Bacteroidales T6SS minor Hcp subunits form heteromers recognising effectors.

Sergio G. San-Miguel[1*], Manal Kamal Saleh Al-Ammari[1*], Anna Kowalska[1], Uwe Sauer[2,3], Paulo Ricardo Batista[4], Linda Sandblad[2,3,5], Bernt Eric Uhlin[1,3,5], David A. Cisneros[1,6]

[1]Department of Molecular Biology, Umeå University, Sweden.

[2]Department of Chemistry, Umeå University, Sweden.

[3] Umeå Centre for Microbial Research (UCMR), Umeå University, Sweden.

[4] Programa de Computação Científica, Vice-Presidência de Educação Informação e Comunicação, Fundação Oswaldo Cruz, 21040-900 Rio de Janeiro, Brasil.

[5]The Laboratory for Molecular Infection Medicine Sweden (MIMS), Umeå University, Sweden.

[6] Wellcome-Wolfson Institute for Experimental Medicine, School of Medicine, Dentistry and Biomedical Sciences, Queen's University Belfast, United Kingdom.

*Equally contributed

Corresponding author: David A. Cisneros

Abstract

The type VI secretion system (T6SS) is a macromolecular protein complex found in Gram-negative bacteria that mediates intercellular antagonism and gut Bacteroidales use this system to increase their competitiveness. Furthermore, T6SS is a recognized gut colonization factor that influences gut biodiversity. Bacteroidales T6SSs are divergent from their well-characterized Pseudomonadota counterparts. One conserved feature in many Bacteroidales T6SSs is the presence of multiple Hcp proteins with low sequence similarity, encoded within the secretion system's locus and adjacent to variable effector cassettes. Here we provide evidence that in the Bacteroides fragilis T6SS, these Hcp subunits, similar to other protein export systems function as high and low abundance major and minor subunits. First, we showed that these minor Hcp proteins form a network of interactions. Biochemical characterization of one of the minor Hcp complexes formed between Hcp1 and Hcp2 showed that they form hetero-hexamers with variable stoichiometry. Furthermore, we showed that this complex recognizes the secreted effector Bte1. Our results suggest that these minor subunits function as recognition particles for effectors to mediate secretion and that this is a conserved feature in Bacteroidales T6SS. The functionality of these Hcp-effector modules is encoded in a high-synteny region within T6SS loci, and bioinformatic analyses imply that this arrangement participates in establishing effector variability. Exploiting this feature could pave the way for characterizing numerous Bacteroidales effectors with unknown functions through copurification with their cognate Hcps, unlocking new insights into bacterial interactions and into mechanisms of gut biodiversity establishment.

INTRODUCTION

39 The human gut microbiome comprises thousands of species that compete for resources and
40 space. Although the metabolic capability of individual species is a major determinant of stable
41 niche colonization in the gut (1), genes that contribute to bacterial antagonism are also well
42 represented in the microbiota (reviewed in (2)). However, little is known about the molecular
43 mechanisms underlying bacterial antagonism in the gut. Elucidating such mechanisms may
44 contribute to understanding how a biodiverse microbiota is established and how it contributes to
45 intestinal homeostasis.

46 The type VI secretion system (T6SS) is a macromolecular protein complex encompassing a
47 contractile injection system related to bacteriophages (3,4) and forming structures that can span
48 the width of gram-negative bacteria (3,5). Soon after its discovery in *Vibrio cholerae* and
49 *Pseudomonas aeruginosa* (6,7), it was established that this system participates in bacterial
50 competition by injecting lethal protein effectors. Encoded in T6SS loci are antitoxins, known as
51 immunity proteins, which mediate protection against bacterial kin attacks by neutralizing the
52 activities of effectors (8). For bacterial pathogens, T6SSs contribute to disease by injecting
53 antieukaryotic effectors subverting host defences (reviewed in (9)).

54 The architecture of the T6SS has mostly been studied in Pseudomonadota and is formed by three
55 main complexes (reviewed recently in (10–12)). A membrane complex, formed by TssJ, TssL and
56 TssM, anchors the system to the cell envelope. The base plate formed by TssE, TssF, TssG and
57 TssK is attached to the membrane complex. A tubular structure homologous to phage tails,
58 constitutes the contractile complex and is formed by an inner tube complex formed by ring-
59 shaped, hexameric Hcp (TssD) subunits stacked on top of each other, which is enveloped by the
60 contractile tubular sheath formed by TssB and TssC. Sitting atop the inner tube, the VgrG-PAAR
61 tip complex forms a puncturing device, which is housed inside the membrane complex. Upon
62 sheath contraction, the inner tube and tip are injected into target cells. Injected effectors can
63 travel both attached to the tip complex or as cargo inside Hcp rings (13,14). However, the
64 structural and functional features of Hcp-effector complexes is poorly studied.

65 Gut Bacteroidales include prominent intestinal microbiota species that contribute to
66 polysaccharide degradation in the gut, but individual species such as *Bacteroides fragilis* also
67 participate in gut homeostasis (15). T6SS-encoding loci are conserved in gut Bacteroidales and
68 are organized into three highly syntenic genetic architectures called GA1, GA2 and GA3 (16).
69 Compared with the better characterised Pseudomonadota T6SSs, Bacteroidales T6SSs are
70 divergent in sequence and have important differences (16,17). For example, the TssJLM complex
71 is not present and is substituted by another megadalton structure constituted by novel proteins

72 named TssN, TssO, TssP, TssQ, and TssR (16,18). Another important difference in Bacteroidales,
73 is that effectors are encoded in variable regions within T6SS loci, which differ at the strain level
74 (16,19). Consequently, gut Bacteroidales are capable of engaging in interspecific and interstrain
75 competition (19–21). Therefore, the T6SS is considered a *bona fide* colonization factor for
76 *Bacteroides* spp. and has been shown to contribute to shaping biodiversity in the gut, especially
77 in infants (22).

78 A conserved feature among Bacteroidales T6SSs, important to this study, is the presence of a
79 multiplicity of Hcp subunits with low sequence similarity in GA2 and GA3 (16,23). Although extra
80 *hcp* genes have been observed in other organisms, they are frequently found outside T6SS loci in
81 auxiliary modules (24–26). In all Bacteroidales GA2 or GA3 T6SS five Hcp subunits are
82 consistently present, which suggests specific and important roles for these proteins.
83 Interestingly, genes encoding variable effector cassettes are adjacent to these *hcp* genes and
84 bioinformatic analyses suggest that such variability may be specifically regulated (23). However,
85 we do not know how multiple Hcp subunits interact or whether they participate in assembly or
86 secretion reactions.

87 In the present study, we probed the existence of protein–protein interactions among
88 Bacteroidales Hcp proteins using the T6SS of *Bacteroides fragilis*. Using biochemical
89 approaches, we found that a subset of Hcp subunits, which we define here as minor Hcp subunits
90 based on their lower abundance, establish heteromeric interactions with cognate partners that
91 through cooperative binding form particles that recognize effectors for secretion. Our results
92 demonstrate that the characterisation of numerous Bacteroides T6SS effectors bound to cognate
93 Hcp heteromeres can unlock the study of molecular mechanisms mediating bacterial
94 competition in the gut microbiota.

95 **RESULTS**

96 ***B. fragilis* Hcp subunits establish heteromeric interactions**

97 The T6SS locus of *B. fragilis* encodes five Hcp genes (Fig. 1A). We hypothesised that, similarly to
98 other Hcp proteins studied to date (reviewed in (2)), *B. fragilis* Hcps would form ring-shaped
99 hexamers. Therefore, we set out to test their organisation and, specifically, to determine whether
100 the Hcp rings would assemble or stack in a specific order. We performed bacterial two-hybrid
101 assays (27) for the five *B. fragilis hcp* genes with N- and C- terminal fusions of the *Bordetella*
102 *pertussis* T18 and T25 *cyaA* fragments. For each *hcp* gene, four fusion genes were generated,
103 resulting in four possible experiments to test all pairwise interactions (Fig. 1B). We cotransformed

104 *Escherichia coli* BTH101 with the two plasmids carrying the fusion genes and analysed the
105 appearance of transformants on plates supplemented with X-gal, after growth. To define the best
106 parameters at which these interactions occur, we used various IPTG concentrations and growth
107 on plates at 30 °C or 37 °C. To our surprise, only the Hcp protein (UniProt Q5LDT1, gene
108 BF9343_1943) showed positive interactions with itself in all four two-hybrid combinations and in
109 all conditions, suggesting that not all combinations lead to the formation of multimers (Fig. 1C).
110 The finding that only this subunit can form homomers, together with the fact that it is the most
111 upstream of the T6SS locus, suggests that, as previously suspected, this gene encodes a
112 structural Hcp protein (23). The subsequent genes located downstream are named *hcp1*, *hcp2,*
113 *Hcp3, and Hcp4* (23). Another interaction, namely, Hcp1-Hcp2 (BF9343_1939 with
114 BF9343_1938), appeared consistently in all four tested combinations. The Hcp3–Hcp4
115 interaction (BF9343_1935 with BF9343_1934) was detected in most combinations (Fig. 1C).
116 Other heterointeractions, such as sHcp-Hcp3, Hcp1-Hcp4 and Hcp2-Hcp3, appeared only in
117 some tested combinations (Fig. 1C, magenta, cyan and blue dashed squares, respectively). The
118 Hcp2-Hcp2 interaction occurred in only one combination (Fig. 1C, C-term *vs*. N-term). To
119 determine the frequency and strength of interaction between Hcp proteins in all four
120 experiments, we performed densitometric analysis to quantify the appearance of blue
121 hydrolysed X-gal normalized to the positive control (Fig. S1 and S2). Averaging across all the
122 tested IPTG concentrations and growth conditions and all four fusion combinations, we obtained
123 a map of interactions, which confirmed our visual observations (Fig. 1D). Moreover, the frequency
124 map revealed a hierarchical interaction cascade in which sHcp interacted with Hcp1, which
125 interacted with Hcp2, and so on. We did not observe a large effect of varying the IPTG
126 concentration on the plate, but we inexplicably observed that some interactions seemingly were
127 lacking at 30 °C or 37 °C or inhibited bacterial growth (Fig. S1 and S2). We also observed that a
128 few interactions were detectable only after prolonged (36 h) incubation, perhaps because of their
129 weak binding strength (Figs. S1 and S3). Taken together, these results suggest that only sHcp has
130 a strong tendency to form homo-oligomeres but that the other Hcps hierarchically form hetero-
131 oligomers.

132

133 **Hcp subunits interact through cooperative binding to cognate pairs**

134 Two of the most frequently observed interactions occurred between the products of adjacent
135 genes, namely, Hcp1-Hcp2 and Hcp3-Hcp4. Because there is a highly syntenic arrangement in
136 these two pairs of *hcp genes,* conserved in many *Bacteroides* T6SSs with GA2 or GA3, we were

137 interested in testing whether we could observe more interactions when coexpressing these two
138 pairs of genes fused to T18 or T25. We cloned two contiguous *hcp* genes and inserted an *E. coli*
139 Shine–Dalgarno sequence between them to ensure translation. In this arrangement, only one *hcp*
140 gene is fused to the T18 or T25 fragment (Fig. 1E). First, we tested the interactions between sHcp
141 (Fig. 1F) and the other Hcps in the presence of their adjacent genes via two-hybrid assays (Fig.
142 1G). We did not find any new interactions for sHcp in this two-hybrid combination. Averaged
143 across the four experimental combinations, only sHcp and Hcp1 were able to interact in the
144 presence of Hcp2 (Fig 1H), as previously observed in the pairwise setting (Fig. 1C). However,
145 sHcp-Hcp3, which showed a level of interaction in the pairwise setting, did not interact on
146 average (see blue dashed squares in Figs. 1C and 1H).

147 We also wanted to determine whether new interactions would occur when adjacent pairs of Hcp
148 proteins were tested via two-hybrid assays. We observed an Hcp1-Hcp1 interaction when T18
149 and T25 Hcp1 fusions were expressed in the presence of Hcp2 (Fig. 1I and 1K). Similarly, we
150 observed an Hcp3-Hcp3 interaction when the T18 and T25 Hcp3 fusions were expressed in the
151 presence of Hcp4 (Fig. 1I and 1K). These results suggest that Hcp1 and Hcp3 can partake in
152 homointeractions but that they exclusively and cooperatively occur in the presence of another
153 Hcp. Our results also showed that some observed heterointeractions are outcompeted by the
154 expression of a second Hcp: specifically, Hcp1-Hcp4 disappear when Hcp3 and Hcp4 are present
155 and Hcp2-Hcp3 disappears when Hcp1 and Hcp2 are produced (Fig. 1C, 1I and 1J, cyan and
156 magenta squares). These results suggest that Bacteroidales Hcp proteins form a complex
157 network of interactions (schematic in Fig. 1K) but that neighboring *hcp* genes encode cognate
158 pairs that preferentially interact with each other, so that Hcp1 preferentially binds to Hcp2 and
159 Hcp3 to Hcp4.

160

161 **Hcp heterocomplexes form Hcp rings**

162 To obtain structural insight into the *B. fragilis* Hcp proteins, we purified two of the complexes
163 described above. As a control for hexamer formation, we purified a His-tagged version of sHcp
164 (Fig. 2A, sHcp-His_6) which, observed by negative staining electron microscopy, appeared as rings
165 (Fig. 2B). We obtained a crystal structure of sHcp-His_6 (Fig. 2C and S4), which was identical to the
166 unbound (to VgRG) structure of He et al. (28). We subsequently coexpressed Hcp1 and Hcp2-His_6
167 for affinity purification, which revealed that Hcp1 readily copurifies with Hcp2 (Fig. 2A).
168 Observation of the Hcp1-Hcp2-His_6 complex via negative staining electron microscopy also
169 revealed the formation of rings that appeared similar to those formed by sHcp (Fig 2B). Cryo-EM

170 observations provided a 2D structure and revealed that the fold of Hcp1-Hcp2-His_6 was that of
171 canonical Hcps, as shown by the presence of an elongated structure corresponding to β-strands
172 and a circular structure corresponding to the α-helix (Fig. 2D, 2E asterisks and S4). AlphaFold
173 models of monomeric Hcp1 and Hcp2 (Fig. 2F and 2G) confirmed a classic Hcp fold. However, a
174 few structural differences were notable. In the case of Hcp1, an insertion between residues G11
175 and S25 forms a short alpha helix that protrudes from the external beta sheet, and in the case of
176 Hcp2, an N-terminal insertion between residues M1 and L13 is unstructured. In both structures,
177 compared with the sHcp sequence, insertions presented lower PLDDT scores. Purification of the
178 tagged versions of Hcp1 or Hcp2 alone did not yield soluble proteins, which suggests that the
179 interactions between these proteins are necessary for their folding and assembly into rings. These
180 results suggest that Hcp1 and Hcp2, are binding partners that cooperatively assemble into
181 hexameric rings.

