01% sodium azide, 50 μM 2,2-dimethyl-2-
621 silapentane-5-sulfonate (DSS) (Sigma) and 8% D₂O.

622 Backbone sequential assignments were carried out using standard 3D NMR experiments including
623 HNCO, HN(CA)CO, HNCACB, CACB(CO)NH. These experiments were collected using non-
624 uniform sampling (NUS). Data processing of NUS data was performed using NMRPipe⁵⁸.
625 Chemical shift referencing was done using DSS as reference. Backbone assignments were
626 facilitated using NMRFAM-Sparky software⁵⁹. Calculation of secondary chemical shifts for
627 secondary structure estimation was done using random coil chemical shifts calculated for the XcpH
628 amino acid sequence using ncIDP library⁶⁰. Secondary chemical shifts were calculated as $\Delta\delta\text{Ca} -$
629 $\Delta\delta\text{Cb}$, where $\Delta\delta\text{Ca} = (\delta\text{Ca}_{\text{XcpH}} - \delta\text{Ca}_{\text{random coil}})$ and $\Delta\delta\text{Cb} = (\delta\text{Cb}_{\text{XcpH}} - \delta\text{Cb}_{\text{random coil}})$. This step
630 effectively removes any error in chemical shift referencing.

631 Chemical shift perturbation assays were carried out using 100 μM ¹⁵N-U XcpH and increasing
632 amounts of unlabeled XcpJ (20 μM, 50 μM, and 80 μM). An independent ¹H-¹⁵N HSQC spectrum
633 was collected for each sample. Total change of the amide proton and nitrogen chemical shift
634 ($\Delta\delta\text{NH}$) was calculated for each residue as $((\Delta\delta\text{H})^2 - 1/6 (\Delta\delta\text{N})^2)^{1/2}$, where $\Delta\delta\text{H}$ and $\Delta\delta\text{N}$ correspond
635 to the difference in chemical shift of proton and nitrogen respectively. $\Delta\delta\text{NH}$ was calculated using
636 the control spectrum as the initial point and the spectrum with 80 μM XcpJ as the endpoint.

637 To perform paramagnetic relaxation enhancement (PRE) experiments, four different XcpJ_p
638 cysteine mutant constructs (R46C, R53C, T178C, E180C) were used to attach the spin label (1-
639 Oxyl-2,2,5,5-tetramethylpyrroline-3-methyl) methanethiosulfonate (MTSL) (Santa Cruz
640 Biotechnology). XcpJ_p variants were purified as above in Buffer A containing 1 mM DTT. MTSL
641 labeling of XcpJ_p was performed in NMR buffer and 0.5 mM DTT. XcpJ_p samples were incubated
642 with a 20x molar excess of MTSL for 4 h at room temperature. Additional 20x molar excess of
643 MTSL was added to the sample which was further incubated at room temperature overnight.
644 Excess MTSL was removed by dialysis against the NMR buffer. The NMR sample contained 80
645 μM ¹⁵N-U XcpH_p and 24 μM of MTSL labeled XcpJ_p. The ratio of XcpH to XcpJ_p was selected
646 based on the chemical shift perturbation assay to avoid signal loss due to complex formation. A
647 2D ¹H-¹⁵N HSQC spectrum was acquired for each sample. After collecting the HSQC experiment
648 of the oxidized or paramagnetic form of MTSL, 2 mM sodium ascorbate (Sigma-Aldrich) was

649 added to the sample and incubated at room temperature for at least one hour before collecting the
650 spectrum of the reduced or diamagnetic MTSL form. Effects of the paramagnetic label were
651 quantified as the ratio of signal intensity of the spectrum collected with MTSL in paramagnetic
652 versus diamagnetic form ($I_{\text{para}}/I_{\text{dia}}$).

653 ***Quaternary complex computational modeling***

654 Modeling of the XcpH_pI_pJ_pK_p quaternary complex was achieved using high ambiguity driven
655 protein-protein Docking (HADDOCK)⁶¹. Structures used for HADDOCK corresponded to the
656 XcpI_pJ_pK_p crystal structure (PDB: 5VTM) and an XcpH_p model prepared based on the structure of
657 EpsH_p from *Vibrio cholerae* (PDB: 2QV8) using the Phyre2 server⁶². Allowed distances for acidic
658 crosslinking restraints between C_α carbons were set between 8-18 Å, while 6-18 Å were used for
659 PRE restraints (**Supplemental Tables 3 & 4**). Docking the XcpH_p model onto the XcpI_pJ_pK_p
660 structure was performed using the HADDOCK 2.2 server⁶¹. The best structure obtained from
661 HADDOCK was selected for further analysis of the XcpH_pI_pJ_pK_p complex.

662

663 ***Biological filament modeling***

664 Modeling of the XcpG Type II secretion system pseudopilus was based on the PulG filament
665 structure from *Klebsiella oxytoca*¹⁵ (PDB code 5WDA) using Pymol. Modeling was divided in the
666 following steps: (1) adding missing transmembrane helices to XcpH_pI_pJ_pK_p proteins, (2)
667 positioning the XcpHIJK complex onto the PulG filament (3) fitting XcpHIJK helices to the
668 position of PulG transmembrane helices within the filament (4) fitting XcpG⁶³ units to the PulG
669 filament and (5) relaxing the whole system to remove atomic clashes. Missing transmembrane
670 helices in the XcpH_pI_pJ_pK_p complex were initially modeled using the Phyre2 server using the
671 respective amino acid sequence of the complete alpha helix. This modeled alpha helix was aligned
672 to the soluble domain alpha helix structure in the HADDOCK model. Specific to XcpH, the helix
673 was modeled on the PulG template from the EM reconstruction to include the unraveled section.
674 Finally, the modeled helices were added to the XcpH_pI_pJ_pK_p complex PDB file using custom
675 Python scripts. For XcpG a similar procedure was performed, in this case a model of the
676 transmembrane helix was based on PulG helix (including the unraveled helix section), which was
677 added to the XcpG_p solution NMR structure⁶³ (PDB code 2KEP). The XcpHIJK complex with the
678 helices added was positioned at the tip of PulG filament by aligning the XcpH soluble domain with

679 the first PulG unit. Then, XcpHIJK model transmembrane helices were modeled semi-manually
680 to fit the interior cavity of the PulG filament. Custom Python scripts were used to modify ϕ and ψ
681 for this purpose. To guide the modeling, the transmembrane helices of PulG units were used as a
682 reference. Finally, PulG subunits were replaced with XcpG to complete the pseudopilus filament.
683 To achieve this, XcpG with the added helix was aligned to a PulG unit and XcpG helix was
684 modeled to closely match the PulG transmembrane helix. Then, the modeled XcpG unit was copied
685 and aligned to the PulG units to complete the filament. Finally, the whole model was relaxed with
686 pyRosetta⁵². During this procedure, the known hydrogen bond between E5 carboxyl group and N-
687 terminal amino group of the next subunit in the filament was enforced as a distance restraint of 2.8
688 Å. Other restraints to maintain the quaternary tip complex and XcpG filament assembly are shown
689 **Supplemental Tables 5 & 6**. Ten filament models were produced, from which the best was
690 selected based on favourable XcpG subunit packing and the lowest RMSD of the XcpH_pI_pJ_pK_p tip
691 structure with respect the HADDOCK model.

692

693 **ACKNOWLEDGMENTS**

694 NMR experimental data have been deposited to the BMRB with accession number 50449. We
695 thank Dr. Eric Durand for pET-XcpG_p plasmid construction, Mrs. Nanou Tanteliarisoa for her
696 help in the preparation of the samples dedicated to the acidic crosslinking, and Dr. Lisa Craig for
697 valuable comments on the manuscript. We acknowledge funding from the ANR to RV (ANR-14-
698 CE09-0027-01 & ANR-19-CE11-0020-01) and from the US National Science Foundation to
699 KTF (IOS 1353674). Mass spectrometers used in this work were supported by the Walloon
700 Region (Belgium) and the European Regional Development Fund. This study made use of the
701 National Magnetic Resonance Facility at Madison, which is supported by NIH grant
702 P41GM103399 (NIGMS), old number: P41RR002301. Equipment was purchased with funds
703 from the University of Wisconsin-Madison, the NIH P41GM103399, S10RR02781,
704 S10RR08438, S10RR023438, S10RR025062, S10RR029220), the NSF (DMB-8415048, OIA-
705 9977486, BIR-9214394), and the USDA.

