Bacteroidales T6SS minor Hcp subunits form heteromers recognising effectors.

Sergio G. San-Miguel[1], Manal Kamal Saleh Al-Ammari[1], Emma Johansson[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.

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. T6SSs are gut colonization factors that influence gut biodiversity. Hemolysin-coregulated proteins (Hcp) are major structural proteins in these systems. Hcp forms hexameric rings that stack to create an inner tube structure essential for translocating effector proteins into target cells. In Bacteroidales, T6SS loci encode multiple Hcp proteins with unknown function. The gut commensal *Bacteroides fragilis* encodes five Hcp subunits (sHcp and Hcp1-4) that have low sequence similarity. In this study, we investigated the roles of these proteins. Interaction studies showed that sHcp forms homohexamers, which is consistent with a major role of forming the bulk of the inner tube. In contrast, the less abundant minor Hcp1-4 were shown to form an interaction network involving heteromeric complexes. Biochemical analyses demonstrated that Hcp1 and Hcp2 assemble into heterohexamers and that this complex recognizes the secreted effector Bte1. Finally, we showed that Hcp modules, which are encoded in highly syntenic regions in T6SS loci of Bacteroidales, cluster with effectors. These results imply that the minor Hcps genetically cosegregate with cognate effectors, contributing to effector cassette variability. Thus, minor Hcp subunits function as recognition particles for effectors to mediate secretion, which appears to be a conserved trait in Bacteroidales T6SSs. Exploiting these features could facilitate the characterization of unknown effectors by copurifying them with their cognate Hcps. This approach may reveal new insights into bacterial interactions and the mechanisms that establish gut biodiversity.

38 **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
45 to 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
55 three main complexes (reviewed recently in [10–12]). A membrane complex, formed by TssJ, TssL
56 and TssM, anchors the system to the cell envelope. The base plate formed by TssE, TssF, TssG
57 and 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 as attached to the tip complex or as cargo inside Hcp rings [13,14]. However, the
64 structural and functional features of Hcp-effector complexes are 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 (GA): GA1, GA2 and GA3 [16].
69 Compared with the better characterised Pseudomonadota T6SSs, Bacteroidales T6SSs diverge
70 in sequence and exhibit important differences [16,17]. For example, the TssJLM complex is not

present and is substituted by another megadalton structure constituted by novel proteins named TssN, TssO, TssP, TssQ, and TssR [16,18]. Another difference in Bacteroidales, is that effectors are encoded in variable regions within T6SS loci, which differ at the strain level [16,19]. Consequently, gut Bacteroidales are capable of engaging in interspecific and interstrain competition [19–21]. Therefore, the T6SS is considered a *bona fide* colonization factor for *Bacteroides* spp. and has been shown to contribute to shaping biodiversity in the gut, especially in infants [22].

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

In the present study, we probed the existence of protein–protein interactions among Bacteroidales Hcp proteins using the T6SS of *Bacteroides fragilis* as a model. Using biochemical approaches, we found that a subset of Hcp subunits, which we define here as minor Hcp subunits on the basis of their lower abundance, establish heteromeric interactions, forming chimeric Hcp particles. These particles were shown to interact with cognate partners, recognising effectors for secretion. Our results demonstrate that the characterisation of numerous Bacteroides T6SS effectors bound to cognate Hcp heteromeres can unlock the study of the molecular mechanisms mediating bacterial competition in the gut microbiota.

RESULTS

B. fragilis Hcp subunits establish heteromeric interactions

The T6SS locus of *B. fragilis* encodes five Hcp genes: *sHcp* (gene BF9343_1943), *hcp1* (BF9343_1939), *hcp2* (BF9343_1938), *hcp3* (gBF9343_1935), and *hcp4* (BF9343_1934) (Figure 1A). We hypothesised that, similar to other Hcp proteins (reviewed in [2]), *B. fragilis* Hcps would form ring-shaped hexamers. Therefore, we set out to test their organisation and, specifically, to determine whether the Hcp rings would assemble or stack in a specific order. We performed bacterial two-hybrid assays [27] for the five *B. fragilis hcp* genes with N- and C- terminal fusions

103 with the *Bordetella pertussis* T18 and T25 *cyaA* fragments. For each *hcp* gene, four fusion genes
104 were generated, resulting in four possible combinations to test all pairwise interactions (Figure
105 1B). We cotransformed *Escherichia coli* BTH101 with two plasmids carrying the fusion genes
106 and analysed the appearance of transformants on plates supplemented with X-gal. To define
107 the best parameters at which these interactions occur, we used various IPTG concentrations
108 and growth on plates at 30 °C or 37 °C. To our surprise, only fusions with sHcp showed positive
109 interactions with itself in all four two-hybrid combinations and conditions, suggesting that not
110 all combinations lead to the formation of multimers (Figures 1C and S1). The finding that only
111 this subunit can form homomers, together with the fact that it is the most upstream of the T6SS
112 locus, suggests that, as previously suspected, this gene encodes a structural Hcp protein [23].
113 Hcp1-Hcp2 also appeared consistently in all four tested combinations. An interaction between
114 Hcp3 and Hcp4 was detected in most combinations (Figure 1C). Other heterointeractions, such
115 as sHcp-Hcp3, Hcp1-Hcp4 and Hcp2-Hcp3, appeared only in some tested combinations
116 (Figure 1C, magenta, cyan and blue dashed squares, respectively). The Hcp2-Hcp2 interaction
117 occurred in only one combination (Figure 1C, C-term *vs*. N-term). To quantify an average level of
118 interaction between Hcp proteins in all four combinations, we performed densitometric
119 analysis to quantify the appearance of blue hydrolysed X-gal normalized to the positive control
120 (Figures S1 and S2). Averaging across all the tested IPTG concentrations and growth conditions
121 and all four fusion combinations, we obtained a map of interactions, which confirmed our visual
122 observations (Figure 1D). Moreover, the frequency map revealed a hierarchical interaction
123 cascade in which sHcp interacted with Hcp1, which interacted with Hcp2, and so on. Varying
124 the IPTG concentration had little effect on the level of interaction. However, we inexplicably
125 observed that some interactions were seemingly lacking at 30 °C or 37 °C or inhibited bacterial
126 growth (Figures S1, S3D and S3E). Of note, a few interactions were detectable only after
127 prolonged (36 h) incubation (Figures S1 and S2). Taken together, these results suggest that only
128 sHcp has a strong tendency to form homo-oligomeres but that the other Hcps hierarchically
129 form hetero-oligomers.

130

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

132 Two of the most frequently observed interactions occurred between the products of adjacent
133 genes, namely, Hcp1-Hcp2 and Hcp3-Hcp4. Moreover, there is a highly syntenic arrangement in
134 these two pairs of *hcp* genes across Bacteroides T6SSs with GA2 or GA3, suggesting that the
135 formation of these heterocomplexes has a functional link. Therefore, we aimed to test whether

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

147 We also wanted to determine whether new interactions would occur when adjacent pairs of
148 Hcp proteins were tested via two-hybrid assays. We observed an Hcp1-Hcp1 interaction when
149 T18 and T25 Hcp1 fusions were expressed in the presence of Hcp2 (Figure 1I and 1J), which was
150 not observed in Hcp2-Hcp2 pairwise two-hybrid assays. Similarly, we observed an Hcp3-Hcp3
151 interaction when the T18 and T25 Hcp3 fusions were expressed in the presence of Hcp4 (Figure
152 1I and 1J). These results suggest that Hcp1 and Hcp3 can partake in homointeractions but that
153 they exclusively and cooperatively occur in the presence of another Hcp. Our results also
154 showed that some observed heterointeractions are outcompeted by the expression of a second
155 Hcp: specifically, Hcp1-Hcp4 disappears when Hcp3 and Hcp4 are present and Hcp2-Hcp3
156 disappears when Hcp1 and Hcp2 are produced (Figure 1C, 1I and 1J, cyan and magenta
157 squares). These results suggest that Bacteroidales Hcp proteins form a complex network of
158 interactions (schematic in Figure 1K) but that neighbouring *hcp* genes encode cognate pairs
159 that preferentially interact with each other, so that Hcp1 preferentially binds to Hcp2 and Hcp3
160 to Hcp4.

161

162 **Hcp heterocomplexes form Hcp rings**

163 To obtain structural insight into the *B. fragilis* Hcp proteins, we purified two of the complexes
164 described above. As a control for hexamer formation, we purified a His-tagged version of sHcp
165 (Figure 2A, sHcp-His_6) which, as observed by negative staining electron microscopy, appeared
166 as rings (Figure 2B). We obtained a crystal structure of sHcp-His_6 (Figures 2C, S4A and Table
167 S1), which was identical to the VgrG-free apo-structure of He *et al.* [28]. We subsequently
168 coexpressed Hcp1 and Hcp2-His_6 for affinity purification, which revealed that Hcp1 readily

169 copurifies with Hcp2 (Figure 2A). Observation of the Hcp1-Hcp2-His_6 complex via negative
170 staining electron microscopy also revealed the formation of rings that appeared similar to those
171 formed by sHcp (Figure 2B). Cryo-EM observations provided a 2D structure and revealed that
172 the fold of Hcp1-Hcp2-His_6 was that of canonical Hcps, as shown by the presence of an
173 elongated structure corresponding to β-strands and a circular structure corresponding to the α-
174 helix (Figures 2C, 2D, 2E asterisks and S4B). AlphaFold models of monomeric Hcp1 and Hcp2
175 (Figures 2F, 2G and S4C) showed a classic Hcp fold. However, a few structural differences were
176 notable. In the case of Hcp1, an insertion between residues G11 and S25 forms a short alpha
177 helix that protrudes from the external beta sheet, and in the case of Hcp2, an N-terminal
178 insertion between residues M1 and L13 is unstructured. In both structures, compared with the
179 sHcp sequence, insertions presented lower PLDDT scores. Purification of the tagged versions of
180 Hcp1 or Hcp2 alone did not yield soluble proteins, which suggests that the interactions
181 between these proteins are necessary for their folding and assembly into rings. These results
182 show that Hcp1 and Hcp2, are binding partners that cooperatively assemble into hexameric
183 rings.

184

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

186 To determine the spatial arrangement of individual protomers in the Hcp1–Hcp2 complex, we
187 performed cross-linking mass spectrometry (XL-MS). Copurified Hcp1-Hcp2-His_6 was
188 subjected to crosslinking with DSS, DSG or EGS at three concentrations (Figure 2H and Tables
189 S2-S4). The crosslinked samples were observed via SDS–PAGE, proteolyzed and subsequently
190 analysed via mass spectrometry. SDS–PAGE analysis of the control sample and crosslinked
191 samples revealed the presence of a band migrating at the 15 kDa marker, corresponding to
192 Hcp1 or Hcp2-His_6 with either noncrosslinked subunits or with intraprotein crosslinks (Figures
193 2H lanes 1-3, 2I and S5). In the DSS crosslinked samples, a prominent band migrated between
194 25 kDa and 37 kDa, corresponding to two crosslinked Hcp proteins (Figure 2H, marked areas).
195 At a low DSS concentration (0.25 mM), the intensity profile of the band showed faster migrating
196 shoulder peaks, which became the most prominent band at higher concentrations (Figure 2I,
197 arrows). This suggested that at higher crosslinker concentrations, two Hcp proteins were
198 crosslinked by multiple linkers, which caused the two proteins to decrease their Stokes radii
199 and migrate faster. To a certain extent, the band migrating close to the 50 kDa marker had the
200 same appearance and showed the same behaviour at higher concentrations of the crosslinker.
201 Although it was less clearly visible, we suggest that three Hcp proteins could also be

202 crosslinked by multiple crosslinkers. Similar observations were made in the analyses of
203 samples treated with DSG (Figure S5 and Table S3). At the tested concentrations, EGS did not
204 efficiently generate crosslinked peptides (Figure S5 and Table S4). Above the 50 kDa marker,
205 there were more discrete bands in the crosslinked samples, which decreased in intensity
206 towards the higher-MW markers. This suggested that more than three Hcp protomers were
207 crosslinked but at lower apparent frequencies. Quantification of both intra- and interprotein
208 crosslinked peptides revealed that residues K119 in Hcp1 and K90 in Hcp2 were the most
209 crosslinked residues (Figure 2J, histogram), suggesting that these residues are exposed.
210 Filtering by residues that appear in multiple replicates (Table S5), the Hcp2(K83)-Hcp1(K113)
211 crosslink was the most consistent crosslink (Figure 2K) among the hetero-interprotein
212 crosslinks. This result confirms the expected interaction between Hcp1 and Hcp2. Consistent
213 with our hypothesis that Hcp1 could form homointeractions in the presence of Hcp2, based on
214 cooperative two-hybrid assays, we found multiple peptides across replicates showing
215 interprotein Hcp1 crosslinks occurring between two identical residues in two Hcp1 subunits,
216 *e.g.,* Hcp1(K44)-Hcp1(K44) and Hcp1(K113)-Hcp1(K113) (Figures 2L, S5 and Table S4). Peptides
217 between residues K111 or K128 crosslinked to residues at the same position also occurred
218 (Table S4). Interestingly, we were unable to observe any Hcp2 interprotein crosslinks, even
219 though there was a network of exposed residues that consistently formed intraprotein
220 crosslinks, namely, K74, K81, K83 and K90 (Figure 2M). In particular, Hcp2(K90) is very exposed,
221 as shown by the total number of crosslinks, but Hcp2(K90) was never found in crosslinks with
222 another Hcp2(K90). This result suggests that Hcp2 is not found close to another Hcp2 molecule
223 or at least less frequently compared withs Hcp1 being close to another Hcp1.

