CapP mediates the structural formation of biofilm-specific pili in the opportunistic human pathogen *Bacillus cereus*

Ana Álvarez-Mena[1], Muhammed Bilal Abdul Shukkoor[2], Joaquín Caro-Astorga[1,3], Melanie Berbon[2], María Luisa Antequera-Gómez[1], Axelle Grélard[2], Brice Kauffmann[4], Estelle Morvan[4], Antonio de Vicente[1], Birgit Habenstein[2], Antoine Loquet[2,5], Diego Romero[1,5,6]

[1]Departamento de Microbiología, Instituto de Hortofruticultura Subtropical y Mediterránea La Mayora, Universidad de Málaga-Consejo Superior de Investigaciones Científicas, Universidad de Málaga, Málaga, Spain

[2]Institut de Chimie et de Biologie des Membranes et des Nano-objets, Institut Européen de Chimie et Biologie, CNRS, UMR5248, Université de Bordeaux, Pessac, France

[3]The Francis Crick Institute, London, NW1 1AT, United Kingdom

[4]Institut Européen de Chimie et Biologie, UAR3033 CNRS, Université de Bordeaux, Inserm US01, Pessac F-33600, France

[5]Corresponding authors

[6]Lead contact: diego_romero@uma.es

X: @diegoromerohi

SUMMARY

Polymeric proteinaceous filaments are structural scaffolds that diversify the functionality of the bacterial extracellular matrix. Here, we report a previously uncharacterized bacterial factor called *bc1280* that is exclusive to *B. cereus* group and indispensable for the establishment of a biofilm lifestyle. We propose that BC1280 is an essential chaperone for the assembly of the filamentous platform that tightly controls the polymerization of heteropili containing CalY and TasA as major subunits in a concentration-dependent manner. Additionally, BC1280 modulates the expression of EPS via an uncharacterized pathway that is activated by a protease and an ECF-type sigma factor. The pilus biogenesis system described in this work highlights the complexity of extracellular matrix assembly in *B. cereus* and introduces a singular three-part structuration mechanism during biofilm formation and maturation.

KEYWORDS: fibrillary proteins, chaperone, pili, biofilm, extracellular matrix, *Bacillus cereus*

INTRODUCTION

Bacillus cereus, a bacterium commonly found in soil, is capable of colonizing the gastrointestinal tracts of mammals or arthropods; *B. cereus* is contracted through the ingestion of contaminated food, including fresh vegetables and ready-to-eat meals.[1–4] Consequently, *B. cereus* strains are frequently identified as the etiological agent of gastrointestinal diseases outbreaks, and the diseases can progress through emetic or diarrheal symptoms.[5,6] As approximately 70% of microbial infections in humans are reportedly associated with biofilms, much effort has been dedicated to understanding the genetic basis and structural components that, upon activation, lead to the formation of these bacterial communities.[7,8] The transition from a planktonic to a sessile lifestyle is metabolically focused on the synthesis of proteins, polysaccharides and extracellular DNA (eDNA), which ultimately organize the extracellular matrix (ECM).[9,10]The ECM performs a wide range of functions, including adhesion to biotic and abiotic surfaces, regulation of nutrient and signaling molecule transport, defense against physical-chemical stress, and virulence, and provides support for biofilm architecture.[11,12] One of the most attractive components of the ECM is the fibril-forming proteins that serve as structural scaffolds, facilitating the assembly of the microbial community and contributing to surface adhesion, among other biological roles.[13,14] These proteins, termed functional amyloid-like proteins, share remarkable similarities with their pathogenic counterparts in humans, highlighted by their propensity to polymerize into highly ordered fibrillary aggregates, resistance to diverse environmental challenges, and their enrichment in β-sheet content.[15–18]

59 The development of *B. cereus* biofilms involves genetic pathways and structural
60 components of the ECM identical to those found in *B. subtilis*, a closely related species.
61 However, differences arise and are predicted to have ecological and pathological
62 implications.[19,20] Complementary reports have described the role of flagella, matrix
63 proteins, eDNA and exopolysaccharides in the formation of biofilms on *B. cereus.*[21–24] A
64 previous study identified a specific genomic region in *B. cereus* that is orthologous to the
65 corresponding region in *B. subtilis*, and it is expressed mainly in biofilm-associated
66 cells.[25,26] This region encodes an ortholog of *sipW;* two orthologs of *tasA*, called *tasA* and
67 *calY*; the main biofilm regulators *sinI* and *sinR*; and an uncharacterized locus called
68 *bc1280*. Interestingly, no orthologs for *tapA* are present in *B. cereus*. SipW is the signal
69 peptidase that processes TasA and TapA for efficient secretion, and both TasA and CalY
70 of *B. cereus* conserve the signal peptide with the canonical cleavage site recognized by
71 the cognate peptidase SipW.[27,28] The genes *sipW* and *tasA* constitute an operon, and
72 *calY* is expressed independently.[25] The deletion of *tasA* or *calY* resulted in differential
73 defects in biofilm formation, suggesting that the product of each gene might be involved
74 in cell-to-abiotic surface or cell-to-cell interactions, respectively.[25] TasA from *B. subtilis*
75 has been described as a globular protein in its monomeric state that polymerizes into
76 filamentous structures, which can be amyloid or non-amyloid, depending on the
77 experimental context.[29–33] TasA native fibers, observed in vivo in biofilm samples and also
78 in vitro upon extraction, exhibit an enrichment in β-sheet conformation while preserving
79 the β-rich globular structure of the monomer. Consequently, these TasA monomers
80 polymerize into filaments via a donor strand complementation mechanism, and as a
81 result, their structure diverges from the canonical cross-β organization associated with

82 amyloid proteins.[33] Nevertheless, recombinant TasA purified by heterologous expression
83 in *Escherichia coli* has been shown to form fibrils with typical amyloid properties, such as
84 the cross-beta signature.[29,32] TasA might therefore adapt its secondary and tertiary
85 structure in a context-dependent manner, forming either an amyloid conformation or a β-
86 rich globular fold. Extensive structural studies on TasA of *B. subtilis* exist, but only El-
87 Mammeri et al. have reported singularities in recombinantly produced TasA and CalY
88 from *B. cereus,* based on their secondary structure propensity, despite the cross-β
89 amyloid pattern displayed by these paralogous proteins.[32]

90 In this study, we provide the first evidence that the *bc1280* product is indispensable
91 for the growth of *B. cereus* biofilms, relying on two complementary functions. The first is
92 a structural role, i.e., the promotion of CalY polymerization, which is related to the correct
93 polymerization of TasA-CalY heterofibrils, essential for the formation and stability of the
94 ECM scaffold. The second is a regulatory function in which the expression of biofilm-
95 related EPSs is controlled via an undescribed sigma-antisigma pathway. Moreover, our
96 findings indicate that when BC1280 is expressed heterologously in *E. coli*, the N-domain
97 displays a predominantly β-sheet conformation within its rigid core, while the tandem
98 repetitions found in the C-domain potentially facilitate the interaction with CalY. From its
99 localization in the cell wall, BC1280 defines and triggers the polymerization of CalY, which
100 in turn incorporates TasA into nascent heterofilaments. Based on these findings, BC1280
101 was renamed CapP, which is a CalY-assisting polymerization protein. Our study provides
102 a new paradigm for the complex assembly of this proteinaceous appendage in *B. cereus*
103 biofilms, identifies CapP as a structural orchestrator for efficient ECM development, and
104 highlights CapP as a promising target for precisely targeted antibiofilm therapies.

105 **RESULTS**

106 ***bc1280* is an essential locus in the developmental program of biofilm formation**

107 First, a phylogenetic analysis of BC1280 was performed using the amino acid sequence
108 of the type strain *B. cereus* ATCC14579 as a reference. The results revealed that BC1280
109 was present within the *B. cereus* group but not within the phylogenetically related species
110 *B. subtilis* (Figure S1A). The *bc1280* product was annotated as a hypothetical protein with
111 the following putative regions (Figure 1A): i) a signal peptide from position 1 to 39
112 predicted to be processed by a signal peptidase type I (SignalP5.0) (Figure S1B); ii) an
113 N-terminal domain from position 40 to 189 (the N-domain from now on) that contains a
114 highly conserved domain of unknown function (from I^{33} to E^{157}) annotated as DUF4047;
115 and iii) a C-terminal domain (the C-domain from now on) from position 190 to the C-
116 terminal end of the protein. This domain is characterized by a variable number of perfect
117 repetitions of QKKV/AEE motifs, depending on the *B. cereus* strains (Figure S1C).

118 A previous transcriptome analysis of the genetic features that distinguish
119 planktonic and biofilm cells in *B. cereus* showed that *bc1280* is specifically expressed in
120 biofilm-associated cells, and a complementary study revealed that the expression of
121 *bc1280* comprised the regulon of the biofilm-specific master regulator SinR.[34] We
122 confirmed the biofilm-specific expression of *bc1280* via RT–qPCR analysis of samples
123 taken at 48 hours (Figure S1D) and determined that *bc1280* is transcribed as a
124 monocistronic unit independent of the previously reported transcriptional units *sipW-tasA*
125 or *calY* (Figure S1E).[25] The biofilm of the *B. cereus* wild-type strain ATCC14579 was
126 characterized by a white biomass ring that adhered to the plate wells; this ring was usually
127 visible at 24 hours and reached maturity at 72 hours (Figure 1B). A *bc1280* knockout

128 mutant strain (*Δbc1280*) was completely arrested during biofilm formation, a phenotype
129 that differed from the less adhesive or thinner biomass of *ΔtasA* or *ΔcalY,* respectively
130 (Figure 1B).[25] According to the distribution and conservation observed in our phylogenetic
131 analysis, deletion of *bc1280* in the emetic strain of *B. cereus* AH187, which belongs to a
132 different *B. cereus* subgroup, resulted in the complete elimination of pellicle formation at
133 the air–liquid interface compared to that in the wild-type strain (Figure S1F).