706

707 **AUTHOR CONTRIBUTIONS**

708 RV, LQ, and KTF conceived the studies. CAE and BD subcloned, made site directed mutants, and
709 purified soluble proteins. CAE carried out CSP and PRE experiments, and all quaternary structure

710 and filament modeling. BD performed cysteine cross linking experiments. GB constructed *P.*
711 *aeruginosa* strains and carried out secretion assays. SA carried out NMR experiments in Supp. Fig
712 S4. BB was responsible for coevolution calculations in Supp. Fig S5. LQ performed all acidic
713 cross linking methods development and experiments. RV, KTF, and CAE drafted the manuscript.
714 All authors discussed and edited manuscript and prepared figures.

715

716 **COMPETING INTERESTS** The authors declare no competing interests.

717

718 REFERENCES

- 719 1. Korotkov, K. V. & Sandkvist, M. Architecture, Function, and Substrates of the Type II
720 Secretion System. *EcoSal Plus* **8**, (2019).
- 721 2. Cianciotto, N. P. & White, R. C. Expanding Role of Type II Secretion in Bacterial
722 Pathogenesis and Beyond. *Infect. Immun.* **85**, (2017).
- 723 3. Sysoeva, T. A., Zepeda-Rivera, M. A., Huppert, L. A. & Burton, B. M. Dimer recognition and
724 secretion by the ESX secretion system in *Bacillus subtilis*. *Proc. Natl. Acad. Sci. U. S. A.* **111**,
725 7653–7658 (2014).
- 726 4. Lasica, A. M., Ksiazek, M., Madej, M. & Potempa, J. The Type IX Secretion System (T9SS):
727 Highlights and Recent Insights into Its Structure and Function. *Front. Cell. Infect. Microbiol.*
728 **7**, 215 (2017).
- 729 5. Evdokimov, A. G. *et al.* Similar modes of polypeptide recognition by export chaperones in
730 flagellar biosynthesis and type III secretion. *Nat. Struct. Biol.* **10**, 789–793 (2003).
- 731 6. Trokter, M. & Waksman, G. Translocation through the Conjugative Type IV Secretion
732 System Requires Unfolding of Its Protein Substrate. *J. Bacteriol.* **200**, (2018).
- 733 7. Thomassin, J.-L., Santos Moreno, J., Guilvout, I., Tran Van Nhieu, G. & Francetic, O. The
734 trans-envelope architecture and function of the type 2 secretion system: new insights raising
735 new questions. *Mol. Microbiol.* **105**, 211–226 (2017).

- 736 8. Gu, S., Shevchik, V. E., Shaw, R., Pickersgill, R. W. & Garnett, J. A. The role of intrinsic
737 disorder and dynamics in the assembly and function of the type II secretion system. *Biochim*
738 *Biophys Acta Proteins Proteom* **1865**, 1255–1266 (2017).
- 739 9. Hu, N.-T. *et al.* XpsG, the major pseudopilin in *Xanthomonas campestris* pv. *campestris*,
740 forms a pilus-like structure between cytoplasmic and outer membranes. *Biochem. J.* **365**, 205–
741 211 (2002).
- 742 10. Berry, J. L. & Pelicic, V. Exceptionally widespread nanomachines composed of type IV
743 pilins: the prokaryotic Swiss Army knives. *FEMS Microbiol Rev* **39**, 134–54 (2015).
- 744 11. Yan, Z., Yin, M., Xu, D., Zhu, Y. & Li, X. Structural insights into the secretin
745 translocation channel in the type II secretion system. *Nat Struct Mol Biol* **24**, 177–183 (2017).
- 746 12. Hay, I. D., Belousoff, M. J. & Lithgow, T. Structural Basis of Type 2 Secretion System
747 Engagement between the Inner and Outer Bacterial Membranes. *mBio* **8**, e01344-17 (2017).
- 748 13. Hay, I. D., Belousoff, M. J., Dunstan, R. A., Bamert, R. S. & Lithgow, T. Structure and
749 Membrane Topography of the Vibrio-Type Secretin Complex from the Type 2 Secretion
750 System of Enteropathogenic *Escherichia coli*. *J Bacteriol* **200**, (2018).
- 751 14. Chernyatina, A. A. & Low, H. H. Core architecture of a bacterial type II secretion
752 system. *Nat. Commun.* **10**, 5437 (2019).
- 753 15. Lopez-Castilla, A. *et al.* Structure of the calcium-dependent type 2 secretion pseudopilus.
754 *Nat Microbiol* **2**, 1686–1695 (2017).
- 755 16. Durand, E. *et al.* XcpX controls biogenesis of the *Pseudomonas aeruginosa* XcpT-
756 containing pseudopilus. *J Biol Chem* **280**, 31378–89 (2005).

- 757 17. Vignon, G. *et al.* Type IV-Like Pili Formed by the Type II Secretion: Specificity,
758 Composition, Bundling, Polar Localization, and Surface Presentation of Peptides. *J. Bacteriol.*
759 **185**, 3416–3428 (2003).
- 760 18. Nivaskumar, M. *et al.* Pseudopilin residue E5 is essential for recruitment by the type 2
761 secretion system assembly platform. *Mol. Microbiol.* **101**, 924–941 (2016).
- 762 19. Craig, L., Forest, K. T. & Maier, B. Type IV pili: dynamics, biophysics and functional
763 consequences. *Nat. Rev. Microbiol.* (2019) doi:10.1038/s41579-019-0195-4.
- 764 20. Parge, H. E. *et al.* Structure of the fibre-forming protein pilin at 2.6 Å resolution. *Nature*
765 **378**, 32–38 (1995).
- 766 21. Pasloske, B. L., Scraba, D. G. & Paranchych, W. Assembly of mutant pilins in
767 *Pseudomonas aeruginosa*: formation of pili composed of heterologous subunits. *J. Bacteriol.*
768 **171**, 2142–2147 (1989).
- 769 22. McLaughlin, L. S., Haft, R. J. F. & Forest, K. T. Structural insights into the Type II
770 secretion nanomachine. *Curr. Opin. Struct. Biol.* **22**, 208–216 (2012).
- 771 23. Ghosal, D. *et al.* In vivo structure of the Legionella type II secretion system by electron
772 cryotomography. *Nat. Microbiol.* **4**, 2101–2108 (2019).
- 773 24. Nunn, D. N. & Lory, S. Cleavage, methylation, and localization of the *Pseudomonas*
774 *aeruginosa* export proteins XcpT, -U, -V, and -W. *J. Bacteriol.* **175**, 4375–82 (1993).
- 775 25. Bally, M. *et al.* Protein secretion in *Pseudomonas aeruginosa*: characterization of seven
776 xcp genes and processing of secretory apparatus components by prepilin peptidase. *Mol.*
777 *Microbiol.* **6**, 1121–1131 (1992).
- 778 26. Bleves, S. *et al.* The secretion apparatus of *Pseudomonas aeruginosa*: identification of a
779 fifth pseudopilin, XcpX (GspK family). *Mol. Microbiol.* **27**, 31–40 (1998).