224 To gain insight into the possible arrangements of the Hcp1-Hcp2 complex, we obtained
225 hexameric models via the AlphaFold server (Figure S6). Using a Cα-Cα cut-off distance of 30 Å
226 [29], we classified the observed crosslinks as permissive or non-permisisve in the models. We
227 observed that in an Hcp1 homohexamer model, interprotein residues between K44 and K113
228 fell within the 30 Å distance in adjacent protomers (Figure 2L, S5 and Table S6) but not between
229 residues two protomers away (P+2). Only the distance between Hcp1 K111 residues in
230 protomer P and P+1 or P+2 was permissive for crosslinking. Conversely, in a Hcp1-Hcp2 (3:3)
231 model in which Hcp1 protomers are two protomers away, only K111 was in the permissive
232 distance. These results suggest that a structural model in which Hcp1 protomers are more
233 frequently found adjacent to each other is more compatible with our crosslinking data.
234 Interprotein crosslinking between Hcp1 and Hcp2, such as Hcp1(K113)-Hcp2(K83) (Figure 2K
235 and Table S6), occurred only within the permissive distance when Hcp1 and Hcp2 were

236 adjacent, confirming that neighbouring Hcp1–Hcp2 proteins interact within the hexameric
237 assembly. The distance between the observed Hcp2 crosslinked residues, with the exception of
238 Hcp2(K74)-Hcp2(K90), was permissive only within the same protomer and not between
239 adjacent protomers (Figure 2M and Table S6). This finding suggests that the majority, if not all, of
240 the crosslinks observed between Hcp2 lysine residues were intraprotein. We measured the
241 theoretical distance between all the residues found in crosslinked peptides and the same
242 residue in the adjacent protomer in an Hcp2 homohexamer model (Table S6), and these
243 residues were not permissive to crosslinking. Therefore, we cannot confirm that Hcp2 directly
244 interacts with itself, as observed by two-hybrid assays, because even if Hcp2 assembles
245 adjacent to another Hcp2, it could not be crosslinked at the same residue. However, the lack of
246 interprotein cross-links in Hcp2 suggests that Hcp1 interprotein cross-links occurred among
247 adjacent protomers and not randomly between Hcp protomers in two different hexamers.
248 Interestingly, Hcp2(K40) did not show any crosslinks and indeed appeared buried in all the
249 AlphaFold models when Hcp2 was adjacent to Hcp1 or Hcp2. K40 had a change in accessible
250 surface area of -66.7 $Å^2$ upon hexamer assembly, whereas K90 resulted in a change of only 0.4
251 Å2, indicating that K40 but not K90 is buried upon hexamerization (Figure S5). We were not able
252 to explain two cross-linked residues involving Hcp1, K128 residues modeled at non allowed
253 distances (Figure S5 and Table S6). However, the distance was close enough to be permissible,
254 and these residues could be considered within range of other residues previously reported [29].
255 Finally, data obtained from the published proteome of *B. fragilis* [30] revealed that Hcp1 is ~53x
256 more abundant than Hcp2 is (Figure S6), suggesting that Hcp2 is limiting. Notably, sHcp is
257 ~1000 fold more abundant than Hcp2 is, which is consistent with an inner tube structural role.
258 Based on these differences in abundance, we refer to sHcp as the majour subunit and Hcp-1-4
259 as minor subunits. All these results, are consistent with an arrangement of the Hcp1–Hcp2
260 hexamer, in which Hcp1 molecules are adjacent at higher frequencies and Hcp2 can be found
261 in at least one or two copies, considering that the interaction between Hcp1 and Hcp1 depends
262 on the presence of Hcp2, as shown by two-hybrid assays, and that Hcp1 and Hcp2 do not
263 appear to be equimolar inside the bacterial cell. Therefore, we obtained a Hcp1-Hcp2 structure
264 from the AlphaFold server with a 4:2 stoichiometry (Figure 2N and S6). The model showed an
265 asymmetric structure with three adjacent Hcp1 protomers and the two Hcp2 protomers
266 separated by another Hcp1. Notably, the Hcp2 protomers were not adjacently assembled.
267 Analysing non-crosslinked samples (Table S7), showed that there were more Hcp1 peptides
268 found compared to Hcp2. All these results are consistent with an Hcp1-Hcp2 complex model
269 formed primordially of Hcp1 with fewer Hcp2 subunits.

270

271

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

273 Having stated that Hcp1-Hcp2 forms ring-shaped hexamers, we set out to determine the
274 stoichiometry of this complex. To this end, we repeated the purification of Hcp1-Hcp2-His_6, this
275 time eluted with an imidazole step gradient and analysed the fractions via SDS–PAGE
276 densitometric analysis. Three fractions were collected at each imidazole concentration. As a
277 control, we constructed a plasmid expressing sHcp-sHcp-His_6, which should promote the
278 formation of mixed hexamers of tagged and untagged monomers. Indeed, SDS–PAGE analysis
279 of the eluted sHcp-sHcp-His_6 bands revealed two fast-migrating bands corresponding to the
280 tagged and untagged proteins (Figure 3A). Two slower migrating bands appeared just below the
281 50 kDa marker with densities proportional to the corresponding higher-migrating bands, which
282 we interpret to be SDS-resistant trimers of sHcp (labelled sHcp*) because single bands
283 appeared in analysis of purified sHcp-His6 samples just above the 15 kDa and just below the 50
284 kDa markers (Figure S4D). Higher proportions of sHcp-His_6 eluted with increasing
285 concentrations of imidazole, possibly because hexamers containing more His_6-tagged
286 protomers bind to the Ni-NTA resin with more avidity and required more imidazole to elute
287 (Figure 3A). Different results were obtained for the Hcp1-Hcp2-His_6 complex (Figure 3B). First,
288 no SDS-resistant proteins were eluted. Second, most of the protein peak was eluted at 150 mM
289 imidazole rather than at 250 mM imidazole as for the sHcp-sHcp-His6 complexes. Given that
290 homohexameric sHcp complexes could contain up to six His-tagged subunits, whilst in the
291 case of untagged Hcp1 with tagged Hcp2 six His-tagged subunits (and thereby the same binding
292 strength to the resin) would be observed only if hexamers devoid of Hcp1 formed. Therefore,
293 lower imidazole elution suggested that Hcp1-Hcp2- His_6 complexes did not have as many His_6-
294 tagged protomers as sHcp-sHcp-His_6, which is consistent with the formation of Hcp1-Hcp2
295 heterohexamers. Third, at all elution concentrations, untagged Hcp1 appeared to be present in
296 higher proportions than Hcp2- His_6.

297 To obtain quantitative measurements of the stoichiometry of the Hcp complexes, we performed
298 densitometric analysis of the SDS–PAGE SYPRO-ruby-stained gels. We calculated the
299 proportions of tagged and untagged proteins and normalized them to six (under the assumption
300 that each eluted particle is a hexamer). For sHcp-sHcp-His_6, only the fast-migrating bands were
301 analysed, which revealed that the number of eluted sHcp-His_6 steadily increased in proportion
302 to the imidazole concentration (Figure 3C). However, for the Hcp1-Hcp2-His_6 complex, up to

303 120 mM imidazole, the eluted complexes had a 4:2, Hcp1:Hcp2 ratio (Figure 3D). Above 120
304 mM, the complexes had fractional ratios, which we interpret to be mixtures of mostly 4:2 with
305 some 3:3 complexes, although we cannot exclude other ratios or trace amounts of
306 supracomplexes. Additionally, we quantified the total fraction of protein eluted at each
307 imidazole concentration by adding the density of both the untagged and tagged proteins and
308 dividing by the sum of all the densities (Figure 3C and 3D, dashed lines). This analysis revealed
309 that approximately half of the eluted Hcp1-Hcp2-His_6 had a 4:2 ratio and confirmed that the
310 protein peak of elution was at 150 mM imidazole.

311 To obtain full insight into the stoichiometry of the Hcp1-Hcp2-His_6 complex, we subjected it to
312 analytical size exclusion chromatography (SEC) with sHcp-His_6 as a control. Using molecular
313 weight standards, we observed that sHcp-His_6 eluted as a single peak from the column after 15
314 ml between the 150 and 44.3 kDa markers, similarly to the Hcp1-Hcp2-His_6 complex (Figure
315 3E). To better quantify the size of the complex, we selected three markers from the molecular
316 weight (MW) standards and plotted the MW *vs*. the elution volume, which showed a logarithmic
317 relationship. The MW values of sHcp-His_6 and Hcp1-Hcp2-His_6 extrapolated from the plot were
318 85.1 and 86.8 kDa, respectively (Figure 3F). The calculated MWs from the sequence for one
319 sHcp-His_6 hexamer are 93.6 kDa, while 97 kDa for a 4:2 Hcp1-Hcp2-His_6 complex. This shows
320 that the Hcp1-Hcp2-His_6 complex is formed by hexamers. Together with the findings described
321 above, these results suggest that the Hcp1-Hcp2-His_6 complex preferentially assembles *in vitro*
322 into hexamers with a stoichiometry of 4:2. This implies that hexamers of Hcp1-Hcp2 complexes
323 must contain adjacent Hcp1 protomers, which were observed in two-hybrid assays, but only if
324 Hcp2 was present. Overall, we propose that the Hcp1-Hcp2 complex may form through
325 cooperative interactions via Hcp1 homointeractions mediated by Hcp2 (Figure 3G). Moreover,
326 we conclude that *in vitro* assembled Hcp1–Hcp2 complexes have variable stoichiometries.

327

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

329 Our data show that the product of *hcp1* forms chimeric complexes with the product of *hcp2*,
330 which is the gene located downstream. Interestingly, the effector *bte1* is downstream of *hcp2*.
331 Bte1 was recently described as the first T6SS effector with the capacity to kill bacteria by
332 altering protein folding [31]. As reported by Coyne *et al.* [16], in the high-synteny arrangement of
333 *Bacteroidales* GA3, two effectors are encoded in variable cassettes but are always arranged in
334 the same order. In most strains one effector is downstream of an Hcp2 homologue. Therefore,
335 we hypothesised that the Hcp1-2 complex could interact with Bte1. First, we tested the

336 interactions of Bte1 with all minor Hcp proteins via bacterial two-hybrid assays. Our results
337 revealed that Bte1 fused to T18 or T25 fragments interacted with both Hcp1 and Hcp2 fused to
338 T18 or T25 fragments when Hcp2 and Hcp1 were coexpressed, but not with the Hcp3 or Hcp4
339 T18 and T25 fusions (Figure 4A). To confirm this interaction, we copurified a His-tagged version
340 of Bte1 with Hcp1 and Hcp2 and analysed it via size exclusion chromatography. Two major
341 peaks were observed (Figure 4B). Pooled fractions from each peak were subsequently analysed
342 by SDS–PAGE, which revealed that peak 1 corresponded to the Hcp1–Hcp2–Bte1 ternary
343 complex, whereas peak 2 corresponded to Bte1-His_6 alone (Figure 4C). Negative staining
344 electron microscopy images of the ternary complex showed particles in various orientations
345 (Figure 4D). A few particles were filled with a density in the centre (Figure 4D, asterisks), and
346 other particles appeared as empty rings (Figure 4D, empty circles). These results suggest that
347 the minor Hcp heterohexamers can act as recognition particles in the case of *Bacteroidales*
348 T6SS effectors.

349 Three more effectors were described for the T6SS GA3 in other strains in addition to Bte1
350 encoded in the minor *hcp*-adjacent variable region. Two proteins were initially annotated using
351 Hidden Markov models (HMMs) and categorized as cluster 51 and cluster 115 (using Coyne *et*
352 *al*. clustering nomenclature [16]), which were upstream of immunity proteins named clusters 107
353 and 53, respectively. The third type of effector, present in *B. fragilis* 638R, is called *bfe1* [20] and
354 has been shown to target other *Bacteroides* spp. To gain an overview of the probability that the
355 *Bacteroidales* minor Hcp proteins function as effector recognition particles in general, we
356 generated a phylogenetic tree of all GA3 available *Hcp2* genes described in Coyne *et.al* [16]. The
357 Hcp2 homologues had a high degree of identity, but they were separated into five clusters
358 (Figure 4E and S8A). We then analysed the effectors adjacent to each cluster and observed that
359 the four types of effectors segregated in different clusters (Figure 4E). Only Bfe1-like effectors,
360 similar to the effector found in *B. fragilis* 638R, which were the most common, segregated in two
361 clusters, which were closer to each other. This cosegregation did not occur when sHcp was
362 used to generate the phylogenetic tree, suggesting that sHcp is not a recognition factor for T6SS
363 GA3 effectors (Figure S7A). Hcp1 and Hcp3 also cosegregated with their effector–immunity
364 pairs (Figure S7B-C). The clustering of Hcp2 with effectors was also present in T6SSs with GA2
365 clusters (Figure S8B) and segregation of Hcp sequences with effectors was described by Bao *et*
366 *al*. [23] Considering that Bte1, Hcp1 and Hcp2 were found to be secreted [19], it is reasonable to
367 propose that they are secreted as a complex. Taken together, our results are consistent with the
368 idea that minor Hcp proteins are cognate to effectors and participate in secretion, and that this
369 feature is conserved in *Bacteroidales* T6SSs.

370

371 **DISCUSSION**

372 This study revealed that in Bacteroidales T6SSs there are different types of Hcp subunits with
373 specific assembly properties and functions. Firstly, a role for sHcp as major subunit
374 constituting the bulk of the inner tube is consistent with data showing its high abundance and
375 its capacity to form homohexamers. In the case of Hcp1 and Hcp2, we showed here that these
376 minor subunits form heterohexamers and have a specialized function of recognition particles
377 for effector secretion.

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

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

401 Two-hybrid assays showed a high level of interaction among Hcp subunits. However, in
402 competition assays Hcp1 preferentially bound to Hcp2 and Hcp3 preferentially bound to Hcp4.
403 These interactions coincide with the corresponding two pairs of genes being adjacent and are
404 consistent with a model in which these proteins coevolve to form heterohexamers that
405 recognize other proteins. This finding is further supported by a high level of synteny in gut
406 Bacteroidales T6SS loci. We also identified other interactions that await biochemical
407 confirmation. For example, sHcp could bind to Hcp1 or Hcp3. These minor Hcps may
408 intersperse with major subunits in the assembled inner tube. However, the difference in
409 abundance and the greater tendency of sHcp to form homohexamers may increase the
410 difficulty of detecting these interactions both *in vivo* and *in vitro*. We also observed binding
411 between Hcp1 and Hcp4. Hcp4 was not detected in proteomic data [30], suggesting that, like
412 Hcp2, Hcp4 may be present in low abundance. It could be that Hcp1 may carry both Hcp2 and
413 Hcp4 protomers to mediate specific interactions. Furthermore, it is feasible that these
414 complexes have additional interaction partners *in vivo*. How the minor Hcp subunits assemble
415 within the inner tube and bind to VgrG remains to be elucidated, but the proteomic abundance
416 data suggest that the formation of the Hcp1–Hcp2 complex and effector recognition is limiting.
417 Therefore, there may be a specific assembly pathway for minor Hcp subunits and their binding
418 partners to assemble at the tip of the inner tube that ensures their function in secreting cognate
419 effectors.

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

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

433 to VgrG [28]. The question of how the Hcp1-Hcp2 complex is assembled in the context of the
434 whole secretion system remains open.

435 In polymorphic toxins, the N-termini typically carry secretion specificity domains and the C-
436 termini carry toxin domains. Toxins are frequently encoded in gene fragments located adjacently
437 upstream of genes encoding cognate antitoxin proteins [32]. Such synteny has resulted in toxin-
438 antitoxin (immunity) pairs being acquired by horizontal gene transfer as gene fusions of N-
439 terminal secretion specificity domains [33]. The high level of synteny of the *Bacteroides* T6SSs
440 translates into function by linking the position of the gene and the hierarchy of interaction of
441 their products as shown here by protein-protein interactions between Hcp1, Hcp2 and Bte1
442 (which in turn would interact with its antitoxin, Bti1). Therefore, our results suggest that in the
443 case of *Bacteroidales* T6SSs, the minor *hcp* genes partly encode domains that mirror
444 transcriptional fusions of polymorphic toxins encoding secretion specificity domains, although
445 without being fused. Indeed, Hcp2 homologues clustered according to their cognate effectors
446 in the phylogenetic tree, suggesting that Hcp2 is a cosegregated specificity factor. In a previous
447 Hcp study, Howard *et al.* concluded that the interactions between effectors and cognate Hcp
448 proteins depend on the overall Hcp structure and not on specific residues on the basis of the
449 clustering of specialized and core Hcps performed on phylogenetic analyses of a large subset of
450 proteins [14]. However, our phylogenetic analysis using a specific subset of *Bacteroidales* T6SS
451 GA3 Hcps clearly segregated effectors with cognate minor Hcps but not with major sHcp
452 proteins. This model is consistent with the minor Hcps participating in specific recognition of
453 effectors. If this is true, it could mean that minor Hcps can recognize different effectors with a
454 minimum of variability, encoded in very few residues. Although this specificity could be also
455 encoded through selective steric hindrance as suggested by Howard *et al.* [14]. This sort of
456 specificity could be aided by an asymmetric assembly with limiting Hcp2 subunits as
457 recognition factors, which could accommodate effectors in the minor Hcp complex with
458 specific stoichiometry.