182

183 **The Hcp1–Hcp2 complex forms heterohexamers of variable stoichiometry**

184 Having stated that Hcp1-Hcp2 forms ring-shaped hexamers, we set out to determine the
185 stoichiometry of this complex. To this end, we repeated the purification of Hcp1-Hcp2-His_6, but
186 this time, we eluted with an imidazole step gradient and analysed the fractions via SDS–PAGE
187 densitometric analysis. Three fractions were collected at each imidazole concentration. As a
188 control, we constructed a plasmid expressing sHcp-sHcp-His_6, which should promote the
189 formation of mixed hexamers of tagged and untagged monomers. SDS–PAGE analysis of the
190 eluted sHcp-sHcp-His_6 bands revealed two fast-migrating bands corresponding to the tagged
191 and untagged proteins (Fig. 3A). Two slower migrating bands appeared just below the 50 kDa
192 marker with densities proportional to the corresponding higher-migrating bands, which we
193 interpret to be SDS-resistant trimers of sHcp (labelled sHcp*) because single bands appeared in
194 analysis of purified sHcp-His6 samples just above the 15 kDa and just below the 50 kDa markers
195 (Fig. S4). Higher proportions of sHcp-His_6 eluted with increasing concentrations of imidazole,
196 possibly because hexamers containing more His_6-tagged protomers bind to the Ni-NTA resin with
197 more avidity and required more imidazole to elute (Fig 3A). Different results were obtained for the
198 Hcp1-Hcp2-His_6 complex (Fig. 3B). First, no SDS-resistant proteins were eluted. Second, most of
199 the protein peak was eluted at 150 mM imidazole rather than at 250 mM imidazole as for the
200 sHcp-sHcp-His6 complexes. Given that homohexameric sHcp complexes could contain up to six
201 His-tagged subunits, whilst in the case of untagged Hcp1 with tagged Hcp2 six His-tagged
202 subunits (and thereby the same binding strength to the resin) would be observed only if hexamers

203 devoid of Hcp1 would form. Therefore, lower imidazole elution suggested that Hcp1-Hcp2- His_6
204 complexes did not have as many His_6-tagged protomers as sHcp-sHcp-His_6, which is consistent
205 with the formation Hcp1-Hcp2 heterohexamers. Third, at all elution concentrations, untagged
206 Hcp1 appeared to be present in higher proportions than Hcp2- His_6.

207 To obtain quantitative measurements of the stoichiometry of the Hcp complexes, we performed
208 densitometric analysis of the SDS–PAGE SYPRO-ruby-stained gels. We calculated the
209 proportions of tagged and untagged proteins and normalized them to six (under the assumption
210 that each eluted particle is a hexamer). For sHcp-sHcp-His_6, only the fast-migrating bands were
211 analysed, which revealed that the number of eluted sHcp-His_6 steadily increased in proportion
212 to the imidazole concentration (Fig. 3C). However, for the Hcp1-Hcp2-His_6 complex, up to 120
213 mM imidazole, the eluted complexes had a 4:2, Hcp1:Hcp2 ratio (Fig. 3D). Above 120 mM, the
214 complexes had fractional ratios, which we interpret to be mixtures of mostly 4:2 with some 3:3
215 complexes, although we cannot exclude other ratios or trace amounts of supracomplexes.
216 Additionally, we quantified the total fraction of protein eluted at each imidazole concentration by
217 adding the density of both the untagged and tagged proteins and dividing by the sum of all the
218 densities (Fig. 3C and 3D, dashed lines). This analysis revealed that approximately half of the
219 eluted Hcp1-Hcp2-His_6 had a 4:2 ratio and confirmed that the protein peak of elution was at 150
220 mM imidazole.

221 To obtain full insight into the stoichiometry of the Hcp1-Hcp2-His_6 complex, we subjected it to
222 analytical size exclusion chromatography with sHcp-His_6 as a control. Using molecular weight
223 standards, we observed that sHcp-His_6 eluted as a single peak from the column after 15 ml
224 between the 150 and 44.3 kDa markers, similarly to the Hcp1-Hcp2-His_6 complex (Fig. 3E). To
225 better quantify the size of the complex, we selected three markers from the molecular weight
226 (MW) standards and plotted the MW *vs.* the elution volume, which showed a logarithmic
227 relationship. The MW values of sHcp-His_6 and Hcp1-Hcp2-His_6 extrapolated from the plot were
228 85.1 and 86.8 kDa, respectively (Fig. 3F). The calculated MWs from the sequence for one sHcp-
229 His_6 hexamer are 93.6 kDa, while 97 kDa for a 4:2 Hcp1-Hcp2-His_6 complex. This shows that the
230 Hcp1-Hcp2-His_6 complex is formed by hexamers. Together with the findings described above,
231 these results suggest that the Hcp1-Hcp2-His_6 complex assembles *in vitro* into hexamers
232 preferentially, but not exclusively with a stoichiometry of 4:2. This implies that hexamers of Hcp1-
233 Hcp2 complexes must contain adjacent Hcp1 protomers, which were observed in two-hybrid
234 assays, but only if Hcp2 was present. Overall, we propose that the Hcp1-Hcp2 complex may form
235 through cooperative interactions via Hcp1 homointeractions mediated by Hcp2 (Fig. 3G).

236 Moreover, we conclude that *in vitro* assembled Hcp1–Hcp2 complexes have variable
237 stoichiometries.

238

239 **XL-MS shows cooperatively formed adjacent Hcp1 protomers**

240 To determine the spatial arrangement of individual protomers in the Hcp1–Hcp2 complex, we
241 performed cross-linking mass spectrometry (XL-MS). Copurified Hcp1-Hcp2-His_6 was subjected
242 to crosslinking with DSS, DSG or EGS at three concentrations (Tables S1-S3). The crosslinked
243 samples were observed via SDS–PAGE, proteolyzed and subsequently analysed via mass
244 spectrometry. SDS–PAGE analysis of the control sample and crosslinked samples revealed the
245 presence of a band migrating at the 15 kDa marker, corresponding to Hcp1 or Hcp2-His_6 with
246 either noncrosslinked subunits or with intraprotein crosslinks (Figs. 4A and S5). In the DSS
247 crosslinked samples, a prominent band migrated between 25 kDa and 37 kDa, corresponding to
248 two crosslinked Hcp proteins (Fig. 4A, marked areas). At a low DSS concentration (0.25 mM), the
249 intensity profile of the band showed faster migrating shoulder peaks, which became the most
250 prominent band at higher concentrations (Fig. 4B, arrows). This suggested that at higher
251 crosslinker concentrations, two Hcp proteins were crosslinked by multiple linkers, which caused
252 the two proteins to decrease their Stokes radii and migrate faster. To a certain extent, the band
253 migrating close to the 50 kDa marker had the same appearance and showed the same behaviour
254 at higher concentrations of the crosslinker. Although it was less clearly visible, we suggest that
255 three Hcp proteins could also be crosslinked by multiple crosslinkers. Similar observations were
256 made in analyses of samples treated with DSG (Fig. S5). At the tested concentrations, EGS did
257 not efficiently generate crosslinked peptides (Fig. S5). Above the 50 kDa marker, there were more
258 discrete bands in the crosslinked samples, which decreased in intensity towards the higher-MW
259 markers. This suggested that more than three Hcp protomers were crosslinked but at lower
260 apparent frequencies. Quantification of both intra- and interprotein crosslinked peptides
261 revealed that residues K119 in Hcp1 and K90 in Hcp2 were the most crosslinked residues (Figure
262 4C, histogram), suggesting that these residues are exposed. Filtering by residues that appear in
263 multiple replicates (Table S4), the Hcp2(K83)-Hcp1(K113) crosslink was the most consistent
264 crosslink (Fig. 4D) among the hetero-interprotein crosslinks. This result confirms the expected
265 interaction between Hcp1 and Hcp2. Consistent with our hypothesis that Hcp1 could form
266 homointeractions in the presence of Hcp2, based on cooperative two-hybrid assays, we found
267 multiple peptides across replicates showing interprotein Hcp1 crosslinks occurring between two
268 identical residues in two Hcp1 subunits, *e.g.,* Hcp1(K44)-Hcp1(K44) and Hcp1(K113)-Hcp1(K113)

269 (Fig. 4E, S5 and Table S4). Peptides between residues K111 or K128 crosslinked to residues at the
270 same position also occurred (Table S4). Interestingly, we were unable to observe any Hcp2
271 interprotein crosslinks, even though there was a network of exposed residues that consistently
272 formed intraprotein crosslinks, namely, K74, K81, K83 and K90 (Figure 4F). In particular,
273 Hcp2(K90) is very exposed, as shown by the total number of crosslinks, but Hcp2(K90) was never
274 found in crosslinks with another Hcp2(K90). This result suggests that Hcp2 is not found close to
275 another Hcp2 molecule or at least less frequently compared withs Hcp1 being close to another
276 Hcp1.

277 To gain insight into the possible arrangements of the Hcp1-Hcp2 complex, we obtained
278 hexameric models via the AlphaFold server. Using a Cα-Cα cut-off distance of 30 Å (29), we
279 classified the observed crosslinks as permissive or non-permisisve in the models. We observed
280 that in an Hcp1 homohexamer model, interprotein residues between K44 and K113 fell within the
281 30 Å distance in adjacent protomers (Fig. 4E, 5S and Table S5) but not between residues two
282 protomers away (P+2). Only the distance between Hcp1 K111 residues in protomer P and P+1 or
283 P+2 was permissive for crosslinking. Conversely, in a Hcp1-Hcp2 (3:3) model in which Hcp1
284 protomers are two protomers away, only K111 was in the permissive distance. These results
285 suggest that a structural model in which Hcp1 protomers are more frequently found adjacent to
286 each other is more compatible with our crosslinking data. Interprotein crosslinking between
287 Hcp1 and Hcp2, such as Hcp2(K113)-Hcp2(K83) (Fig. 4F and Table S5), occurred only within the
288 permissive distance when Hcp1 and Hcp2 were adjacent, confirming that neighbouring Hcp1–
289 Hcp2 interact within the hexameric assembly. The distance between the observed Hcp2
290 crosslinked residues, with the exception of Hcp2(K74)-Hcp2(K90), was permissive only within the
291 same protomer and not between adjacent protomers (Table S5). This finding suggests that the
292 majority, if not all, of the crosslinks observed between Hcp2 lysine residues were intraprotein. We
293 measured the theoretical distance between all the residues found in crosslinked peptides and
294 the same residue in the adjacent protomer in an Hcp2 homohexamer model (Fig. 4F and Table
295 S5), and these residues were not permissive to crosslinking. Therefore, we cannot confirm that
296 Hcp2 directly interacts with itself, as observed by two-hybrid assays, because even if Hcp2
297 assembled adjacently to another Hcp2, it could not be crosslinked at the same residue. However,
298 the lack of interprotein cross-links in Hcp2 suggests that Hcp1 interprotein cross-links occurred
299 among adjacent protomers and not randomly between Hcp protomers in two different hexamers.
300 Interestingly, Hcp2(K40) did not show any crosslinks and indeed appeared buried in all the
301 AlphaFold models when Hcp2 was adjacent to Hcp1 or Hcp2. K40 had a change in accessible
302 surface area of -66.7 $Å^2$ upon hexamer assembly, whereas K90 resulted in a change of only 0.4

303 Å2, indicating that K40 but not K90 is buried upon hexamerization (Figure S5). We were not able
304 to explain two cross-linked residues involving Hcp1, K128 residues modeled at non allowed
305 distances (Fig. S5 and Table S5). However, the distance was close enough to be permissible, and
306 these residues could be considered within range of other residues previously reported (29).
307 Finally, data obtained from the published proteome of *B. fragilis* (30) revealed that Hcp1 is ~53x
308 more abundant than Hcp2 is (Fig S6), suggesting that Hcp2 is limiting. Notably, sHcp is ~1000 fold
309 more abundant than Hcp2, consistent with an inner tube structural role. Based on these
310 differences in abundance, we refer to sHcp as the majour subunit and Hcp-1-4 as minor subunits.
311 All these results, combined with our stoichiometric analysis, are consistent with an arrangement
312 of the Hcp1–Hcp2 hexamer, in which Hcp1 molecules are adjacent at higher frequencies and
313 Hcp2 can be found in at least one or two copies, considering that the interaction between Hcp1
314 and Hcp1 depends on the presence of Hcp2, as shown by two-hybrid assays, and that Hcp1 and
315 Hcp2 do not appear to be equimolar inside the bacterial cell. Therefore, we obtained a Hcp1-
316 Hcp2 structure from the AlphaFold server with a 4:2 stoichiometry (Figs. 4G and S6). The model
317 showed an asymmetric structure with three adjacent Hcp1 protomers and the two Hcp2
318 protomers separated by another Hcp1. Notably, the Hcp2 protomers were not adjacently
319 assembled. Analysing non-crosslinked samples (Table S6), showed that there were more Hcp1
320 peptides found compared to Hcp2. All these results are consistent with SDS-PAGE densitometric
321 analysis showing an Hcp1-Hcp2 complex formed primordially of Hcp1 with fewer Hcp2 subunits.

322

323 **Minor Hcps are linked to effector recognition and secretion**

324 Our data show that the product of *Hcp1* forms complexes with the product of *Hcp2*, which is the
325 gene located downstream. Interestingly, the effector *bte1* is downstream of *Hcp2*. As reported by
326 Coyne *et al.* (16), in the high-synteny arrangement of *Bacteroidales* GA3, two effectors are
327 encoded in variable cassettes but are always arranged in the same order. In most strains one
328 effector is downstream of an Hcp2 homologue. Therefore, we hypothesised that the Hcp1-3
329 complex could interact with Bte1. First, we tested the interactions of Bte1 with all minor Hcp
330 proteins via bacterial two-hybrid assays. Our results revealed that Bte1 fused to T18 or T25
331 fragments positively interacted with both Hcp1 and Hcp2 fused to T18 or T25 fragments when
332 Hcp2 and Hcp1 were coexpressed, but not with the Hcp3 or Hcp4 T18 and T25 fusions (Fig. 5A).
333 To confirm this interaction, we copurified a His-tagged version of Bte1 with Hcp1 and Hcp2 and
334 analysed it via size exclusion chromatography. Two major peaks were observed (Fig. 5B). Pooled
335 fractions from each peak were subsequently analysed by SDS–PAGE, which revealed that peak 1

336 corresponded to the Hcp1–Hcp2–Bte1 ternary complex, whilst peak 2 corresponded to Bte1-His_6
337 alone (Fig. 5C). Negative staining electron microscopy images of the ternary complex showed
338 particles in various orientations (Fig. 5D). However, a few particle structures were filled with a
339 density in the centre, other particles appeared as empty rings (Fig 5D insets). These results
340 suggest that the minor Hcp heterohexamers can be involved as recognition particles in case of
341 *Bacteroidales* T6SS effectors.

342 Three more effectors were described for the T6SS GA3 in other strains in addition to Bte1 encoded
343 in the minor *hcp*-adjacent variable region. Two proteins were initially annotated using Hidden
344 Markov models (HMM) and categorized as cluster 51 and cluster 115 (using Coyne *et al.*
345 clustering nomenclature (16)), which were upstream of immunity proteins named clusters 107
346 and 53, respectively. The third type of effector, present in *B. fragilis* 638R, is called *bfe1* (20) and
347 has been shown to target other *Bacteroides* spp. To gain an overview of the probability that the
348 *Bacteroidales* minor Hcp proteins function as effector recognition particles in general, we
349 generated a phylogenetic tree of all GA3 available *Hcp2* genes described in Coyne *et.al* (16). The
350 Hcp2 homologues had a high degree of identity (Fig. S8), but they were separated into six clusters,
351 three of which were closer to each other (Fig. 5E). We then analysed the effectors adjacent to
352 each cluster and observed that the four types of effectors segregated in different clusters (Fig.
353 5E). Only Bfe1-like effectors, similar to the effector found in *B. fragilis* 638R, which were the most
354 common, segregated in three clusters, which were closer to each other. This cosegregation did
355 not occur when sHcp was used to generate the phylogenetic tree, suggesting that sHcp is not a
356 recognition factor for T6SS GA3 effectors. Hcp1 and Hcp4 only partially cosegregated with their
357 effector–immunity pairs (Fig. S7). Hcp3 also cosegregated with the other pairs involved in the
358 immune system, and although we did not detect any protein–protein interactions, this result may
359 suggest that Hcp3–Hcp4 could recognize Bte1 under different conditions (Fig. S7). The clustering
360 of Hcp2 with effectors was also present in twelve T6SS GA2 clusters described by Coyne *et al.*
361 (16)(Fig. S8). Considering that Bte1, Hcp1 and Hcp2 were found to be secreted (19), it is
362 reasonable to propose that they are secreted as a complex. Taken together all of our results are
363 consistent with the idea that the minor Hcp proteins are cognate to effectors and participate in
364 secretion and suggesting that this feature is conserved in the *Bacteroidales* T6SSs.