- 780 27. Douzi, B. *et al.* The XcpV/GspI pseudopilin has a central role in the assembly of a
781 quaternary complex within the T2SS pseudopilus. *J. Biol. Chem.* **284**, 34580–34589 (2009).
- 782 28. Zhang, Y. *et al.* Structure-guided disruption of the pseudopilus tip complex inhibits the
783 Type II secretion in *Pseudomonas aeruginosa*. *PLoS Pathog.* **14**, e1007343 (2018).
- 784 29. Lucas, C. E., Brown, E. & Fields, B. S. Type IV pili and type II secretion play a limited
785 role in *Legionella pneumophila* biofilm colonization and retention. *Microbiol. Read. Engl.*
786 **152**, 3569–3573 (2006).
- 787 30. Maruno, T. *et al.* Homo-trimeric Structure of the Type IVb Minor Pilin CofB Suggests
788 Mechanism of CFA/III Pilus Assembly in Human Enterotoxigenic *Escherichia coli*. *J. Mol.*
789 *Biol.* **428**, 1209–1226 (2016).
- 790 31. Ng, D. *et al.* The *Vibrio cholerae* Minor Pilin TcpB Initiates Assembly and Retraction of
791 the Toxin-Coregulated Pilus. *PLoS Pathog.* **12**, e1006109 (2016).
- 792 32. Douzi, B., Ball, G., Cambillau, C., Tegoni, M. & Voulhoux, R. Deciphering the Xcp
793 *Pseudomonas aeruginosa* type II secretion machinery through multiple interactions with
794 substrates. *J. Biol. Chem.* **286**, 40792–40801 (2011).
- 795 33. Reichow, S. L., Korotkov, K. V., Hol, W. G. J. & Gonen, T. Structure of the cholera
796 toxin secretion channel in its closed state. *Nat. Struct. Mol. Biol.* **17**, 1226–1232 (2010).
- 797 34. Treuner-Lange, A. *et al.* PilY1 and minor pilins form a complex priming the type IVa
798 pilus in *Myxococcus xanthus*. *Nat. Commun.* **11**, 5054 (2020).
- 799 35. Winther-Larsen, H. C. *et al.* A conserved set of pilin-like molecules controls type IV
800 pilus dynamics and organelle-associated functions in *Neisseria gonorrhoeae*. *Mol. Microbiol.*
801 **56**, 903–917 (2005).

- 802 36. Giltner, C. L., Habash, M. & Burrows, L. L. *Pseudomonas aeruginosa* minor pilins are
803 incorporated into type IV pili. *J. Mol. Biol.* **398**, 444–461 (2010).
- 804 37. Cisneros, D. A., Bond, P. J., Pugsley, A. P., Campos, M. & Francetic, O. Minor
805 pseudopilin self-assembly primes type II secretion pseudopilus elongation. *EMBO J.* **31**,
806 1041–1053 (2012).
- 807 38. Nguyen, Y. *et al.* *Pseudomonas aeruginosa* minor pilins prime type IVa pilus assembly
808 and promote surface display of the PilY1 adhesin. *J. Biol. Chem.* **290**, 601–611 (2015).
- 809 39. Yanez, M. E., Korotkov, K. V., Abendroth, J. & Hol, W. G. Structure of the minor
810 pseudopilin EpsH from the Type 2 secretion system of *Vibrio cholerae*. *J Mol Biol* **377**, 91–
811 103 (2008).
- 812 40. Korotkov, K. V. & Hol, W. G. J. Structure of the GspK-GspI-GspJ complex from the
813 enterotoxigenic *Escherichia coli* type 2 secretion system. *Nat. Struct. Mol. Biol.* **15**, 462–468
814 (2008).
- 815 41. Yanez, M. E., Korotkov, K. V., Abendroth, J. & Hol, W. G. The crystal structure of a
816 binary complex of two pseudopilins: EpsI and EpsJ from the type 2 secretion system of *Vibrio*
817 *vulnificus*. *J Mol Biol* **375**, 471–86 (2008).
- 818 42. Lam, A. Y., Pardon, E., Korotkov, K. V., Hol, W. G. J. & Steyaert, J. Nanobody-aided
819 structure determination of the EpsI:EpsJ pseudopilin heterodimer from *Vibrio vulnificus*. *J.*
820 *Struct. Biol.* **166**, 8–15 (2009).
- 821 43. Franz, L. P. *et al.* Structure of the minor pseudopilin XcpW from the *Pseudomonas*
822 *aeruginosa* type II secretion system. *Acta Crystallogr Biol Crystallogr* **67**, 124–30 (2011).
- 823 44. Denise, R., Abby, S. S. & Rocha, E. P. C. The Evolution of Protein Secretion Systems by
824 Co-option and Tinkering of Cellular Machineries. *Trends Microbiol.* **28**, 372–386 (2020).

- 825 45. Sheppard, D. *et al.* The major subunit of widespread competence pili exhibits a novel and
826 conserved type IV pilin fold. *J. Biol. Chem.* **295**, 6594–6604 (2020).
- 827 46. Dubnau, D. & Blokesch, M. Mechanisms of DNA Uptake by Naturally Competent
828 Bacteria. *Annu. Rev. Genet.* **53**, 217–237 (2019).
- 829 47. Albers, S.-V. & Jarrell, K. F. The Archaeallum: An Update on the Unique Archaeal
830 Motility Structure. *Trends Microbiol.* **26**, 351–362 (2018).
- 831 48. Braun, T. *et al.* Archaeal flagellin combines a bacterial type IV pilin domain with an Ig-
832 like domain. *Proc. Natl. Acad. Sci. U. S. A.* **113**, 10352–10357 (2016).
- 833 49. Poweleit, N. *et al.* CryoEM structure of the *Methanospirillum hungatei* archaeallum
834 reveals structural features distinct from the bacterial flagellum and type IV pilus. *Nat.*
835 *Microbiol.* **2**, 16222 (2016).
- 836 50. Leitner, A. *et al.* Chemical cross-linking/mass spectrometry targeting acidic residues in
837 proteins and protein complexes. *Proc. Natl. Acad. Sci. U. S. A.* **111**, 9455–9460 (2014).
- 838 51. Fajardo, J. E. *et al.* Assessment of chemical-crosslink-assisted protein structure modeling
839 in CASP13. *Proteins Struct. Funct. Bioinforma.* **87**, 1283–1297 (2019).
- 840 52. Chaudhury, S., Lyskov, S. & Gray, J. J. PyRosetta: a script-based interface for
841 implementing molecular modeling algorithms using Rosetta. *Bioinforma. Oxf. Engl.* **26**, 689–
842 691 (2010).
- 843 53. Kolappan, S. *et al.* Structure of the *Neisseria meningitidis* Type IV pilus. *Nat. Commun.*
844 **7**, 13015 (2016).
- 845 54. Ball, G., Chapon-Hervé, V., Bleves, S., Michel, G. & Bally, M. Assembly of XcpR in the
846 Cytoplasmic Membrane Is Required for Extracellular Protein Secretion in *Pseudomonas*
847 *aeruginosa*. *J. Bacteriol.* **181**, 382–388 (1999).

- 848 55. Jeong, J.-Y. *et al.* One-step sequence- and ligation-independent cloning as a rapid and
849 versatile cloning method for functional genomics studies. *Appl. Environ. Microbiol.* **78**, 5440–
850 5443 (2012).
- 851 56. Viarre, V. *et al.* HxcQ Liposecretin Is Self-piloted to the Outer Membrane by Its N-
852 terminal Lipid Anchor. *J. Biol. Chem.* **284**, 33815–33823 (2009).
- 853 57. Lima, D. B. *et al.* SIM-XL: A powerful and user-friendly tool for peptide cross-linking
854 analysis. *J. Proteomics* **129**, 51–55 (2015).
- 855 58. Delaglio, F. *et al.* NMRPipe: a multidimensional spectral processing system based on
856 UNIX pipes. *J. Biomol. NMR* **6**, 277–293 (1995).
- 857 59. Lee, W., Tonelli, M. & Markley, J. L. NMRFAM-SPARKY: enhanced software for
858 biomolecular NMR spectroscopy. *Bioinforma. Oxf. Engl.* **31**, 1325–1327 (2015).
- 859 60. Tamiola, K., Acar, B. & Mulder, F. A. A. Sequence-specific random coil chemical shifts
860 of intrinsically disordered proteins. *J. Am. Chem. Soc.* **132**, 18000–18003 (2010).
- 861 61. van Zundert, G. C. P. *et al.* The HADDOCK2.2 Web Server: User-Friendly Integrative
862 Modeling of Biomolecular Complexes. *J. Mol. Biol.* **428**, 720–725 (2016).
- 863 62. Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N. & Sternberg, M. J. E. The Phyre2
864 web portal for protein modeling, prediction and analysis. *Nat. Protoc.* **10**, 845–858 (2015).
- 865 63. Alphonse, S. *et al.* Structure of the *Pseudomonas aeruginosa* XcpT pseudopilin, a major
866 component of the type II secretion system. *J. Struct. Biol.* **169**, 75–80 (2010).
- 867 64. Zhang, Y. & Jin, C. *Solution structure of E.Coli GspH.* doi: 10.2210/pdb2KNQ/pdb
868 (2009).
- 869
- 870