459 It was recently reported that Bte1 is an effector that disrupts protein folding in the periplasm
460 and is used for antagonistic interactions [31]. Lim *et al.* showed that heterologous expression of
461 Bte1 in *Bacteroides thetaiotaomicron* causes toxicity by targeting the chaperone-acting PpiD-
462 YfgM complex, thus suggesting that Bte1 requires no other factors for its activity. Moreover,
463 Hcp1, Hcp2 and Bte1 have been shown to be present in the secreted fraction [19,21]. Therefore, it is
464 possible that Bte1 is co-translocated with Hcp1 and Hcp2 as a complex. Alternatively, a
465 chaperone function could be necessary to position Bte1 for secretion.

466 The separation of major and minor subunit proteins with high and low abundance, respectively,
467 in gut Bacteroidales T6SSs appears akin to macromolecular export systems that assemble
468 filamentous structures. For example, type II secretion systems include one major pseudopilin
469 (or endopilin) and four minor pseudopilins. The minor subunits form a tip complex that
470 functions in priming the assembly of the major pseudopilin into a filament [34,35]. This type of
471 assembly and function of minor pilins is conserved in type IV pili [36–38]. In type 1 fimbria minor
472 pilins mediate adhesion and appear at the tip of fimbriae because of their higher affinity to the
473 usher subunit [39,40]. Further studies are needed to elucidate the *in vivo* assemblies of minor Hcp
474 subunits in the *Bacteroides* T6SSs. Finally, our results shed light on how Bacteroidales T6SSs
475 evolved to exhibit a high diversity and variability of effectors. The presence of major and
476 specialized minor Hcp subunits that recognize these effectors is an evolved feature that
477 facilitates this variability, contributing to *Bacteroides* biodiversity.

478

479 **MATERIALS AND METHODS**

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

499 plate. The two pictures were treated as above before quantification and placed in a stack in FIJI.
500 The two images were aligned via the linear stack alignment (SIFT) plugin and stack-averaged.
501 This process greatly reduced differences in the background caused by deviations in light
502 capture at certain pixel position in the camera's sensor, which caused some spots to appear to
503 have less signal than the negative controls.

504 **Protein purification.** Protein purification was carried out by expressing *hcp* genes from pET28
505 constructs in BL21(DE3) cells. Co-expression of *Hcp1, Hcp2* and *bte1* was performed using
506 plasmid pETDUET-1. Cells were grown in LB medium at 37°C until exponential phase and
507 induced with 0.4 mM IPTG after cooling the cells. Induction was performed overnight at 20°C or
508 for three hours at 30°C. Cells were harvested by centrifugation and resuspended in 50mM NaP
509 8.0. 500mM NaCl, 10% glycerol, 10mM Imidazole, 1mM β-mercaptoethanol. Lysis was
510 performed by sonication and the cleared lysate was prepared by centrifugation at 40,000 g for
511 30 minutes. The cleared lysate was incubated with a Ni-NTA resin (ThermoScientific) for 1 h,
512 washed with resuspension buffer and eluted with resuspension buffer containing 300 mM
513 imidazole. For preparative size exclusion chromatography (SEC) a HiLoad
514 16/600 Superdex200pg column (Cytiva) was used. For analytical SEC, a Superdex200 10/300gl
515 column (Cytiva) was used and a Protein Standard Mix 15 - 600 kDa (Merck). Proteins samples
516 were characterised using SDS-PAGE stained with blue Coomassie or SYPRO-Ruby.

517 **Densitometric analysis of the SDS–PAGE gel (fluorescence intensity).** The gels were stained
518 with SYPRO-Ruby and scanned with a GE Typhoon FLA 9000 Gel Imager. The tiff images were
519 then analysed in FIJI. For each quantified protein band, two ROIs were created. The first ROI
520 contained the band. The second ROI was the same size but encompassed an immediately
521 adjacent area without stain as a blank. The 'integrated density' of each ROI was measured, and
522 the blank was subtracted to obtain the corrected fluorescence intensity ($fluo_{corr}$) of the protein
523 band. To obtain the normalized SYPRO-ruby fluorescence intensity at each lane ($fluo_{norm}$) of
524 each band of His-tagged ($fluo_{norm}1$) protein *vs*. non-His-tagged protein ($fluo_{norm}2$), the corrected
525 fluorescence intensity of each band was normalized by the following equation: $fluo_{norm}1=$
526 $fluo_{corr}1/(fluo_{corr}1 + fluo_{corr}2)*6$ or $fluo_{norm}2= fluo_{corr}2/(fluo_{corr}1 + fluo_{corr}2)*6$. The obtained number
527 represents the fraction multiplied by six of one or the other protein normalized by the total
528 intensity.

529 **Electron microscopy.** For transmission electron microscopy (TEM) the samples were
530 resuspended in purification buffer and 3.5 µL adsorbed to formvar and carbon-coated glow
531 discharged grids (cupper, 300 mesh, Ted Pella Inc.) and were negatively stained with 1.5% w/v

532 uranyl acetate. The samples were imaged with a Talos L120 TEM using the Ceta detector and
533 Velox software (Thermo Fisher). For Cryogenic TEM (Cryo-EM) 4 µL of the samples were
534 adsorbed to holey carbon film on 300 mech cupper grids (Quantifoil). The Grids were blotted
535 and plunge frozen in liquid ethane using the Vitrobot (Thermo Fisher). Cryo-EM imaging was
536 performed with Titan Krios G2 Cryo-EM using Gatan K2 BioQuantom detector and EPU software
537 (Thermo Fisher). The Cryo-EM images were analysed by single particle analyse approaches with
538 the CryoSPARC software package.

539 **Protein alignments and phylogenetic analyses.** Protein alignments were performed at the
540 MAFFT website [41] using the G-INS-I algorithm and UPGMA. The output phylogenetic guide trees
541 with effector or Hcp alignments were plotted with TreeViewer [42]. Effectors were clustered by
542 UPGMA distance in JalView.

543 **Crosslinking of the Hcp1-Hcp2 complexes.** A total of 20 µG of Hcp1-Hcp2 complexes in 20
544 mM NaP, pH 8.0, 150 mM NaCl, 1 mM DTT were crosslinked in duplicates with three different
545 concentrations (0.25 mM, 0.5 mM and 1.0 mM) of heavy/light DSS (DSS-H12/D12, Creative
546 Molecules Inc., 001S), heavy/light DSG (DSG-H6/D6, Creative Molecules Inc., 010S) or
547 heavy/light EGS (EGS-H12/D12, Creative Molecules Inc., 003S), respectively in a total reaction
548 volume 20 uL. Non-crosslinked samples were kept as controls. All samples were incubated for
549 30 min at 37°C, 1000 rpm. The cross-linking reaction was quenched with a final concentration
550 of 50 mM of ammonium bicarbonate for 15 min at 37°C, 1000 rpm. For SDS-PAGE analysis of
551 the crosslinked samples, 5 uL of the above reactions were mixed with 5 uL of 2x sample buffer
552 (BioRAD) and boiled at 95°C, 5 min. The samples were analyzed on a 4-20% BioRAD Criterion
553 TGX gel.

554 **Sample preparation for mass spectrometry.** The remaining samples crosslinked with DSS
555 and DSG were denatured using an 8 M urea– 100 mM ammonium bicarbonate solution. For the
556 EGS-crosslinked samples, 8 M urea without ammonium bicarbonate was used. The cysteine
557 bonds were reduced with a final concentration of 5 mM Tris (2-carboxyethyl) phosphine
558 hydrochloride (TCEP, Sigma, 646547) for 60 min at 37°C, 1000 rpm and subsequently alkylated
559 using a final concentration 10 mM 2-iodoacetamide for 30 min at 22°C in the dark. For
560 digestion, 1 µG of lysyl endopeptidase (LysC, Wako Chemicals, 125-05061) was added, and the
561 samples incubated for 2 h at 37°C, 100 rpm. The samples crosslinked with DSS and DSG were
562 diluted with 100 mM ammonium bicarbonate to a final urea concentration of 1.5 M, whereas
563 the samples crosslinked with EGS were diluted using LC-H_2O. For digestion, 1 µG of sequencing

564 grade trypsin (Promega, V5111) was added for 16 h at 37°C, 1000 rpm. The digested samples
565 were acidified with 10% formic acid to a final pH of 3.0. Peptides were purified and desalted
566 using C18 reverse phase columns (The Nest Group, Inc.) following the manufacturer's
567 recommendations. Dried peptides were reconstituted in 50 µL of 2% acetonitrile and 0.1%
568 formic acid prior to MS analysis.

569 **Liquid chromatography tandem mass spectrometry (LC-MS/MS).** A total of 1 µL of peptides
570 were analyzed on an Orbitrap Eclipse mass spectrometer connected to an ultra-high
571 performance liquid chromatography Dionex Ultra300 system (both Thermo Scientific). The
572 peptides were loaded and concentrated on an Acclaim PepMap 100 C18 precolumn (75 µm × 2
573 cm,) and then separated on an Acclaim PepMap RSLC column (75 µm × 25 cm, nanoViper,C18,
574 2 µm, 100 Å) (both columns Thermo Scientific), at a column temperature of 45°C and a
575 maximum pressure of 900 bar. A linear gradient of 2% to 25% of 80% acetonitrile in aqueous
576 0.1% formic acid was run for 100 min followed by a linear gradient of 25% to 40% of 80%
577 acetonitrile in aqueous 0.1% formic acid for 20 min. One full MS scan (resolution 120,000; mass
578 range of 400-1600 *m/z*) was followed by MS/MS scans (resolution 15,000) with a 3 sec cycle
579 time. Precursors with a charge state of 3-8 were included. The precursor ions were isolated with
580 1.6 *m/z* isolation window and fragmented using higher-energy collisional-induced dissociation
581 (HCD) at normalized collision energies (NCE) of 21, 26, 31. The dynamic exclusion was set to 45
582 s.

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

596 ions was set to 10 ppm and for fragment ions to 0.02 Da. The protein false discovery rate was
597 set to 1%.

598 **Circos plots and structural analysis.** Circos plots were obtained in Rstudio using the Circlize
599 library [43]. The AlphaFold server [44] was used to create hexamer structures using default
600 parameters and a random seed number. Visualization of the structures, distance
601 measurements between residues and structural alignments were performed in Chimera [45].

602 **ACKNOWLEDGEMENTS**

603 We acknowledge support from Mikael Lindberg and Lotta Happonen at the the Umeå node of
604 the Protein Expertise Platform (PEP) /Protein Production Sweden (PPS) and the Swedish
605 National Infrastructure for Biological Mass Spectrometry (BioMS)/ Integrated Structural Biology
606 (ISB) platform and the SciLifeLab, respectively. We are thankful to Dr. Andrea Puhar, WWIEM,
607 Queen's University Belfast, UK for critical reading. Electron microscopy instruments and
608 methods were supported by Michael Hall at the Umeå Centre for Electron Microscopy (UCEM)/
609 SciLifeLab national Cryo-EM infrastructure supported by the Kempe foundations and Knut and
610 Alice Wallenberg foundations. B.E.U. received support from the Swedish Research Council
611 (2019-01720, 2016-06598), Kempestiftelserna (JCK-1527.1, JCK-1724) and the Faculty of
612 Medicine, Umeå University (Insamlingsstiftelsen grant 2021-2023). D.A.C. received support
613 from the Kempestiftelserna (SMK-1860), Carl Tryggers Stiftelse för Vetenskaplig Forskning (CTS
614 15-96 and CTS 18-65) and the Faculty of Medicine, Umeå University (Insamlingsstiftelsen 2019-
615 2021).
616

617 **CONTRIBUTIONS (CRediT taxonomy)**

618 Conceptualization, data curation, formal analysis and writing original draft: DAC. Funding
619 acquisition, supervision, resources and project administration: DAC & BEU. Visualization:
620 SGSM, MKSAA, DAC. Methodology, Investigation, review and editing: SGSM, MKSAA, EJ, AK, US,
621 PB, LS, BEU & DAC

622

623 **The authors declare no competing interests**

624

625 **REFERENCES**

1. Shepherd, E.S., DeLoache, W.C., Pruss, K.M., Whitaker, W.R., and Sonnenburg, J.L. (2018). An exclusive metabolic niche enables strain engraftment in the gut microbiota. Nature *557*, 434–438. https://doi.org/10.1038/s41586-018-0092-4.

2. García-Bayona, L., and Comstock, L.E. (2018). Bacterial antagonism in host-associated microbial communities. Science *361*, eaat2456. https://doi.org/10.1126/science.aat2456.

3. Basler, M., Pilhofer, M., Henderson, G.P., Jensen, G.J., and Mekalanos, J.J. (2012). Type VI secretion requires a dynamic contractile phage tail-like structure. Nature *483*, 182–186.

4. Böck, D., Medeiros, J.M., Tsao, H.-F., Penz, T., Weiss, G.L., Aistleitner, K., Horn, M., and Pilhofer, M. (2017). In situ architecture, function, and evolution of a contractile injection system. Science *357*, 713–717. https://doi.org/10.1126/science.aan7904.

5. Brunet, Y.R., Henin, J., Celia, H., and Cascales, E. (2014). Type VI secretion and bacteriophage tail tubes share a common assembly pathway. EMBO reports *15*, 315–321.

6. Pukatzki, S., Ma, A.T., Sturtevant, D., Krastins, B., Sarracino, D., Nelson, W.C., Heidelberg, J.F., and Mekalanos, J.J. (2006). Identification of a conserved bacterial protein secretion system in Vibrio cholerae using the Dictyostelium host model system. Proceedings of the National Academy of Sciences *103*, 1528–1533. https://doi.org/10.1073/pnas.0510322103.

7. Mougous, J.D., Cuff, M.E., Raunser, S., Shen, A., Zhou, M., Gifford, C.A., Goodman, A.L., Joachimiak, G., Ordoñez, C.L., Lory, S., et al. (2006). A Virulence Locus of Pseudomonas aeruginosa Encodes a Protein Secretion Apparatus. Science *312*, 1526–1530. https://doi.org/10.1126/science.1128393.

8. Hood, R.D., Singh, P., Hsu, F., Güvener, T., Carl, M.A., Trinidad, R.R.S., Silverman, J.M., Ohlson, B.B., Hicks, K.G., Plemel, R.L., et al. (2010). A Type VI Secretion System of Pseudomonas aeruginosa Targets a Toxin to Bacteria. Cell Host & Microbe *7*, 25–37. https://doi.org/10.1016/j.chom.2009.12.007.

9. Monjarás Feria, J., and Valvano, M.A. (2020). An Overview of Anti-Eukaryotic T6SS Effectors. Front Cell Infect Microbiol *10*, 584751. https://doi.org/10.3389/fcimb.2020.584751.

10. Wang, J., Brodmann, M., and Basler, M. (2019). Assembly and Subcellular Localization of Bacterial Type VI Secretion Systems. Annu Rev Microbiol *73*, 621–638. https://doi.org/10.1146/annurev-micro-020518-115420.

11. Cherrak, Y., Flaugnatti, N., Durand, E., Journet, L., and Cascales, E. (2019). Structure and Activity of the Type VI Secretion System. Microbiology Spectrum *7*, 10.1128/microbiolspec.psib-0031–2019. https://doi.org/10.1128/microbiolspec.psib-0031-2019.

12. Allsopp, L.P., and Bernal, P. (2023). Killing in the name of: T6SS structure and effector diversity. Microbiology *169*, 001367. https://doi.org/10.1099/mic.0.001367.

13. Silverman, J.M., Agnello, D.M., Zheng, H., Andrews, B.T., Li, M., Catalano, C.E., Gonen, T., and Mougous, J.D. (2013). Haemolysin Coregulated Protein Is an Exported Receptor and Chaperone of Type VI Secretion Substrates. Molecular Cell *51*, 584–593. https://doi.org/10.1016/j.molcel.2013.07.025.