134 The wild-type phenotype for biofilm formation was completely restored through
135 complementation of *Δbc1280*, which was achieved through ectopic integration of the
136 *bc1280 locus* (*bc1280*::Δ*bc1280*) under the control of its own putative promoter at a
137 neutral locus (Figure 1B). Moreover, these findings suggested that the promoter for the
138 expression of *bc1280* must be contained within the noncoding sequence between the
139 open reading frame of *tasA* and *bc1280*. BC1280 contains a distinctive C-domain with
140 perfect tandem repetition. As previously described, these characteristic regions often
141 have significant structural and functional importance for the stabilization of protein folding
142 overall.[35] Therefore, we investigated whether the C-domain is involved in biofilm
143 formation. The biofilm formation phenotype of *Δbc1280* was still rescued by a version of
144 BC1280 that lacked the C-domain from Q^{191}-E^{279} (Δ*C-domain*) (Figure 1B). Quantitative
145 measurements at 575 nm of the biomass stained with crystal violet reinforced the visual
146 differences in biofilm thickness (Figure 1C). No significant differences were observed
147 between the strains at 24 hours, when initial cell attachment precedes biofilm maturation;
148 however, differences became evident at 48 hours, coinciding with the initiation of ECM
149 assembly and the growth of biofilms. After cells are initially attached to surfaces, cell-to-
150 cell interactions occur that facilitate the subsequent growth of the biofilm, a cellular

151 phenomenon that can be monitored through auto-aggregation experiments. The wild-type
152 and *ΔC-domain* strains exhibited sharp aggregation kinetics; in contrast, *Δbc1280* did not
153 form sediments at the bottom of the tube at the end of the 24-hour experiment (Figure
154 1D). Taken together, these results indicate that the product encoded by *bc1280* is
155 involved in the developmental stages of bacterial cell-to-cell adhesion and the subsequent
156 assembly and maturation of the ECM, a function that may rely on the N-domain of the
157 protein.
158

159 **The biofilm-defective phenotype of *bc1280* is independent of the expression level**
160 **of *tasA* or *calY***

161 The *Δbc1280* phenotype may result from changes in the expression levels of ECM
162 components or structural malfunctioning in the polymerization of TasA-CalY
163 heterofilaments; these two alternative explanations are not mutually exclusive. A whole-
164 transcriptome analysis of the *Δbc1280* strain revealed 35 genes that were overexpressed
165 and 4 genes that were repressed compared to those in the wild-type strain at 24 h.
166 However, none of these genes were predicted to be associated with biofilm formation
167 (Table S1). Notable transcriptional changes were observed at 48 hours, with 21 genes
168 overexpressed and 28 genes repressed in *Δbc1280*, including *tasA* and *calY* (log2FC
169 values of -3.13 and -3.14, respectively) (Table S2). This result was further supported by
170 RT–qPCR analysis, which revealed lower expression levels for the genes located within
171 the *eps1* and *eps2* genomic regions (Figure 2A). Our transcriptomic analysis additionally
172 revealed that the expression of *bc2793* and *bc2794* pairs (log2FC values of -8.11 and -
173 8.43, respectively) was significant repressed. These uncharacterized loci are annotated

174 as a putative Clp protease and an RNA polymerase ECF-type sigma factor, respectively.
175 ECF sigma factors are known to transduce changes in the cell envelope into a bacterial
176 genetic response[36,37]; thus, we hypothesized that under normal conditions, *bc2793* and
177 *bc2794* positively regulate the gene expression of the main ECM components. The
178 overexpression of *bc2793* and *bc2794* in *Δbc1280* and induction with a 10 μM solution of
179 isopropyl β-D-thiogalactopyranoside (IPTG) significantly increased the biofilm biomass
180 (Figure S2A). Considering that the putative protease (*bc2793*) could regulate the activity
181 of *bc2794*, we overexpressed only the ECF factor, resulting in a six-fold increase in crystal
182 violet staining at 100 μM IPTG compared to the strain without IPTG induction (Figure
183 S2A). In contrast to these findings, the deletion of *bc2794* did not significantly affect
184 biofilm formation, whereas the deletion of *bc2793-bc2794* resulted in a slight reduction in
185 biofilm production compared to the wild-type strain, though this reduction was minimal
186 relative to the biofilm-impaired phenotype observed in *Δbc1280* (Figure S2A). RT–qPCR
187 analysis revealed no significant changes in the expression levels of *tasA* or *calY* in
188 *Δbc1280*-overexpressing *bc2794* or *bc2793-bc2794* compared to those in *Δbc1280*
189 (Figure 2B). However, the relative expression levels of *eps1* and *eps2* were significantly
190 upregulated in response to both genetic combinations. Thus, the ECF factor and its anti-
191 sigma factor were associated with the regulation of exopolysaccharide expression in *B.*
192 *cereus*, rather than the biofilm-related genes *tasA* and *calY*.

193 In accordance with our hypothesis on the structural role of BC1280 in the
194 polymerization of TasA or CalY, no reversion of phenotypes was observed in *Δbc1280*
195 cells overexpressing *tasA, calY*, or both loci (Figure S2B). TEM analysis of negatively
196 stained samples and immunolabeled with anti-TasA or anti-CalY antibodies revealed

197 (Figure 2C and D): i) the presence of filaments from both proteins on the surfaces of WT
198 cells, ii) aggregates reactive to anti-TasA antibodies in *Δbc1280* cells overexpressing
199 *tasA,* and iii) the absence of filaments in *Δbc1280* cells overexpressing *calY* but with an
200 extracellular signal associated with the anti-CalY antibody. We next examined whether
201 TasA or CalY, which were produced by heterologous expression in *E. coli* and added
202 exogenously to *Δbc1280*, might rescue the wild-type biofilm phenotype. The monomeric
203 conformation of TasA or CalY failed to rescue biofilm formation; however, ring biomass
204 formation was only partially restored when the proteins were incubated in vitro for one
205 week to promote filaments assembly and then added at a final concentration of 6 μM (Fig
206 2D). In line with these findings, TEM analysis revealed that upon addition of TasA and
207 CalY in their already assembled conformation, the cell surfaces were decorated with
208 specific immunoreactive aggregates of these proteins (Figure 2E). Taken together, these
209 data suggest that BC1280 is crucial for ECM maturation and raise the hypothesis that it
210 might be involved in the polymerization of CalY or TasA. However, to fully understand the
211 downregulation of *tasA* and *calY* over a longer period, additional research is required to
212 ascertain whether this is a direct result of the absence of *bc1280* or an indirect
213 consequence of an accumulation of TasA and CalY in the cytoplasm that are not being
214 secreted.

215 Our previous findings indicated that the N-domain might partially fulfill the function
216 of BC1280 in biofilm formation by specifically facilitating cell-to-cell interactions.
217 Therefore, this domain may influence the final localization of TasA or CalY within the
218 ECM. TEM analysis and immunolabeling of the *ΔC-domain* strain revealed that nanogold
219 particles associated with TasA extracellular filaments were exposed on the cell surfaces;

220 nevertheless, the morphologies of these filaments were different from those formed in the
221 wild-type cells (Figure 2C and F). Nanogold particles corresponding to CalY monomers
222 or oligomers, but not filaments, were observed in the extracellular matrix of the *ΔC-*
223 *domain* strain (Figure 2C and F). Based on the alterations observed in the polymerization
224 of TasA and especially CalY in the *ΔC-domain* strain, the C-domain of BC1280 might be
225 involved in the polymerization of these proteins or their initial association with the cell
226 surface.
227
228 **BC1280 is a cell wall-associated protein that partially colocalizes with CalY in the**
229 **ECM**
230 To precisely determine how BC1280 contributes to the assembly of the ECM, we further
231 studied the localization of the protein at the cellular level. Our preliminary data indicated
232 that BC1280 is likely expressed at very low levels and is undetectable by
233 immunocytochemistry assays. Therefore, we conducted these experiments using a
234 derivative *Δbc1280* strain that expressed a version of BC1280 fused with a histidine tag
235 at the carboxyl terminal domain. The biofilm formation of *Δbc1280* was rescued by
236 plasmid leakage without IPTG induction or upon induction with a 10 μM IPTG solution
237 (Figure S3B). Furthermore, fractionation and immunodetection experiments with an anti-
238 histidine antibody confirmed the presence of BC1280 in the extracellular matrix, cell wall,
239 and cellular fractions (membrane and cytosolic content) (Figure 3A). TEM analysis and
240 immunolabeling using anti-histidine antibodies revealed that foci of nanogold particles
241 related to BC1280 were distributed regularly along filaments that emerged from the cell
242 surfaces (Figure 3A). To better determine the localization of BC1280 in the cell wall, we

243 conducted confocal laser scanning microscopy (CLSM) and immunocytochemistry using
244 anti-histidine and secondary antibodies conjugated to Atto 488 and double staining with
245 wheat germ agglutinin (WGA) and Hoechst to visualize the cell wall and DNA,
246 respectively. According to the results of the fractionation experiments, a limited number
247 of foci corresponding to BC1280 proteins fused to the histidine tag per cell were observed
248 in association with the cell surface (Figure 3C and S3C). Considering the relevance of the
249 N-domain to the functionality of BC1280, we also studied its contribution to the cellular
250 localization of whole BC1280. Based on crystal violet staining of the biomass that had
251 adhered to the well, the amount of N-domain expressed without IPTG induction or after
252 induction with 10 μM IPTG solution was sufficient to rescue the biofilm phenotype of
253 *Δbc1280* (Figure S3B). Cellular fractionation and immunoblotting assays demonstrated a
254 specific reactive band in the extracellular matrix and cellular fractions and an undetectable
255 signal associated with the cell wall (Figure 3B). In addition, visualization via TEM
256 confirmed the presence of N-domain-related foci along the filaments emerging from the
257 cells, as previously observed for the entire protein, but no detectable fluorescent signals
258 associated with the cell wall were observed via CLSM (Figure 3B and C). The disparities
259 noted between the full-length and the truncated form of the protein suggest that the C-
260 domain contributes to the retention of BC1280 in the bacterial cell wall, likely before
261 BC1280 reaches the ECM.

262 To further define the connection between BC1280 and CalY or TasA at the cellular
263 level, we studied their colocalization via immunocytochemical analysis of cells expressing
264 pUTE657-*bc1280* fused to a histidine tag using specific primary antibodies. CLMS
265 analysis revealed that CalY and BC1280 colocalized along the fibrils that compose the

266 ECM (Pearson's coefficient of 0.58 and a p value of 0.0002) (Figures 3D and S3D). In
267 most cases, the signal corresponding to CalY was homogeneously distributed, and
268 BC1280 was detected as discrete foci dispersed along the filaments. However, no
269 colocalization signal was detected for TasA or BC1280, with a Pearson's coefficient of
270 0.19 and a p value of 0.0002 (Figure S3E and F). These findings were supported by
271 Mander's coefficient, which represents the percentage of overlapping pixels for each
272 fluorescent channel. For the BC1280-CalY pair, approximately 65% of the pixels
273 overlapped in both channels; for BC1280-TasA, the overlap decreased to 26% for Atto
274 488 and 21% for Alexa Fluor 647 (Figure S3F). The observation of BC1280 and CalY
275 colocalizing within the ECM hints at a potential structural relationship, including possible
276 direct interactions, between these proteins.

277 **Heterologously produced BC1280 oligomerizes into macromolecular complexes**
278 **driven by the N-domain**

279 BC1280 and the truncated version comprising the N-terminal domain (from A^{39} to G^{190})
280 were purified by heterologous expression in *E. coli* Lemo21(DE3) and subjected to further
281 structural analyses (Figure 4A). The two versions of the protein were purified under
282 soluble conditions by affinity chromatography followed by size exclusion chromatography,
283 and by tandem mass spectrometry analyses, the amino acid sequences of each protein
284 were confirmed (Figure S4A). The two fractions collected at different elution volumes in
285 the size exclusion chromatogram suggested that in buffer containing 20 mM Tris and 50
286 mM NaCl at pH 7, BC1280 and N-domain$_{39\text{-}190}$ could assemble into two main molecular
287 entities with distinct oligomeric states (Figure 4B and E). Dynamic light scattering (DLS)
288 analysis showed that both protein assemblies exhibited a degree of polydispersity, with

289 particles ranging in diameter from ~10 to 600 nm (Figure 4C-D and F-G for BC1280 and
290 N-domain$_{39-190}$, respectively; for further details, see Table S3). Remarkable differences
291 were observed in the second elution sample; specifically, the N-domain$_{39-190}$ exhibited a
292 greater proportion (~53%) of small oligomers than did BC1280 (~18%) (Figure 4D and
293 G). To complement the DLS results, each fraction was resolved in a native gel (Figure
294 4H). The sample corresponding to fraction 1 of each protein, which presumably contained
295 larger aggregates, failed to enter the gel, suggesting that the molecular complex was
296 larger than 1000 kDa. The second elution fraction was resolved as a band at 480 or 240
297 kDa for BC1280 or N-domain$_{39-190}$, respectively, which corresponds to oligomers of
298 approximately 15 subunits each. TEM analysis confirmed that fraction 1 contained large
299 aggregates of varying sizes for both proteins (Figure S4B). The samples corresponding
300 to the second elution fraction contained macromolecular complexes that were similar but
301 smaller in diameter (Figure S4B). Taken together, these results indicate that the variability
302 in the oligomerization of the full-length protein is preserved in the truncated version that
303 contains only the N-terminal region.