871 **FIGURE LEGENDS**

872

873 **Fig 1.** *Minor pseudopilins are required for Type II Secretion.* a. Secretion of the principal Xcp
 874 T2SS-dependent exoprotein (LasB) assayed by Coomassie blue stained SDS-PAGE of equal OD
 875 equivalent loading of culture supernatants of the wild type PAO1, Δxcp (DZQ40), $\Delta xcpH$, $\Delta xcpI$,
 876 $\Delta xcpJ$ and $\Delta xcpK$ strains harboring empty vector (EV), or a complementing plasmid encoding
 877 XcpH (*pxcpH*), XcpI (*pxcpI*), XcpJ (*pxcpJ*) or XcpK (*pxcpK*). b. LasB protease activity on a skim
 878 milk plate is indicated by a halo around a colony, corresponding to casein degradation by secreted
 879 LasB. Colonies representing WT (PAO1), Δxcp , or complemented Δxcp strains are labeled
 880 accordingly. Note $\Delta xcpH$ is complemented by both *pxcpH* and by pJN105 expressing the XcpH
 881 deletion of the $\beta 3$ - $\beta 4$ loop, *pxcpH _{$\Delta 34$}* . c. Secretion assay, as in panel a. showing secretion of LasB
 882 by $\Delta xcpH/pxcpH_{\Delta 34}$.

883

884 **Fig 2.** *XcpH_p NMR assignments and secondary structure determination.* a. XcpH_p assignments by
 885 solution NMR. 2D ¹H-¹⁵N HSQC spectrum showing amino acid assignments for XcpH_p. NMR
 886 experiments were performed at 37°C using a sample of 500 μ M ¹³C-¹⁵N labeled XcpH_p in NMR
 887 buffers. b. Secondary structure analysis of XcpH_p. Secondary chemical shifts were calculated for
 888 XcpH_p using α and β carbon chemical shifts obtained after backbone sequential assignments.
 889 Stretches of amino acids with positive values indicate the presence of α helix while negative values
 890 correspond to β strands. Amino acids not assigned or overlapped are represented by a value of
 891 zero. Secondary structure elements in the XcpH_p model structures are shown above.

892

893 **Fig 3.** *XcpH_p chemical shift perturbations due to interaction with XcpJ_p.* a. Plot of CSPs calculated
 894 for XcpH_p amino acids. Each value was calculated using NH chemical shift differences obtained
 895 from the XcpH_p spectrum collected without XcpJ_p and or with 80 μ M XcpJ_p. CSP values above
 896 0.06 were considered significant. Amino acids presented with value zero correspond to residues
 897 unassigned or overlapped. b. CSP values were mapped onto an XcpH_p model based on *E. coli*
 898 GspH (PDB: 2KNQ⁶⁴). Non-significant changes are shown in white, while more significant values
 899 are depicted in a gradient of red. Amino acids with no data are shown in grey.

900 **Fig 4.** *Cysteine cross linking suggests XcpH_p:XcpJ_p α1C packing.* Cysteine substitutions were
901 made in indicated positions within XcpH_p (by column) and in XcpJ_p (by row). Purified proteins
902 were incubated under oxidizing conditions and visualized by Coomassie blue staining of non-
903 reducing SDS PAGE. In each of the six pairwise combinations, Lane 1 is XcpH_p alone, Lane 2 is
904 XcpH_p+XcpJ_p, and Lane 3 is XcpJ_p alone. Monomer, dimer, and heterodimer sizes are indicated
905 along the right-hand side for all three gels in each row.

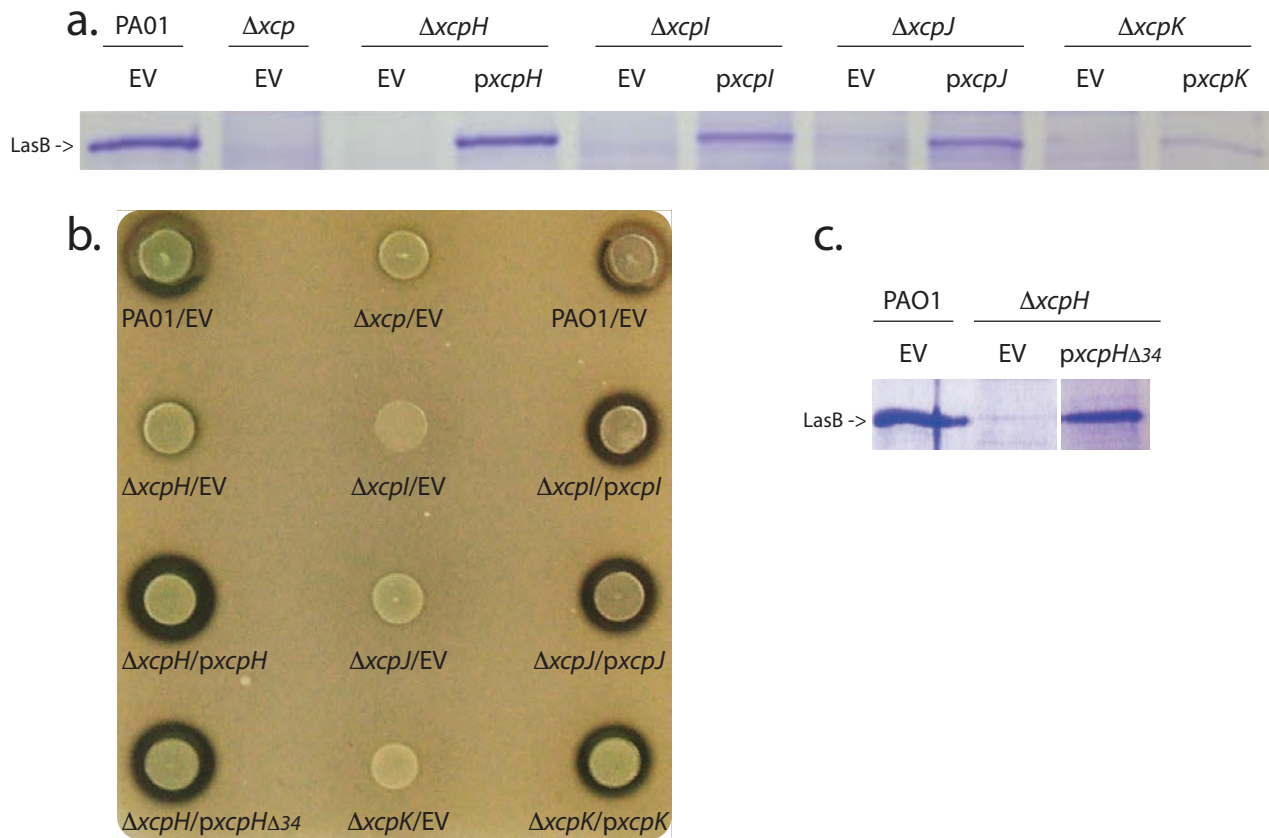
906
907 **Fig 5.** *PRE experiments of XcpH_p in the presence of MTSL labeled XcpJ_p.* a. PRE effect of the two
908 α1C MTSL-labeled XcpJ_p cysteine mutants on the XcpH_p spectra. Spectra were collected using
909 80 μM ¹⁵N XcpH_p and 24 μM XcpJ_{pR46C}-MTSL or XcpJ_{pR53C}-MTSL. Significant signal intensity
910 changes are designated with a gradient of red. All amino acids with displayed data have an intensity
911 ratio above zero, while residues with no assignments or overlapped have values of zero. b. PRE
912 effects observed with XcpJ_{pR46C}-MTSL or XcpJ_{pR53C}-MTSL were mapped onto the structural
913 model of XcpH_p.