14. Howard, S.A., Furniss, R.C.D., Bonini, D., Amin, H., Paracuellos, P., Zlotkin, D., Costa, T.R.D., Levy, A., Mavridou, D.A.I., and Filloux, A. (2021). The Breadth and Molecular Basis of Hcp-Driven Type VI Secretion System Effector Delivery. mBio *12*, 10.1128/mbio.00262-21. https://doi.org/10.1128/mbio.00262-21.

15. Chu, H., Khosravi, A., Kusumawardhani, I.P., Kwon, A.H.K., Vasconcelos, A.C., Cunha, L.D., Mayer, A.E., Shen, Y., Wu, W.L., Kambal, A., et al. (2016). Gene-microbiota interactions contribute to the pathogenesis of inflammatory bowel disease. Science *352*, 1116–1120.

16. Coyne, M.J., Roelofs, K.G., and Comstock, L.E. (2016). Type VI secretion systems of human gut Bacteroidales segregate into three genetic architectures, two of which are contained on mobile genetic elements. BMC genomics *17*, 58.

17. Russell, A.B., Wexler, A.G., Harding, B.N., Whitney, J.C., Bohn, A.J., Goo, Y.A., Tran, B.Q., Barry, N.A., Zheng, H., Peterson, S.B., et al. (2014). A type VI secretion-related pathway in Bacteroidetes mediates interbacterial antagonism. Cell host & microbe *16*, 227–236.

18. Bongiovanni, T.R., Latario, C.J., Le Cras, Y., Trus, E., Robitaille, S., Swartz, K., Schmidtke, D., Vincent, M., Kosta, A., Orth, J., et al. (2024). Assembly of a unique membrane complex in type VI secretion systems of Bacteroidota. Nat Commun *15*, 429. https://doi.org/10.1038/s41467-023-44426-1.

19. Wexler, A.G., Bao, Y., Whitney, J.C., Bobay, L.-M., Xavier, J.B., Schofield, W.B., Barry, N.A., Russell, A.B., Tran, B.Q., Goo, Y.A., et al. (2016). Human symbionts inject and neutralize antibacterial toxins to persist in the gut. Proc. Natl. Acad. Sci. U.S.A. *113*, 3639–3644. https://doi.org/10.1073/pnas.1525637113.

20. Chatzidaki-Livanis, M., Geva-Zatorsky, N., and Comstock, L.E. (2016). Bacteroides fragilis type VI secretion systems use novel effector and immunity proteins to antagonize human gut Bacteroidales species. pnas.org.

21. Hecht, A.L., Casterline, B.W., Earley, Z.M., Goo, Y.A., Goodlett, D.R., and Bubeck Wardenburg, J. (2016). Strain competition restricts colonization of an enteric pathogen and prevents colitis. EMBO reports *17*, 1281–1291.

22. Verster, A.J., Ross, B.D., Radey, M.C., Bao, Y., Goodman, A.L., Mougous, J.D., and Borenstein, E. (2017). The Landscape of Type VI Secretion across Human Gut Microbiomes Reveals Its Role in Community Composition. Cell host & microbe *22*, 411-419.e4.

23. Bao, H., Coyne, M.J., García-Bayona, L., and Comstock, L.E. (2022). Analysis of Effector and Immunity Proteins of the GA2 Type VI Secretion Systems of Gut Bacteroidales. Journal of Bacteriology *204*, e00122-22. https://doi.org/10.1128/jb.00122-22.

24. Blondel, C.J., Jiménez, J.C., Contreras, I., and Santiviago, C.A. (2009). Comparative genomic analysis uncovers 3 novel loci encoding type six secretion systems differentially distributed in Salmonella serotypes. BMC Genomics *10*, 354. https://doi.org/10.1186/1471-2164-10-354.

25. Ma, J., Pan, Z., Huang, J., Sun, M., Lu, C., and Yao, H. (2017). The Hcp proteins fused with diverse extended-toxin domains represent a novel pattern of antibacterial effectors in type VI secretion systems. Virulence *8*, 1189–1202. https://doi.org/10.1080/21505594.2017.1279374.

26. Anderson, A.J.G., Morrell, B., Lopez Campos, G., and Valvano, M.A. (2023). Distribution and diversity of type VI secretion system clusters in Enterobacter bugandensis and Enterobacter cloacae. Microbial Genomics *9*, 001148. https://doi.org/10.1099/mgen.0.001148.

27. Karimova, G., Pidoux, J., Ullmann, A., and Ladant, D. (1998). A bacterial two-hybrid system based on a reconstituted signal transduction pathway. Proceedings of the National Academy of Sciences *95*, 5752–5756. https://doi.org/10.1073/pnas.95.10.5752.

28. He, W., Wu, K., Ouyang, Z., Bai, Y., Luo, W., Wu, D., An, H., Guo, Y., Jiao, M., Qin, Q., et al. (2023). Structure and assembly of type VI secretion system cargo delivery vehicle. Cell Reports *42*, 112781. https://doi.org/10.1016/j.celrep.2023.112781.

29. Merkley, E.D., Rysavy, S., Kahraman, A., Hafen, R.P., Daggett, V., and Adkins, J.N. (2014). Distance restraints from crosslinking mass spectrometry: Mining a molecular dynamics simulation database to evaluate lysine–lysine distances. Protein Science *23*, 747–759. https://doi.org/10.1002/pro.2458.

30. Müller, J.B., Geyer, P.E., Colaço, A.R., Treit, P.V., Strauss, M.T., Oroshi, M., Doll, S., Virreira Winter, S., Bader, J.M., Köhler, N., et al. (2020). The proteome landscape of the kingdoms of life. Nature *582*, 592–596. https://doi.org/10.1038/s41586-020-2402-x.

31. Lim, B., Xu, J., Wierzbicki, I.H., Gonzalez, C.G., Chen, Z., Gonzalez, D.J., Gao, X., and Goodman, A.L. (2025). A human gut bacterium antagonizes neighboring bacteria by altering their protein-folding ability. Cell Host & Microbe *33*, 200-217.e24. https://doi.org/10.1016/j.chom.2025.01.008.

32. Zhang, D., de Souza, R.F., Anantharaman, V., Iyer, L.M., and Aravind, L. (2012). Polymorphic toxin systems: Comprehensive characterization of trafficking modes, processing, mechanisms of action, immunity and ecology using comparative genomics. Biology Direct *7*, 18.

33. Ruhe, Z.C., Low, D.A., and Hayes, C.S. (2020). Polymorphic Toxins and Their Immunity Proteins: Diversity, Evolution, and Mechanisms of Delivery. Annual Review of Microbiology *74*, 497–520. https://doi.org/10.1146/annurev-micro-020518-115638.

34. Korotkov, K.V., and Hol, W.G.J. (2008). Structure of the GspK–GspI–GspJ complex from the enterotoxigenic Escherichia coli type 2 secretion system. Nat Struct Mol Biol *15*, 462–468. https://doi.org/10.1038/nsmb.1426.

35. Cisneros, D.A., Bond, P.J., Pugsley, A.P., Campos, M., and Francetic, O. (2012). Minor pseudopilin self-assembly primes type II secretion pseudopilus elongation. EMBO J *31*, 1041–1053. https://doi.org/10.1038/emboj.2011.454.

36. Cisneros, D.A., Pehau-Arnaudet, G., and Francetic, O. (2012). Heterologous assembly of type IV pili by a type II secretion system reveals the role of minor pilins in assembly initiation. Mol Microbiol *86*, 805–818. https://doi.org/10.1111/mmi.12033.

37. Nguyen, M., Wu, T.-H., Danielson, K.J., Khan, N.M., Zhang, J.Z., and Craig, L. (2023). Mechanism of secretion of TcpF by the Vibrio cholerae toxin-coregulated pilus. Proceedings of the National Academy of Sciences *120*, e2212664120. https://doi.org/10.1073/pnas.2212664120.

38. McCallum, M., Burrows, L.L., and Howell, P.L. (2019). The Dynamic Structures of the Type IV Pilus. Microbiology Spectrum *7*, 10.1128/microbiolspec.psib-0006–2018. https://doi.org/10.1128/microbiolspec.psib-0006-2018.

39. Saulino, E.T., Thanassi, D.G., Pinkner, J.S., and Hultgren, S.J. (1998). Ramifications of kinetic partitioning on usher-mediated pilus biogenesis. The EMBO Journal *17*, 2177–2185. https://doi.org/10.1093/emboj/17.8.2177.

40. Busch, A., Phan, G., and Waksman, G. (2015). Molecular mechanism of bacterial type 1 and P pili assembly. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences *373*, 20130153. https://doi.org/10.1098/rsta.2013.0153.

41. Katoh, K., Rozewicki, J., and Yamada, K.D. (2019). MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization. Briefings in Bioinformatics *20*, 1160–1166. https://doi.org/10.1093/bib/bbx108.

42. Bianchini, G., and Sánchez-Baracaldo, P. (2024). TreeViewer: Flexible, modular software to visualise and manipulate phylogenetic trees. Ecology and Evolution *14*, e10873. https://doi.org/10.1002/ece3.10873.

43. Gu, Z., Gu, L., Eils, R., Schlesner, M., and Brors, B. (2014). circlize implements and enhances circular visualization in R. Bioinformatics *30*, 2811–2812. https://doi.org/10.1093/bioinformatics/btu393.

44. Abramson, J., Adler, J., Dunger, J., Evans, R., Green, T., Pritzel, A., Ronneberger, O., Willmore, L., Ballard, A.J., Bambrick, J., et al. (2024). Accurate structure prediction of biomolecular interactions with AlphaFold 3. Nature *630*, 493–500. https://doi.org/10.1038/s41586-024-07487-w.

45. Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., and Ferrin, T.E. (2004). UCSF Chimera—A visualization system for exploratory research and analysis. Journal of Computational Chemistry *25*, 1605–1612. https://doi.org/10.1002/jcc.20084.

FIGURE LEGENDS

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 Hcp proteins fused to the *Bordetella pertussis* T18 and T25 two-hybrid fragments in the *cyaA- E. coli* strain BTH101. Interaction of fused proteins leads to the expression of β-galactosidase and hydrolysis of X-gal. **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). Asterisks denote interactions inhibited in panels G-J (squares). Hcp proteins are labelled above. The positive control (+) is the dimerization of the yeast GCN4 leucine zipper and the negative controls (-) are the empty vectors in four different combinations. Pictures were taken at 24h from plates containing 250 mM IPTG at 30 ºC. **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

787 normalized to both the GCN4 leucine zipper (Zip) positive and negative controls. The color scale
788 between black and yellow shows the minimum and maximum, respectively in that plot. The
789 data represents the average of data obtained with varying concentrations of IPTG, at 24h and
790 36h and at 30 °C and 37 °C. **E)** Scheme of genes coding simultaneous expression of two Hcp
791 proteins fused to T18 and/or T25 fragments used for two-hybrid assays shown in G-J. An artificial
792 *E. coli* ribosome binding site (RBS) was placed upstream of genes encoding a second Hcp
793 subunit. **F)** Scheme of genes coding for N- and C- terminal sHcp proteins fused to the T18 and
794 T25 fragments used for two-hybrid assays shown in G-H. **G)** Photographic images of competitive
795 two-hybrid assay with plasmids encoding sHcp (s) fused to N- or C- terminal T18 or T25
796 fragments tested for pairwise interactions against T18 or T25 fusions of Hcp1 (1), Hcp2 (2),
797 Hcp3 (3) and Hcp4 (4) in the presence of Hcp2, Hcp1, Hcp4 and Hcp3, respectively. Pictures
798 were taken at 36h. **H)** Quantification of G. **I)** Photographic images of competitive and
799 cooperative two-hybrid assay for pairwise interactions among T18 or T25 translational fusions of
800 Hcp1 (1), Hcp2 (2), Hcp3 (3) and Hcp4 (4) in the presence of Hcp2, Hcp1, Hcp4 and Hcp3,
801 respectively. Pictures were taken at 36h. **J)** Quantification of I **K)** Scheme of interactions among
802 Hcp subunits derived for two-hybrid assays. The wider arrows represent the most frequently
803 observed interactions. Grey arrows represent interactions that are inhibited by other pairwise
804 interactions. Coloured arrows represent cooperative interactions that depend on another
805 interaction.

806 **Figure 2. Purification and biochemical characterization of the major and minor Hcp**
807 **subunits. A)** Phosphoimager scan of SYPRO Ruby stained SDS-PAGE gels of purified sHcp-His6
808 and Hcp1 and Hcp2-His6 complexes. **B)** Negative staining electron microscopy image of sHcp-
809 His6 and Hcp1 and Hcp2-His6 complex. Scale bars: 50 nm (left) and 15 nm (right). **C)** Crystall
810 structure of sHcp. Asterisks denote two alpha helices in the crystal structure observed also by
811 Cryo-EM data averaging. D) Cryo-EM data by single particle analysis of Hcp1-Hcp2-His6. **E)**
812 Main class of particles isolated after a two-dimensional image single particle analysis). F)
813 PLDDT colour-coded AlphaFold model of Hcp1, AF-Q5LDT5-F1. **G)** PLDDT colour-coded
814 AlphaFold model of Hcp2, AF-A0A380Z0E1-F1. **H)** SDS-PAGE of two replicates (1 or 2) of X-
815 linked Hcp proteins using DSS. **I)** Inset of SDS-PAGE with intensity line profile showing X-linked
816 Hcp dimers at three DSS concentrations. **J)** Circos plot of all X-linked peptides found. Gray
817 track contains a histogram of the total number of peptides found for each residue. **K)** Network of
818 close residues in Hcp2 showing possible intramolecular X-links. **L)** K44 inter-protein Hcp1 X-
819 links between residues at the same position modelled with AlphaFold as a Hcp1 homo-
820 hexamer. **M)** Residues found in X-linked peptides between Hcp1 and Hcp2 modelled with
821 AlphaFold as a 3:3 complex. **N)** AlphaFold model of an Hcp1-Hcp2 complex in a 4:2 ratio.

822 **Figure 3. Stoichiometry of purified Hcp complexes as determined by an imidazole step**
823 **gradient. A-B)** SYPRO Ruby stained SDS-PAGE gels of samples obtained from the affinity co-
824 purification of sHcp and sHcp-His6 or Hcp1 and Hcp2-His6. Four elution steps were performed
825 for each imidazole concentration (black rectangles on top of the gels). Labels: C (clear fraction),
826 FT (flow-trough), W (wash), 1-40 (elution fractions). **C- D)** Densitometric quantification of
827 SYPRO Ruby fluorescence intensity from the gels shown in A and B. The average fractional
828 density of fluorescence intensity of the two regions of interest (ROI) corresponding to each
829 purified Hcp protein (with and without His6 tag) normalized to the total density multiplied by 6
830 (x6) is plotted against the imidazole concentration. The percent of eluted protein is shown in the
831 secondary *y*-axis as the sum of the two ROIs in each lane, divided by the sum of all ROIs and
832 multiplied by 100. The possible stoichiometric arrangements of sHcp-sHcp-His6 and Hcp1-

833 Hcp2 complexes are indicated by color coded filed hexamers: Light red (sHcp-His6), dark red
834 (sHcp), yellow (Hcp2-His6) and skyblue (Hcp1). **E)** SEC analysis of a protein-standard mixture
835 (top) purified sHcp-His6 (midddle, red) and the co-purified complex of Hcp2-His6 with Hcp1
836 (bottom, blue). **F)** Calibration of size exclusion chromatography molecular weight standards.
837 Elution volumes of the SEC molecular weight protein-standards (black circles) are plotted
838 against their molecular weights. A linear fit using three of the log10 MW of the standards (dotted
839 line) and the extrapolated MW of the Hcp complexes are also shown. The intrapolated MW of
840 sHcp-His6 was 85.1 kDa (red square). The intrapolated MW for the Hcp2-His6 co-purified with
841 Hcp1 was 86.8 kDa (yellow-blue triangles). **G)** Assembly of possible radially asymmetric
842 heterohexamers of Hcp1 and Hcp2.