365

366 **DISCUSSION**

367 This study revealed that in Bacteroidales T6SSs there are different types of Hcp subunits with
368 specific assembly properties and functions. Firstly, a role for sHcp as major subunit constituting

369 the bulk of the inner tube is consistent with data showing its high abundance and its capacity to
370 form homohexamers. In the case of Hcp1 and Hcp2, we showed here that these minor subunits
371 form heterohexamers and have a specialized function of recognition particles for effector
372 secretion.

373 Our results are consistent with and extend the findings of Silverman *et al.* and Howard *et al.*, that
374 delineated a role for Hcp proteins in effector recognition (13,14). In the case of *Pseudomonas*
375 *aeruginosa,* Hcp subunits have been shown to recognize effectors as chaperones but are also
376 required for export (13), and different T6SSs present in this bacterium can specifically recognize
377 cognate effectors (14). However, our results suggest a clear division of functions between major
378 and minor Hcp subunits. Considering that effectors in gut Bacteroidales are encoded in variable
379 chromosomal regions within T6SS loci, our results imply that the recombination of effector genes
380 require co-segregation of genes encoding cognate Hcp proteins that would facilitate export of
381 multiple effectors by the same secretion system, suggesting a bypass of the evolution of specific
382 recognition motifs only to major Hcp, VgrG or PAAR subunits. Furthermore, our results imply that
383 the functional characterization of many Bacteroidales T6SS effectors may require the
384 characterization of their cognate Hcp protein complexes.

385 Consistent with a structural role for sHcp, proteomic data showed that this subunit is ~1000-fold
386 more abundant compared to Hcp2 (30), which is compatible with this protein constituting the
387 bulk of the inner tube and that Hcp1 and Hcp2 act as minor Hcps where Hcp2 is an effector-
388 specific limiting factor. These data are consistent with previously published RNA-seq data (20)
389 showing that *shcp*'s mRNA is up to ~43 fold more abundant than the minor *hcp's* mRNAs,
390 suggesting that rate of transcription and/or mRNA lifetime modulation may contribute to inner
391 tube subunit abundance. Moreover, sHcp is the only subunit present in all three genetic
392 architectures of Bacteroidales T6SSs, whereas only GA2 and GA3 have multiple minor Hcp
393 subunits (16). T6SSs with GA1 have two copies of sHcp belonging to the same HMM cluster with
394 no adjacent genes coding for effectors (16,23), suggesting that effector recognition and
395 recombination may function differently in organisms with these T6SSs.

396 Two-hybrid assays showed a high level of interaction among Hcp subunits. However, in
397 competition assays Hcp1 preferentially bound to Hcp2 and Hcp3 preferentially bound to Hcp4.
398 These interactions coincide with the corresponding two pairs of genes being adjacent and are
399 consistent with a model in which these proteins coevolve to form heterohexamers that recognize
400 other proteins. This finding is further supported by a high level of synteny in gut Bacteroidales
401 T6SS loci. We also identified other interactions that await biochemical confirmation. For

402 example, sHcp could bind to Hcp1 or Hcp3. These minor Hcps may intersperse with major
403 subunits in the assembled inner tube. However, the difference in abundance and the greater
404 tendency of sHcp to form homohexamers may increase the difficulty of detecting these
405 interactions both *in vivo* and *in vitro*. We also observed binding between Hcp1 and Hcp4. Hcp4
406 was not detected in proteomic data (30), suggesting that, like Hcp2, Hcp4 may be present in low
407 abundance. It could be that Hcp1 may carry both Hcp2 and Hcp4 protomers to mediate specific
408 interactions. Furthermore, it is feasible that these complexes have additional interaction partners
409 *in vivo*. How the minor Hcp subunits assemble within the inner tube and bind to VgrG remains to
410 be elucidated, but the proteomic abundance data suggest that the formation of the Hcp1–Hcp2
411 complex and effector recognition is limiting. Therefore, there may be a specific assembly pathway
412 for minor Hcp subunits and their binding partners to assemble at the tip of the inner tube that
413 ensures their function in secreting cognate effectors.

414 Copurification, electron microscopy imaging and SEC of the Hcp1-Hcp2 complex demonstrated
415 that it formed heterohexamers. This type of assembly has implications for T6SS containing
416 specialized or evolved Hcp subunits with effector-carrying C-terminal extensions (24,25), which
417 are also present in gut *Bacteroides* (16). Our results imply that in systems carrying evolved Hcps,
418 major and minor Hcps carry distinct functions forming distinct assemblies for secretion.

419 Our results also highlight the question of how hexameric Hcp complexes are assembled. The
420 SDS-resistant Hcp trimers observed during sHcp purification could represent a stable
421 intermediate in its assembly pathway. Therefore, our results hint that two stable trimer
422 intermediate structures assemble into hexamers. The heteromeric assembly of the Hcp1-Hcp2
423 complexes suggests a mechanism by which effectors are concomitantly assembled with minor
424 Hcps. Indeed, the homointeractions between Hcp1 and its purification required the presence of
425 Hcp2, and it is reasonable to think that Bte1 recognition may benefit from such a cooperative
426 assembly pathway. He *et al.* discovered a conformational change associated with sHcp binding
427 to VgrG (28). The question of how the Hcp1-Hcp2 complex is assembled in the context of the
428 whole secretion system remains open.

429 In polymorphic toxins, the N-termini typically carry secretion specificity domains and the C-
430 termini carry toxin domains. Toxins are frequently encoded in gene fragments located adjacently
431 upstream of genes encoding cognate antitoxin proteins (31). Such synteny has resulted in toxin-
432 antitoxin (immunity) pairs being acquired by horizontal gene transfer as gene fusions of N-
433 terminal secretion specificity domains (32). The high level of synteny of the *Bacteroides* T6SSs
434 translates into function by linking the position of the gene and the hierarchy of interaction of their

435 products as shown here by protein-protein interactions between Hcp1, Hcp2 and Bte1 (which in
436 turn would interact with its antitoxin, Bti1). Therefore, our results suggest that in the case of
437 *Bacteroidales* T6SSs, the minor *hcp* genes partly encode domains that mirror transcriptional
438 fusions of polymorphic toxins encoding secretion specificity domains, although without being
439 fused. Indeed, Hcp2 homologues clustered according to their cognate effectors in the
440 phylogenetic tree, suggesting that Hcp2 is a cosegregated specificity factor. Moreover, Hcp1,
441 Hcp2 and Bte1 have been shown to be present in the secreted fraction (19,21), suggesting that
442 the function of the Hcp1-Hcp2 complex is to guide effector for secretion. In a previous Hcp study,
443 Howard *et al*. concluded that the interactions between effectors and cognate Hcp proteins
444 depend on the overall Hcp structure and not on specific residues on the basis of the clustering of
445 specialized and core Hcps performed on phylogenetic analyses of a large subset of proteins (14).
446 However, our phylogenetic analysis using a specific subset of *Bacteroidales* T6SS GA3 Hcps
447 clearly segregated effectors with cognate minor Hcps but not with major sHcp proteins. This
448 model is consistent with the minor Hcps participating in specific recognition of effectors. If this
449 is true, it could mean that minor Hcps can recognize different effectors with a minimum of
450 variability, encoded in very few residues. Although this specificity could be also encoded through
451 selective steric hindrance as suggested by Howard *et al*. (14). This sort of specificity could be
452 aided by an asymmetric assembly with limiting Hcp2 subunits as recognition factors, which could
453 accommodate effectors in the minor Hcp complex with specific stoichiometry.

454 The separation of major and minor subunit proteins with high and low abundance, respectively,
455 in gut Bacteroidales T6SSs appears akin to macromolecular export systems that assemble
456 filamentous structures. For example, type II secretion systems include one major pseudopilin (or
457 endopilin) and four minor pseudopilins. The minor subunits form a tip complex that functions in
458 priming the assembly of the major pseudopilin into a filament (33,34). This type of assembly and
459 function of minor pilins is conserved in type IV pili (35–37). In type 1 fimbria minor pilins mediate
460 adhesion and appear at the tip of fimbriae because of their higher affinity to the usher subunit
461 (38,39). Therefore, studying the biophysical characteristics shared by these proteins diverse in
462 function and sequence but similar in abundance differences could contribute to understanding
463 how these large macromolecular complexes are assembled. In all these examples, hierarchical
464 assembly assures the function of minor subunits. In the case of the *Bacteroides* minor Hcps, a
465 mechanism by which the Hcp1-Hcp2 effector complex would randomly incorporate into the inner
466 tube (with a ratio of 1:1000 sHcp to Hcp2) would not be very efficient to assure effector secretion,
467 but the formation of a minor Hcp-effector particles proximal to the tip complex would greatly
468 facilitate effector export. Further studies are needed to elucidate the *in vivo* assemblies of minor

469 Hcp subunits. Finally, our results shed light on how Bacteroidales T6SSs evolved to exhibit a high
470 diversity and variability of effectors. The presence of major and specialized minor Hcp subunits
471 that recognize these effectors is an evolved feature that facilitates this variability, contributing to
472 *Bacteroides* biodiversity.

473

474 **MATERIALS AND METHODS**

475 **Bacterial two-hybrid assay.** Twenty-five microlitres of ultracompetent *E. coli* BTH101 (*cyaA*$^-$)
476 cells were cotransformed with plasmids carrying T18 and T25 translational fusions via heat
477 shock. After recovery in 2X YT medium, 10 microlitres of the bacterial suspension were spotted
478 on Luria broth agar plates containing 40 µg/ml and 0.125, 0.25 or 0.5 mM IPTG. Drops of bacterial
479 suspensions cotransformed with control plasmids were included in each plate. The plates were
480 incubated at 30 or 37 °C for 24 and 36 h. For image analysis (see below), the density of bacteria
481 must be similar. We found that the best results were obtained when the drops of bacteria grew
482 between single colonies and a small lawn. The plates were photographed at each time point via
483 a Fas-DIGI system (NIPPON genetics) equipped with a white screen and a PENTAX RICOH-
484 IMAGING camera with the following settings: ISO 200, F number 2.5, and an exposure time of
485 1/800 s. The original images look dimmed, but this dimness is important for image analysis. The
486 images were analysed with FIJI by separating the colour image into its three RGB components and
487 taking the red channel only. The image was inverted, and the average value of the background
488 outside the plates was manually subtracted. To quantify each spot, a region of interest (ROI) that
489 was only slightly smaller than the growing bacteria excluding the borders but equal for each spot
490 was measured for each sample via Fiji to obtain the average intensity. The normalized intensity at
491 each spot on the plate was calculated by subtracting the value of the negative control from the
492 intensity value of each spot and dividing it by the difference between the positive and negative
493 controls. For some experiments, two pictures were taken for each plate. The two pictures were
494 treated as above before quantification and placed in a stack in FIJI. The two images were aligned
495 via the linear stack alignment (SIFT) plugin and stack-averaged. This process reduced any
496 differences in the background caused by deviations in light capture at each pixel position in the
497 camera's sensor.

498 **Protein purification.** Protein purification was carried out by expressing *hcp* genes from pET28
499 constructs in BL21(DE3) cells. Co-expression of *Hcp1, Hcp2* and *bte1* was performed using
500 plasmid pETDUET-1. Cells were grown in LB medium at 37°C until exponential phase and induced
501 with 0.4 mM IPTG after cooling the cells. Induction was performed overnight at 20°C or for three

502 hours at 30°C. Cells were harvested by centrifugation and resuspended in 50mM NaP 8.0.
503 500mM NaCl, 10% glycerol, 10mM Imidazole, 1mM β-mercaptoethanol. Lysis was performed by
504 sonication and the cleared lysate was prepared by centrifugation at 40,000 g for 30 minutes. The
505 cleared lysate was incubated with a Ni-NTA resin (ThermoScientific) for 1 h, washed with
506 resuspension buffer and eluted with resuspension buffer containing 300 mM imidazole. For
507 preparative size exclusion chromatography (SEC) a HiLoad 16/600 Superdex200pg column
508 (Cytiva) was used. For analytical SEC, a Superdex200 10/300gl column (Cytiva) was used and a
509 Protein Standard Mix 15 - 600 kDa (Merck). Proteins samples were characterised using SDS-PAGE
510 stained with blue Coomassie or SYPRO-Ruby.

511 **Densitometric analysis of the SDS–PAGE gel (fluorescence intensity).** The gels were stained
512 with SYPRO-Ruby and scanned with a GE Typhoon FLA 9000 Gel Imager. The tiff images were then
513 analysed in FIJI. For each quantified protein band, two ROIs were created. The first ROI contained
514 the band. The second ROI was the same size but encompassed an immediately adjacent area
515 without stain as a blank. The 'integrated density' of each ROI was measured, and the blank was
516 subtracted to obtain the corrected fluorescence intensity ($fluo_{corr}$) of the protein band. To obtain
517 the normalized SYPRO-ruby fluorescence intensity at each lane ($fluo_{norm}$) of each band of His-
518 tagged ($fluo_{norm}1$) protein *vs.* non-His-tagged protein ($fluo_{norm}2$), the corrected fluorescence
519 intensity of each band was normalized by the following equation: $fluo_{norm}1= fluo_{corr}1/(fluo_{corr}1 +$
520 $fluo_{corr}2)*6$ or $fluo_{norm}2= fluo_{corr}2/(fluo_{corr}1 + fluo_{corr}2)*6$. The obtained number represents the
521 fraction multiplied by six of one or the other protein normalized by the total intensity.

522 **Electron microscopy.** For transmission electron microscopy (TEM) the samples were
523 resuspended in purification buffer and 3.5 μL adsorbed to formvar and carbon-coated glow
524 discharged grids (cupper, 300 mesh, Ted Pella Inc.) and were negatively stained with 1.5% w/v
525 uranyl acetate. The samples were imaged with a Talos L120 TEM using the Ceta detector and
526 Velox software (Thermo Fisher). For Cryogenic TEM (Cryo-EM) 4 μL of the samples were adsorbed
527 to holey carbon film on 300 mech cupper grids (Quantifoil). The Grids were blotted and plunge
528 frozen in liquid ethane using the Vitrobot (Thermo Fisher). Cryo-EM imaging was performed with
529 Titan Krios G2 Cryo-EM using Gatan K2 BioQuantom detector and EPU software (Thermo Fisher).
530 The Cryo-EM images were analysed by single particle analyse approaches with the CryoSPARC
531 software package.