914
915 **Fig 6.** *Acidic crosslinking of XcpH_pI_pJ_pK_p soluble complex.* a. SDS-PAGE of the cross-linked
916 proteins with a box around the band identified by the mean of proteomics as the cross-linked
917 XcpH_pI_pJ_pK_p tetramer. b. MS/MS spectrum of a particularly relevant cross-link pair between
918 XcpH_p and XcpJ_p. The spectrum shows intense ions identified as y- and b-type ions, allowing an
919 unambiguous identification of the two partners (mass accuracy for the parent < 2ppm and for the
920 fragment <10ppm). c. Circular representation of all significant cross-linked pairs detected in the
921 tetramer.

922
923 **Fig 7.** *3D model of XcpH_pI_pJ_pK_p soluble complex.* A model of the XcpH_pI_pJ_pK_p quaternary complex
924 (XcpH_p gray, XcpI_p blue, XcpJ_p green and XcpK_p wheat) was obtained via HADDOCK docking.
925 Acidic cross-linking derived restraints are shown in yellow dotted lines. In the case of PRE
926 restraints, XcpJ_p spin labeled residues R46 and R53 are shown as sticks while most affected XcpH_p
927 residues N55 and D49 are represented as blue spheres at the amide nitrogen. XcpH_pJ_p residues
928 forming salt bridges or hydrogen bonds are shown in stick model connected by yellow dotted lines.
929 Residues from XcpH_p and XcpJ_p are labeled in dark grey and dark green respectively.

930

931 **Fig 8.** *Model of Type II secretion system pseudopilus.* a. Model of XcpHIJK tip complex and
932 XcpG filament shown as a solvent-accessible surface. (above) Top and bottom views of the
933 models (below) Side views of the pseudopilus model. Color coding of each XcpG protofilament
934 is a lighter shade of the respective tip subunit color from whence it originates. b. Transmembrane
935 α 1N contacts between neighboring subunits include canonical F1-E5 salt bridge (included as
936 restraint during minimization), and hydrophobic residues at position 16 and 10 previously
937 described³⁷ but importantly not used as a modeling restraint. Images were generated using Pymol
938 software.



Escobar & Douzi, *et al.*, Figure 1

Fig 1. Minor pseudopilins are required for Type II Secretion. a. Secretion of the principal Xcp T2SS-dependent exoprotein (LasB) assayed by Coomassie blue stained SDS-PAGE of equal OD equivalent loading of culture supernatants of the wild type PAO1, Δxcp (DZQ40), $\Delta xcpH$, $\Delta xcpI$, $\Delta xcpJ$ and $\Delta xcpK$ strains harboring empty vector (EV), or a complementing plasmid encoding XcpH (*pxcpH*), XcpI (*pxcpI*), XcpJ (*pxcpJ*) or XcpK (*pxcpK*). b. LasB protease activity on a skim milk plate is indicated by a halo around a colony, corresponding to casein degradation by secreted LasB. Colonies representing WT (PAO1), Δxcp , or complemented Δxcp strains are labeled accordingly. Note $\Delta xcpH$ is complemented by both *pxcpH* and by pJN105 expressing the XcpH deletion of the beta3- beta4 loop, *pxcpH* Δ_{34} . c. Secretion assay, as in panel a. showing secretion of LasB by $\Delta xcpH$ /*pxcpH* Δ_{34} .

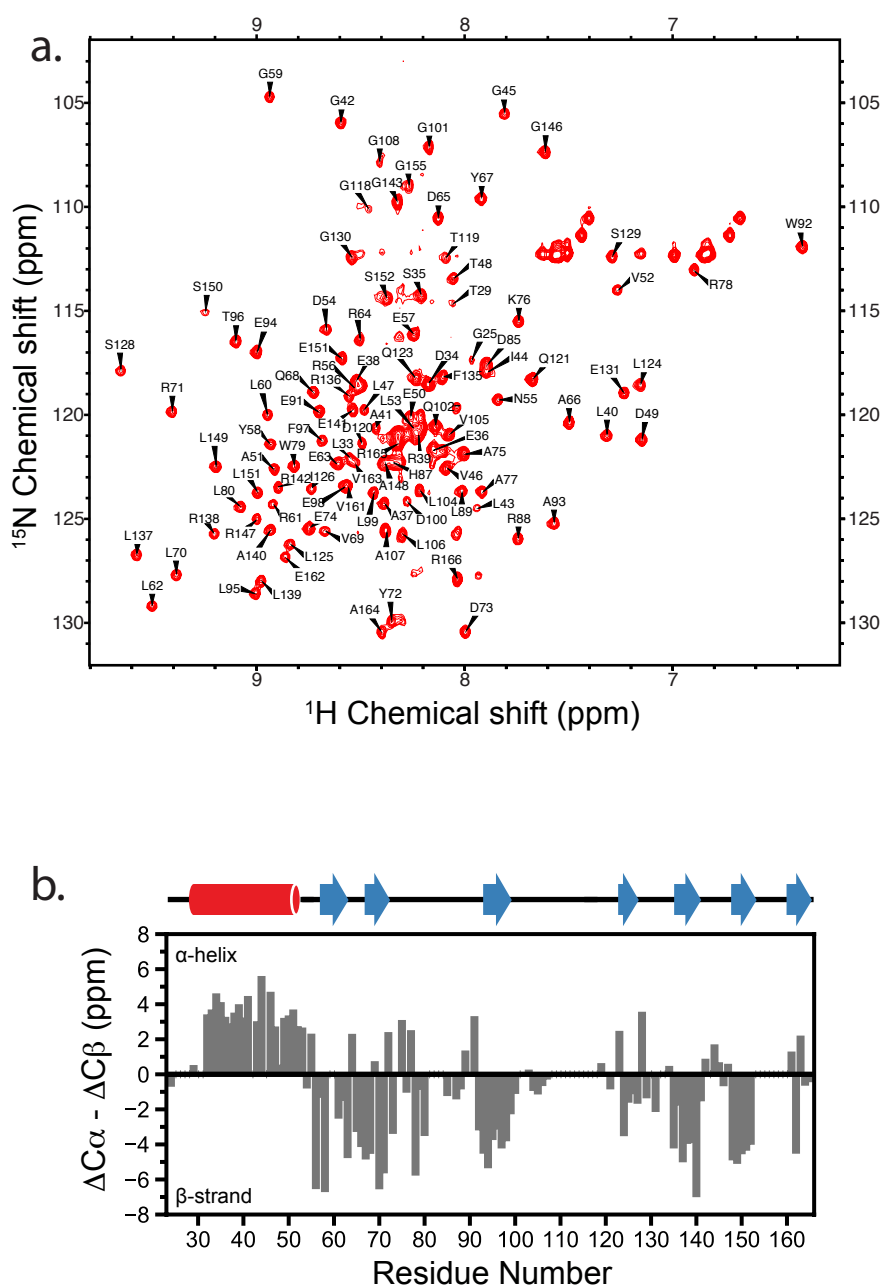
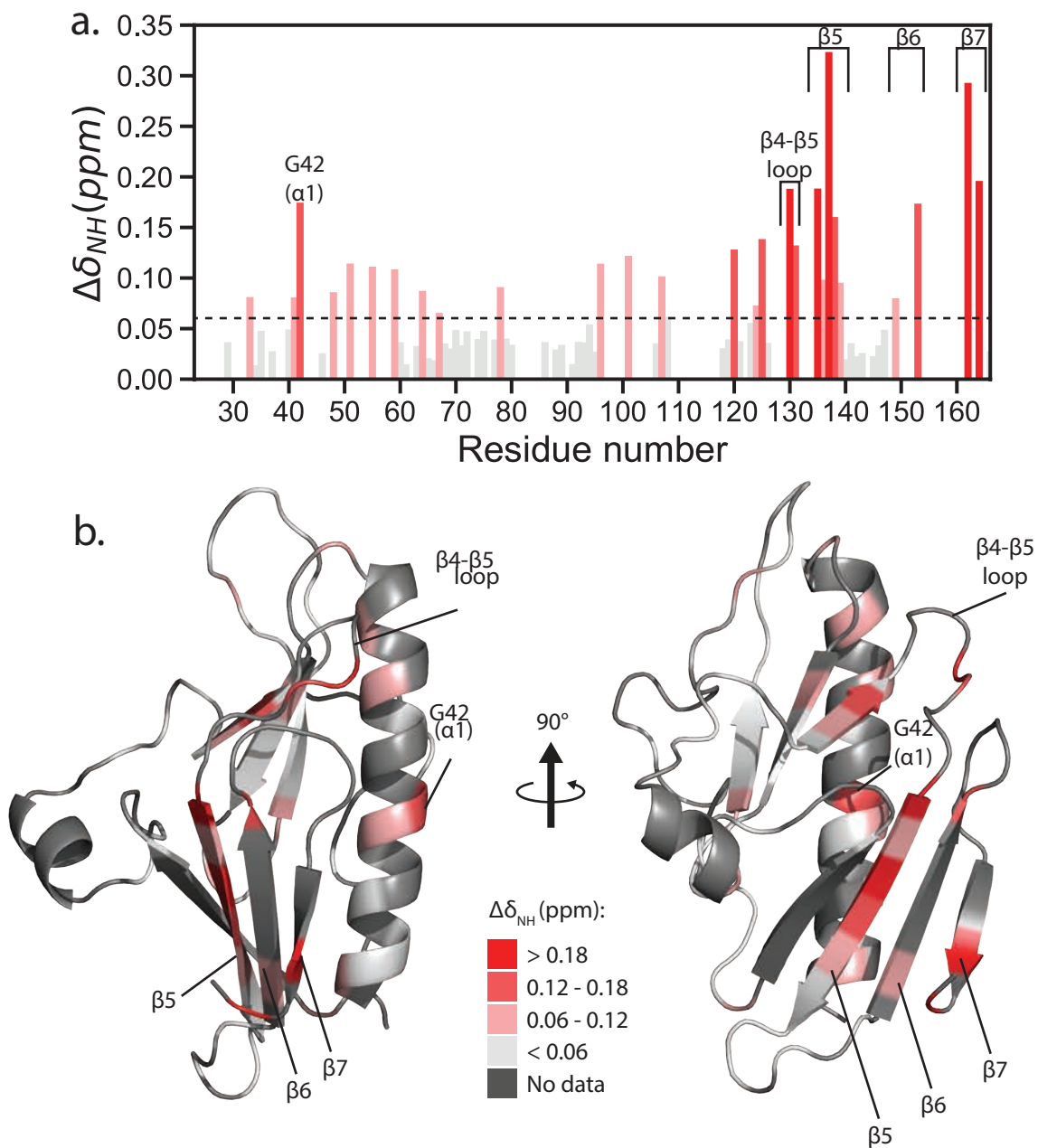
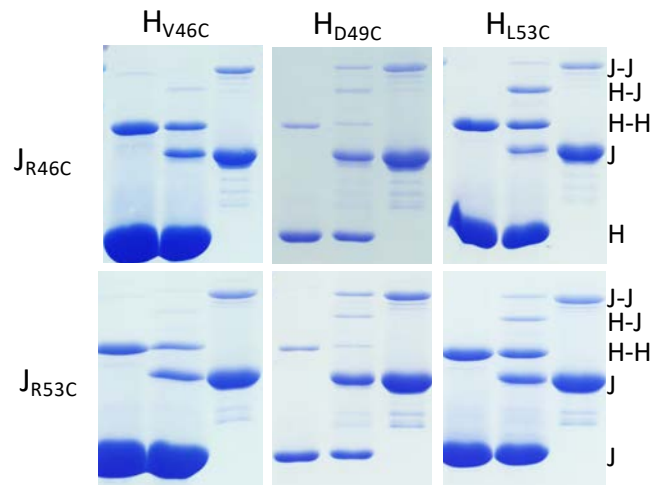