843 **Figure 4. Hcp2 determines binding to GA3 T6SS effectors.** **A)** Photographic images of
844 cooperative two-hybrid assay for pairwise interactions between T18 or T25 translational fusions
845 of Bte1 with Hcp1 in the presence of Hcp2 (2(3)), Hcp2 in the presence of Hcp1 (3(2)), and so
846 on. Pictures were taken at 36h. **B-C)** Size exclusion chromatography of co-purified Hcp-effector
847 complex. Two peaks were characterized by SDS-PAGE. The first and largest peak (1) contained
848 Hcp1 (16.0kDa), Hcp2 (15.3 kDa) and Bte1-His6. The second peak (2) only contained the
849 effector. **D)** Negative staining electron microscopy of the Hcp1-3-Bte1 effector complex. Filled
850 structures are marked with asterisks (*), empty structures are marked with empty circles (○).
851 Scale bar 50 nm (insets 20 nm). **E)** Clustering of phylogenetic tree of Hcp2 homologues from
852 GA3 T6SS with their cognate effector immunity pairs. Bte1 with Bti1 (blue), cluster 115 with
853 immunity gene (magenta), Bfe1 with Bfi1 (orange) and cluster 51 with immunity gene (green).

854

Figure 1

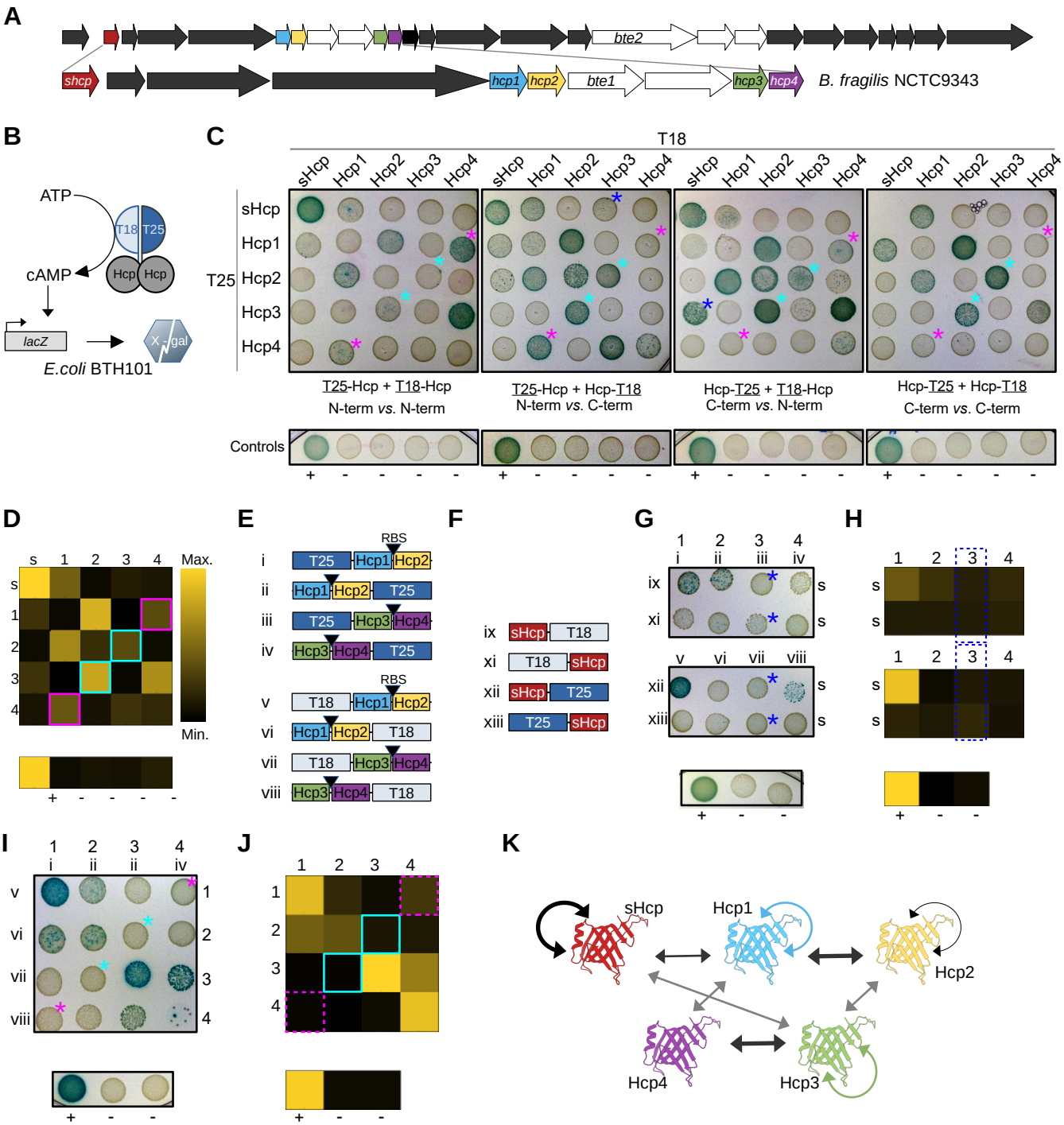

Figure 1. Two-hybrid assays of the major and minor Hcp proteins from *B. fragilis*' T6SS. A) Scheme of *B. fragilis* NCTC9343 T6SS locus. **B)** Scheme of Hcp proteins fused to the *Bordetella pertussis* T18 and T25 two-hybrid fragments in the *cyaA*$^-$ *E. coli* strain BTH101. Interaction of fused proteins leads to the expression of β-galactosidase and hydrolysis of X-gal. **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). Asterisks denote interactions inhibited in panels G-J (squares). Hcp proteins are labelled above. The positive control (+) is the dimerization of the yeast GCN4 leucine zipper and the negative controls (-) are the empty vectors in four different combinations. Pictures were taken at 24h from plates containing 250 mM IPTG at 30 °C. **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 that 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 simultaneous expression of two Hcp proteins fused to T18 and/or T25 fragments used for two-hybrid assays shown in G-J. An artificial *E. coli* ribosome binding site (RBS) was placed upstream of genes encoding a second Hcp subunit. **F)** 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. **G)** Photographic images of competitive two-hybrid assay with plasmids encoding sHcp (s) 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)** Quantification of G. **I)** 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. **J)** Quantification of I **K)** Scheme of interactions among Hcp subunits derived for two-hybrid assays. The wider arrows represent the most frequently observed interactions. Grey arrows represent interactions that are inhibited by other pairwise interactions. Coloured arrows represent cooperative interactions that depend on another interaction.

Figure 2

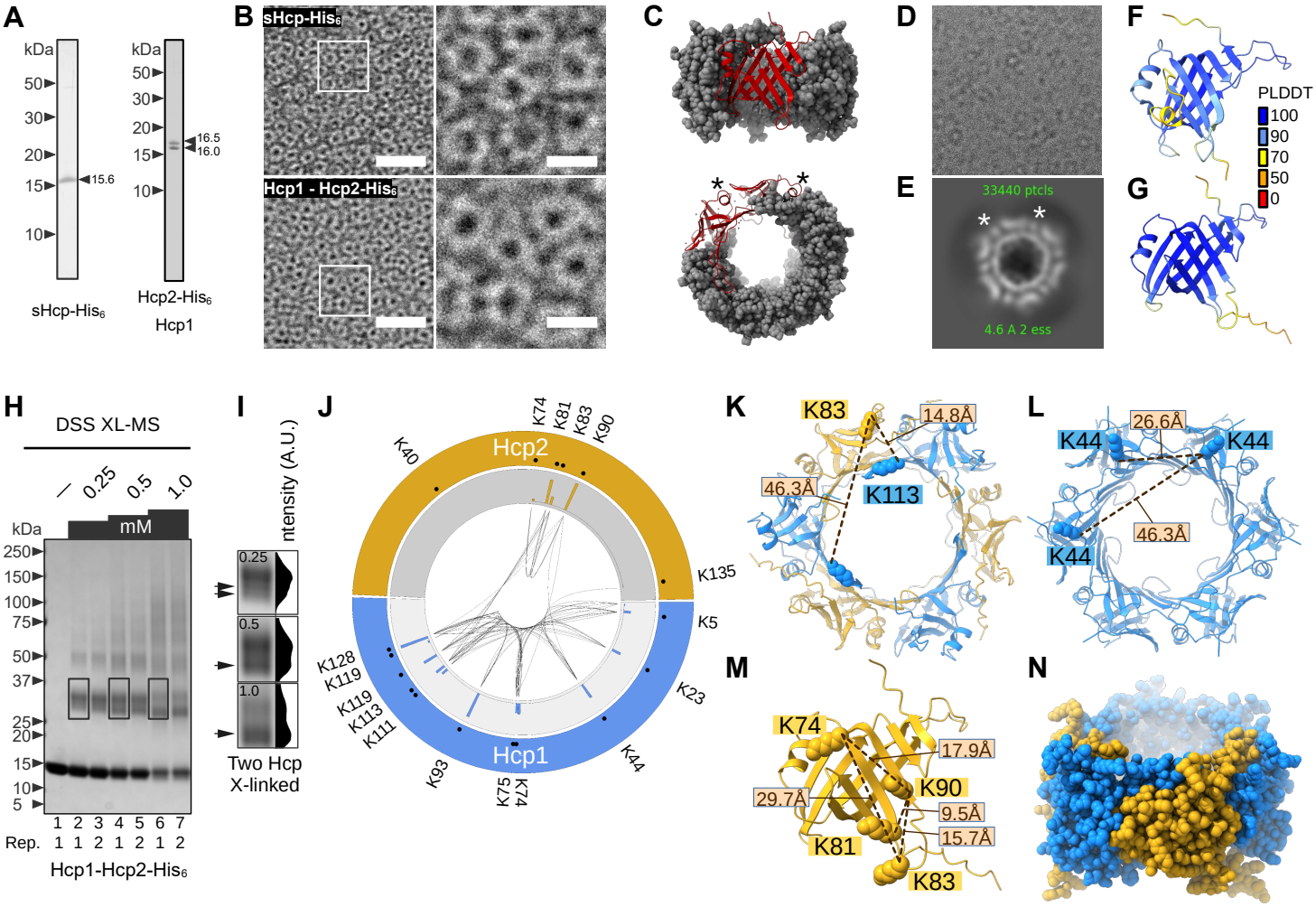

Figure 2. Biochemical characterization of the major and minor Hcp subunits. A) Phosphoimager scan of SYPRO Ruby stained SDS-PAGE gels of purified sHcp-His$_6$ and Hcp1 and Hcp2-His$_6$ complexes. **B)** Negative staining electron microscopy image of sHcp-His$_6$ and Hcp1 and Hcp2-His$_6$ complex. Scale bars: 50 nm (left) and 15 nm (right). **C)** Crystall structure of sHcp. Asterisks denote two alpha helices in the crystal structure observed also by Cryo-EM data averaging. **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)** PLDDT colour-coded AlphaFold model of Hcp1, AF-Q5LDT5-F1. **G)** PLDDT colour-coded AlphaFold model of Hcp2, AF-A0A380Z0E1-F1. **H)** SDS-PAGE of two replicates (1 or 2) of X-linked Hcp proteins using DSS. **I)** Inset of SDS-PAGE with intensity line profile showing X-linked Hcp dimers at three DSS concentrations. **J)** Circos plot of all X-linked peptides found. Gray track contains a histogram of the total number of peptides found for each residue. **K)** Network of close residues in Hcp2 showing possible intramolecular X-links. **L)** K44 inter-protein Hcp1 X-links between residues at the same position modelled with AlphaFold as a Hcp1 homo-hexamer. **M)** Residues found in X-linked peptides between Hcp1 and Hcp2 modelled with AlphaFold as a 3:3 complex. **N)** AlphaFold model of an Hcp1-Hcp2 complex in a 4:2 ratio.

Figure 3

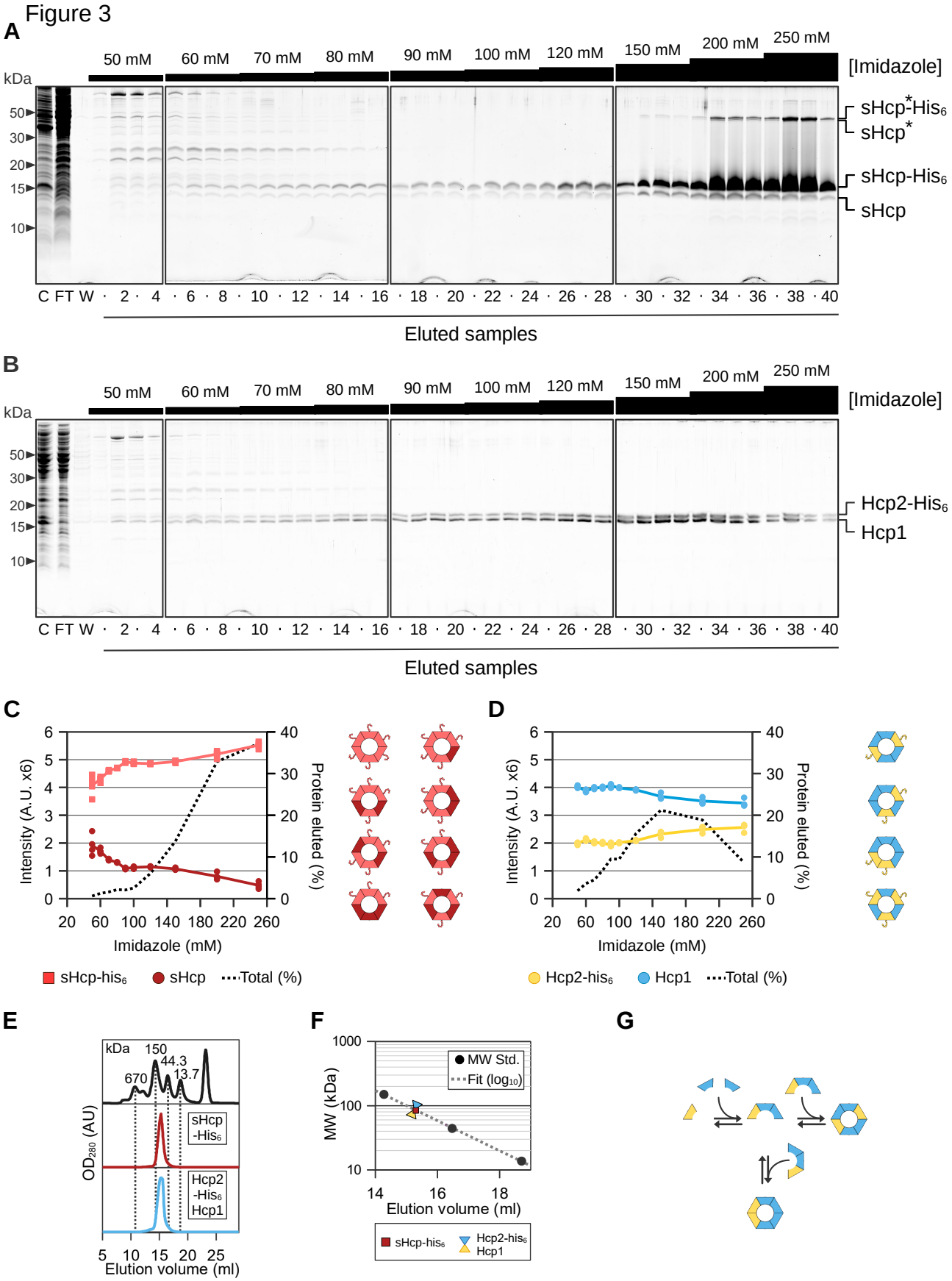

Figure 3. Stoichiometry of purified Hcp complexes as determined by an imidazole step gradient. A-B) SYPRO Ruby stained SDS-PAGE gels of samples obtained from the affinity co-purification of sHcp and sHcp-His$_6$ or Hcp1 and Hcp2-His$_6$. Four elution steps were performed for each imidazole concentration (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). **E)** 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 (black 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). **G)** Assembly of possible radially asymmetric heterohexamers of Hcp1 and Hcp2.