532 **Protein alignments and phylogenetic analyses.** Protein alignments were performed at the
533 MAFFT website (40) with default parameters unless specified otherwise and phylogenetic trees
534 were subsequently obtained using the NJ method and the JTT substitution model with

535 bootstrapping (1000 iterations). Visualization of phylogenetic trees with alignments was
536 performed using TreeViewer (41).

537 **Crosslinking of the Hcp1-Hcp2 complexes.** A total of 20 µG of Hcp1-Hcp2 complexes in 20 mM
538 NaP, pH 8.0, 150 mM NaCl, 1 mM DTT were crosslinked in duplicates with three different
539 concentrations (0.25 mM, 0.5 mM and 1.0 mM) of heavy/light DSS (DSS-H12/D12, Creative
540 Molecules Inc., 001S), heavy/light DSG (DSG-H6/D6, Creative Molecules Inc., 010S) or
541 heavy/light EGS (EGS-H12/D12, Creative Molecules Inc., 003S), respectively in a total reaction
542 volume 20 uL. Non-crosslinked samples were kept as controls. All samples were incubated for
543 30 min at 37°C, 1000 rpm. The cross-linking reaction was quenched with a final concentration of
544 50 mM of ammonium bicarbonate for 15 min at 37°C, 1000 rpm. For SDS-PAGE analysis of the
545 crosslinked samples, 5 uL of the above reactions were mixed with 5 uL of 2x sample buffer
546 (BioRAD) and boiled at 95°C, 5 min. The samples were analyzed on a 4-20% BioRAD Criterion TGX
547 gel.

548 **Sample preparation for mass spectrometry.** The remaining samples crosslinked with DSS and
549 DSG were denatured using an 8 M urea– 100 mM ammonium bicarbonate solution. For the EGS-
550 crosslinked samples, 8 M urea without ammonium bicarbonate was used. The cysteine bonds
551 were reduced with a final concentration of 5 mM Tris (2-carboxyethyl) phosphine hydrochloride
552 (TCEP, Sigma, 646547) for 60 min at 37°C, 1000 rpm and subsequently alkylated using a final
553 concentration 10 mM 2-iodoacetamide for 30 min at 22°C in the dark. For digestion, 1 µG of lysyl
554 endopeptidase (LysC, Wako Chemicals, 125-05061) was added, and the samples incubated for
555 2 h at 37°C, 100 rpm. The samples crosslinked with DSS and DSG were diluted with 100 mM
556 ammonium bicarbonate to a final urea concentration of 1.5 M, whereas the samples crosslinked
557 with EGS were diluted using LC-H_2O. For digestion, 1 µG of sequencing grade trypsin (Promega,
558 V5111) was added for 16 h at 37°C, 1000 rpm. The digested samples were acidified with 10%
559 formic acid to a final pH of 3.0. Peptides were purified and desalted using C18 reverse phase
560 columns (The Nest Group, Inc.) following the manufacturer's recommendations. Dried peptides
561 were reconstituted in 50 µL of 2% acetonitrile and 0.1% formic acid prior to MS analysis.

562 **Liquid chromatography tandem mass spectrometry (LC-MS/MS).** A total of 1 µL of peptides
563 were analyzed on an Orbitrap Eclipse mass spectrometer connected to an ultra-high
564 performance liquid chromatography Dionex Ultra300 system (both Thermo Scientific). The
565 peptides were loaded and concentrated on an Acclaim PepMap 100 C18 precolumn (75 µm × 2
566 cm,) and then separated on an Acclaim PepMap RSLC column (75 µm × 25 cm, nanoViper,C18,

567 2 µm, 100 Å) (both columns Thermo Scientific), at a column temperature of 45°C and a maximum
568 pressure of 900 bar. A linear gradient of 2% to 25% of 80% acetonitrile in aqueous 0.1% formic
569 acid was run for 100 min followed by a linear gradient of 25% to 40% of 80% acetonitrile in
570 aqueous 0.1% formic acid for 20 min. One full MS scan (resolution 120,000; mass range of 400-
571 1600 *m/z*) was followed by MS/MS scans (resolution 15,000) with a 3 sec cycle time. Precursors
572 with a charge state of 3-8 were included. The precursor ions were isolated with 1.6 *m/z* isolation
573 window and fragmented using higher-energy collisional-induced dissociation (HCD) at
574 normalized collision energies (NCE) of 21, 26, 31. The dynamic exclusion was set to 45 s.

575 **Crosslinking data analysis.** All spectra from the crosslinked samples were analyzed using pLink
576 2 (version 2.3.11). The target protein database contained the sequence for Hcp1 and Hcp2 and
577 common contaminants. pLink2 was run using default settings for conventional HCD DSS-
578 H12/D12, DSG-H8/D8 or EGS-H12/D12 crosslinking, with trypsin as the protease and up to 3
579 missed cleavages allowed. Peptides with a mass range of 600–6000 *m/z* were selected (peptide
580 length 6–60 residues) and the precursor and fragment tolerance were set to 20 and 20 ppm,
581 respectively. The results were filtered with a filter tolerance of 10 ppm and a 5% FDR. The mass
582 spectrometry data has been deposited to the ProteomeXchange consortium via the MassIVE
583 partner repository (https://massive.ucsd.edu/) with the dataset identifier PXD055833. The non-
584 crosslinked control samples were analyzed in MaxQuant (v. 2.6.2.0) against the same sequence
585 database as above. Fully tryptic digestion was used allowing two missed cleavages.
586 Carbamidomethylation (C) was set to static and protein N-terminal acetylation and oxidation (M)
587 to variable modifications. Mass tolerance for precursor ions was set to 10 ppm and for fragment
588 ions to 0.02 Da. The protein false discovery rate was set to 1%.

589 **Circos plots and structural analysis.** Circos plots were obtained in Rstudio using the Circlize
590 library (42). The AlphaFold server (43) was used to create hexamer structures using default
591 parameters and a random seed number. Visualization of the structures, distance measurements
592 between residues and structural alignments were performed in Chimera (44).

593 **ACKNOWLEDGEMENTS**

594 • We acknowledge support from Mikael Lindberg and the Protein Expertise Platform (PEP),
595 which is the Umeå node of the national infrastructure Protein Production Sweden (PPS),
596 is gratefully acknowledged.

- We acknowledge support from Lotta Happonen and the Swedish National Infrastructure for Biological Mass Spectrometry (BioMS) and the SciLifeLab, Integrated Structural Biology (ISB) platform, is gratefully acknowledged.
- We acknowledge Dr. Andrea Puhar, Wellcome-Wolfson Institute for Experimental Medicine, Queen's University Belfast, United Kingdom for critical reading.
- Electron microscopy instruments and methods were supported by Umeå Centre for Electron Microscopy (UCEM), host of the SciLifeLab national Cryo-EM infrastructure supported by the Kempe foundations and Knut and Alice Wallenberg foundations.
- B.E.U. received support from the Swedish Research Council (2019-01720, 2016-06598), Kempestiftelserna (JCK-1527.1, JCK-1724) and the Faculty of Medicine, Umeå University (Insamlingsstiftelsen grant 2021-2023).
- D.A.C. received support from the Kempestiftelserna (SMK-1860), Carl Tryggers Stiftelse för Vetenskaplig Forskning (CTS 15-96 and CTS 18-65) and the Faculty of Medicine, Umeå University (Insamlingsstiftelsen).

CONTRIBUTIONS (CRediT taxonomy)

Conceptualization, data curation, formal analysis and writing original draft: DAC

Funding acquisition, supervision, resources and project administration: DAC & BEU

Visualization: SGSM, MKSAA, DAC

Methodology, Investigation, review and editing: SGSM, MKSAA, AK, US, MH, PB, LS, BEU & DAC

The authors declare no competing interests

REFERENCES

1. Shepherd ES, DeLoache WC, Pruss KM, Whitaker WR, Sonnenburg JL. An exclusive metabolic niche enables strain engraftment in the gut microbiota. Nature. 2018 May;557(7705):434–8.

2. García-Bayona L, Comstock LE. Bacterial antagonism in host-associated microbial communities. Science. 2018 Sep 21;361(6408):eaat2456.

3. Basler M, Pilhofer M, Henderson GP, Jensen GJ, Mekalanos JJ. Type VI secretion requires a dynamic contractile phage tail-like structure. Nature. 2012 Feb;483(7388):182–6.

4. Böck D, Medeiros JM, Tsao HF, Penz T, Weiss GL, Aistleitner K, et al. In situ architecture, function, and evolution of a contractile injection system. Science. 2017 Aug 18;357(6352):713–7.

629 5. Brunet YR, Henin J, Celia H, Cascales E. Type VI secretion and bacteriophage tail tubes share a
630 common assembly pathway. EMBO reports. 2014 Mar;15(3):315–21.

631 6. Pukatzki S, Ma AT, Sturtevant D, Krastins B, Sarracino D, Nelson WC, et al. Identification of a
632 conserved bacterial protein secretion system in Vibrio cholerae using the Dictyostelium host
633 model system. Proceedings of the National Academy of Sciences. 2006 Jan 31;103(5):1528–33.

634 7. Mougous JD, Cuff ME, Raunser S, Shen A, Zhou M, Gifford CA, et al. A Virulence Locus of
635 Pseudomonas aeruginosa Encodes a Protein Secretion Apparatus. Science. 2006 Jun
636 9;312(5779):1526–30.

637 8. Hood RD, Singh P, Hsu F, Güvener T, Carl MA, Trinidad RRS, et al. A Type VI Secretion System
638 of Pseudomonas aeruginosa Targets a Toxin to Bacteria. Cell Host & Microbe. 2010 Jan
639 21;7(1):25–37.

640 9. Monjarás Feria J, Valvano MA. An Overview of Anti-Eukaryotic T6SS Effectors. Front Cell
641 Infect Microbiol. 2020;10:584751.

642 10. Wang J, Brodmann M, Basler M. Assembly and Subcellular Localization of Bacterial Type VI
643 Secretion Systems. Annu Rev Microbiol. 2019 Sep 8;73:621–38.

644 11. Cherrak Y, Flaugnatti N, Durand E, Journet L, Cascales E. Structure and Activity of the Type VI
645 Secretion System. Microbiology Spectrum. 2019 Jul 12;7(4):10.1128/microbiolspec.psib-0031–
646 2019.

647 12. Allsopp LP, Bernal P. Killing in the name of: T6SS structure and effector diversity. Microbiology.
648 2023;169(7):001367.

649 13. Silverman JM, Agnello DM, Zheng H, Andrews BT, Li M, Catalano CE, et al. Haemolysin
650 Coregulated Protein Is an Exported Receptor and Chaperone of Type VI Secretion Substrates.
651 Molecular Cell. 2013 Sep 12;51(5):584–93.

652 14. Howard SA, Furniss RCD, Bonini D, Amin H, Paracuellos P, Zlotkin D, et al. The Breadth and
653 Molecular Basis of Hcp-Driven Type VI Secretion System Effector Delivery. mBio. 2021
654 Jun;12(3):10.1128/mbio.00262-21.

655 15. Chu H, Khosravi A, Kusumawardhani IP, Kwon AHK, Vasconcelos AC, Cunha LD, et al. Gene-
656 microbiota interactions contribute to the pathogenesis of inflammatory bowel disease. Science.
657 2016 May;352(6289):1116–20.

658 16. Coyne MJ, Roelofs KG, Comstock LE. Type VI secretion systems of human gut Bacteroidales
659 segregate into three genetic architectures, two of which are contained on mobile genetic elements.
660 BMC genomics. 2016;17(1):58.

661 17. Russell AB, Wexler AG, Harding BN, Whitney JC, Bohn AJ, Goo YA, et al. A type VI secretion-
662 related pathway in Bacteroidetes mediates interbacterial antagonism. Cell host & microbe. 2014
663 Aug;16(2):227–36.

664 18. Bongiovanni TR, Latario CJ, Le Cras Y, Trus E, Robitaille S, Swartz K, et al. Assembly of a
665 unique membrane complex in type VI secretion systems of Bacteroidota. Nat Commun. 2024 Jan
666 10;15(1):429.

667 19. Wexler AG, Bao Y, Whitney JC, Bobay LM, Xavier JB, Schofield WB, et al. Human symbionts
668 inject and neutralize antibacterial toxins to persist in the gut. Proc Natl Acad Sci USA. 2016 Mar
669 29;113(13):3639–44.

670 20. Chatzidaki-Livanis M, Geva-Zatorsky N, Comstock LE. Bacteroides fragilis type VI secretion
671 systems use novel effector and immunity proteins to antagonize human gut Bacteroidales species.
672 pnas.org. 2016;

673 21. Hecht AL, Casterline BW, Earley ZM, Goo YA, Goodlett DR, Bubeck Wardenburg J. Strain
674 competition restricts colonization of an enteric pathogen and prevents colitis. EMBO reports.
675 2016 Sep;17(9):1281–91.

676 22. Verster AJ, Ross BD, Radey MC, Bao Y, Goodman AL, Mougous JD, et al. The Landscape of
677 Type VI Secretion across Human Gut Microbiomes Reveals Its Role in Community Composition.
678 Cell host & microbe. 2017 Sep;22(3):411-419.e4.

679 23. Bao H, Coyne MJ, García-Bayona L, Comstock LE. Analysis of Effector and Immunity Proteins
680 of the GA2 Type VI Secretion Systems of Gut Bacteroidales. Journal of Bacteriology. 2022 Jun
681 23;204(7):e00122-22.

682 24. Blondel CJ, Jiménez JC, Contreras I, Santiviago CA. Comparative genomic analysis uncovers 3
683 novel loci encoding type six secretion systems differentially distributed in Salmonella serotypes.
684 BMC Genomics. 2009 Aug 4;10(1):354.

685 25. Ma J, Pan Z, Huang J, Sun M, Lu C, Yao H. The Hcp proteins fused with diverse extended-toxin
686 domains represent a novel pattern of antibacterial effectors in type VI secretion systems.
687 Virulence. 2017 Oct 3;8(7):1189–202.

688 26. Anderson AJG, Morrell B, Lopez Campos G, Valvano MA. Distribution and diversity of type VI
689 secretion system clusters in Enterobacter bugandensis and Enterobacter cloacae. Microbial
690 Genomics. 2023;9(12):001148.

691 27. Karimova G, Pidoux J, Ullmann A, Ladant D. A bacterial two-hybrid system based on a
692 reconstituted signal transduction pathway. Proceedings of the National Academy of Sciences.
693 1998 May 12;95(10):5752–6.

694 28. He W, Wu K, Ouyang Z, Bai Y, Luo W, Wu D, et al. Structure and assembly of type VI secretion
695 system cargo delivery vehicle. Cell Reports. 2023 Jul 25;42(7):112781.

696 29. Merkley ED, Rysavy S, Kahraman A, Hafen RP, Daggett V, Adkins JN. Distance restraints from
697 crosslinking mass spectrometry: Mining a molecular dynamics simulation database to evaluate
698 lysine–lysine distances. Protein Science. 2014 Jun;23(6):747–59.

699 30. Müller JB, Geyer PE, Colaço AR, Treit PV, Strauss MT, Oroshi M, et al. The proteome landscape
700 of the kingdoms of life. Nature. 2020 Jun;582(7813):592–6.

701 31. Zhang D, de Souza RF, Anantharaman V, Iyer LM, Aravind L. Polymorphic toxin systems:
702 Comprehensive characterization of trafficking modes, processing, mechanisms of action,
703 immunity and ecology using comparative genomics. Biology Direct. 2012 Jun;7(1):18.