Fig 2. XcpHp NMR assignments and secondary structure determination. a. XcpHp assignments by solution NMR. 2D ¹H-¹⁵N HSQC spectrum showing amino acid assignments for XcpHp. NMR experiments were performed at 37°C using a sample of 500 μM ¹³C-¹⁵N labeled XcpHp in NMR buffers. **b.** Secondary structure analysis of XcpHp. Secondary chemical shifts were calculated for XcpHp using alpha and beta carbon chemical shifts obtained after backbone sequential assignments. Stretches of amino acids with positive values indicate the presence of helix while negative values correspond to strands. Amino acids not assigned or overlapped are represented by a value of zero. Secondary structure elements in the XcpHp model structures are shown above.



Escobar & Douzi, *et al.*, Figure 3

Fig 3. XcpHp chemical shift perturbations due to interaction with XcpJp. **a.** Plot of CSPs calculated for XcpHp amino acids. Each value was calculated using NH chemical shift differences obtained from the XcpHp spectrum collected without XcpJp and or with 80 μM XcpJp. CSP values above 0.06 were considered significant. Amino acids presented with value zero correspond to residues unassigned or overlapped. **b.** CSP values were mapped onto an XcpHp model based on *E. coli* GspH (PDB: 2KNQ64). Non-significant changes are shown in white, while more significant values are depicted in a gradient of red. Amino acids with no data are shown in grey.



Escobar & Douzi, *et al.*, Figure 4

Fig 4. Cysteine cross linking suggests XcpHp:XcpJp 1:1C packing. Cysteine substitutions were made in indicated positions within XcpHp (by column) and in XcpJp (by row). Purified proteins were incubated under oxidizing conditions and visualized by Coomassie blue staining of non-reducing SDS PAGE. In each of the six pairwise combinations, Lane 1 is XcpHp alone, Lane 2 is XcpHp+XcpJp, and Lane 3 is XcpJp alone. Monomer, dimer, and heterodimer sizes are indicated along the right-hand side for all three gels in each row.