304 To further delineate the dynamics of oligomerization, we focused on the second
305 fraction, which was composed mainly of small oligomers and could promote the formation
306 of larger molecular entities. The fluorescence signal emitted upon binding to thioflavin T
307 (ThT)[38] showed a fast polymerization curve without a lag phase, aligning with the
308 presence of oligomers under native conditions (Figure 4I). Interestingly, the slope of the
309 exponential phase was more pronounced for the N-domain$_{39-190,}$ and the fluorescence
310 intensity signal was twofold greater than that of the full-length protein after 150 hours of
311 incubation. These results correspond with the higher percentage of small oligomers

312 formed by the N-domain$_{39-190}$ (via DLS analysis), which suggested that the truncated
313 protein seeds rapidly. Furthermore, TEM analysis of negatively stained samples revealed
314 macroscopic particles of variable sizes and an amorphous aggregated morphology for
315 both proteins, resembling the aggregates observed for CalY but not the fibrillar structures
316 formed by TasA (Figure 4J).[32] These results strongly indicate that the N-domain is
317 involved in protein oligomerization and suggest that the absence of the C-domain in the
318 truncated protein might facilitate a faster aggregation rate.

319 **The assembled N-domain$_{39-190}$ contains a rigid core with two β-strand-rich regions,**
320 **while the C-domain increases global rigidity**

321 The amyloid-like nature of TasA and CalY[32], when heterologously produced in *E.coli*,
322 along with the association of BC1280 with these two proteins *in vivo* and its tendency to
323 oligomerize *in vitro*, led us to suspect an amyloid-like nature for BC1280. Since the N-
324 domain retains the functionality of BC1280 and the ability to form oligomers, this version
325 of the protein was considered for our initial structural studies. The X-ray diffraction pattern
326 of the centrifuged pellet of the N-domain$_{39-190}$ assemblies displayed typical features
327 reminiscent of the cross-β structure of amyloids, with strong signals at 4.7 Å and a slight
328 signal at ~10-11 Å, corresponding to the inter-strand and inter-sheet spacings,
329 respectively[39] (Figure 5A). In light of these findings, we predicted protein aggregation and
330 amyloidogenicity based on the amino acid sequence of BC1280 using the bioinformatics
331 tools AmylPred2,[40] FoldAmyloid,[41] MetAmyl,[42] PASTA 2.0[43] and TANGO.[44] At least two
332 algorithms predicted the following amino acid stretches as putative aggregative regions
333 and amyloid-forming segments: i) L^{51}STAIIFP^{58}, ii) I^{98}S^{99}, iii) L^{115}QNVYT^{120} and iv)
334 V^{144}SITVN^{149} (Figure S4C; letters in bold). Based on bioinformatics analysis using

335 FoldUnfold,[45] the N-domain appears intrinsically folded but the C-domain is predicted to
336 be unstructured, possibly due to the presence of repetitive amino acid sequences rich in
337 Gln, Lys, and Glu, which are disorder-promoter residues[35,46] (Figure S4C). As the
338 propensity to form β-sheet-rich self-assembled structures is often associated with regions
339 enriched in hydrophobic residues, we analyzed the hydrophobicity pattern of BC1280
340 using ProtScale.[47] The peptide T^{50}LSTAIIFPKT^{60}, previously identified as a putative
341 amyloidogenic segment, was found to be highly hydrophobic (Figure S4D). Finally,
342 AlphaFold[48,49] was used to obtain a tentative secondary structure. The model
343 predominantly exhibits an α-helix conformation with low confidence value, except for a
344 small central helical sandwich. This contrasts with the enrichment in β-sheet structural
345 elements derived from ThT polymerization kinetics and the cross-β structure observed by
346 X-ray diffraction (Figure S4E). We further performed fine structural characterization of the
347 molecular conformation of these protein assemblies using magic-angle spinning (MAS)
348 SSNMR, which has previously been employed to study non-crystalline self-assemblies,
349 such as amyloids.[50–52] Uniform ^{13}C/^{15}N-labeled N-domain$_{39\text{-}190}$ assemblies were produced
350 to improve NMR sensitivity and perform multidimensional MAS SSNMR studies. To
351 investigate the presence of a central rigid core within the assemblies, we used dipolar-
352 based polarization transfer (cross-polarization) and two-dimensional ^{13}C-^{13}C
353 spectroscopy. A 50 ms mixing time proton-driven spin-diffusion (PDSD) experiment was
354 used to detect intra-residue ^{13}C-^{13}C correlations from the rigid core of the assembly. For
355 the spectral analysis, we first attempted to exploit the secondary structure model
356 predicted by AlphaFold,[48,49] and the Cα and Cβ chemical shifts were predicted using
357 SPARTA.[53] Subsequently, the predicted spectrum was superimposed with the

358 experimental 2D ^{13}C-^{13}C spectrum acquired for N-domain$_{39-190}$ (Figure S5A). No
359 correlation was found between the experimental and predicted NMR spectral patterns,
360 which reflects the discrepancies between the AlphaFold model and the structure of the
361 experimentally obtained multimers. These findings highlight that, while bioinformatics
362 predictions might reflect the relevant structures under certain conditions, differences
363 between predicted and experimentally observed structures can result from factors such
364 as the nature of the sample, the buffer composition (pH or salt), or oligomerization. The
365 NMR data provide evidence for the propensity of BC1280 to undergo structural changes
366 upon assembly under certain *in vitro* or environmental conditions. Thus, we manually
367 inspected the SSNMR cross-peaks and identified ~35-40 spin systems (i.e.,
368 corresponding to ~35-40 residues), as exemplified for an isoleucine in Figure 5B.
369 Following well-established statistical data for Cα and Cβ chemical shifts according to their
370 secondary structure,[54] we noticed that these residues are involved mainly in β-sheet
371 secondary structure elements. This finding was illustrated by the isoleucine spectral
372 pattern observed with Cα and Cβ chemical shift values of ~60 and 42 ppm, respectively,
373 which are indicative of a β-sheet conformation (Figure 5B). Approximately 35-40 out of
374 the 151 amino acids comprising the N-domain$_{39-190}$ were detected within the rigid core by
375 SSNMR, suggesting that the rigid core encompasses only a portion of the N-domain$_{39-190}$
376 sequence. However, the presence of residues undetectable in dipolar-based SSNMR
377 experiments is not surprising, as this observation has been made for most pathological
378 and functional amyloids and self-assemblies to date.[55,56]

379 In light of these findings, we observed that the spectral resolution, as measured by
380 the ^{13}C line width, is ~60-120 Hz (full width at half-height), which indicates a relatively

381 well-ordered of the subunits within the complex. These observations are consistent with
382 our previously reported ^{13}C line widths for *B. subtilis* TasA (~30-100 Hz), *B. subtilis* TapA
383 (~200 Hz) and *B. cereus* CalY (~200 Hz), which, respectively form well-ordered fibrils
384 (TasA) and polymorphic aggregates (TapA and CalY) in vitro.[32] To gain more site-specific
385 insights into the rigid core of the N-domain$_{39-190}$, the protein was incubated with proteinase
386 K, the bands corresponding to the rigid regions were resolved via SDS–PAGE, and the
387 amino acid sequences were determined through tandem mass spectrometry analysis
388 (Figure 5C and D). Interestingly, proteinase K digestion suggested that the following
389 resistant regions were present: the highly hydrophobic region V^{47}-K^{59}, which contains the
390 Ala-Ile-Ile segment, and the region A^{88}-K^{134}. Further examination of the SSNMR data
391 using an additional 200 ms PDSD experiment (Figure S5C) led us to identify several rigid
392 protein segments. Specifically, these segments correspond to the regions identified
393 through proteinase K digestion analysis, namely, $K^{46}VEA^{49}$ and $E^{93}IKQAISV^{100}$. Notably,
394 the total number of residues identified as resistant to proteinase K was slightly greater
395 than the number of spin systems identified by SSNMR. This difference suggests that the
396 core visible in SSNMR contains fewer residues compared to the regions identified as
397 proteinase K resistant. Alternatively, this observation may result from insufficient spectral
398 resolution to detect all correlations. Based on these results, it can be inferred that the N-
399 domain$_{39-190}$ contains a structured core with a significant presence of rigid residues in a
400 β-sheet conformation.

401 As the full-length protein exhibits distinct polymerization kinetics from those of the
402 N-domain$_{39-190}$, we conducted a comprehensive biophysical characterization of full-length
403 BC1280 to identify putative differences from the truncated N-domain$_{39-190}$ in the context

404 of macromolecular assemblies. The full-length protein sample underwent a similar
405 treatment as the N-domain$_{39-190}$ sample, and upon ultracentrifugation, the following
406 notable differences were observed: i) BC1280 assembled into a transparent gel-like pellet
407 with high viscosity, while ii) the N-domain$_{39-190}$ pellet aggregated into a non-viscous white
408 precipitate. Consistent with previous findings for the N-domain$_{39-190}$, the diffraction pattern
409 of the BC1280 assemblies showed an intense signal at ~4.7 Å and a less pronounced
410 peak at ~10 Å (Figure S5D). The C-domain of BC1280 consists of fifteen tandem repeats
411 that were predicted to be disordered; thus, we wondered whether this domain participates
412 in the overall folding of the full-length protein assemblies. BC1280 was uniformly ^{13}C
413 labeled, and the resulting aggregates were analyzed via MAS SSNMR. To compare the
414 molecular conformation of N-domain$_{39-190}$ to that of full-length protein assemblies, we
415 superimposed 2D ^{13}C-^{13}C SSNMR spectral fingerprints and observed significant overlap,
416 suggesting that the rigid core of the N-domain is highly conserved in full-length protein
417 assemblies (Figure S5E). As only a few additional NMR signals were observed in the full-
418 length protein sample, we hypothesize that the C-domain is not a primary component of
419 the rigid structural fold of the BC1280 assemblies.

420 To complement the SSNMR analysis of the rigid BC1280 core, we used dynamics-
421 based spectral editing via J-coupling polarization transfer to detect dynamic protein
422 segments[57] and gain insights into any dissimilarities between the two protein versions.
423 The superimposition of 2D ^{1}H-^{13}C-insensitive nuclei was enhanced by polarization
424 transfer (INEPT) of the BC1280 and N-domain$_{39-190}$ assemblies, and the results revealed
425 a significant difference in the spectral fingerprint between the two samples (Figure 5E).
426 More INEPT signals were observed in the N-domain$_{39-190}$ than the full-length protein,

427 suggesting that the absence of tandem repetitions in the C-domain results in a more
428 flexible conformation overall. To confirm this hypothesis, we estimated the degree of
429 rigidity by comparing signal intensities from 1D ^{13}C-detected INEPT and CP experiments
430 (Figure 5F), respectively, by probing mobile and rigid protein segments through the
431 INEPT/CP intensity ratio. We observed a different behavior, with a decrease in mobile
432 signals in BC1280 (ratio ~0.1) compared to N-domain$_{39-190}$ (ratio ~0.25), suggesting fewer
433 dynamic regions in the assemblies of the full-length protein.