704 32. Ruhe ZC, Low DA, Hayes CS. Polymorphic Toxins and Their Immunity Proteins: Diversity,
705 Evolution, and Mechanisms of Delivery. Annual Review of Microbiology. 2020 Sep 8;74(Volume
706 74, 2020):497–520.

707 33. Korotkov KV, Hol WGJ. Structure of the GspK–GspI–GspJ complex from the enterotoxigenic
708 Escherichia coli type 2 secretion system. Nat Struct Mol Biol. 2008 May;15(5):462–8.

709 34. Cisneros DA, Bond PJ, Pugsley AP, Campos M, Francetic O. Minor pseudopilin self-assembly
710 primes type II secretion pseudopilus elongation. EMBO J. 2012 Feb 15;31(4):1041–53.

711 35. Cisneros DA, Pehau-Arnaudet G, Francetic O. Heterologous assembly of type IV pili by a type II
712 secretion system reveals the role of minor pilins in assembly initiation. Mol Microbiol. 2012
713 Nov;86(4):805–18.

714 36. Nguyen M, Wu TH, Danielson KJ, Khan NM, Zhang JZ, Craig L. Mechanism of secretion of
715 TcpF by the Vibrio cholerae toxin-coregulated pilus. Proceedings of the National Academy of
716 Sciences. 2023 Apr 18;120(16):e2212664120.

717 37. McCallum M, Burrows LL, Howell PL. The Dynamic Structures of the Type IV Pilus.
718 Microbiology Spectrum. 2019 Mar;7(2):10.1128/microbiolspec.psib-0006–2018.

719 38. Saulino ET, Thanassi DG, Pinkner JS, Hultgren SJ. Ramifications of kinetic partitioning on
720 usher-mediated pilus biogenesis. The EMBO Journal. 1998 Apr 15;17(8):2177–85.

721 39. Busch A, Phan G, Waksman G. Molecular mechanism of bacterial type 1 and P pili assembly.
722 Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering
723 Sciences. 2015 Mar 6;373(2036):20130153.

724 40. Katoh K, Rozewicki J, Yamada KD. MAFFT online service: multiple sequence alignment,
725 interactive sequence choice and visualization. Briefings in Bioinformatics. 2019 Jul
726 19;20(4):1160–6.

727 41. Bianchini G, Sánchez-Baracaldo P. TreeViewer: Flexible, modular software to visualise and
728 manipulate phylogenetic trees. Ecology and Evolution. 2024;14(2):e10873.

729 42. Gu Z, Gu L, Eils R, Schlesner M, Brors B. circlize implements and enhances circular
730 visualization in R. Bioinformatics. 2014 Oct 1;30(19):2811–2.

731 43. Abramson J, Adler J, Dunger J, Evans R, Green T, Pritzel A, et al. Accurate structure prediction
732 of biomolecular interactions with AlphaFold 3. Nature. 2024 Jun;630(8016):493–500.

733 44. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, et al. UCSF
734 Chimera—A visualization system for exploratory research and analysis. Journal of Computational
735 Chemistry. 2004;25(13):1605–12.

736

Figure 1

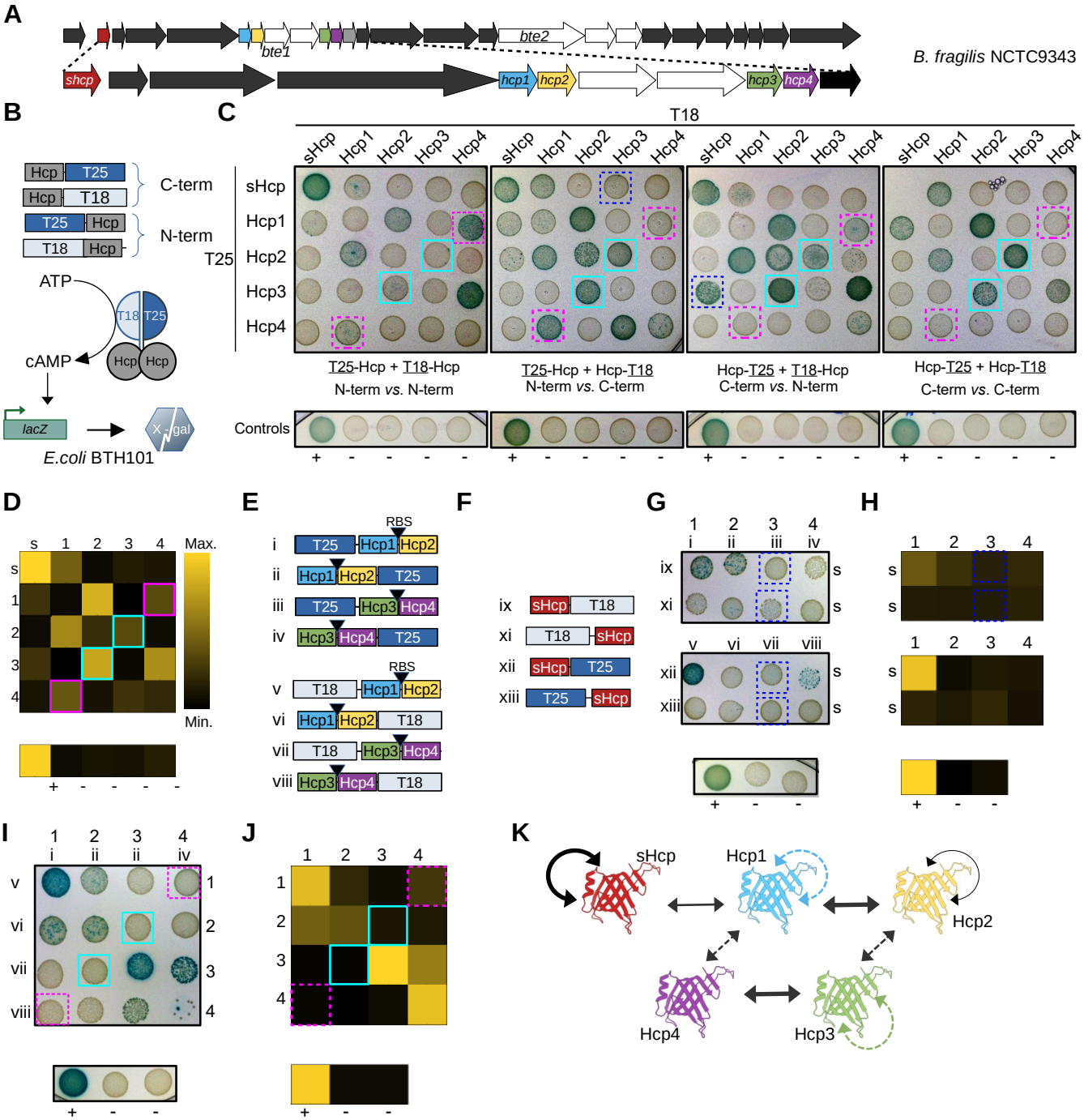

Figure 1. Two-hybrid assays of the major and minor Hcp proteins from *B. fragilis*' T6SS.
A) Scheme of *B. fragilis* NCTC9343 T6SS locus. **B)** Scheme of genes coding for N- and C- terminal Hcp proteins fused to the T18 and T25 fragments of the adenylate cyclase CyaA from *Bordetella pertussis*. Interaction between proteins fused to a T18 and T25 fragment in the *cyaA⁻ E. coli* strain BTH101 leads to the expression of β-galactosidase and hydrolysis of X-gal, which turns the colonies blue. **C)** Photographic images of two-hybrid assays with plasmids carrying genes encoding Hcp proteins fused to either a C-terminal T18 and T25 fragments(left panel) or a N-terminal T25 and a C-terminal T18 fragment (centre left panel) or a C-terminal T25 and a N-terminal T18 fragment (centre-right panel) or a N-terminal T18 and T25 fragments (right panel). Hcp proteins are labeled with characters: sHcp (s), Hcp1 (1), Hcp2 (2), Hcp3 (3) and Hcp4 (4). The positive control (+) is the dimerization of the yeast GCN4 leucine zipper and the negative controls (-) are the empty vectors in four different combinations. Picture were taken at 24h from plates containing 250 mM IPTG. **D)** Quantification of two-hybrid assays. The average hydrolized X-gal of all four experiments shown in B plotted as a color scale in arbitrary units normalized to both the GCN4 leucine zipper (Zip) positive and negative controls. The color scale between black and yellow shows the minimum and maximum, respectively in the plot. The data represents the average of data obtained with varying concentrations of IPTG, at 24h and 36h and at 30 °C and 37 °C. **E)** Scheme of genes coding for N- and C- terminal sHcp proteins fused to the T18 and T25 fragments used for two-hybrid assays shown in G-H. **F)** Scheme of genes coding simultaneous expression of two Hcp proteins fused to T18 and/or T25 fragments used for two-hybrid assays shown in H. An artificial *E. coli* ribosome binding site (RBS) was placed upstream of genes encoding a second Hcp subunit. **G)** Photographic images of competitive two-hybrid assay with plasmids encoding sHcp fused to N- or C- terminal T18 or T25 fragments tested for pairwise interactions against T18 or T25 fusions of Hcp1 (1), Hcp2 (2), Hcp3 (3) and Hcp4 (4) in the presence of Hcp2, Hcp1, Hcp4 and Hcp3, respectively. Pictures were taken at 36h. **H)** Photographic images of competitive and cooperative two-hybrid assay for pairwise interactions among T18 or T25 translational fusions of Hcp1 (1), Hcp2 (2), Hcp3 (3) and Hcp4 (4) in the presence of Hcp2, Hcp1, Hcp4 and Hcp3, respectively. Pictures were taken at 36h. **I-J)** Quantification of two-hybrid assays shown in G and H. The normalized average hydrolized X-gal plotted as a color scale in arbitrary units. The color scale is the same as in C. The data represents the average of data obtained at 250 mM IPTG, at 24h and 36h and at 30 °C and 37 °C. **K)** Scheme of interaction network among Hcp subunits derived from all two-hybrid assays. The wider arrows represent the most frequently observed interactions. Dashed arrows represent interactions that are inhibited by other pairwise interactions. Coloured arrow represent cooperative interactions that depend on another interaction.

Figure S1

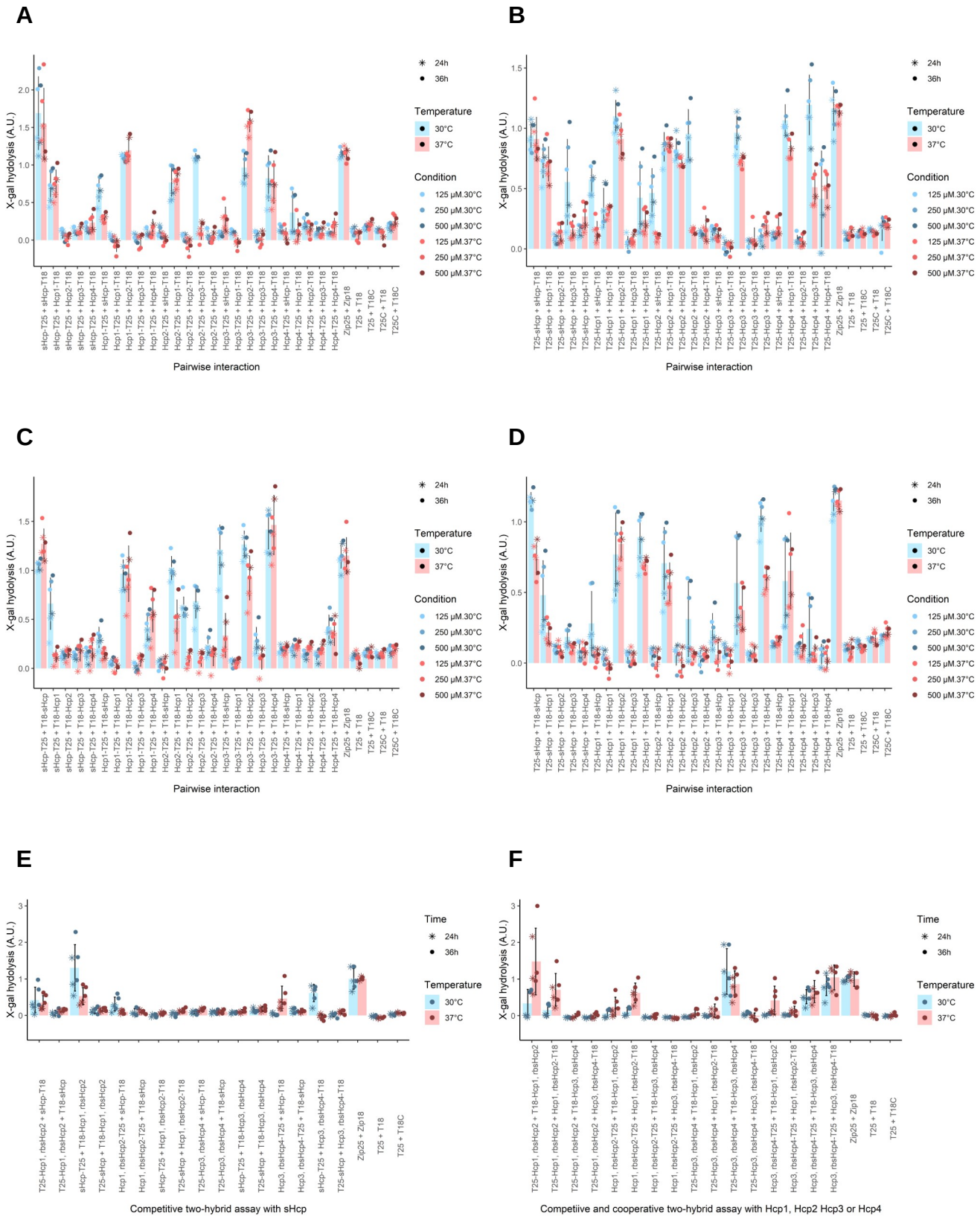

Figure S1. Quantification of normalized X-gal hydrolisis.
Quantification of two-hybrid assays of the major and minor Hcp proteins from *B. fragilis* with plasmids carrying genes encoding Hcp proteins fused to either **A)** a C-terminal T18 and T25 fragments or **B)** a N-terminal T25 and a C-terminal T18 fragment or **C)** a C-terminal T25 and a N-terminal T18 fragment or **D)** a N-terminal T18 and T25 fragments. **E)** Quantification of competitive two-hybrid assay with plasmids encoding sHcp fused to N- or C- terminal T18 or T25 fragments tested for pairwise interactions with T18 or T25 fusions of Hcp1, Hcp2, Hcp3 or Hcp4 in the presence of Hcp3, Hcp2, Hcp5 and Hcp4, respectively. **F)** Quantification of competitive and cooperative two-hybrid assay for pairwise interactions pairwise interactions among T18 or T25 translational fusions of Hcp1, Hcp2, Hcp3 and Hcp4 in the presence of Hcp2, Hcp1, Hcp4 and Hcp3, respectively.