Figure 4

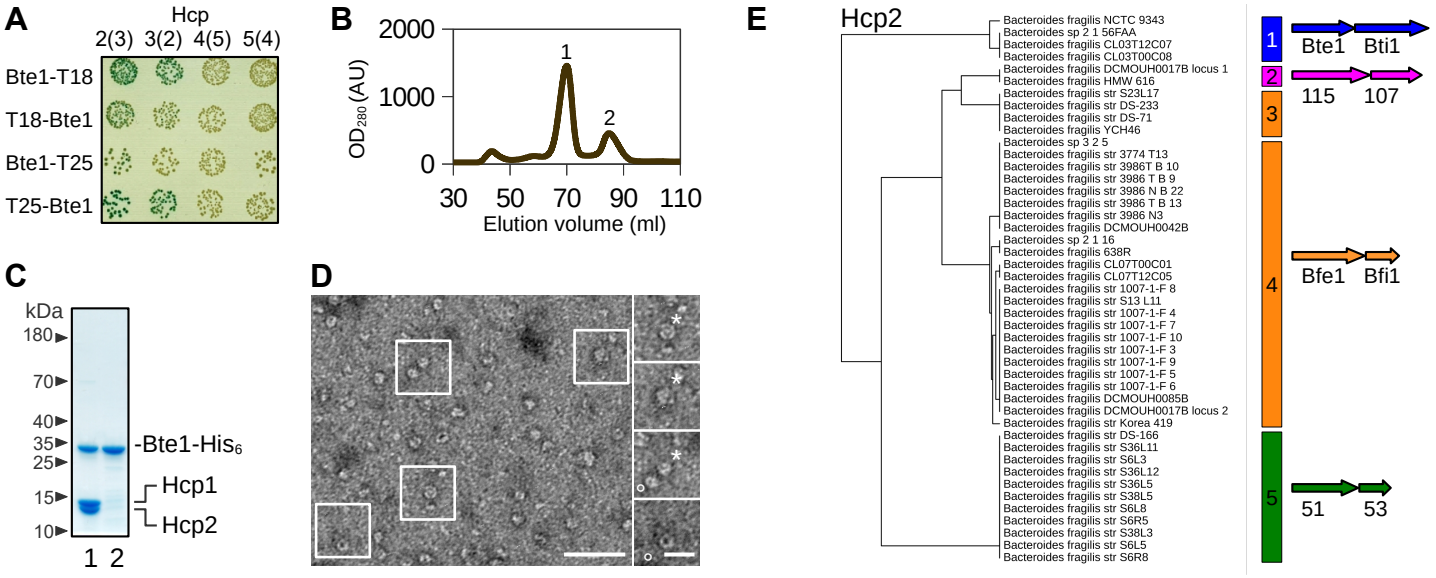

Figure 5. Hcp2 determines binding to GA3 T6SS effectors.
A) Photographic images of cooperative two-hybrid assay for pairwise interactions between T18 or T25 translational fusions of Bte1 with Hcp1 in the presence of Hcp2 (2(3)), Hcp2 in the presence of Hcp1 (3(2)), and so on. Pictures were taken at 36h. **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. Filled structures are marked with asterisks (*), empty structures are marked with empty circles (○). Scale bar 50 nm (insets 20 nm). **E)** Clustering of phylogenetic tree of Hcp2 homologues from GA3 T6SS with their cognate effector immunity pairs. Bte1 with Bti1 (blue), cluster 115 with immunity gene (magenta), Bfe1 with Bfi1 (orange) and cluster 51 with immunity gene (green).

Supplemental Information

Bacteroidales T6SS minor Hcp subunits form heteromers recognising effectors.

Sergio G. San-Miguel[1*], Manal Kamal Saleh Al-Ammari[1], Emma Johansson[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.

Corresponding author: David A. Cisneros

Contents:

- Supplementary methods
- Table S1
- Figures S1 to S8
- Tables S2 to S6

Supplemental methods

Crystallization of Se-Met labelled sHcp protein

The sHcp protein was genetically modified to contain an uncleavable C-terminal 6His-tag for immobilized metal affinity chromatography (IMAC) purification. For X-ray crystal structure determination, the sHcp protein was labelled with Seleno-Methionine (Se-Met) as described elsewhere[1,2] for Single-wavelength Anomalous Diffraction (SAD) data collection.

Crystals of the C-terminal 6His-tagged Se-Met sHcp1 protein (amino acid sequence below) were set-up by mixing equal amounts, typically 2µl + 2µl of ultra-pure sHcp protein, at a concentration of 14 mg/ml, with precipitant solution containing 30% w/v PEG 400, 200 mM trisodium citrate ($Na_3C_6H_5O_7$) and 100 mM Tris pH 8.5. The crystals were grown by the hanging-drop vapour-diffusion method at 18˚C (291 K). Hexagonal crystals of about 0.4 x 0.4 x 0.05 mm^3 grew within one week. They were mounted free standing supported by a film of buffer in nylon loops (Hampton Research), vitrified and stored in liquid nitrogen. Diffraction was initially verified using the in-house X-ray source (X8Proteum system, Bruker AXS) while maintaining the crystals at 100 K in a nitrogen gas stream (Oxford CryoSystems Ltd, UK).

Synchrotron data collection and 3D structure determination

Diffraction data were collected at 100 K and a wavelength ⍰ = 0.9677 Å (12.812 keV) at beam-line ID30A-3 (or MASSIF-3) of the European Synchrotron Facility (ESRF) in Grenoble, France. During data collection, the crystal was rotated by 0.1° per exposure and 3600 images were collected. The X-ray diffraction images were processed with XDS[3]. The integrated X-ray intensities were analyzed with the program Pointless[4] to determine the space group symmetry and then scaled with the program Scala[5] both of which are part of the CCP4 program suite[6,7].

Using the SAD data, experimental phases were determined, which lead to the calculation of an interpretable electron density map. An initial structural model of sHcp comprising amino acids A2 to K129 was built into the electron density. Iterative rounds of manual model building with Coot[8] (version 0.9.8.1 EL) and sHcp structure refinement against data extending to 2.05 Å with the software packages REFMAC5[9] and Phenix[10]. Final refinement as carried out using phenix.refine[11] of the Phenix program package until the R-factor and R-free factor converged. Structure validation was performed with the web services provided by the wwPDB (https://validate-rcsb-1.wwpdb.org/). The data collection and refinement statistics are summarized in supplementary Table 1. The atomic coordinates and structure factors for sHcp have been deposited with the Worldwide Protein Data Bank (wwPDB) under PDB code 9QWL.

Table S1. Data collection and refinement statistics.

sHcp_SeM*	
Wavelength	0.9677
Resolution range	22.44 - 1.95 (2.02 - 1.95)
Space group	P 6

Unit cell	77.366 77.366 47.635 90 90 120
Total reflections	248578 (26002)
Unique reflections	11983 (1189)
Multiplicity	20.7 (21.9)
Completeness (%)	99.64 (99.58)
Mean I/sigma(I)	35.10 (3.28)
Wilson B-factor	46.84
R-merge	0.05516 (1.127)
R-meas	0.05659 (1.154)
R-pim	0.01246 (0.246)
CC1/2	1 (0.943)
CC*	1 (0.985)
Reflections used in refinement	11961 (1185)
Reflections used for R-free	1192 (120)
R-work	0.2092 (0.4084)
R-free	0.2331 (0.3649)
CC(work)	0.946 (0.812)
CC(free)	0.960 (0.682)
Number of non-hydrogen atoms	1034
macromolecules	1015
ligands	34
solvent	5
Protein residues	129
RMS(bonds)	0.017
RMS(angles)	1.80
Ramachandran favored (%)	88.98
Ramachandran allowed (%)	6.30
Ramachandran outliers (%)	4.72
Rotamer outliers (%)	1.79
Clashscore	3.44
Average B-factor	67.34
macromolecules	67.27
ligands	76.43
solvent	55.49
*Statistics for the highest-resolution shell are shown in parentheses.	

>sHcp M1-K129+ASGGGHHHHHH, 129+11=140 aa., structural Hemolysin-coregulated-protein, based on Q5LDT1, Bacteroides fragilis strain NCTC9343 170518
MAFRATLSFAGKEFDVLDCTYSLKRDVDSKGRPSSNIYGGQIRLHVESTDDTSILENMTN
QFKPHSGSIVFKKGDEEAKMKELTWENGYITEFTENIDIVGSQPMTITFVVSAQVIKIGG
AQFEQNWPKASGGGHHHHHH

Supplementary references

1. Van Duyne, G.D., Standaert, R.F., Karplus, P.A., Schreiber, S.L., and Clardy, J. (1993). Atomic structures of the human immunophilin FKBP-12 complexes with FK506 and rapamycin. J Mol Biol *229*, 105–124. https://doi.org/10.1006/jmbi.1993.1012.

2. Doublié, S. (1997). Preparation of selenomethionyl proteins for phase determination. Methods Enzymol *276*, 523–530.

3. Kabsch, W. (2010). XDS. Acta Crystallogr D Biol Crystallogr *66*, 125–132. https://doi.org/10.1107/S0907444909047337.

4. Evans, P.R. (2011). An introduction to data reduction: space-group determination, scaling and intensity statistics. Acta Crystallogr D Biol Crystallogr *67*, 282–292. https://doi.org/10.1107/S090744491003982X.

5. Evans, P. (2006). Scaling and assessment of data quality. Acta Crystallogr D Biol Crystallogr *62*, 72–82. https://doi.org/10.1107/S0907444905036693.

6. Collaborative Computational Project, Number 4 (1994). The CCP4 suite: programs for protein crystallography. Acta Crystallogr D Biol Crystallogr *50*, 760–763. https://doi.org/10.1107/S0907444994003112.

7. Panjikar, S., Parthasarathy, V., Lamzin, V.S., Weiss, M.S., and Tucker, P.A. (2005). Auto-rickshaw: an automated crystal structure determination platform as an efficient tool for the validation of an X-ray diffraction experiment. Acta Crystallogr D Biol Crystallogr *61*, 449–457. https://doi.org/10.1107/S0907444905001307.

8. Emsley, P., Lohkamp, B., Scott, W.G., and Cowtan, K. (2010). Features and development of Coot. Acta Crystallogr D Biol Crystallogr *66*, 486–501. https://doi.org/10.1107/S0907444910007493.

9. Murshudov, G.N., Skubák, P., Lebedev, A.A., Pannu, N.S., Steiner, R.A., Nicholls, R.A., Winn, M.D., Long, F., and Vagin, A.A. (2011). REFMAC5 for the refinement of macromolecular crystal structures. Acta Crystallogr D Biol Crystallogr *67*, 355–367. https://doi.org/10.1107/S0907444911001314.

10. Adams, P.D., Afonine, P.V., Bunkóczi, G., Chen, V.B., Davis, I.W., Echols, N., Headd, J.J., Hung, L.-W., Kapral, G.J., Grosse-Kunstleve, R.W., et al. (2010). PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr D Biol Crystallogr *66*, 213–221. https://doi.org/10.1107/S0907444909052925.

11. Afonine, P.V., Grosse-Kunstleve, R.W., Echols, N., Headd, J.J., Moriarty, N.W., Mustyakimov, M., Terwilliger, T.C., Urzhumtsev, A., Zwart, P.H., and Adams, P.D. (2012). Towards automated crystallographic structure refinement with phenix.refine. Acta Crystallogr D Biol Crystallogr *68*, 352–367. https://doi.org/10.1107/S0907444912001308.

Figure S1

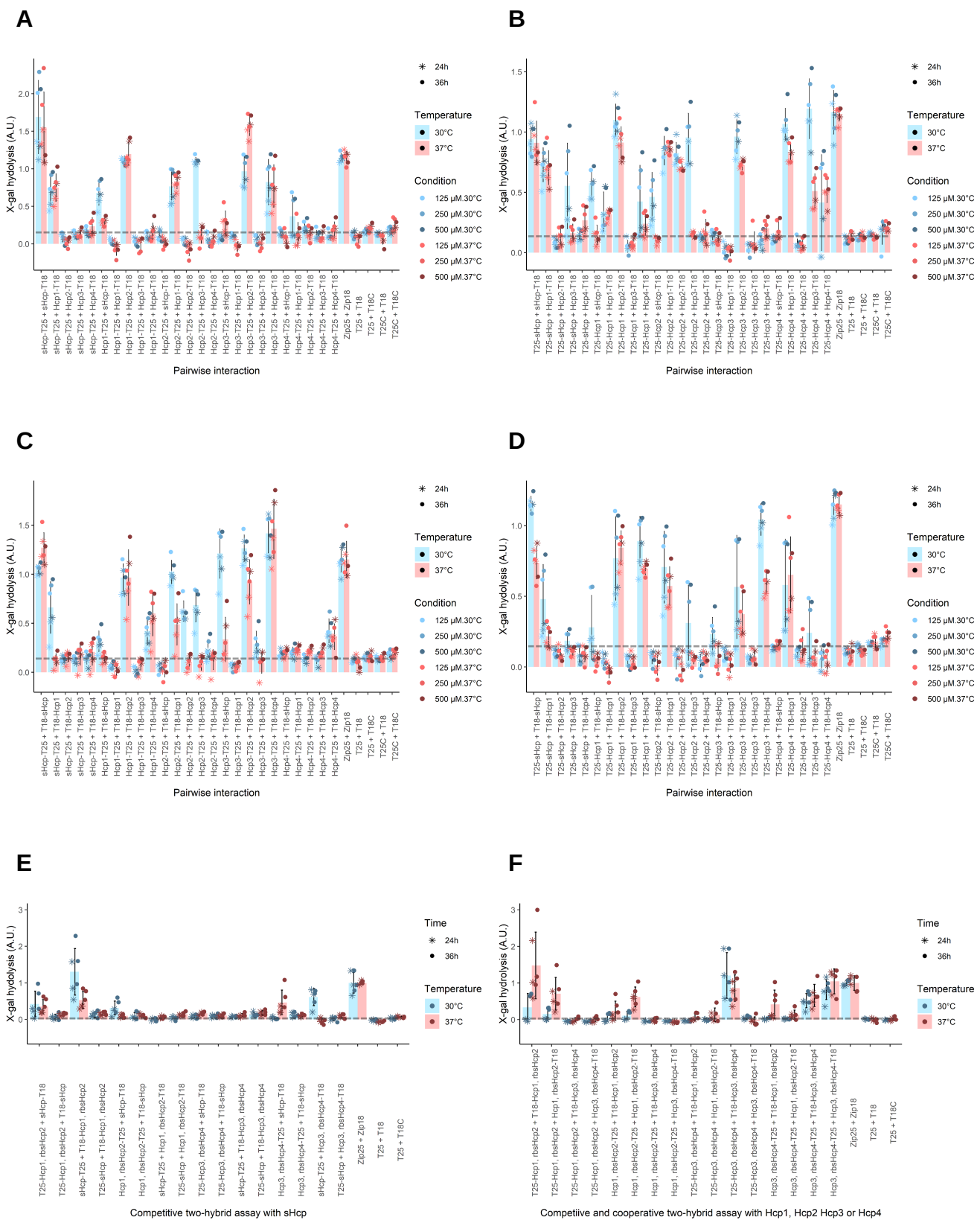

Figure S1. Quantification of normalized X-gal hydrolisis.
Quantification of two-hybrid assays of the major and minor Hcp proteins from *B. fragilis* with plasmids carrying genes encoding Hcp proteins fused to either **A)** a C-terminal T18 and T25 fragments or **B)** a N-terminal T25 and a C-terminal 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. A dashed line is shown at the average of negative controls.