434 Considering the differences in the overall rigidity between BC1280 and N-
435 domain$_{39-190}$, we estimated the relative amount of secondary structural elements using
436 attenuated total reflection-Fourier transform infrared spectroscopy (ATR-FTIR) analysis.
437 To identify the minimum peaks within the amide I range according to well-established
438 assignments,[58] the second derivative was calculated from the raw data, and
439 deconvolution was performed to determine the percentage of contribution for each
440 secondary structure type (Figure 5G and S5F, G and H).Overall, due to the additional
441 structural elements, the ATR-FTIR spectra of BC1280 was more challenging to fit. The
442 N-domain$_{39-190}$ exhibited 33% antiparallel β-sheets, while the full-length protein had 25%
443 antiparallel β-sheets and a slightly greater proportion of parallel β-sheets. Both proteins
444 displayed an equal overall percentage of secondary structure in the random coil
445 conformation, while BC1280 contained slightly more α-helices and turns than the N-
446 domain$_{39-190}$.

447

448 **BC1280 acts as a chaperone for CalY, with its functionality modulated by the C-**
449 **domain**

450 The observed in vivo colocalization of BC1280-CalY in the extracellular matrix and the
451 potential involvement of BC1280 in CalY polymerization led us to ask whether BC1280
452 might act as an accessory protein that influences CalY assembly. We examined how
453 BC1280 affects CalY by studying the aggregation kinetics of CalY at different BC1280
454 molar ratios using ThT fluorescence emission. Surprisingly, we observed that even at
455 ratios (CalY:BC1280) of 70:30 or 50:50, the polymerization kinetics were significantly
456 affected compared to CalY incubated alone (Figure 6A). This negative impact, together
457 with the fact that BC1280 expression is likely tightly regulated, prompted us to investigate
458 the relative levels of BC1280 and CalY during biofilm formation using an isobaric tag for
459 relative and absolute quantification (iTRAQ) analysis (Figure 6B). At 24 hours, the relative
460 accumulation rate of CalY was approximately 3.3 times higher than that of BC1280;
461 nevertheless, BC1280 accumulated by 48 hours, reaching a ratio (CalY:BC1280) of 2:1.
462 This information was used to re-estimate the protein ratios for copolymerization assays
463 of CalY with BC1280. The presence of low concentrations of BC1280 (2 and 10 μM) did
464 not impact the polymerization kinetics of a 40 μM solution of CalY (Figure 6C). In contrast,
465 the presence of only the N-domain$_{39\text{-}190}$ significantly reduced the rate of polymerization
466 kinetics of CalY, resulting in a lower fluorescence intensity at the end of the experiment
467 compared to CalY alone (Figure 6C). In addition, the negative impact on CalY
468 polymerization was amplified in the presence of non-freshly purified full-length protein or
469 N-domain$_{39\text{-}190}$, which are enriched in preformed oligomers (Figure S6B and C). The
470 samples corresponding to CalY incubated with a freshly purified 10 μM solution of
471 BC1280 or N-domain$_{39\text{-}190}$ were visualized via TEM. Consistent with the ThT results,
472 incubation of 10 μM BC1280 with CalY resulted in macromolecular aggregates similar to

473 those formed by CalY (Figure 6D). However, the incubation of CalY with the N-domain$_{39-}$
474 $_{190}$ led to the accumulation of small oligomers (Figure 6D), aligning with the previous *in*
475 *vivo* results where the absence of the C-domain (*ΔC-domain*) arrested the formation of
476 CalY filaments (Figure 2F). The ability of BC1280 to regulate the CalY polymerization rate
477 points to a potential chaperone role for BC1280 in a concentration-dependent manner.
478 This hypothesis was further supported by circular dichroism analysis, which revealed
479 significant structural changes in CalY in the presence of BC1280 in a concentration-
480 dependent manner. After 16 hours of incubation under agitation, a 10 μM solution of
481 BC1280 induced an increase in the β-sheet conformation of CalY compared to CalY alone
482 or after incubation with 2 μM BC1280, which predominantly appeared as a random coil
483 or a disordered conformation (Figure 6E).

484 The experimental data shown in this work strongly suggest a direct interaction
485 between BC1280 and CalY, yet the contribution of TasA within the filaments that support
486 the ECM of *B. cereus* remains uncertain. In vitro co-assembly experiments have
487 demonstrated that CalY can function as an accessory protein in the assembly of TasA,
488 without altering its structural arrangement.[32] Moreover, previous findings from our
489 laboratory indicated that deletions in *tasA* or *calY*, while containing the *bc1280* gene, did
490 not completely arrest biofilm formation; instead, mutant strains displayed biofilms with
491 distinct phenotypes.[25] Therefore, we addressed the distribution of their respective
492 paralogue protein within the ECM in *ΔtasA* and *ΔcalY* mutants compared with the wild-
493 type strain. Immunochemistry and CLMS analysis revealed fibrillary structures reactive to
494 anti-TasA antibodies decorating the surfaces of wild-type cells, whereas discrete foci of
495 TasA fluorescence signals were observed in the ECM of *ΔcalY* (Figure 6F). Remarkably,

496 in *ΔtasA* mutant strain, CalY protein exhibited a higher degree of fibrillary organization
497 compared to the immunofluorescence signal observed in *B. cereus* wild-type cells (Figure
498 6F). These results provide evidence that CalY can polymerize into filaments in *ΔtasA* in
499 the presence of BC1280. In contrast, its paralogue TasA doesn't polymerize into filaments
500 *in vivo* when CalY is absent, highlighting the high complexity of the mechanism underlying
501 the ECM assembly in the human pathogen *B. cereus*.
502

503 **DISCUSSION**

504 Bacterial extracellular filaments are fascinating polymers from structural and biological
505 perspectives. The diversity of structures, sizes, functions and mechanisms of
506 polymerization reflects the specificity of their functions and need for tight regulation. In
507 this work, we pinpointed BC1280 as a specific biofilm orchestrator factor that is exclusive
508 to the *B. cereus* group. In addition, we propose an intriguing molecular mechanism by
509 which this protein could contribute to the assembly of TasA-CalY filaments on the cell
510 surface of this opportunistic human pathogen. The severe defective phenotype observed
511 in *Δbc1280*, when contrasted with the phenotypes of *ΔtasA* and *ΔcalY*, substantiates the
512 indispensable role of BC1280 in biofilm maturation.[25]

513 While BC1280 could be compared with the chaperone TapA from the
514 phylogenetically related strain *B. subtilis*, notable differences between the two proteins
515 highlight their distinct characteristics. First, BC1280 is exclusive to biofilm cells, whereas
516 TapA is expressed in the planktonic population and from the early stages of biofilm
517 formation. Upon *tapA* deletion, altered and fewer filaments are produced due to a drastic
518 decrease in TasA levels, while the transcriptional expression of *tasA* remains

519 unaffected.[59] Remarkably, our findings reveal that *tasA* and *calY* levels are
520 transcriptionally downregulated in *Δbc1280*, resulting in the absence of filaments. It is
521 important to highlight that in a previous study by our group, we observed that the
522 heterologous expression of *B. cereus* alleles in the *B. subtilis ΔtapA* strain only reverts
523 biopellicle formation when *sipW-to-calY*, including *bc1280*, are overexpressed, indicating
524 only partial functional and mechanistic overlap of TapA and the proteins encoded in *sipW-*
525 *to-calY*.[25] Moreover, in the same study, it was demonstrated that in the *B. subtilis ΔtasA*
526 mutant, in which *tapA* is still expressed, when only *sipW-tasA* or *sipW-calY* from *B. cereus*
527 are overexpressed without the *bc1280* allele, the biopellicle is not formed, clearly
528 evidencing that the mechanism of assembly of TasA and CalY is specific of BC1280 and
529 cannot be replaced by TapA from *B. subtilis*.[25] The idea of a partial mechanistic overlap
530 between BC1280 and TapA is further corroborated by the structural differences observed
531 in the predicted model of BC1280 and the solved structure of TapA.[60] While considering
532 that structures exhibit context-dependent variations, they were partially predicted with
533 high confidence, supporting the notion that these structures should occur in a specific in
534 vivo configuration and perform functions such as membrane- or cell-wall interactions.

535 The results of this study demonstrate that BC1280 contains distinct structural and
536 functional regions: the N-terminal half, which is sufficient to fulfill the functionality of
537 BC1280; and the C-terminal half, which appears to be involved in attachment to the cell
538 surface and also controls the interaction of BC1280 with CalY. After establishing that
539 BC1280 is an ECM protein and its association with the cell surface, along with its
540 functional relationship with CalY, we propose its role as a chaperone for CalY based on
541 the following lines of evidence: i) the biofilm phenotype remains unchanged even after

542 the overexpression of *tasA* and/or *calY* in the *Δbc1280* mutant strain; ii) filaments of CalY
543 are not formed in the *ΔC-domain* strain; iii) colocalization of BC1280 and CalY within the
544 ECM occurs *in vivo;* and iv) BC1280 induces a structural rearrangement of CalY in vitro.
545 Collectively, these findings prompted us to rename BC1280 as CapP, a CalY-assisting
546 polymerization protein. In addition to its fundamental role in controlling pilus assembly,
547 CapP functions as a checkpoint of biofilm maturation by regulating the expression of EPS
548 via an uncharacterized pair of Clp-proteases and cognate ECF-type sigma factors.
549 However, this is a novel and specific biofilm regulatory pathway that deserves further
550 investigation.

551 We determined in vitro that the N-domain of CapP might trigger or support protein
552 aggregation, leading to the formation of a rigid core of ~ 40 amino acids enriched in β-
553 sheets. Through SSNMR, we determined at the atomic level that the rigid core contains
554 at least the following rigid regions: $K^{46}VEA^{49}$ and $E^{93}IKQAISV^{100}$. The role of each amino
555 acid stretch remains to be elucidated; nevertheless, we propose that the
556 $K^{46}VEATLSTAIIFP^{58}$ region significantly influences the aggregation propensity of CapP
557 in vitro, considering that i) the sequence is situated in the most hydrophobic region of the
558 protein, ii) the residues within this region are highly conserved, and iii) the region was
559 detected in the proteinase K-resistant fraction. Although the core is conserved in the full-
560 length protein, it exhibits slower polymerization kinetics and a greater polydispersity index
561 than that of the truncated N-domain. In addition, a higher global rigidity is generated by
562 the C-domain in CapP compared to the truncated N-domain. Together, these findings
563 strongly suggest that the entire protein has a more complex folding pattern, in which
564 tandem repetitions may play a role in regulating oligomerization and macromolecular

565 aggregate assembly. Because the in vitro-generated assemblies could represent only a
566 fraction of the structures adopted in vivo, the similar aggregation propensity of the
567 proteins into β-rich filaments supports the hypothesis that such interactions could promote
568 in vivo structuring.