Figure S2

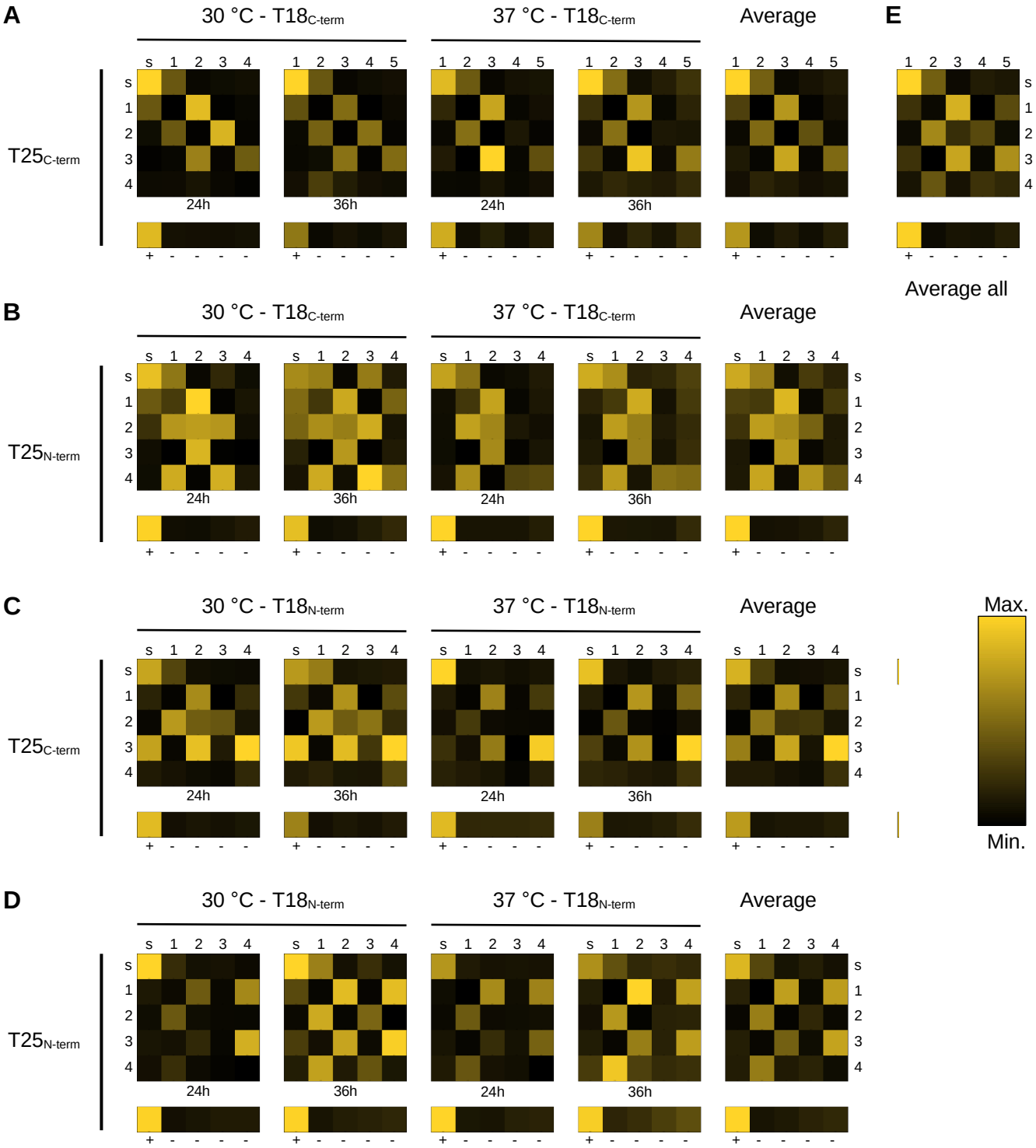

Figure S2. Quantification of two-hybrid assays. The average hydrolized X-gal for each condition tested for two-hybrid assays with plasmids carrying genes encoding Hcp proteins fused to either **A)** a C-terminal T18 and T25 fragments or **B)** a N-terminal T25 and a C-terminal T18 fragment or **C)** a C-terminal T25 and a N-terminal T18 fragment or **D)** a N-terminal T18 and T25 fragments. Hcp proteins are labeled with characters: sHcp (s), Hcp1 (1), Hcp2 (2), Hcp3 (3) and Hcp4 (4). The positive control (+) is the dimerization of the yeast GCN4 leucine zipper and the negative controls (-) are the empty vectors in four different combinations.The data is plotted as a color scale in arbitrary units normalized to both the positive and negative controls. The color scale between black and yellow shows the minimum and maximum, respectively in each individual plot.

Figure S3

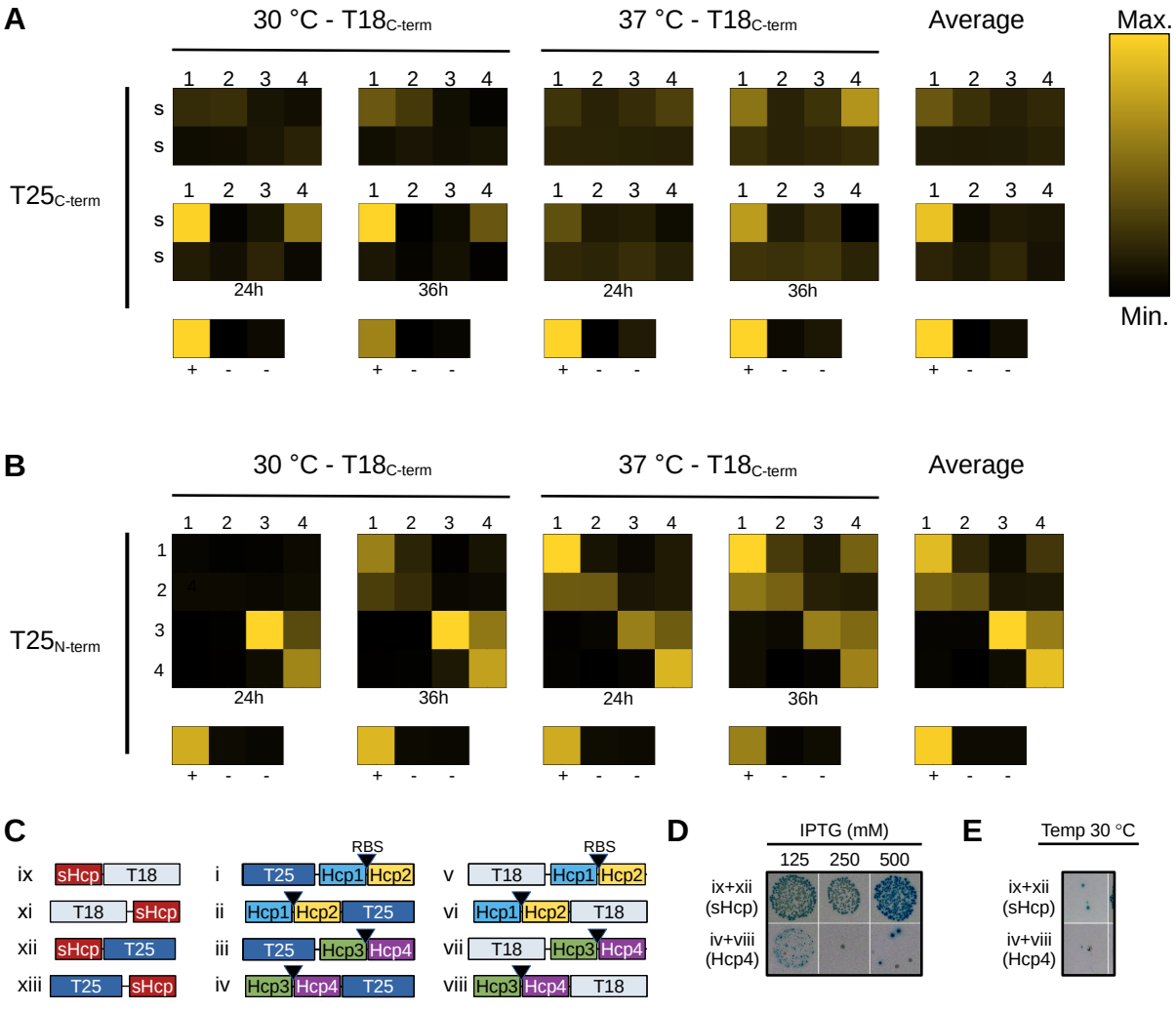

Figure S3. A) The average hydrolized X-gal for each condition tested for competitive two-hybrid assay with plasmids encoding sHcp fused to N- or C-terminal T18 or T25 fragments tested for pairwise interactions against T18 or T25 fusions of Hcp1 (1), Hcp2 (2), Hcp3 (3) and Hcp4 (4) in the presence of Hcp2, Hcp1, Hcp4 and Hcp3, respectively. The color scale between black and yellow shows the minimum and maximum, respectively in each individual plot. **B)** The average hydrolized X-gal for each condition tested in competitive and cooperative two-hybrid assay for pairwise interactions between T18 or T25 translational fusions of Hcp1 (1), Hcp2 (2), Hcp3 (3) and Hcp4 (4) in the presence of Hcp2, Hcp1, Hcp4 and Hcp3, respectively. Color scale as in A). **C)** Construct of translational T18 or T25 fusions used for the assay shown in A-B) and they follow the same scheme shown in Figure 1. **D-E)** Effect of IPTG concentration and temperature on two hybrid assays of interaction between sHcp C-terminal translational fusions and between constructs carrying Hcp4 T18 and T25 translational fusions in the rpesence of Hcp3. **D**) Photographic images of two-hybrid assays performed at 37 °C and at different IPTG concentrations. **E)** Photographic images of two-hybrid assays performed at 30 °C at a concentration of 125 mM IPTG.

Figure 3

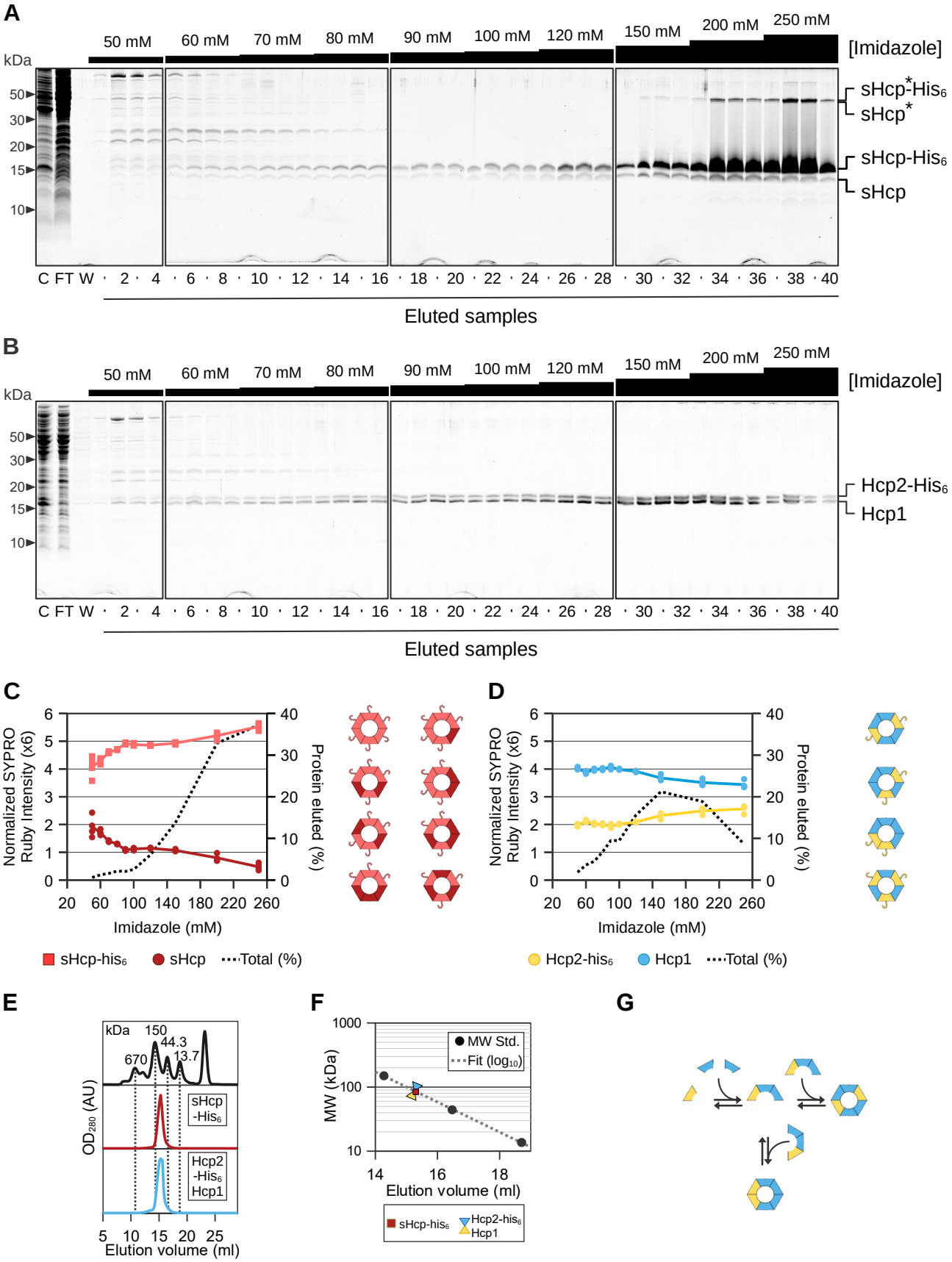

Figure 3. Stoichiometry of purified Hcp complexes as determined by an imidazole step gradient. A-B) SYPRO Ruby stained SDS-PAGE gels of samples obtained from the affinity co-purification of sHcp and sHcp-His$_6$ or Hcp1 and Hcp2-His$_6$. Four elution steps were performed for each imidazole concentration (black rectangles on top of the gels). Labels: C (clear fraction), FT (flow-trough), W (wash), 1-40 (elution fractions). **C- D)** Densitometric quantification of SYPRO Ruby fluorescence intensity from the gels shown in A and B. The average fractional density of fluorescence intensity of the two regions of interest (ROI) corresponding to each purified Hcp protein (with and without His$_6$ tag) normalized to the total density multiplied by 6 (x6) is plotted against the imidazole concentration. The percent of eluted protein is shown in the secondary *y*-axis as the sum of the two ROIs in each lane, divided by the sum of all ROIs and multiplied by 100. The possible stoichiometric arrangements of sHcp-sHcp-His6 and Hcp1-Hcp2 complexes are indicated by color coded filed hexamers: Light red (sHcp-His$_6$), dark red (sHcp), yellow (Hcp2-His$_6$) and skyblue (Hcp1). **D)** SEC analysis of a protein-standard mixture (top) purified sHcp-His6 (midddle, red) and the co-purified complex of Hcp2-His6 with Hcp1 (bottom, blue). **F)** Calibration of size exclusion chromatography molecular weight standards. Elution volumes of the SEC molecular weight protein-standards (balck circles) are plotted against their molecular weights. A linear fit using three of the log$_{10}$ MW of the standards (dotted line) and the extrapolated MW of the Hcp complexes are also shown. The intrapolated MW of sHcp-His$_6$ was 85.1 kDa (red square). The intrapolated MW for the Hcp2-His$_6$ co-purified with Hcp1 was 86.8 kDa (yellow-blue triangles).