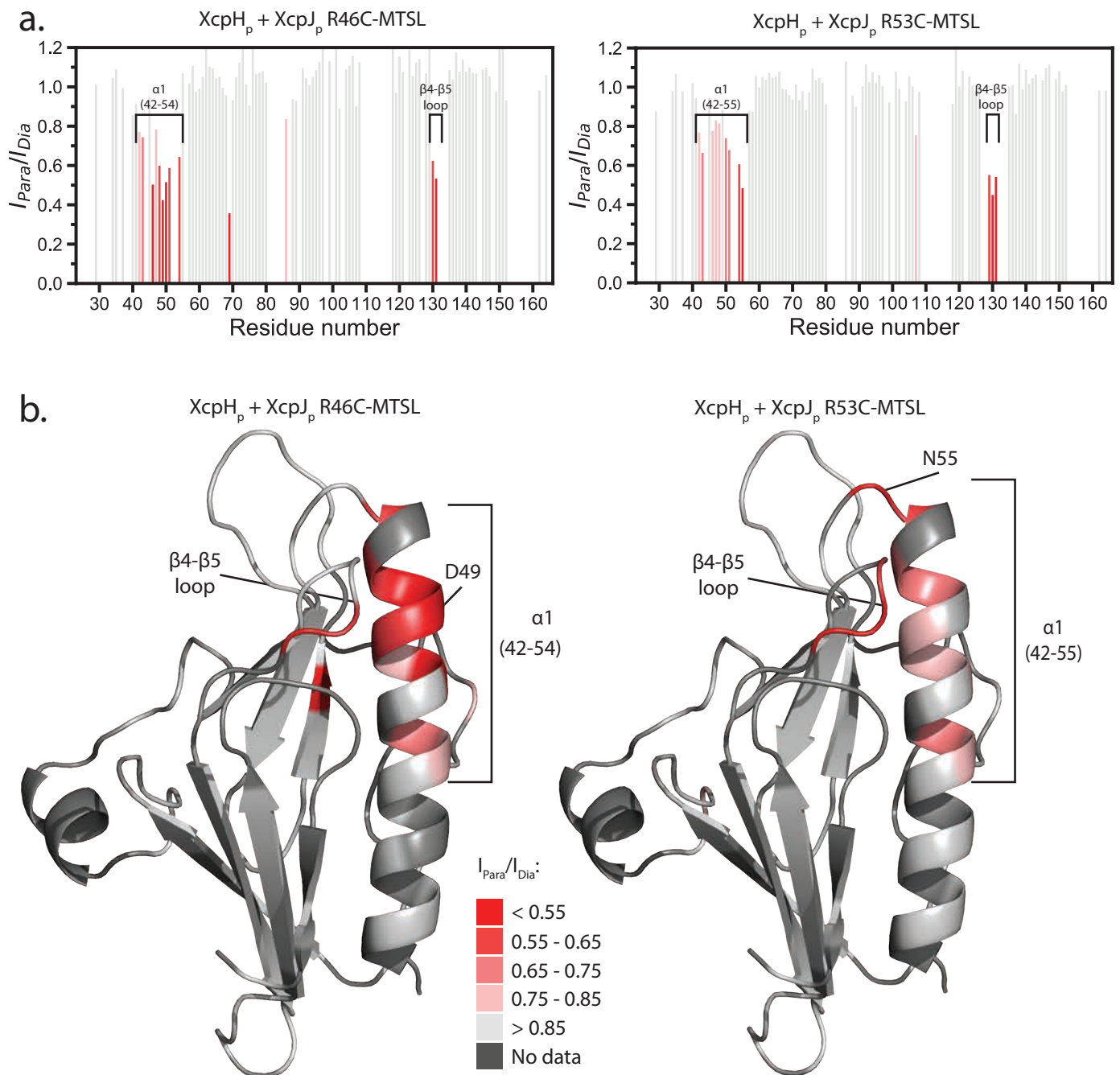
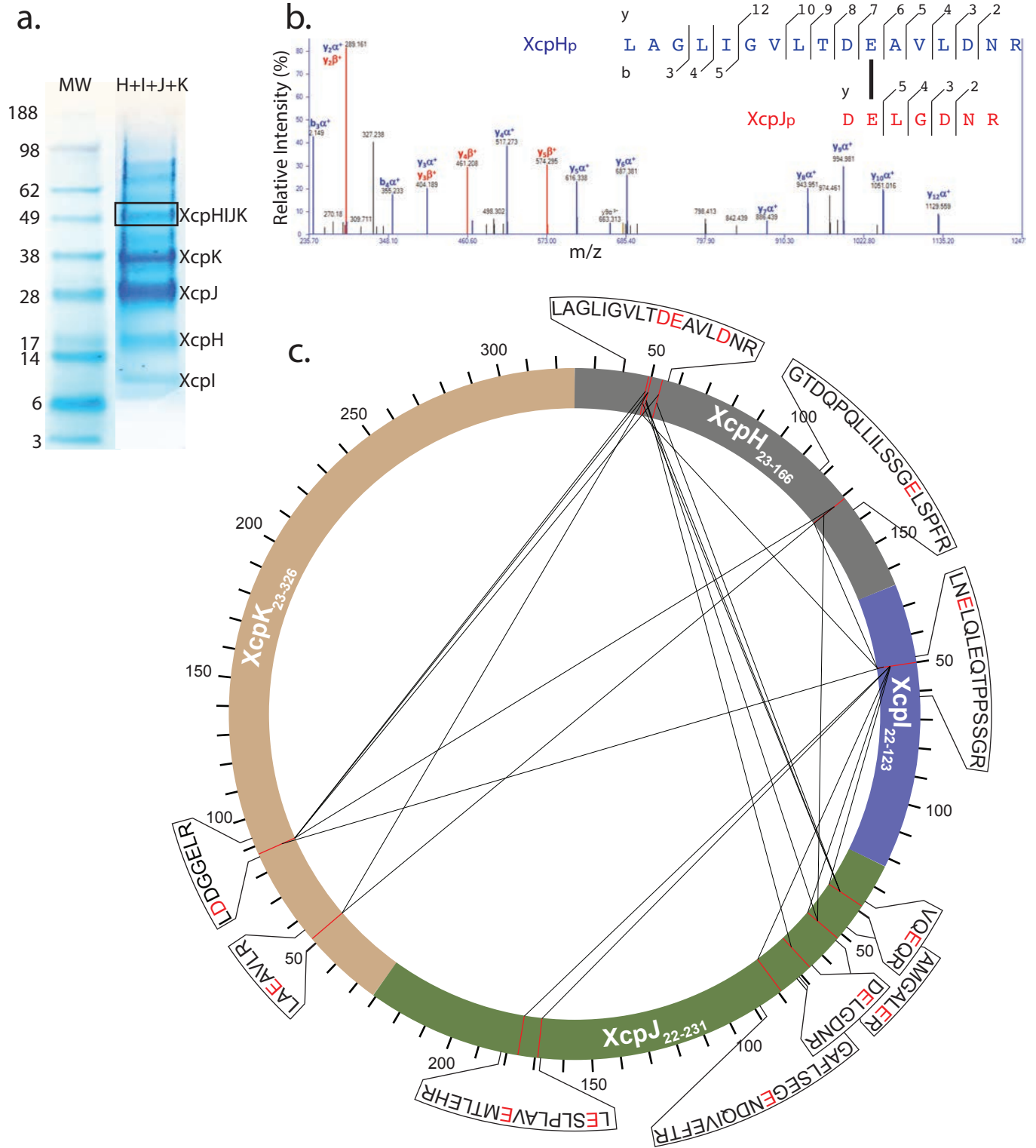
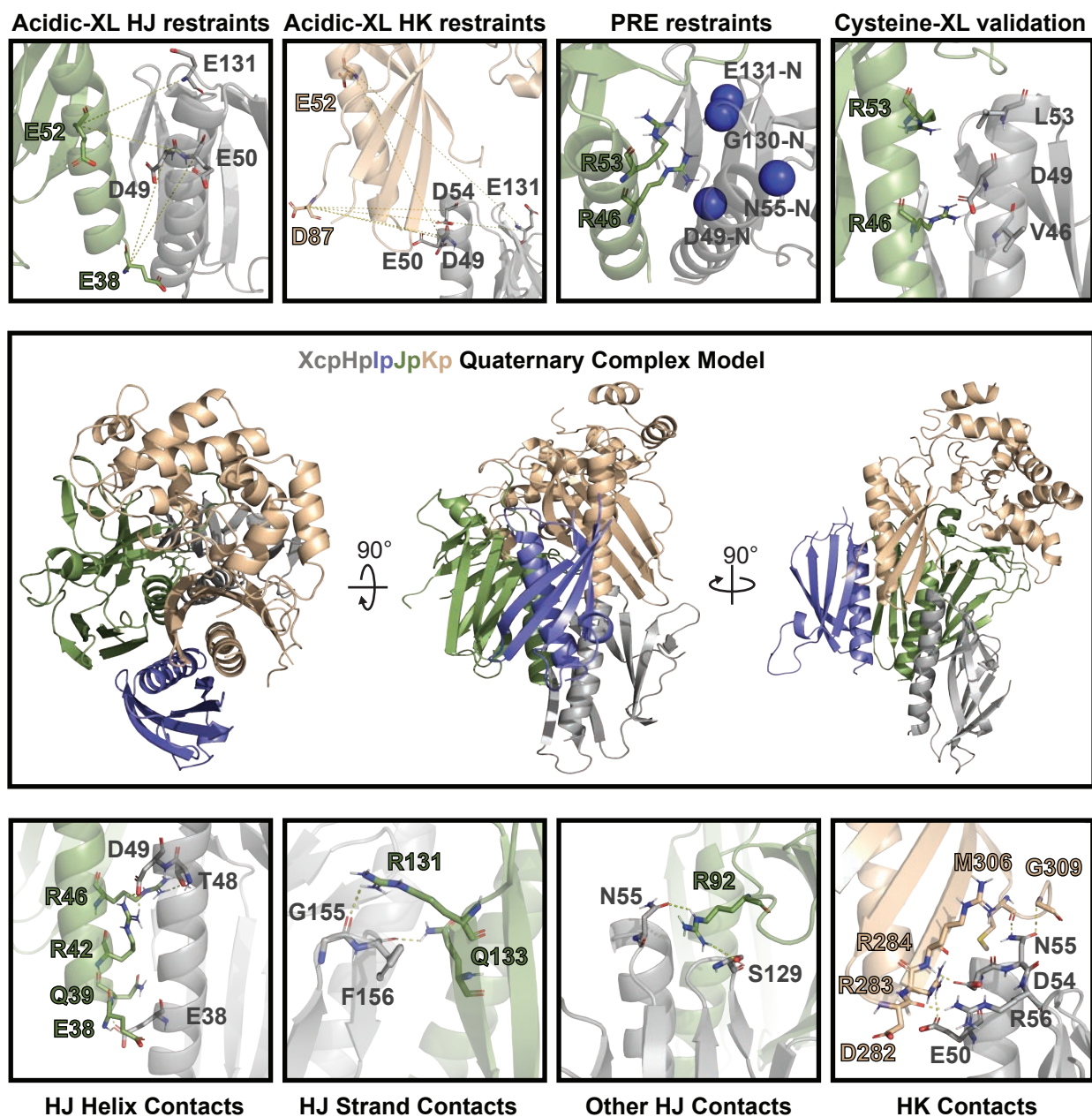