Figure S2

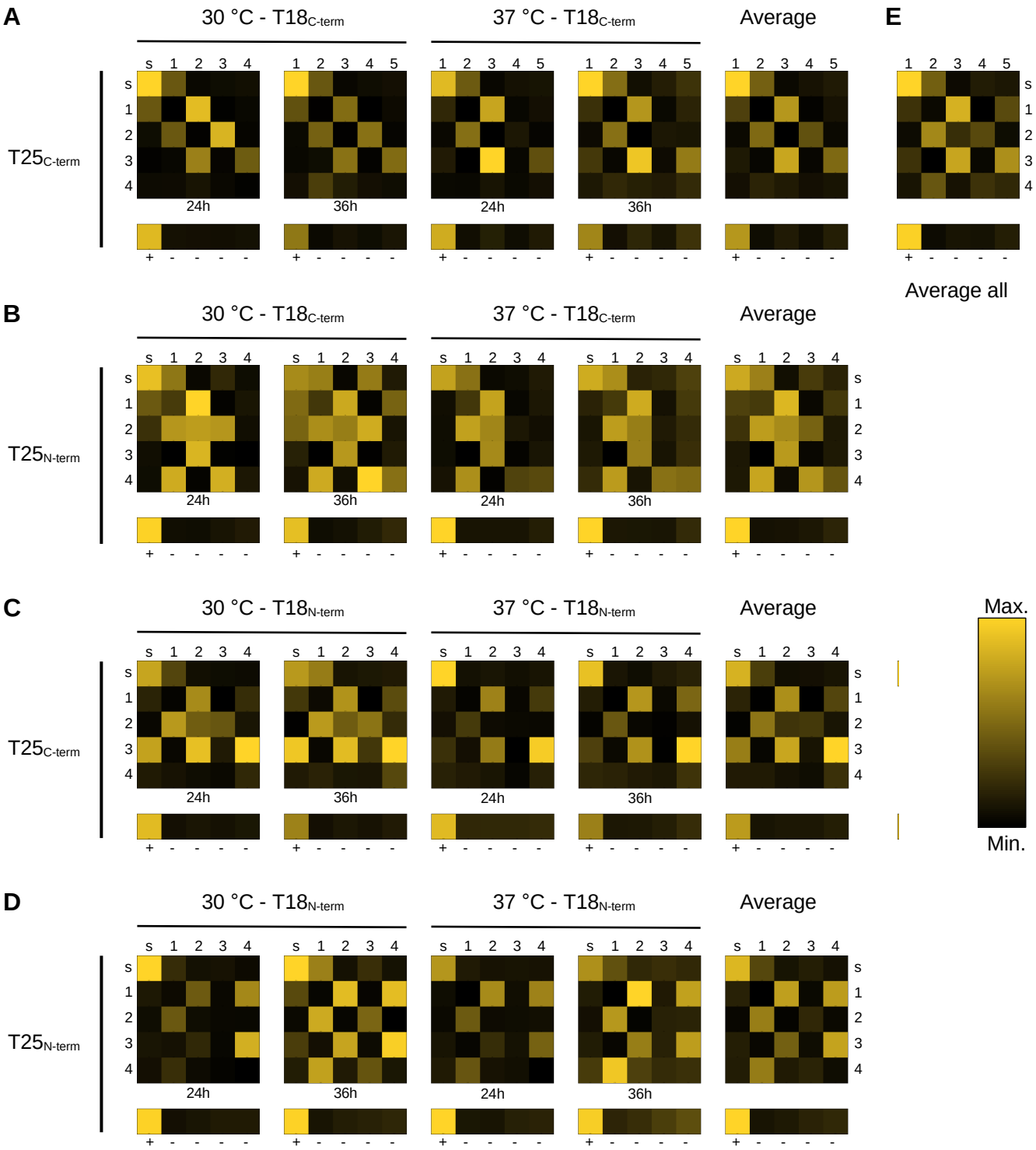

Figure S2. Quantification of two-hybrid assays. The average hydrolized X-gal for each condition tested for two-hybrid assays with plasmids carrying genes encoding Hcp proteins fused to either **A)** a C-terminal T18 and T25 fragments or **B)** a N-terminal T25 and a C-terminal T18 fragment or **C)** 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

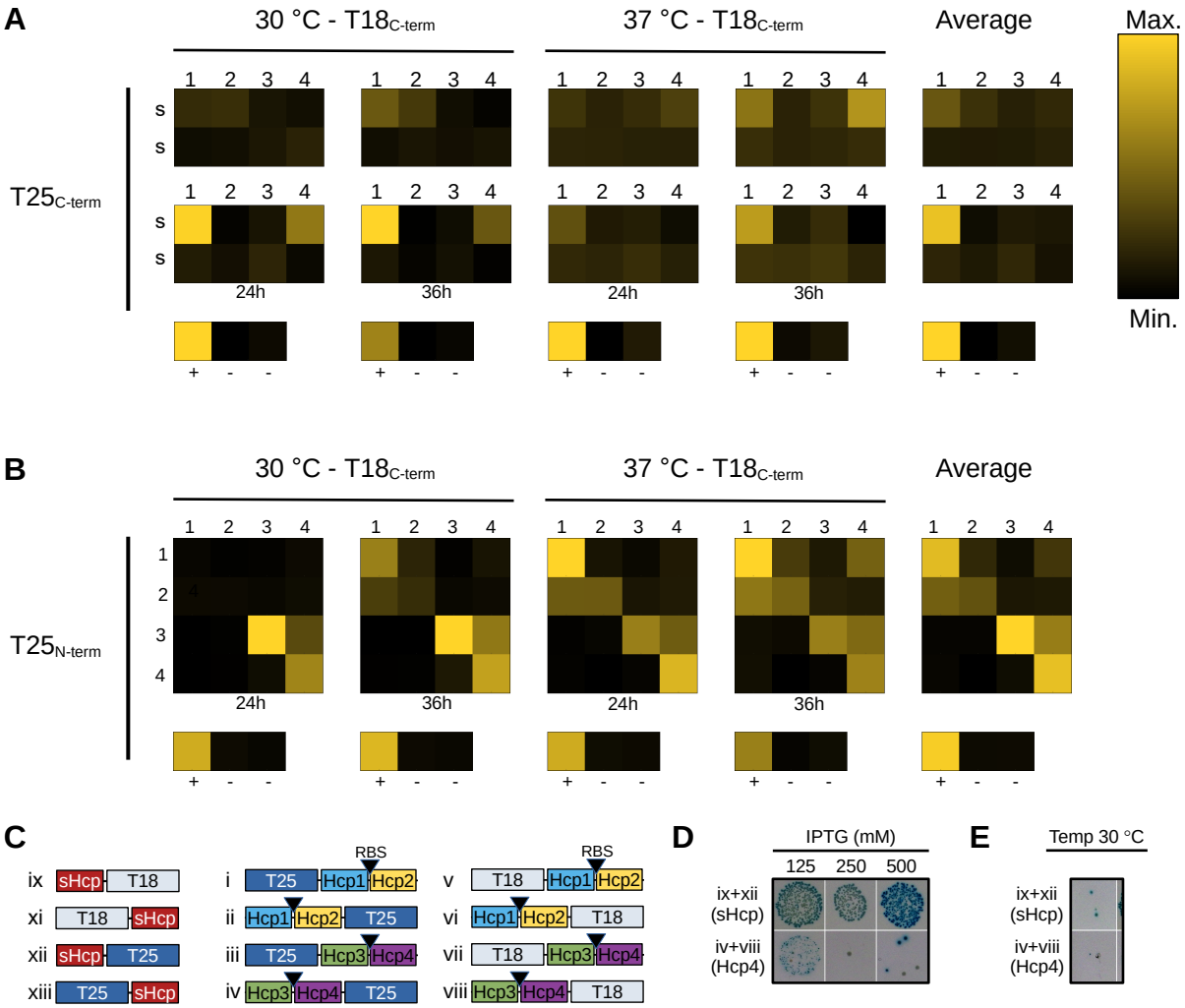

Figure S3. A) The average hydrolized X-gal for each condition tested for competitive two-hybrid assay with plasmids encoding sHcp fused to N- or C-terminal T18 or T25 fragments tested for pairwise interactions against T18 or T25 fusions of Hcp1 (1), Hcp2 (2), Hcp3 (3) and Hcp4 (4) in the presence of 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 S4

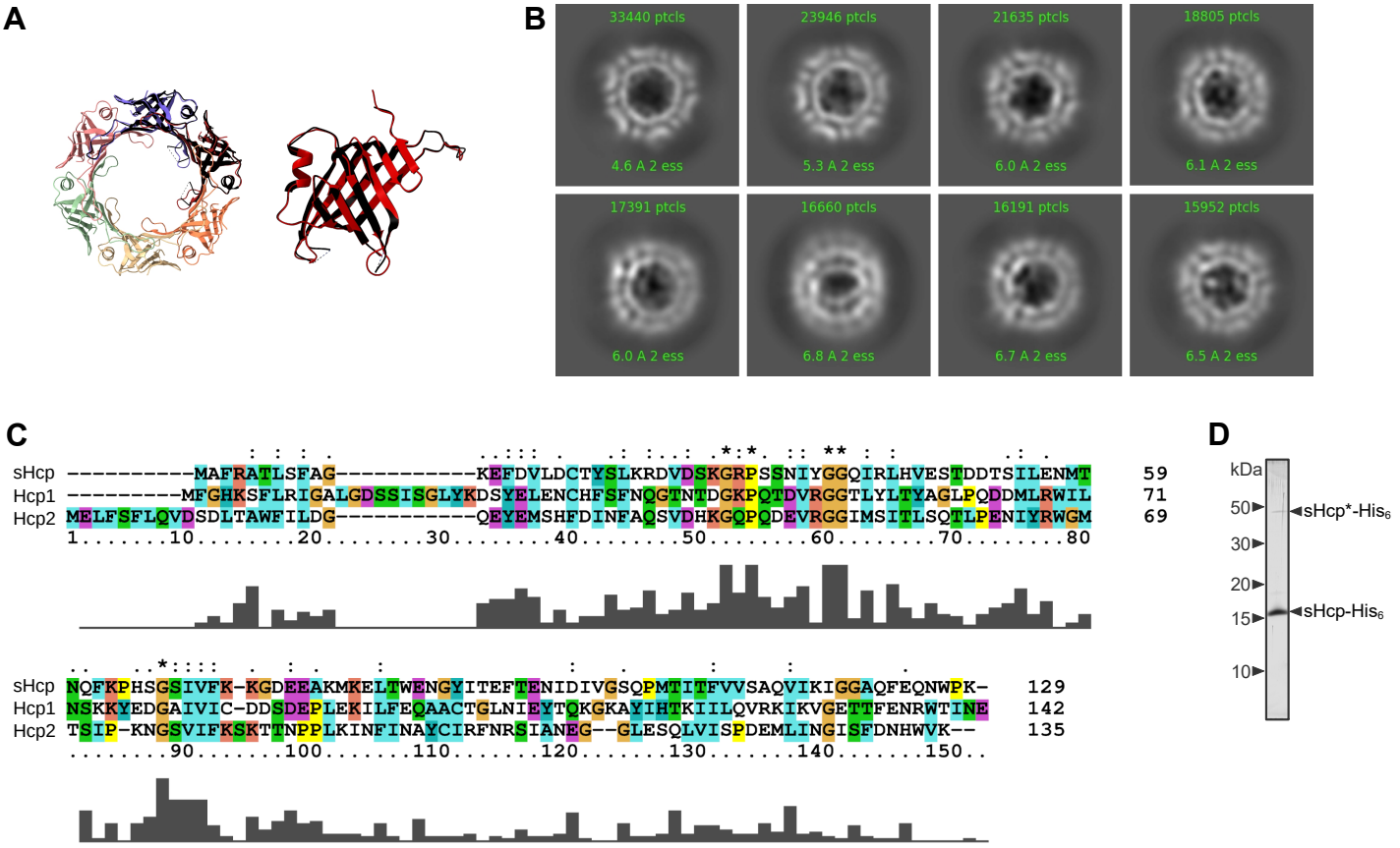

Figure S4. A) Structural alignment of to the He *et al.* unbound structure (black) to our sHcp hexameric crustal structure (left) and monomer (right). **B)** Top six 2-D classes averaged from Cryo-EM data from Hcp1-Hcp2-His_6. **C)** Chimera structural sequence alignment of sHcp with Hcp1 and Hcp2. **D)** Phosphoimager scan of SYPRO Ruby stained SDS-PAGE gels of purified sHcp-His_6 at higher exposure.