Figure 2

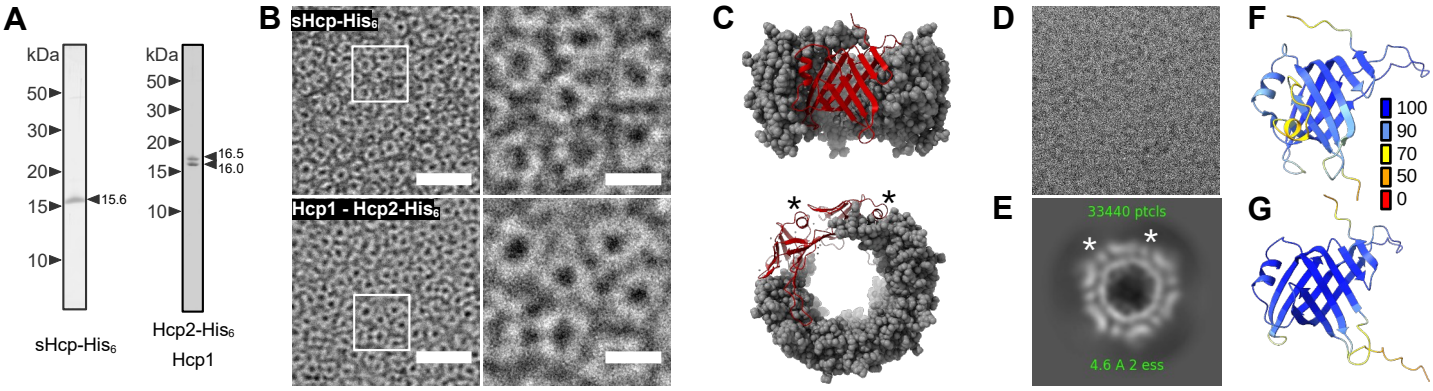

Figure 2. Purification and biochemical characterization of the major and minor Hcp subunits. A) Phosphoimager scan of SYPRO Ruby stained SDS-PAGE gels of purified sHcp-His$_6$ and Hcp1 and Hcp2-His$_6$ complex. **B)** Negative staining electron microscopy image of sHcp-His$_6$ and Hcp1 and Hcp2-His$_6$ complex. Scale bars: 50 nm (left) and 15 nm (right). **C)** Crystall structure of sHcp.**D)** Cryo-EM data by single particle analysis of Hcp1-Hcp2-His$_6$. **E)** Main class of particles isolated after a two dimensional image single particle analysis). **F)** AlphaFold model of Hcp1, AF-Q5LDT5-F1. **G)** AlphaFold model of Hcp2, AF-A0A380Z0E1-F1. The confidence PLDDT between 0 and 100 scale is shown.

Figure S4

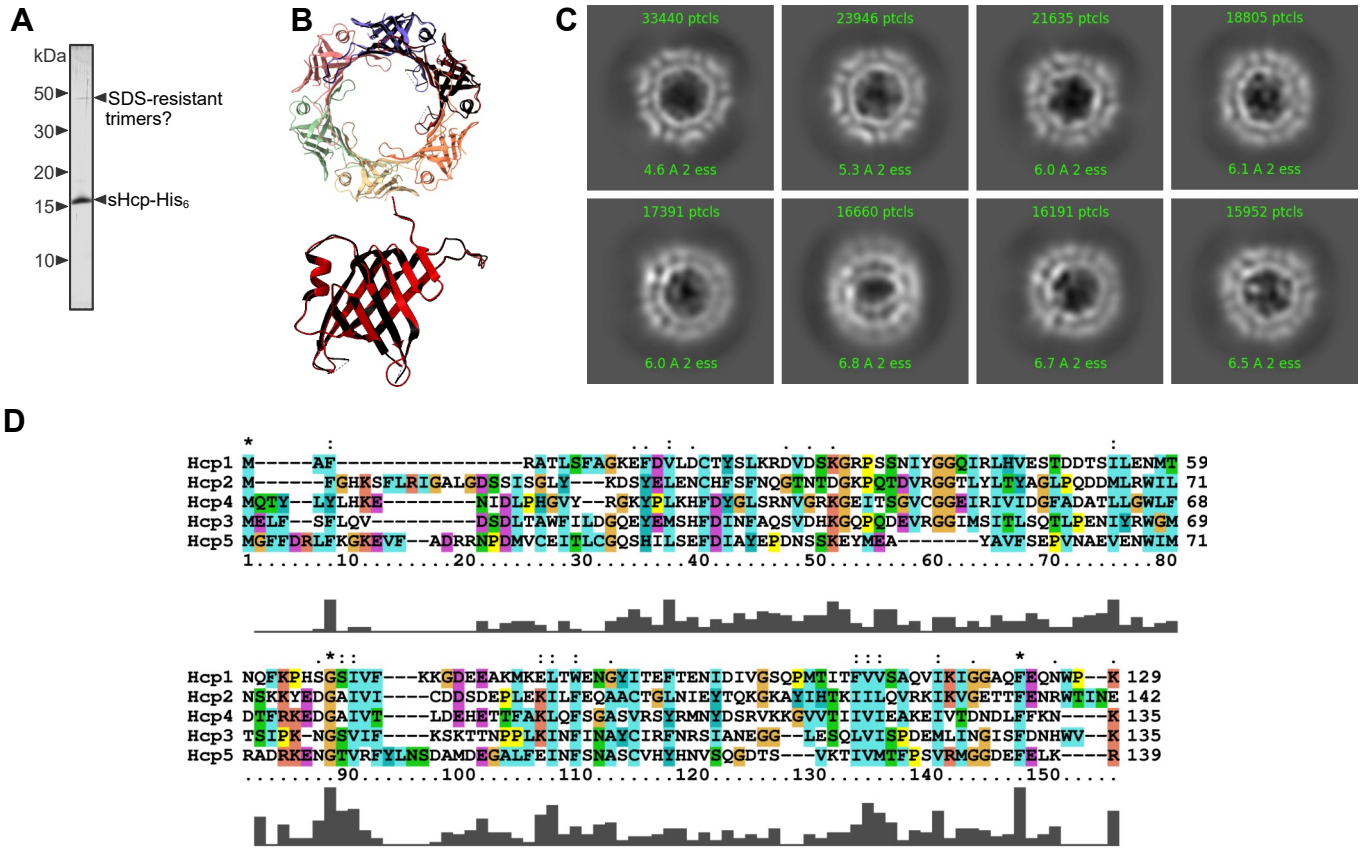

Figure S4. A) Phosphoimager scan of SYPRO Ruby stained SDS-PAGE gels of purified sHcp-His$_6$ at higher exposure. **B)** Structural alignment of to the He *et al.* unbound structure (black) to our sHcp crustal structure. **C)** Top six 2-D classes averaged from Cryo-EM data. D) G-INS-1 MAFFT alignment of *B fragilis* NCTC9343 Hcp proteins..

Figure 4

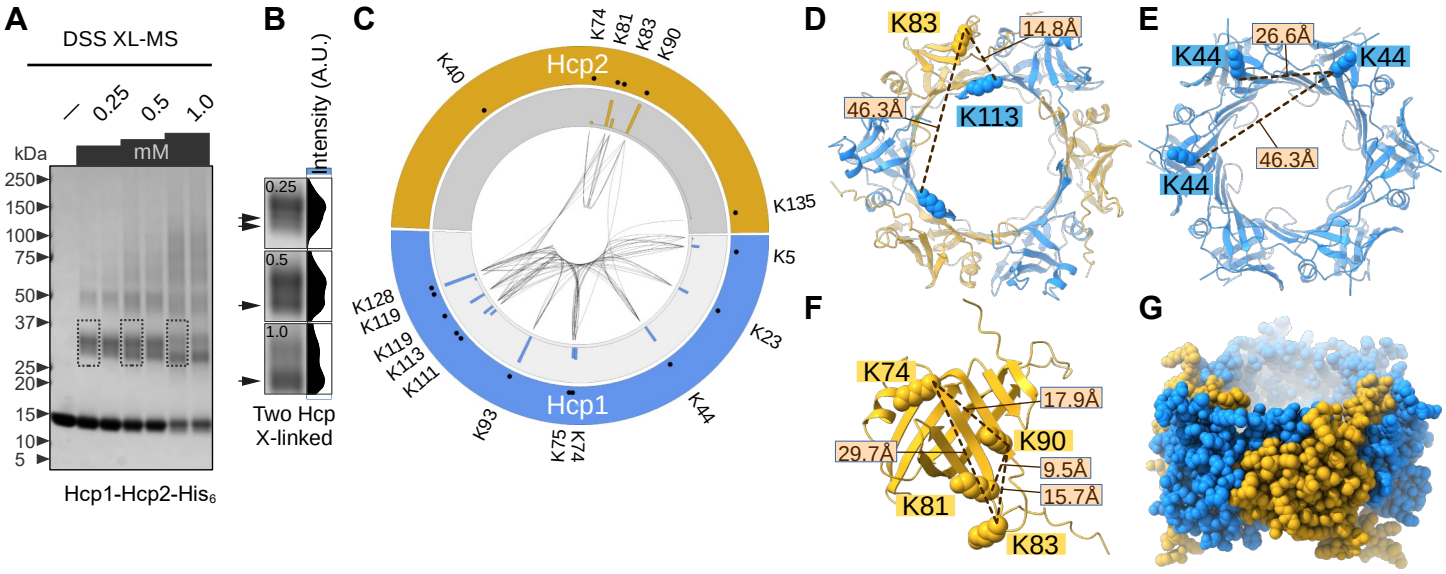

Figure 4. Crosslinking MS of the Hcp1-Hcp2 complex show homo-Hcp1 interactions.
A) SDS-PAGE of two replicates of X-linked Hcp proteins using DSS. **B)** Inset of SDS-PAGE with intensity line profile showing X-linked Hcp dimers at three DSS concentrations. **C)** Circos plot of all X-linked peptides found. Gray track contains a histogram of the total number of peptides found for each residue. **D)** Network of close residues in Hcp2 showing possible intramolecular X-links. **E)** K44 inter-protein Hcp1 X-links between residues at the same position modelled with AlphaFold as a Hcp1 homo-hexamer. **F)** .Residues found in X-linked peptides between Hcp1 and Hcp2 modelled with AlphaFold as a 3:3 complex. **G)** AlphaFold model of an Hcp1-Hcp2 complex in a 4:2 ratio.

Figure S5

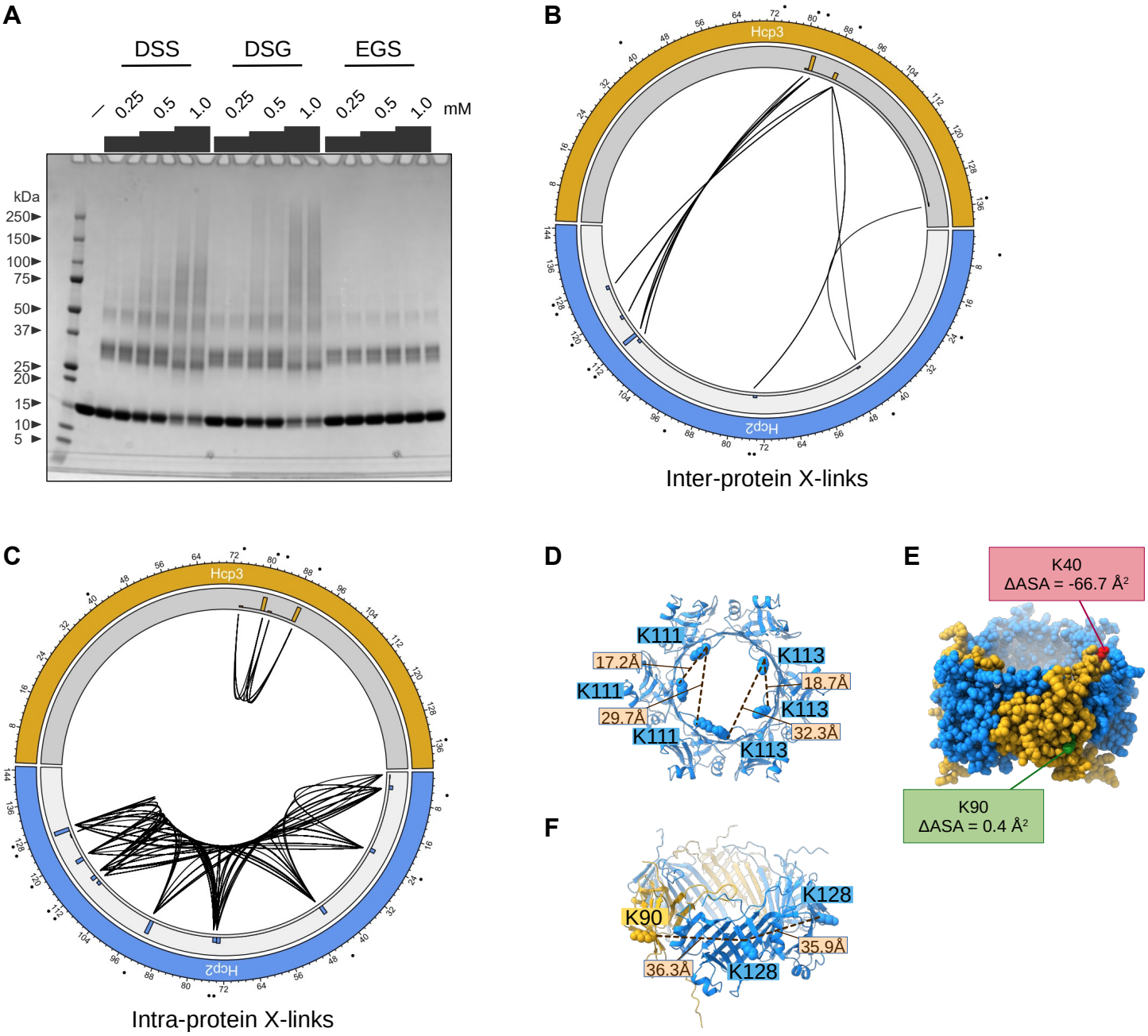

Figure S5. Crosslinking MS of the Hcp1-Hcp2 complex show homo-Hcp1 interactions.
A) SDS-PAGE of X-linked Hcp proteins using DSS, DSG and EGS. **B-C)** Circos plot of all x-linked peptides found in inter- and intra- protein X-links. Gray track contains a histogram of the total number of peptides found for each residue. **D)** 4:2 complex showing K40 in Hcp2, which was not found in any X-link. E) 4:2 complex showing lysines in Hcp1 involved in interprotein X-links as determined by their X-linking to the same residue in another Hcp1 protein.