Fig 5. PRE experiments of XcpHp in the presence of MTSL labeled XcpJp. **a.** PRE effect of the two α 1C MTSL-labeled XcpJp cysteine mutants on the XcpHp spectra. Spectra were collected using 80 μ M 15N XcpHp and 24 μ M XcpJpR46C-MTSL or XcpJpR53C-MTSL. Significant signal intensity changes are designated with a gradient of red. All amino acids with displayed data have an intensity ratio above zero, while residues with no assignments or overlapped have values of zero. **b.** PRE effects observed with XcpJpR46C-MTSL or XcpJpR53C-MTSL were mapped onto the structural model of XcpHp.



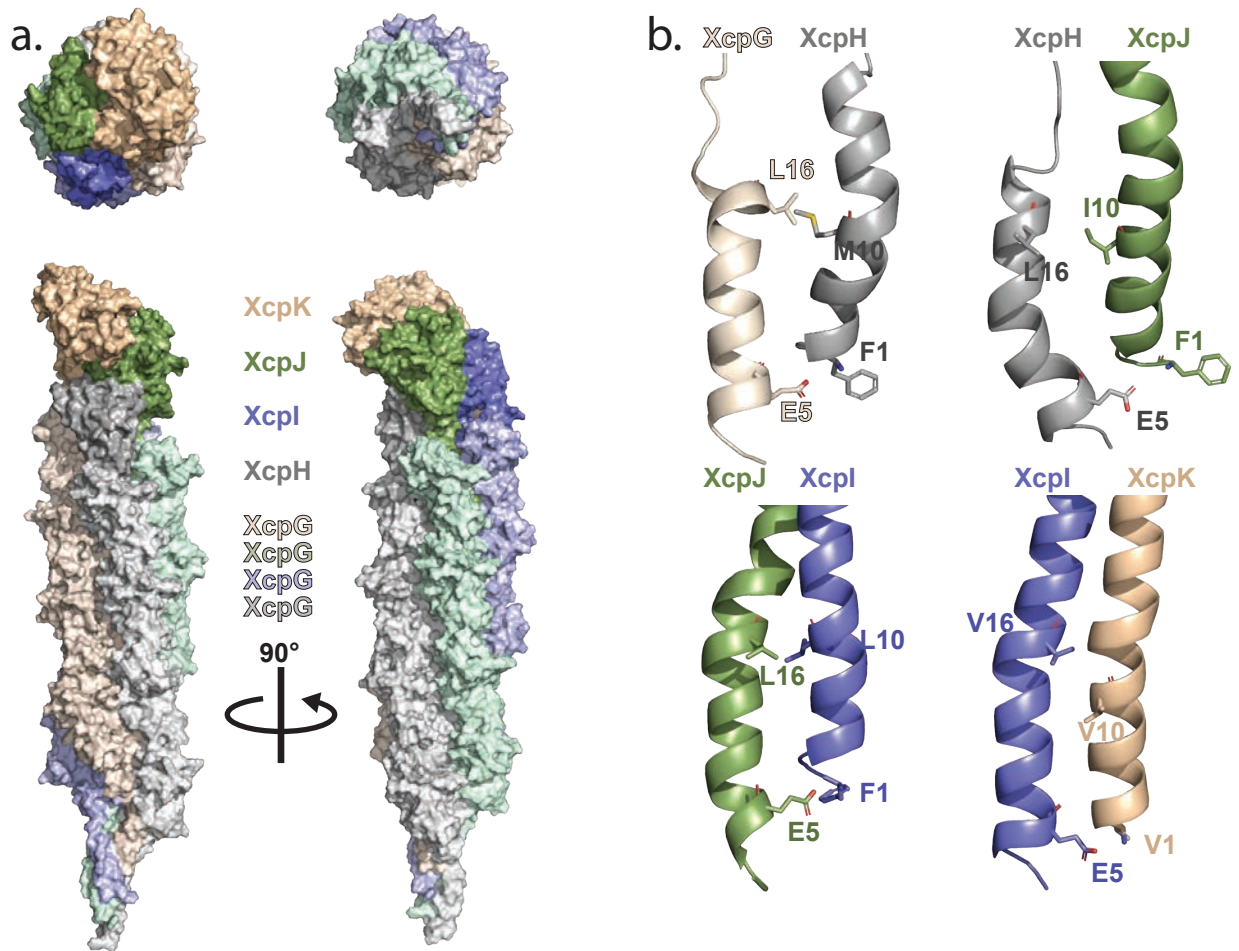
Escobar & Douzi, et al., Figure 6

Fig 6. Acidic crosslinking of XcpHlpJpKp soluble complex. **a.** SDS-PAGE of the cross-linked proteins with a box around the band identified by the mean of proteomics as the cross-linked XcpHlpJpKp tetramer. **b.** MS/MS spectrum of a particularly relevant cross-link pair between XcpHp and XcpJp. The spectrum shows intense ions identified as y - and b -type ions, allowing an unambiguous identification of the two partners (mass accuracy for the parent < 2ppm and for the fragment < 10ppm). **c.** Circular representation of all significant cross-linked pairs detected in the tetramer.



Escobar & Douzi, et al., Figure 7

Fig 7. 3D model of XcpHlpJpKp soluble complex. A model of the XcpHlpJpKp quaternary complex (XcpHp gray, XcpIp blue, XcpJp green and XcpKp wheat) was obtained via HADDOCK docking. Acidic cross-linking derived restraints are shown in yellow dotted lines. In the case of PRE restraints, XcpJp spin labeled residues R46 and R53 are shown as sticks while most affected XcpHp residues N55 and D49 are represented as blue spheres at the amide nitrogen. XcpHlpJp residues forming salt bridges or hydrogen bonds are shown in stick model connected by yellow dotted lines. Residues from XcpHp and XcpJp are labeled in dark grey and dark green respectively



Escobar & Douzi, et al., Figure 8

Fig 8. Model of Type II secretion system pseudopilus. a. Model of XcpHIJK tip complex and XcpG filament shown as a solvent-accessible surface. (above) Top and bottom views of the models (below) Side views of the pseudopilus model. Color coding of each XcpG protofilament is a lighter shade of the respective tip subunit color from whence it originates. **b.** Transmembrane $\alpha 1N$ contacts between neighboring subunits include canonical F1-E5 salt bridge (included as restraint during minimization), and hydrophobic residues at position 16 and 10 previously described (37) but importantly not used as a modeling restraint. Images were generated using Pymol software.