569 These findings complement previously conducted structural studies on the
570 amyloid-like TasA and CalY proteins[32] and expand our knowledge of the coordinated
571 assembly of these proteins into filaments in the extracellular matrix of *B. cereus*.
572 Interestingly, CapP adds another level of structural complexity compared to the
573 orthologous pair TasA and CalY because it contains a C-domain region that is not
574 involved in the rigid core but could still mediates CalY polymerization. Moreover, CapP
575 (Figure S5B), similar to TasA and CalY,[32] is disordered in vitro prior to aggregation, a
576 notable difference from the TasA-TapA system of *B. subtilis*, which adopts a folded or
577 partially folded conformation before assembly into fibrils.[30,32,61] These distinctions suggest
578 that this mechanism in *B. cereus* depends on an external trigger, which could be chemical
579 or biophysical, such as the specific action of accessory proteins. This additional
580 requirement has been previously proposed for the *in vivo* polymerization of CalY, and
581 moreover, CalY has been described to enhance the polymerization of TasA without
582 perturbing the global filament architecture.[22,32] Based on these observations and our
583 findings, it is reasonable to propose that the three proteins cooperate in assembling
584 external pili and, consequently, organizing the ECM.

585 The diverse mechanisms that govern the assembly and attachment of non-flagellar
586 proteinaceous filaments to bacterial cell surfaces include various structures, such as
587 chaperone-usher pili, type V pili, type IV pili, curli, fap fibers, conjugative, type IV secretion

588 pili, e-pili in diderm bacteria, and sortase-mediated pili as well as type IV pili in monoderm
589 bacteria.[62,63] In the monoderm *B. cereus*, BcpA-BcpB fimbriae are covalently bound to
590 vegetative cells by sortase, while Ena appendages, exclusively found on spores, are
591 assembled through subunit crosslinking via disulfide bonding.[64–66] Nevertheless, our work
592 describes a new type of fibrillar assembly exclusive to the *B. cereus* group, orchestrated
593 by the chaperone function of CapP, with TasA and CalY serving as principal subunits.
594 The biphasic effect, dependent on the concentration of CapP in CalY polymerization, has
595 been reported for several extracellular chaperones.[67] Under conditions where its levels
596 are lower relative to the substrate, the chaperone acts to stabilize the conformation and
597 promote filament formation. Conversely, at high concentrations, the chaperone forms
598 soluble complexes with the substrate, completely inhibiting its assembly. Finally, it is
599 noteworthy that disordered regions in chaperones are critical for interacting with target
600 proteins.[68–70] Although we have demonstrated the involvement of the C-domain of CapP
601 in CalY polymerization, additional investigation is required to elucidate their precise
602 structural and molecular contacts.

603 Building upon the novel insights unveiled in this study and the current knowledge
604 on pili polymerization, we propose a revised model for the assembly of biofilm-specific pili
605 in *B. cereus* (see Graphical abstract), which involves the paralogous proteins TasA and
606 CalY, as well as the biofilm-specific protein CapP. Once the switch from a motile to sessile
607 lifestyle has occurred, further growth and maturation of the biofilm ECM rely upon several
608 factors, including the products of the operon *sipW-tasA* and the loci *capP* and *calY*, which
609 are expressed under the tight control of the master regulator SinR. TasA, CapP and CalY
610 are processed by the cognate signal peptidase SipW and secreted via the Sec pathway

611 in an unfolded conformation. At 24 hours, the relative levels of CapP compared to CalY
612 and TasA are lower, and associated with the cell surface, CapP interacts with unfolded
613 CalY, catalizing the structural rearrangement necessary to initiate fibrillar polymer growth.
614 Subsequently, CalY facilitates the incorporation of TasA monomers into growing pili.
615 When the biofilm reaches a maturation stage at 48 hours, CapP levels rise, and it interacts
616 with CalY within the ECM to form stable chaperone-substrate complexes that arrest the
617 elongation of TasA-CalY heteropolymers. The ECM scaffold, composed of TasA-CalY
618 filaments, exopolysaccharides, and eDNA, will provide *B. cereus* with the chemical and
619 physical properties necessary for the correct assembly of the microbial community upon
620 adhesion to biotic or abiotic surfaces.

621 Our findings contribute to knowledge on the pivotal role of the biofilm-specific
622 protein CapP in orchestrating the formation and regulation of biofilms on *B. cereus* strains;
623 in addition, the results provide insights into a new type of pilus assembly mechanism in
624 monoderm bacteria. These remarkable advancements carry substantial implications for
625 the clinical and food industries, in which innovative and effective strategies are urgently
626 needed to address the critical challenge of *B. cereus* biofilm formation and its associated
627 risks.

628 **ACKNOWLEDGMENTS**

629 We thank Saray Morales Rojas for the technical support. We thank John Pearson from
630 the Nanoimaging Unit of the Plataforma Bionand, Instituto de Investigación Biomédica de
631 Málaga (IBIMA), for his technical support in confocal microscopy, data analysis and
632 sample preparation. We thank Juan Félix López from the Nanoimaging Unit of the
633 Plataforma Bionand, Instituto de Investigación Biomédica de Málaga (IBIMA), for his
634 technical assistance in transmission electron microscopy imaging and analysis. We also
635 thank Josefa Gómez Maldonado from the Ultrasequencing Unit of the SCBI-UMA for RNA
636 sequencing and Luis Díaz Martínez from the SCBI-UMA for the transcriptomic analysis.
637 The authors thank Carolina Lobo García and Casimiro Cárdenas García from the
638 Proteomic Unit of the SCAI-UMA for protein sequencing and MS analysis. We also thank
639 José Luis Zafra Paredes and Jose María Montenegro Martos from the Vibrational
640 Spectroscopy Unit and the Electronic Spectroscopy Unit, respectively, of the SCAI-UMA
641 for their technical support with the ATR-FITR spectroscopy and dynamic light scattering
642 experiments. We are grateful to Theresa M. Koehler (University of Texas Health Science
643 Center at Houston) for kindly providing the plasmid pUTE657 and to Stephen Leppla
644 (National Institute of Health) for kindly providing us with the rat anti-sigmaA.

645 This work was supported by grants from an ERC Starting Grant (BacBio 637971), the
646 Plan Nacional de I+D+i of the Ministerio de Ciencia e Innovación (PID2019-107724GB-
647 I00 and PID2022-141664NB-I00), and Junta de Andalucía (P20_00479). Ana Álvarez-
648 Mena is the recipient of an FPI contract (BES- 2017-081275). This work benefited from
649 the support of the French National Research Agency ANR (BH, grant No. ANR-23-CE11-

0005-01), the Biophysical and Structural Chemistry Platform at IECB, CNRS UAR 3033, and INSERM US001.

AUTHOR CONTRIBUTIONS

D.R. conceived the study. D.R., A.L. and A.A.M. designed the experiments and method analysis. A.A.M. performed the main experimental work and analyzed the data, with the following exceptions: A.A.M. and M.B.A.S. analyzed the SS-NMR data; M.L.A.G. performed some microbiological experiments; A.A.M., M.B. and B.K. performed biophysical analysis; A.G., E.M., and A.L. performed NMR analysis. JCA contributed to setting the initial hypothesis and performing iTRAQ experiments. A.A.M., M.B.A.S., B.H., A.L. and D.R. prepared the figures. A.A.M., B.H., A.A.L. and D.R. wrote the manuscript. A.V. contributed critically to writing the final version of the manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests

673 **FIGURE TITLES AND LEGENDS**

674 **Figure 1. *bc1280* is an essential locus in the developmental program of biofilm**
675 **formation. A)** Schematic domain composition of BC1280. The signal peptide (M1-A39)
676 is shown in red; the domain of unknown function, DUF4047 (I33-E157), which is located
677 within the N-domain, is highlighted in blue; and the C-domain containing the repetitive
678 region (Q190-E 279) is shown in yellow. The amino acid alignment of DUF4047 among
679 different strains of the *B. cereus* group revealed that it is highly conserved, with the
680 residues that are 100% conserved in all the strains marked in color. The color pattern is
681 determined by the amino acid conservation. **B)** Biomasses formed by the *B. cereus*
682 ATCC14579 wild-type strain and the mutants that adhered to abiotic surfaces were
683 evaluated at different time points. Biofilms were stained with crystal violet to enable
684 visualization and subsequent quantitative comparison. **C)** The amount of biofilm mass for
685 each strain was estimated. Three independent experiments were conducted, each with
686 three replicates. Average values are shown, and the error bars represent the standard
687 deviation for each strain. Multiple comparisons were performed at each time point, with
688 the wild-type serving as the control, and ordinary one-way ANOVA was used (** p value
689 < 0.01 and **** p value < 0.0001). **D)** Auto-aggregation kinetics due to cell-to-cell
690 interactions were studied for the wild-type, *Δbc1280* and *ΔC-domain* strains. The optical
691 density at 600 nm at the air-liquid interface of the cell suspensions was measured each
692 hour. The initial optical density was set to 3 at the start of the experiment, and the samples
693 were incubated without agitation at room temperature. The error bars indicate the SDs
694 calculated from three biological replicates.

695 **Figure 2. The biofilm-defective phenotype of *bc1280* is independent of the**
696 **expression level of *tasA* or *calY.* A)** Relative expression levels of *tasA*, *calY*, *eps2* and
697 genes belonging to the *eps1* operon (*bc5263*, *bc5268*, *bc5277* and *bc5279*) obtained by
698 RT–qPCR in *Δbc1280* planktonic cells compared to wild-type cells at 48 hours. The
699 average values of three biological replicates are shown with error bars indicating the
700 standard deviation. **B)** Relative expression levels of *tasA*, *calY*, *eps1*, and *eps2* were
701 measured in planktonic cells at 72 hours in the *Δbc1280* strain that overexpressed
702 *bc2793-bc2794* or *bc2794* and compared to the wild-type strain. Average values from
703 three biological replicates are presented, with error bars representing the standard
704 deviation. **C)** Transmission electron micrographs of the wild-type and *Δbc1280* strains
705 overexpressing *tasA* or *calY*, immunolabeled with specific anti-TasA or anti-CalY
706 antibodies, respectively. Scale bar: 1 μm **D)** Extracellular complementation of *Δbc1280*
707 after the addition of TasA or CalY at 6 μM in monomeric or fibrillar form. The biofilm mass
708 was quantified by crystal violet staining, followed by absorbance measurement at 575 nm
709 and comparison with the *Δbc1280* strain. Statistical analysis was performed using one-
710 way ANOVA with Dunn's multiple comparisons test. Statistical significance is indicated
711 as $p < 0.05$ (*) and $p < 0.01$ (**). **E)** Phenotype reversion of *Δbc1280* upon the addition of
712 6 μM polymerized TasA or CalY, visualized by immunolabeling with anti-TasA and anti-
713 CalY antibodies, respectively. Scale bar: 1 μm **F)** Negative-stained micrographs of TasA
714 and CalY immunodetection in the *ΔC-domain* strain. Scale bar: 1 μm
715 **Figure 3. BC1280 is a cell wall-associated protein that partially colocalizes with**
716 **CalY in the ECM. A)** and **B)** Left: Immunodetection by Western blotting was performed
717 for BC1280 (A) and the N-domain (B) fused to a histidine tag in different cellular fractions