Figure S6

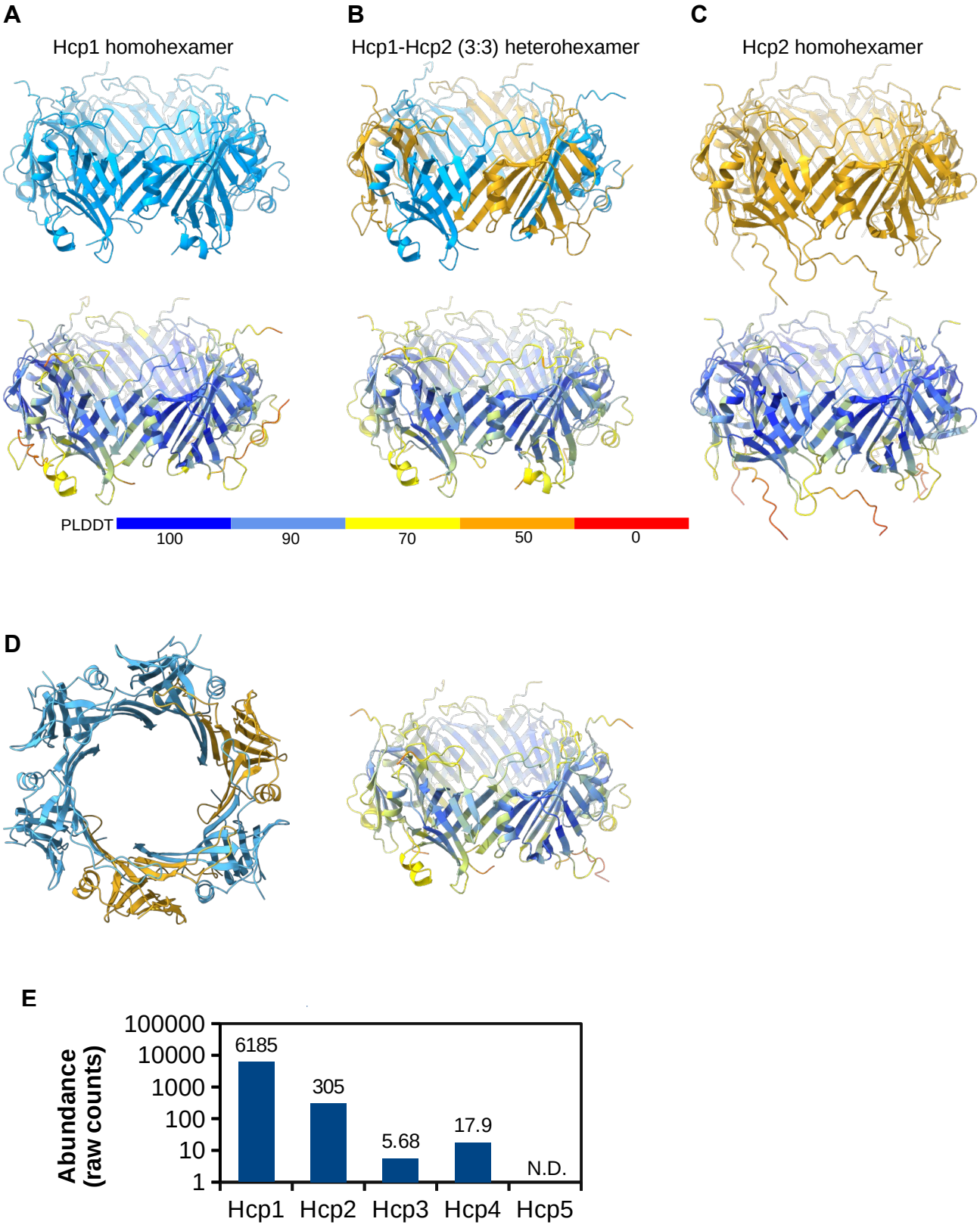

Figure S5. AlphaFold server models of Hcp1 and Hcp2.
A) Homohexameric model of Hcp1 with PLDDT score(bottom). **B)** Heterohexameric model of Hcp1 and Hcp2 with PLDDT score (bottom) in a 3:3 ratio. **C)** Homohexameric model of Hcp2 with PLDDT score (bottom). **D)** AlphaFold model of a Hcp1-Hcp2 complex in a 4:2 complex with PLDDT score (side). Hcp1 shown in blue and Hcp2 shown in yellow. **E)** Abundance of Hcp proteins in *B. fragilis* according to Müller, *et al.* (29).

Figure 5

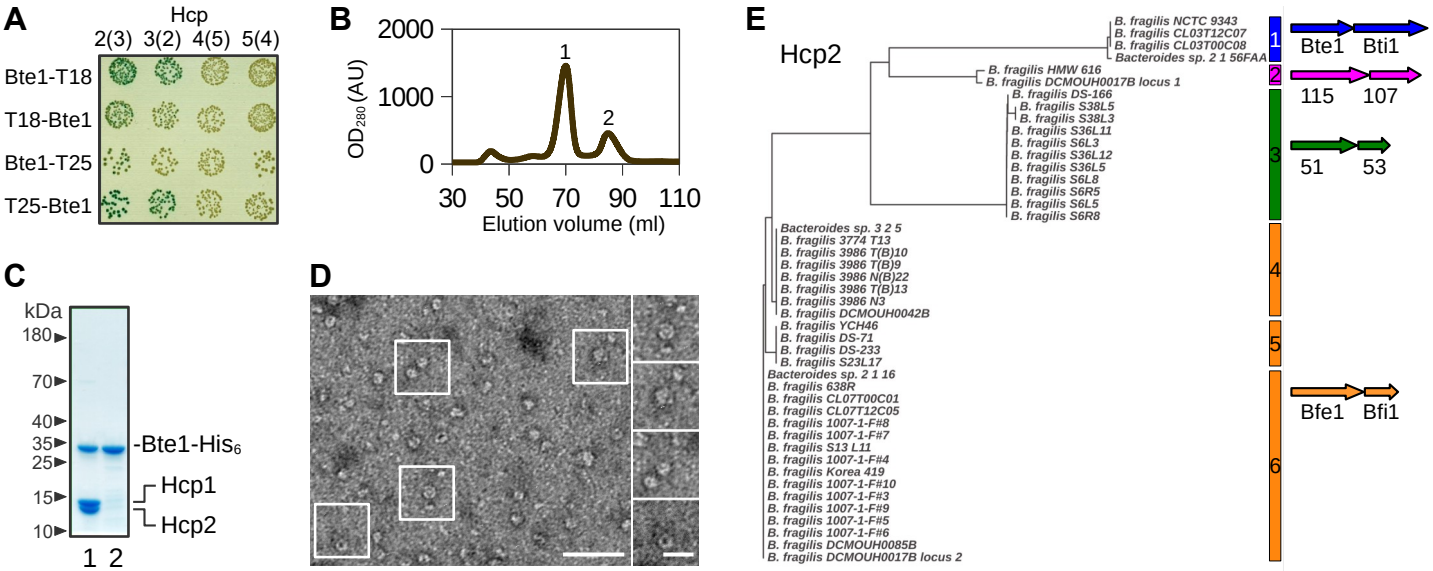

Figure 5. Hcp2 determines binding to GA3 T6SS effectors.
A) Photographic images of cooperative two-hybrid assay for pairwise interactions between T18 or T25 translational fusions of Bte1 with Hcp1 in the presence of Hcp2 (2(3)), Hcp2 in the presence of Hcp1 (3(2)), and so on. Pictures were taken at 36h. **B-C)** Size exclusion chromatography of co-purified Hcp-effector complex. Two peaks were characterized by SDS-PAGE. The first and largest peak (1) contained Hcp1 (16.0kDa), Hcp2 (15.3 kDa) and Bte1-His$_6$. The second peak (2) only contained the effector. **D)** Negative staining electron microscopy of the Hcp1-3-Bte1 effector complex. Scale bar 50 nm (insets 20 nm). . **E)** Clustering of phylogenetic tree of Hcp2 homologues from GA3 T6SS with their cognate effector immunity pair.

Figure S7

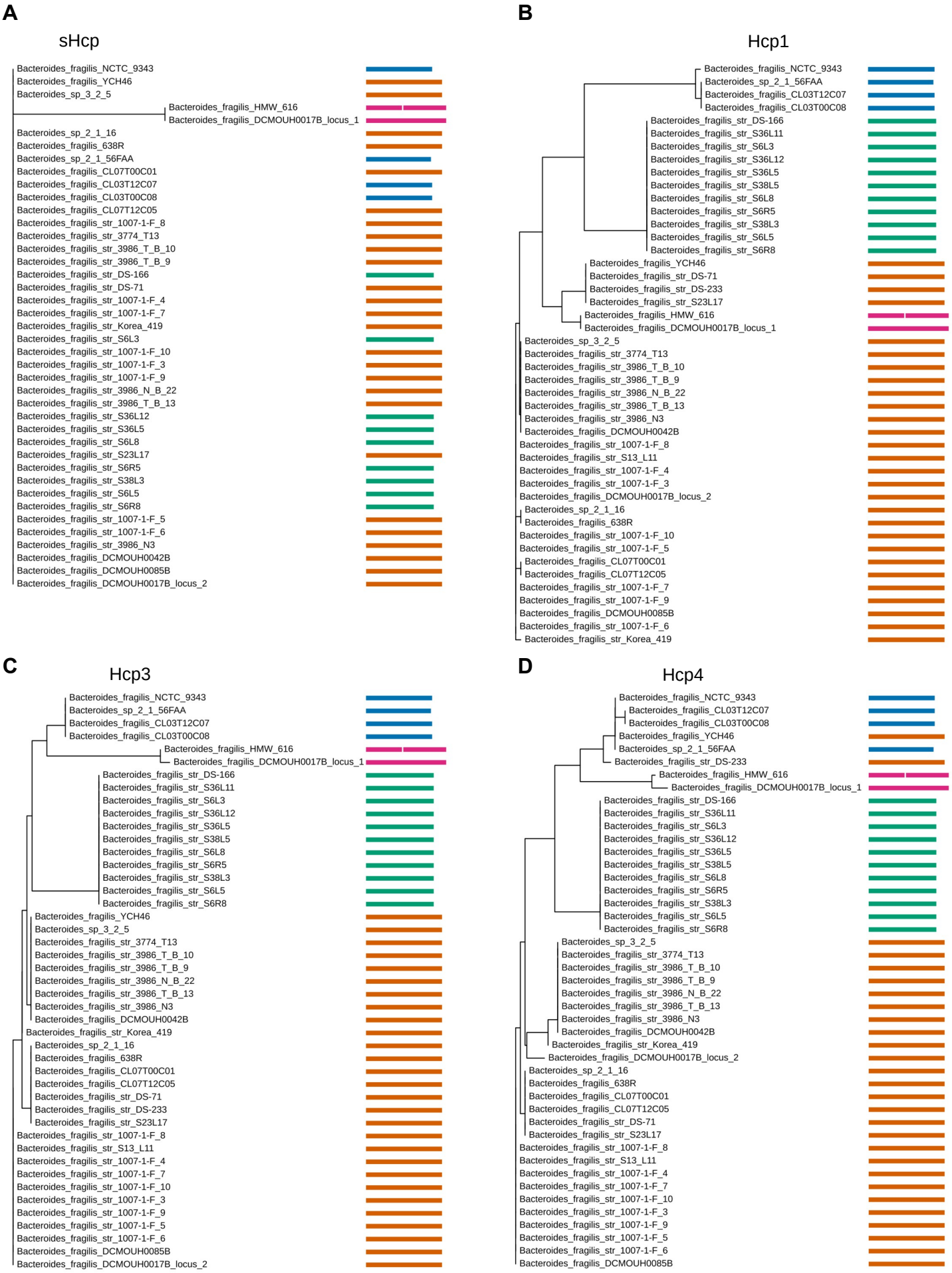

Figure S7. Clustering of phylogenetic trees of Hcp proteins from GA3 T6SS with effector proteins. Bte1 (blue), Bfe1 (orange), Cluster 51 (green), Cluster 115 (magenta). Generated with TreeViewer. Trees were compressed horizontally.

Figure S8

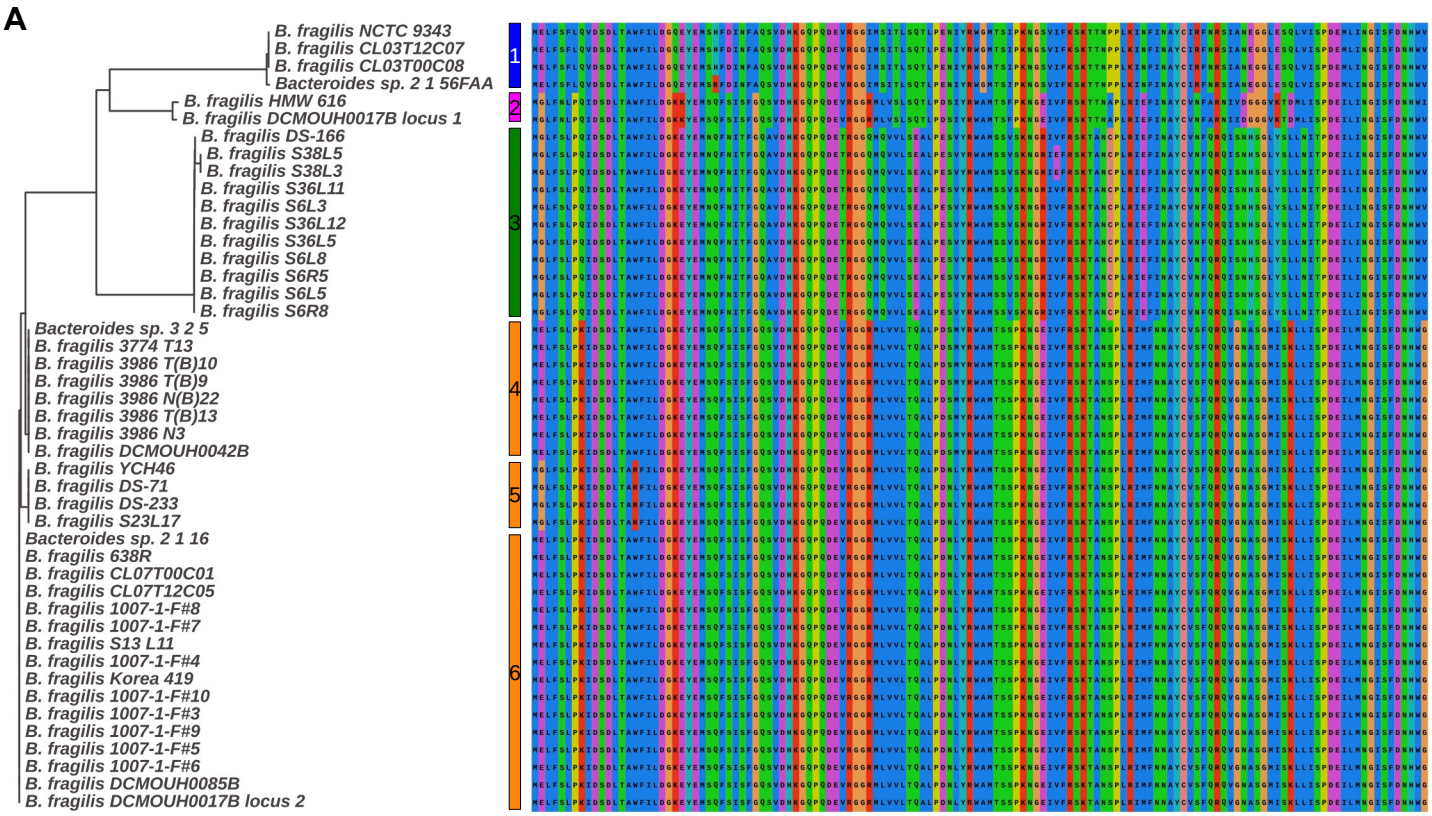

Figure S8. Hcp2 determines binding to Bacteroidales GA2 and GA3 T6SS effectors.
A) Clustering of phylogenetic tree of Hcp2 proteins from GA3 T6SS with their corresponding sequence. **B)** Clustering of phylogenetic trees of Hcp2 proteins from Bacteroidales T6SS GA2 architecture with their effectors (colour bars). The cluster number (78, 146, 147, 152 and 155) determines the HMM model named by Coyne, *et al.* 2016 (16). Two identical effectors from *Parabacteroides distasonis* strains 3999B T(B) 4 and T(B) 6 were labelled here as Pdef1. Effectors were aligned with the MAFFT-FFT-NS-i algorithm before adding the alignment to the phylogenetic tree. Spaces between the sequences are gaps in the global alignment of all effectors. Generated with TreeViewer.