Figure S5

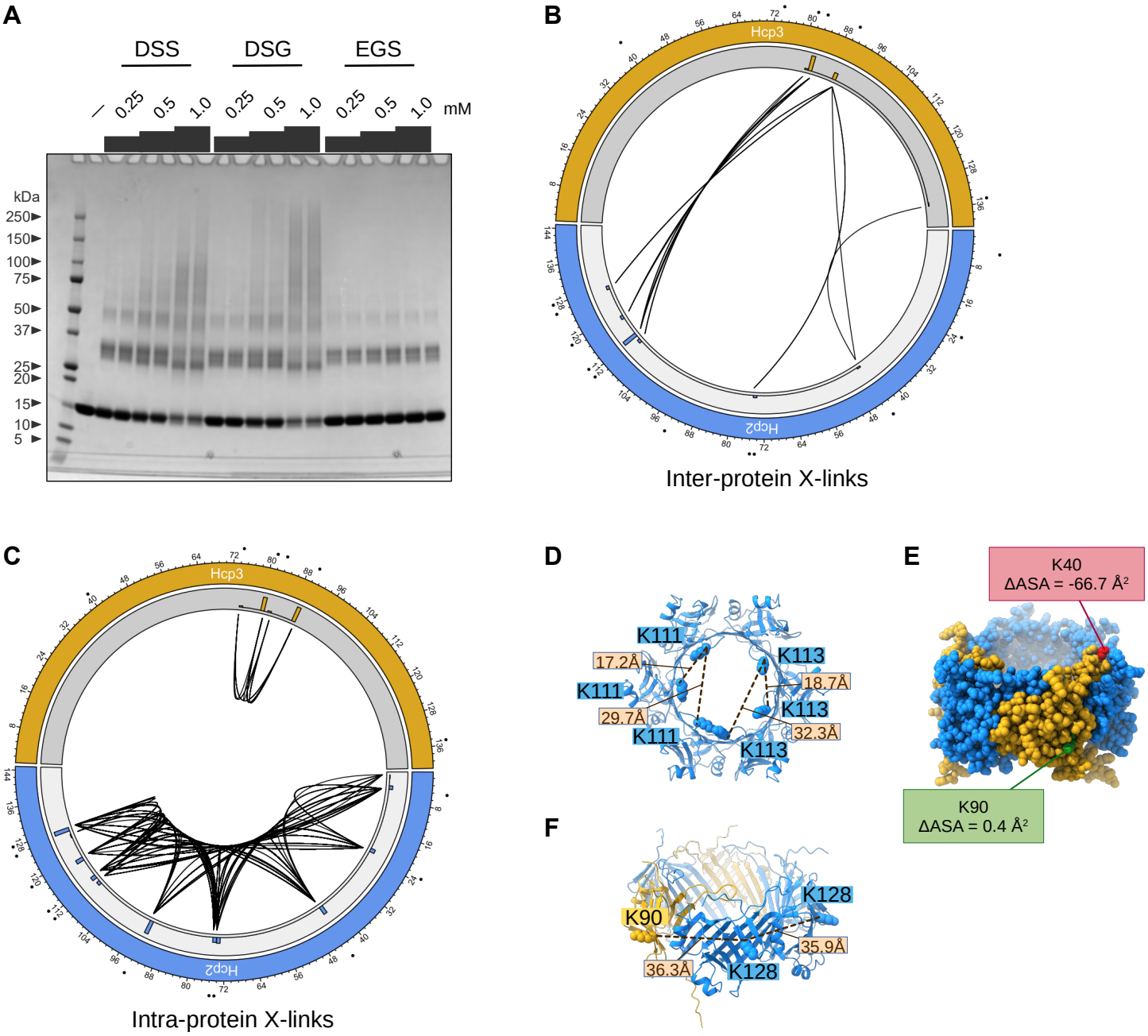

Figure S5. Crosslinking MS of the Hcp1-Hcp2 complex show homo-Hcp1 interactions.
A) SDS-PAGE of X-linked Hcp proteins using DSS, DSG and EGS. **B-C)** Circos plot of all x-linked peptides found in inter- and intra- protein X-links. Gray track contains a histogram of the total number of peptides found 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

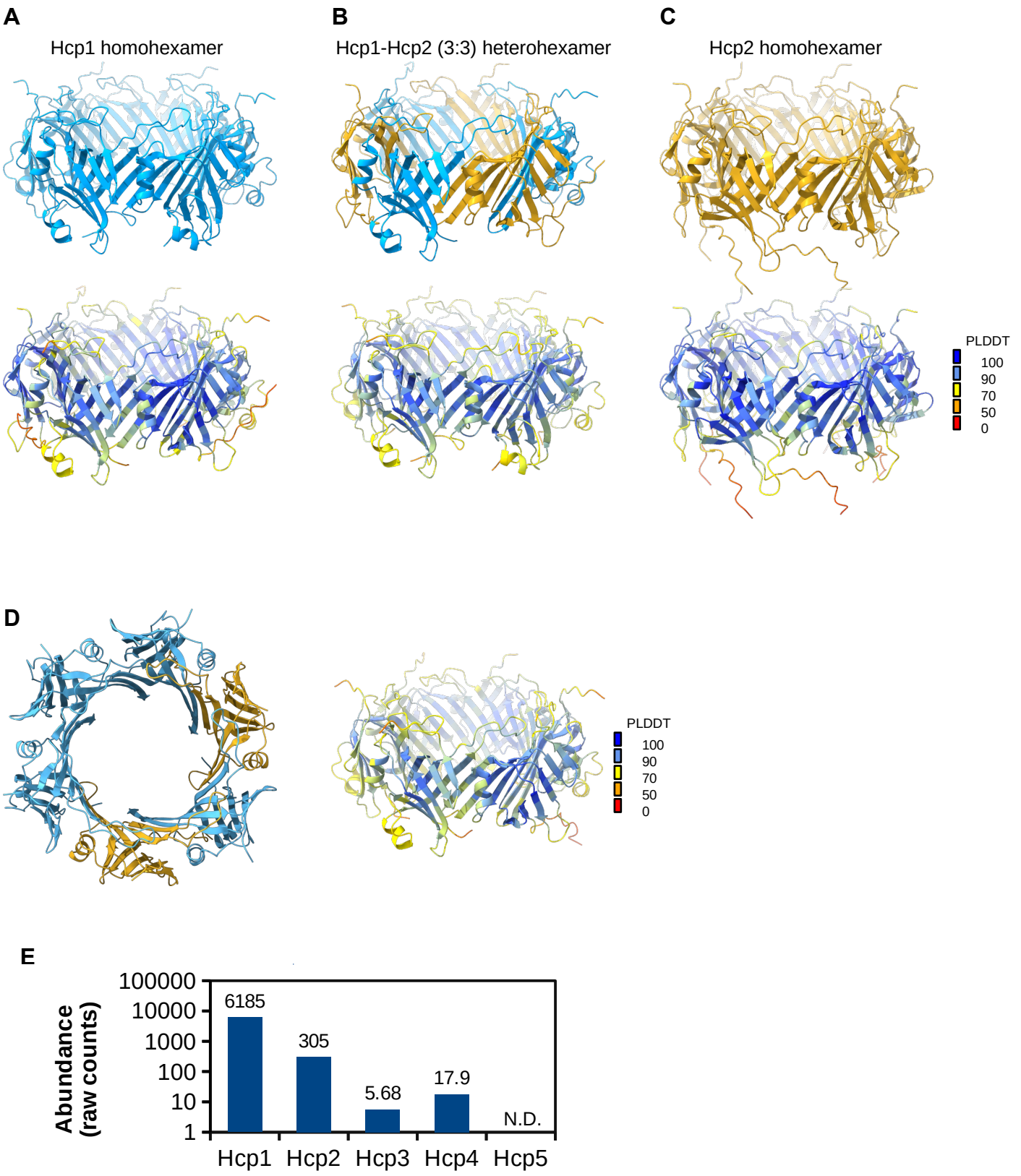

Figure S5. AlphaFold server models of Hcp1 and Hcp2.
A) Homohexameric model of Hcp1 with PLDDT score(bottom). **B)** Heterohexameric model of Hcp1 and Hcp2 with PLDDT colour-coded structure (bottom) in a 3:3 ratio. **C)** Homohexameric model of Hcp2 with PLDDT colour-coded structure (bottom). **D)** AlphaFold model of a Hcp1-Hcp2 complex in a 4:2 complex with PLDDT colour-coded structure (right side). Hcp1 shown in blue and Hcp2 shown in yellow. **E)** Abundance of Hcp proteins in *B. fragilis* according to Müller, *et al.* (https://doi.org/10.1038/s41586-020-2402-x).

Figure S7

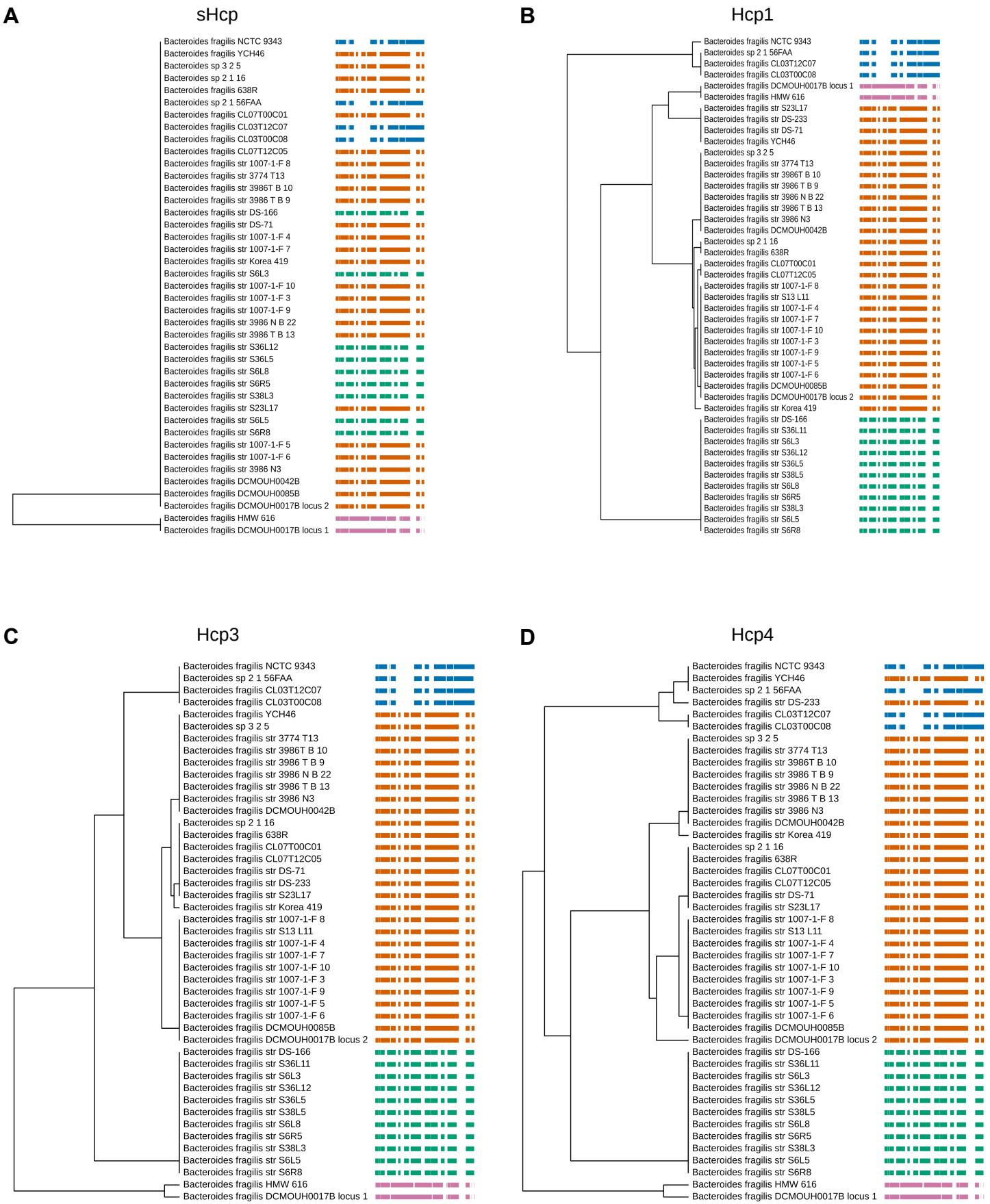

Figure S7. Clustering of phylogenetic trees of sHcp, Hcp1, Hcp3 and Hcp4 proteins from GA3 T6SS with effector proteins. Bte1 (blue), Bfe1 (orange), Cluster 51 (green), Cluster 115 (magenta). Generated with TreeViewer. Trees were compressed horizontally.

Figure S8

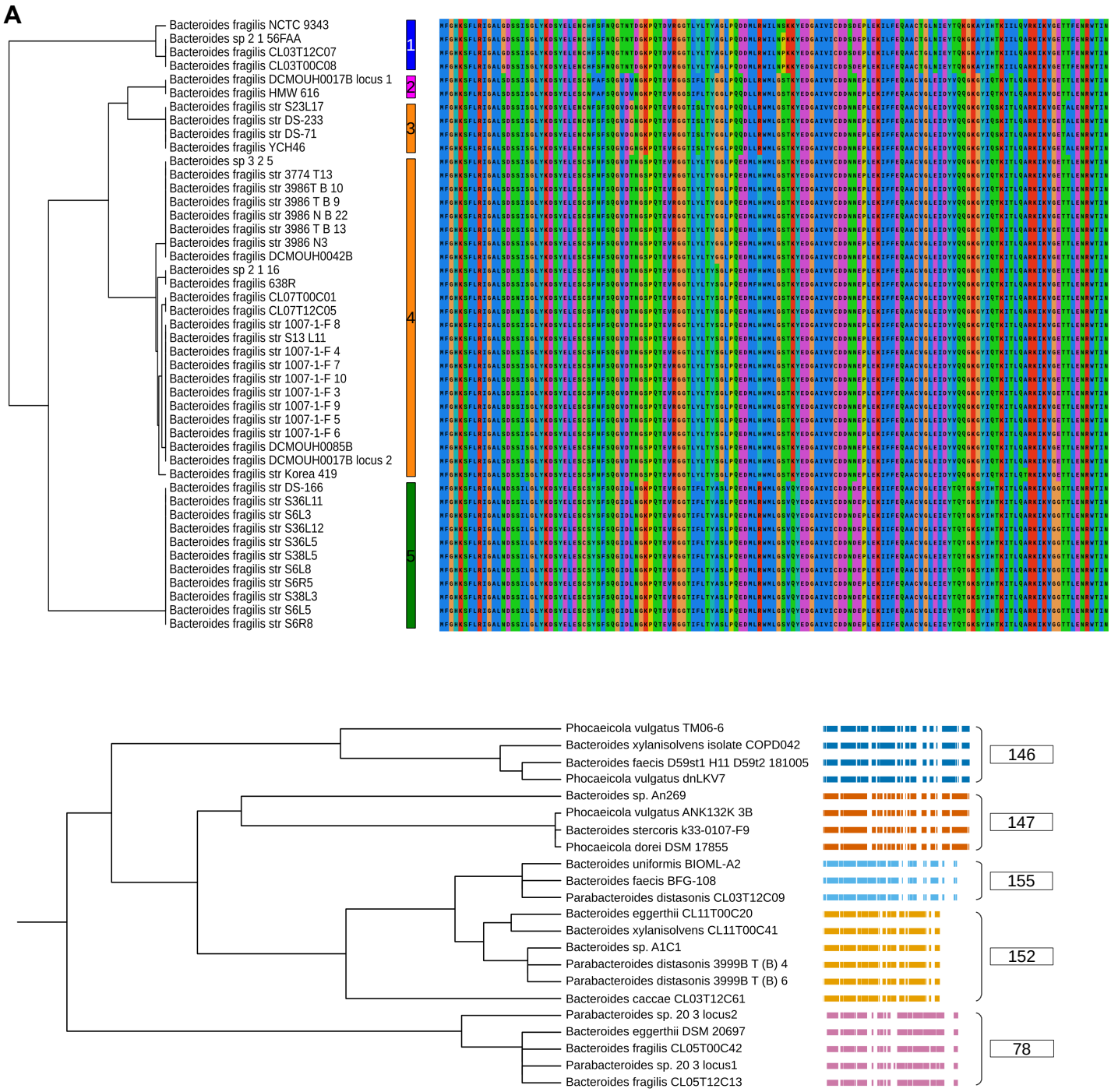

Figure S8. Hcp2 determines binding to Bacteroidales GA2 and GA3 T6SS effectors.
A) Clustering of phylogenetic tree of Hcp2 proteins from GA3 T6SS with their corresponding sequence coloured with Clustal scheme. **B)** Clustering of phylogenetic trees of Hcp2 proteins from Bacteroidales T6SS GA2 architecture with their effectors (colour bars). The cluster number (78, 146, 147, 152 and 155) designates the HMM model named by Coyne, *et al.* 2016 (16). Effectors were aligned with the MAFFT-G-INS-1 algorithm and clustered by average (UPGMA) distance in JalView before adding the different aligned cluster sequences (coloured bars) to the phylogenetic tree. Spaces between the coloured bars (sequences) are gaps in the global alignment of all effectors. Generated with TreeViewer.