718 of *Δbc1280*. This assay was conducted under conditions of overexpression using the
719 vector pUTE657 and with the addition of 10 μM IPTG. Biofilm samples were grown for 48
720 hours, after which the resulting biomass was recovered for fractionation into EM
721 (extracellular medium), CW (cell wall), and C (membrane and cytosol) fractions. The
722 immunoreactive bands were detected using an anti-His antibody (1:2500). The
723 immunoblot images were edited by cropping and splicing for better illustration. Right:
724 Negative stained images of BC1280 (A) and the N-domain (B) in the ECM of the *Δbc1280*
725 strain. The proteins were immunolabeled using an anti-His antibody (1:100) as the
726 primary antibody and a goat-antirabbit secondary antibody conjugated with 20 nm
727 nanogolds (1:100). Scale bar: 500 nm **C)** Immunolocalization of BC1280 and the N-
728 domain determined using CLSM. Biofilm samples were processed using
729 immunochemistry, with anti-His antibody (1:100) serving as the primary antibody and a
730 secondary antibody conjugated to Atto 488. The localization of BC1280 and the N-domain
731 is shown in green. The cell wall is labeled in red using WGA, and the DNA is stained in
732 blue using Hoechst. Scale bar: 5 μm **D)** Immunolocalization images of BC1280 and CalY
733 obtained through CLMS. An anti-His antibody (1:100) and a secondary antibody
734 conjugated with Atto 488 (1:400) were used for the immunolabeling of BC1280. A specific
735 primary anti-CalY antibody (1:100) and a secondary antibody labeled with Alexa Fluor
736 647 (1:400) were used for CalY detection. Additionally, DNA was stained with Hoechst
737 (1:1000) to visualize the bacteria, which are shown in blue. Scale bar: 3 μm

738 **Figure 4. Heterologously produced BC1280 oligomerizes into macromolecular**
739 **complexes driven by the N-domain. A)** Scheme illustrating the two constructs based
740 on the full-length sequence of BC1280. Each construct was fused to a poly-histidine tag

741 at the C-terminal domain, and the resulting two recombinant proteins were produced
742 through recombinant expression in *E. coli*. **B)** Size-exclusion chromatogram obtained
743 after BC1280 was purified under native conditions. **C)** and **D)** Size distribution histograms
744 obtained via DLS showing the diameters of the aggregates formed by BC1280 for peaks
745 1 and 2, respectively, as determined by SEC. **E)** Size-exclusion chromatogram obtained
746 after purifying N-domain$_{39\text{-}190}$ under native conditions. **F)** and **G)** Size distribution
747 histograms obtained via DLS, illustrating the diameters of the aggregates formed by N-
748 domain$_{39\text{-}190}$ for peaks 1 and 2, respectively, as determined by SEC. **H)** Native
749 polyacrylamide gel showing the oligomers formed by BC1280 and N-domain$_{39\text{-}190}$
750 following native purification through heterologous expression in *E. coli*. The numbers on
751 each line correspond to the first and second peaks, respectively, that were detected in
752 the size-exclusion chromatograms of each sample. The gel lanes were cropped and
753 spliced for illustrative purposes. All the lines superimposed on the image delineate the
754 lanes and boundaries of the spliced images. **I)** Aggregation kinetics of BC1280 (black
755 curve) and N-domain$_{39\text{-}190}$ (red curve) determined using ThT fluorescence. **J)**
756 Transmission electron micrographs of BC1280 and N-domain$_{39\text{-}190}$ at the end of the
757 polymerization experiment shown in I. Scale bar: 200 nm

758 **Figure 5. The assembled N-domain$_{39\text{-}190}$ contains a rigid core with two β-strand-rich**
759 **regions, while the C-domain increases global rigidity. A)** X-ray diffraction signals of
760 N-domain$_{39\text{-}190}$ at 4.7 and 10 Å, indicating interstrand and intersheet spacing, respectively.
761 **B)** 2D ^{13}C-^{13}C PDSD experiments at 50 ms for the N-domain$_{39\text{-}190}$. The chemical shifts
762 that correspond to isoleucine residues in the β-sheet conformation are highlighted. **C)** A
763 12% SDS–PAGE gel was obtained after the treatment of N-domain$_{39\text{-}190}$ aggregates with

764 proteinase K for 45 minutes. The three bands marked were analyzed by tandem mass
765 spectrometry. **D)** Determination of the rigid core of N-domain$_{39-190}$ by tandem mass
766 spectrometry. MS data were obtained from the bands selected after digestion of N-
767 domain$_{39-190}$ with proteinase K. The amino acids outlined by a box are the rigid regions
768 identified in the absence of proteinase-K, and letters in red correspond to the residues
769 identified by SSNMR PDSD at 50 and 200 ms. **E)** Superimposition of the ^{1}H-^{13}C INEPT
770 spectra of BC1280 (blue) and N-domain$_{39-190}$ (red), recorded at a 1H frequency of 600
771 MHz (276 K and 96 scans for BC1280; 300 K and 640 scans for N-domain$_{39-190}$) showing
772 signals corresponding to the flexible part. **F)** Estimation of the global rigidity based on the
773 INEPT:CP ratio of chemical shifts for each sample. **G)** Secondary structure content of the
774 BC1280 and N-domain$_{39-190}$ aggregates probed via ATR-FTIR analysis.
775
776 **Figure 6. BC1280 acts as a chaperone for CalY, with its functionality modulated by**
777 **the C-domain. A)** Polymerization kinetics of different CalY/BC1280 molar ratios
778 measured by ThT fluorescence emission. Error bars represent the SEM. **B)** Relative
779 levels of BC1280 and CalY in biofilm cells during biofilm formation using iTRAQ. **C)**
780 Kinetics of ThT binding of CalY in the presence of BC1280 or the N-domain$_{39-190}$ at
781 concentrations of 2 and 10 μM. Error bars represent the SEM. **D)** Transmission electron
782 micrographs of negatively stained samples of CalY aggregates in the presence of
783 BC1280 or N-domain$_{39-190}$ at a concentration of 10 μM after 30 hours of incubation at
784 37°C. Scale bar: 200 nm **E)** Circular dichroism (CD) spectra of CalY incubated for 16
785 hours under agitation alone or in the presence of BC1280 at 2 and 10 μM. **F)**
786 Immunolocalization of TasA and CalY in wild-type and mutant strains using CLSM with

787 specific primary antibodies against TasA and CalY, and a secondary antibody conjugated
788 to Atto 488. Green fluorescence represents TasA and CalY localization, with the cell wall
789 stained red (WGA) and DNA stained blue (Hoescht). Scale bar: 5 μm
790 **Graphical abstract.** During the process of biofilm formation in *B. cereus*, TasA, CapP,
791 and CalY serve as the main components, in addition to other factors, ensuring that the
792 extracellular matrix correctly assembles. The three proteins are produced and processed
793 by the signal peptidase SipW and subsequently secreted through the Sec pathway in an
794 unfolded conformation. At early biofilm formation stages, CapP is initially present at low
795 levels, associates with the cell wall and is secreted into the ECM where it interacts with
796 unfolded subunits of CalY, facilitating CalY folding and initiating fibril growth. CalY serves
797 as a nucleator for incorporating TasA subunits into the pilus, forming TasA-CalY
798 heteropolymers. Once the ECM scaffold is established, CapP levels increase, forming
799 stable complexes with CalY that prevent its folding and arrest filament growth.
800

801 **SUPPLEMENTAL INFORMATION TITLES AND LEGENDS**

802 **Figure S1. *bc1280* is exclusively found in the *B. cereus* group and is essential for**
803 **biofilm formation.** Related to Figure 1. **A)** Phylogenetic conservation of BC1280 among
804 *B. cereus* group strains. Phylogenetic analysis was conducted on *B. cereus* group strains
805 using the amino acid sequence of BC1280 (accession number in the KEGG database:
806 BC1280) from *B. cereus* ATCC14579 as the reference. Evolutionary distances were
807 calculated using the neighbor joining method. **B)** Signal peptide prediction by SignalP 5.0.
808 A potential cleavage site was identified for the BC1280 protein between amino acid
809 residues A^{39} and A^{40}. **C)** Amino acid alignment of the C-domain among different strains
810 of the *B. cereus* group, with color indicating conservation level. **D)** The relative expression
811 of *bc1280* in planktonic and biofilm cells of the wild-type strain at 48 hours is presented.
812 The average values from three biological replicates are displayed, with error bars
813 indicating the standard deviation. Statistical analysis was performed using a two-tailed t
814 test, revealing a significant difference between planktonic and biofilm cells (** p
815 value=0.0055). **E)** Top panel: Bands corresponding to the products of amplification using
816 cDNA from RNA extracted from the wild-type strain grown for 48 hours as a template are
817 shown. Bottom panel: To confirm the functionality of the primers, each fragment was
818 amplified via PCR using genomic DNA from the wild-type strain *B. cereus* ATCC14579
819 as a template. **F)** Biofilm phenotypes formed at the air-liquid interphase by the emetic
820 strain *B. cereus* AH187 wild-type and the *tasA*, *calY* and *bc1280* mutant strains. Biofilm
821 formation was evaluated as the formation of a pellicle at the air-liquid interphase after 3
822 days of incubation at 28°C.

823 **Figure S2. The genes *tasA*, *calY*, and an uncharacterized ECF sigma factor are**
824 **deregulated in the *Δbc1280* mutant at 48 hours.** Related to Figure 2**. A)** Biofilm rings
825 formed after 72 hours by overexpressing *bc2793-bc2794* or *bc2794* in *Δbc1280* at various
826 IPTG concentrations, as well as in *Δbc2793-bc2794* or *Δbc2794* strains, with wild-type
827 and *Δbc1280* strains serving as controls. The biofilms were stained with crystal violet, and
828 the biomass was measured by absorbance at 575 nm. Statistical significance was
829 estimated using one-way ANOVA with Dunn's multiple comparisons test ($p < 0.05$ (*), p
830 < 0.01 (**), $p < 0.001$ (***), and $p < 0.0001$ (****). **B)** Biofilm phenotypes of the wild-type
831 and *Δbc1280* mutant strains were assessed after overexpressing *tasA*, *calY*, or both
832 genes using the plasmid pUTE657 with 10 μM IPTG for induction. As a negative control,
833 each strain containing the plasmid was grown without IPTG. Biofilm formation was
834 visualized by crystal violet staining. The biofilm mass was estimated by absorbance at
835 575 nm, and statistical significance was assessed by Dunn's multiple comparisons test.
836 No significant differences were detected between the control strain and upon induction
837 with IPTG. **C)** Evaluation of the specificity of the nanogold particles in immunolabeling
838 studies was conducted using the wild-type strain, with any primary antibody and the goat-
839 antirabbit secondary antibody conjugated with 20 nm nanogolds (1:100). Scale bar: 1 μm
840 **D)** Negative control showing TasA and CalY at 6 μM in monomer and fibrillar forms, grown
841 in TY medium without bacterial inoculation. The wells were stained with crystal violet, and
842 absorbance at 575 nm was measured to assess biofilm formation.
843 **Figure S3. Biofilm formation is reverted upon overexpression of *bc1280* or *N-***
844 ***domain*$_{39\text{-}190}$ in *Δbc1280*, and BC1280 colocalizes with CalY but not with TasA.**
845 Related to Figure 3. **A)** and **B)** Overexpression of *bc1280* or the *N-domain* in the wild-

846 type and *Δbc1280* strain at 10 and 100 μM IPTG. The resulting biofilms were stained with
847 crystal violet and the absorbance at 575 nm was measured. **C)** Estimation of the
848 proportion of fluorescence from co-localized pixels in each channel was performed to
849 localize BC1280 in the cell wall using the Mander's split coefficient and from five different
850 fields of view. **D)** Pearson's coefficient (*** p value = 0.0002 using a two-tailed t-test with
851 Welch's correction) and Mander's split coefficient values obtained from BC1280-CalY
852 CLSM studies. **E)** For colocalization studies of BC1280 and TasA using CLSM, a specific
853 primary anti-TasA antibody (1:100) was utilized. Scale bar: 5 μm. **F)** Pearson's coefficient
854 (*** p value = 0.0002 using a two-tailed t-test with Welch's correction) and Mander's split
855 coefficient values derived from BC1280-CalY CLSM studies.
856
857 **Figure S4. Preliminary studies indicate that BC1280 in vitro has a high tendency to**
858 **oligomerize, and only the N-domain appears to be folded.** Related to Figure 4. **A)**
859 Mass spectrometry results obtained from the bands that corresponded to the elution
860 fractions of BC1280 and N-domain$_{39\text{-}190}$. **B)** Negative-staining micrographs of the first and
861 second peaks detected by SEC for BC1280 and N-domain$_{39\text{-}190}$. **C)** Prediction of the
862 folding state using FoldUnfold software. The amino acid residues are classified as folded
863 (black), folded (blue), or unfolded (red). The underlined amino acids are the regions
864 identified as putative amyloidogenic by at least two different algorithms. **D)** Hydrophobicity
865 prediction of residues in BC1280 using ProtScale software on the ExPASy server. **E)** The
866 PDB model was constructed with the AlphaFold package for BC1280. The color pattern
867 is displayed based on the confidence value.
868

869 **Figure S5**. **The structural rearrangement of the N-domain is conserved in the full-**
870 **length protein.** Related to Figure 5. **A)** Peaks predicted by SPARTA, based on the
871 structure predicted by AlphaFold, are represented by the cross. The red signal indicates
872 the experimental data obtained by SSNMR at a mixing time of 50 ms for the N-domain$_{39-}$
873 $_{190}$. **B)** Solution NMR ^{15}N-^{1}H spectrum of N-domain$_{39-190}$ protein demonstrates limited
874 chemical shift dispersion. **C)** Cross-peak identification for inter-residue interaction.
875 Overlay of the 2D ^{13}C-^{13}C PDSD correlation spectra obtained by SSNMR for N-domain$_{39-}$
876 $_{190}$ at 50 ms (black signals) and 200 ms (red signals) mixing times. The cross-peaks that
877 correspond to chemical shifts between residues are marked. **D)** X-ray diffraction signals
878 of BC1280 within 0–15 Å. **E)** Superimposition of the 2D ^{13}C-^{13}C PDSD correlation spectra
879 obtained by SSNMR for BC1280 (blue) and N-domain$_{39-190}$ (red) at a mixing time of 50
880 ms. **F)** Second derivative analysis of BC1280 (black) and N-domain$_{39-190}$ in the amide I
881 band using ATR-FTIR. **G)** and **H)** ATR-FTIR spectra of BC1280 and N-domain$_{39-190}$,
882 respectively, in the amide I band. The contribution of each peak was determined through
883 secondary derivative analysis of the raw data, followed by a deconvolution approach.
884
885 **Figure S6. BC1280 levels have an impact on CalY polymerization.** Related to Figure
886 6. **A)** Estimation of the relative levels of TasA, BC1280, and CalY in biofilm cells over 72
887 hours using iTRAQ. **B)** and **C)** Study of the effect of BC1280 and the N-domain39-190
888 without recent purification on CalY polymerization. ThT binding kinetics of CalY in the
889 presence of freshly and non-freshly purified BC1280 (B) or N-domain39-190 (C) samples.
890 Error bars indicate the standard error of the mean. **D)** Circular dichroism (CD) spectra

891 obtained for each sample without subtracting the contribution of BC1280 at 2 and 10 μM

892 when incubated with CalY for 16 hours under agitation.

893

894 **STAR METHODS**

895 **RESOURCE AVAILABILITY**

896 **Lead contact**

897 Further information and requests for resources and reagents should be directed to and
898 will be fulfilled by the lead contact, Diego Romero (diego_romero@uma.es).

899 **Materials availability**

900 Plasmids generated in this study are available at the BacBio Lab. Please contact Diego
901 Romero (diego_romero@uma.es) for access. This study did not generate new unique
902 reagents.

903 **Data and code availability**

904 RNaseq data have been deposited at GEO database and are publicly available as of the
905 date of publication. Accession numbers are listed in the Key Resources Table. The mass
906 spectrometry proteomics data have been deposited to the ProteomeXchange Consortium
907 via the PRIDE partner repository with the dataset identifier PXD010211. Microscopy data
908 reported in this paper will be shared by the lead contact upon request.

909 This paper does not report original code.

910 Any additional information required to reanalyze the data reported in this paper is
911 available from the lead contact upon request.

912 **EXPERIMENTAL MODEL AND STUDENT PARTICIPANT DETAILS**

913 **Bacterial strains and culture conditions**

914 The bacterial strains used in this study are listed in the Key Resources Table. For
915 experiments conducted using *B. cereus* wild-type or mutant strains, cultures were
916 routinely grown at 28°C from frozen stocks on Luria–Bertani (LB) plates containing 1%

917 tryptone (Oxoid), 0.5% yeast extract (Oxoid), 0.5% NaCl and 1.5% agar before the
918 respective experiments were performed. Biofilm assays were performed in TY medium
919 supplemented with 1% tryptone (Oxoid), 0.5% yeast extract (Oxoid), 0.5% NaCl, 10 mM
920 $MgSO_4$ and 1 mM $MnSO_4$.

921 For cloning and plasmid replication, the *E. coli* DH5α strain was used. For
922 recombinant protein expression and purification, *E. coli* BL21(DE3) pLysS (Merck,
923 Kenilworth, NJ, USA) was used. All the strains were cultured in Luria–Bertani (LB) liquid
924 media. For the agar-solid plates, LB media was supplemented with 1.5% bacteriological
925 agar (Oxoid). The final antibiotic concentrations were 100 μg/ml ampicillin, 50 μg/ml
926 kanamycin, 100 μg/ml spectinomycin, and 5 μg/ml erythromycin.

927 For the purification of $^{13}C/^{15}N$-labeled proteins, the cultures were grown in M9
928 minimal medium (the final concentration was 48 mM Na_2HPO_4, 22 mM KH_2PO_4, 8.6 mM
929 NaCl, 1 mM $MgSO_4$, 0.01 mM $ZnCl_2$, 0.001 mM $FeCl_3$, 0.1 mM $CaCl_2$ and 10 ml 100X
930 minimum essential medium vitamin solution) supplemented with 1 g/L $^{15}NH_4Cl$ as nitrogen
931 and 2 g/L $^{13}C_6$-D-glucose or 2-^{13}C glycerol as the carbon source.

932 **METHOD DETAILS**

933 **Plasmid and strain construction**

934 All the primers used to construct the strains used in this study are listed in Table S4.

935 *B. cereus* mutant strains were constructed via homologous recombination using
936 the upstream and downstream regions of the gene of interest, which were subsequently
937 cloned and inserted into the pMAD plasmid.[71] The constructs were designed using NEB
938 Builder HiFi DNA Assembly Master Mix (New England Biolabs, MA, USA) following the
939 manufacturer's instructions and utilizing specific primers. The pMAD vector was linearized

940 using the restriction enzyme SmaI (FastDigest, Thermo Fisher Scientific), and the
941 digested vector was incubated at 50°C for 1 hour along with the respective upstream and
942 downstream fragments. The combined total amount of fragments and plasmid was 0.2
943 pmol, and a proportion of 1:2 (vector: fragments) was maintained. The resulting reaction
944 mixture was subsequently transformed into *E. coli* DH5α, and positive colonies were
945 selected using ampicillin. Subsequently, the plasmids were purified, subjected to PCR
946 amplification and digestion for verification, and sequenced.

947 For the overexpression experiments in *B. cereus,* the gene of interest was cloned
948 and inserted into the plasmid pUTE657[72] under a promoter that was inducible by IPTG
949 using the primers listed in Table S4. Following gene amplification and the inclusion of a
950 specific RBS for *Bacillus*, the resulting PCR product was digested with SalI and SphI
951 (FastDigest, Thermo Fisher Scientific) and subsequently cloned and inserted into the
952 pUTE657 plasmid that was cut with the same restriction enzymes using T4 DNA ligase
953 (Thermo Fisher Scientific). Subsequently, the resulting reaction was transformed into *E.*
954 *coli* DH5α, and colonies were selected with 100 μg/ml ampicillin. After plasmid
955 purification, all the vectors were subjected to PCR amplification, digestion with restriction
956 enzymes, and sequencing.

957 The vectors were transformed into *B. cereus* by electroporation following
958 previously described methods with some modifications. A single colony from a pure
959 culture of *B. cereus* was inoculated in 5 ml of LB medium and incubated under shaking
960 conditions at 30°C for 16 hours. This culture was subsequently used to inoculate a flask
961 containing 100 ml of LB, which was incubated at 30°C with agitation until an OD_{600nm} of
962 0.3 was reached. The cell culture mixture was centrifuged (3000 rpm, 5 minutes, 4°C),

963 and the resulting pellet was resuspended in 12 ml of electroporation buffer (0.5 mM
964 K_2HPO_4, 0.5 mM KH_2PO_4, 0.5 mM $MgCl_2$, and 272 mM saccharose). The suspension was
965 centrifuged again and then resuspended in 250 µl of electroporation buffer. The
966 electrocompetent cells were mixed with 0.5–1 µg of plasmid and incubated on ice for 10
967 minutes. Next, the mixture was electroporated in a 0.2 cm cuvette using a voltage of 1.4
968 kV, a capacitance of 25 µF, and a resistance of 200 Ω. After electroporation, the
969 suspension was recovered and incubated with 1.5 ml of LB for 5 hours at 30°C with
970 agitation. The resulting culture was plated onto LB plates supplemented with the
971 corresponding antibiotic to select the colonies that had been successfully transformed
972 with the vector. For integration of the deletion, *B. cereus* transformants were incubated
973 without antibiotics at 40°C, facilitating the integration of the mutation into the genome
974 through homologous recombination. Finally, the culture was plated on LB agar plates,
975 and the colonies were picked onto LB agar plates supplemented with erythromycin.
976 Colonies that did not grow on LB supplemented with the antibiotic were selected, and the
977 mutation was confirmed through PCR amplification.

978 For the heterologous expression of BC1280, the N-domain$_{39\text{-}190}$ and the C-
979 domain$_{191\text{-}279}$ in *E. coli,* the open reading frames containing the *bc1280* gene, the region
980 from amino acids 39 to 190. The PCR products and the pET24b vector were digested
981 with NdeI and XhoI (FastDigest, Thermo Fisher Scientific) and subsequently cloned using
982 T4 DNA ligase (Thermo Fisher Scientific) following the instructions supplied by the
983 manufacturer. The reaction mixtures were subsequently transformed into *E. coli* DH5α,
984 after which the colonies were successfully transformed with 50 µg/ml kanamycin.
985

986 **Phylogenetic tree**

987 The amino acid sequence of BC1280 was obtained from the KEGG database using *B.*
988 *cereus* ATCC14579 as a reference. BLASTP analysis was performed, and the top 50 hits
989 were aligned using Clustal Omega.[73] The phylogenetic tree was then constructed using
990 neighbor joining distances.

991

992 **RT–qPCR**

993 Quantitative reverse transcription PCR was used to estimate the transcription levels of i)
994 *bc1280* in planktonic and biofilm cells of the wild-type strain; ii) *tasA*, *calY*, *eps1* and *eps2*
995 in *Δb1280* compared to those in the wild-type; and iii) *tasA*, *calY*, *eps1* and *eps2* in
996 *Δb1280* overexpressing *bc2794* or *bc2793-2794* compared to those in the wild-type. RT–
997 qPCR was performed using an iCycler-iQ system (Bio-Rad) and Power SYBR Green
998 Master Mix (Thermo Fisher Scientific). Primer3 software (https://primer3.ut.ee/) was used
999 for primer design, maintaining the default parameters.[74] First, 1 µg of total RNA was
1000 reverse transcribed into cDNA using SuperScript III reverse transcriptase (Invitrogen)
1001 following instructions provided by the manufacturer. The reactions were performed in
1002 triplicate in 96-well plates with a total volume of 20 µl. The RT–qPCR cycle started with
1003 an initial denaturation at 95°C for 3 minutes, followed by 40 amplification cycles (95°C for
1004 20 seconds, 56°C for 20 seconds, and 72°C for 30 seconds), and the final step was 95°C
1005 for 30 seconds. The target genes, including *bc1280*, which encodes the protein BC1280;
1006 *bc1279*, which encodes TasA (also known as spore coat-associated protein N); *bc1281*,
1007 which encodes CalY (annotated as cell envelope-bound metalloprotease, camelysin);
1008 *bc5263*, which encodes UDP-glucose 4-epimerase; *bc5268*, which encodes secreted

1009 polysaccharide polymerase; *bc5277*, which encodes tyrosine-protein kinase; *bc5279*,
1010 which encodes tyrosine-protein kinase; and *bc1583*, which encodes O-acetyl transferase,
1011 were amplified using the primer pairs provided in Table S5. Primer efficiency and
1012 amplification product specificity were assessed as previously described.[75] The estimation
1013 of relative expression levels was calculated using the ΔΔCt threshold (Ct) method.[76] The
1014 housekeeping gene *rpoA* was used as a reference for data normalization. The relative
1015 expression value was calculated as the difference between the qPCR threshold cycle (Ct)
1016 of the gene of interest and the Ct obtained for *rpoA* ($\Delta Ct = Ct_{target\ gene} - Ct_{rpoA}$). The results
1017 obtained for the genes of interest were normalized to the values obtained for the wild-
1018 type strain. The RT–qPCR analyses were conducted three times, each with three
1019 independent biological replicates.
1020

1021 **Determination of the transcriptional unit by PCR**

1022 To determine whether *bc1280* constitutes an operon with *tasA* or *calY*, primers were
1023 designed between the genes and inside each gene to amplify four specific fragments that
1024 were 974 bp, 572 bp, 1668 bp, and 560 bp in size (primers listed in Table S6). cDNA was
1025 used as the template, and PCR was performed using genomic DNA purified with the
1026 commercial JetFlex™ Genomic DNA Purification Kit (Thermo Fisher Scientific, Bremen,
1027 Germany) as a positive control. The fragments were amplified by PCR using Phusion™
1028 High-Fidelity DNA Polymerase (Thermo Fisher Scientific) with the following program:
1029 initial denaturation at 98°C for 30 seconds, followed by a 30-cycle amplification program
1030 (98°C for 10 seconds, 56°C for 30 seconds, and 72°C for 1 minute), and a final extension

1031 step at 72°C for 5 minutes. The samples were loaded onto an agarose gel and compared
1032 to the fragments amplified from genomic DNA.
1033
1034 **Biofilm formation and extracellular complementation assays**
1035 To study the biofilm formation phenotype for each strain or condition, bacterial cultures
1036 were grown in 24-well plates as previously described.[25] The strains were subsequently
1037 grown on LB agar plates at 28°C for 24 hours. The bacterial mass was resuspended in 1
1038 ml of TY medium, and the OD_{600} was adjusted to 1. Subsequently, 10 µl of the suspension
1039 was inoculated into each well containing 1 ml of TY medium. The plates were then
1040 incubated at 28°C without agitation for 72 hours.
1041 For the extracellular complementation assay involving the addition of TasA or CalY
1042 as monomers or in their fibrillated states, the proteins were previously purified *in vitro* from
1043 inclusion bodies through heterologous expression in *E. coli* following a protocol
1044 established by El-Mammeri et al., in 2019.[32] Subsequently, the proteins were dialyzed
1045 against a 20 mM Tris and 50 mM NaCl solution at pH 7.4, after which the protein
1046 concentrations were measured. The assembly of filaments was promoted at 37°C under
1047 agitation at 200 rpm for one week. TasA or CalY protein, in both its monomeric and
1048 fibrillated forms, was added at a final concentration of 6 µM to 24-well plates inoculated
1049 with the *Δbc1280* strain. These plates were then incubated without agitation at 28°C for
1050 72 hours, after which the resulting biomass was stained with crystal violet.
1051
1052 **Crystal violet adhesion assay**

1053 Bacterial adhesion to the well plate surface was quantitatively measured using crystal
1054 violet staining, as previously described.[77] The biofilms were grown for 72 hours, after
1055 which the medium was removed. Then, 1 ml of 1% crystal violet was added to each well
1056 of the 24-well plate, and the cells were incubated for 15 minutes. The plates were washed
1057 multiple times with water and left to dry for 45 minutes. Next, the biomass that was formed
1058 in each well was resuspended in 1 ml of 50% acetic acid, and the absorbance at 575 nm
1059 was measured for each sample using a plate reader (Tecan Infinite M1000 Pro; Tecan,
1060 Männedorf, Switzerland). The statistical significance was evaluated using one-way
1061 ANOVA with Dunnett´s multiple comparisons test. P values less than 0.05 were
1062 considered to indicate statistical significance.
1063
1064 **Auto-aggregation assay**
1065 The effect of *bc1280* and *C-domain* deletion on the aggregation rate of *B. cereus* was
1066 estimated using a kinetic method following a previously described protocol.[24] The strains
1067 were inoculated into a flask containing 20 ml of TY medium and incubated overnight at
1068 28°C with agitation. The OD_{600} was adjusted to 3 for a final volume of 10 ml, after which
1069 the supernatant was centrifuged (5000 rpm, 10 minutes, 20°C) for dilution. The samples
1070 were then incubated in vertical tubes under static conditions at room temperature for 24
1071 hours, after which the OD_{600} was measured every hour at the air–liquid interface.
1072
1073 **RNA extraction**
1074 For RNA extraction, a previously described protocol was followed with several
1075 modifications.[26] Bacterial strains were cultivated in 24-well plates to retrieve the desired

1076 population—either planktonic or biofilm cells—at a specific incubation time. For the biofilm
1077 cells, the supernatant was removed from each well, followed by three wash steps, and
1078 the biomass was subsequently resuspended in 1 ml of PBS. To minimize variability
1079 between replicates, each sample was prepared by extracting 500 µl from 8 distinct wells.
1080 The bacterial suspensions were then centrifuged at 12000 × g for 5 minutes, and the
1081 resulting pellets were frozen at -80°C for 30 minutes or until needed. The pellets were
1082 resuspended in 900 µl of TRI-Reagent (Merck), and the cells were disrupted with 0.1 mm
1083 beads using a TissueLyser (QIAGEN) for 3 minutes, followed by incubation at 55°C for 3
1084 minutes. Then, 200 µl of chloroform was added to each tube, which was vortexed for 10
1085 seconds and incubated at room temperature for 3 minutes. Finally, the samples were
1086 centrifuged at 12000 × g for 10 minutes at 4°C. The resulting aqueous phase was
1087 transferred to a new tube containing 500 µl of ice-cold isopropyl alcohol. The tubes were
1088 inverted a few times, incubated for 10 minutes at room temperature, and then centrifuged
1089 at 12000 × g for 10 minutes at 4°C. The supernatants were subsequently removed, and
1090 the pellets were washed with 1 ml of ice-cold 75% ethanol, followed by centrifugation
1091 (12000 × g, 5 minutes, 4°C). Briefly, the RNA pellets were dried and then resuspended in
1092 50 µl of DEPC-treated water. The residual DNA was eliminated through treatment with
1093 rDNAse, which is included in the Nucleo-Spin RNA Plant Kit (Macherey-Nagel), following
1094 the instructions provided by the manufacturer. The quality and integrity of the total RNA
1095 were assessed using an Agilent 2100 bioanalyzer (Agilent Technologies) and gel
1096 electrophoresis.
1097
1098 **Whole-transcriptome analysis**

1099 A RNA-Seq analysis was performed on *Δbc1280* planktonic cells in comparison to the
1100 wild-type after 24 and 48 hours of incubation at 28°C. RNA extraction was performed
1101 following a previously described protocol, and each sample corresponding to a specific
1102 condition was collected in duplicate. RNA sequencing was performed by the Omics Unit
1103 of Supercomputing and Bioinnovation Center (SCBI, University of Malaga, Spain). rRNA
1104 removal was achieved using the RiboZero rRNA Depletion Kit (Illumina, CA, USA), and
1105 subsequently, 100-bp single-end read libraries were prepared using the TruSeq Stranded
1106 Total RNA Kit (Illumina). Next, the libraries were sequenced using the NextSeq550
1107 instrument (Illumina). To eliminate regions of low quality, ambiguity, and low complexity,
1108 as well as potential contamination, the raw reads were preprocessed using SeqTrimNext
1109 with the specific configuration parameters used for NGS technology.[78] Subsequently, the
1110 clean reads were aligned and annotated using the *B. cereus* ATCC14579 genome
1111 (NC_004722.1) as the reference. This alignment was conducted with Bowtie2, resulting
1112 in BAM files that were subsequently sorted and indexed using the SAMtools v1.484
1113 program.[79,80] To calculate the read count for each gene, the script Sam2counts
1114 (https://github.com/vsbuffalo/sam2counts) was used to determine the number of uniquely
1115 mapped reads. The analysis of differentially expressed genes (DEGs) between *Δbc1280*
1116 and the wild-type strain was conducted using the R script DEgenes Hunter. This script
1117 employs a combined p value, utilizing Fisher´s method to consider the nominal p values
1118 obtained from edgeR and DEseq2.[81,82] The combined p value was adjusted using the
1119 Benjamini–Hochberg (BH) test, which employs the false discovery rate approach. The
1120 adjusted p value was subsequently used to rank all the DEGs obtained. A p value < 0.05
1121 and a log2FoldChange < -1 or > 1 were considered to indicate statistical significance.

1122 **Transmission electron microscopy and immunochemistry studies**

1123 To investigate the location of the proteins within the ECM using transmission electron
1124 microscopy, we followed a protocol that was previously established and described, with
1125 several modifications.[25,29,32] For TasA and CalY immunodetection, the strains were
1126 cultured in multi-well plates at 28°C for 72 hours. Alternatively, for the localization of
1127 BC1280 or the N-domain within *B. cereus* cells, the strains corresponding to *Δbc1280*
1128 containing the plasmids pUTE657-*bc1280-6xHis* or pUTE657-*N-domain-6xHis* were
1129 grown in multiwell plates with TY medium supplemented with a 10 μM IPTG solution. The
1130 biofilm mass was subsequently harvested in 1 ml of PBS after 48 hours.

1131 Then, 20 μl of each sample was applied to a copper grid and incubated for 2 hours.
1132 The grids were subsequently washed with PBS and incubated for 5 minutes to remove
1133 any excess bacterial cells. The cells that adhered to the grid were fixed using a 2%
1134 paraformaldehyde solution diluted in PBS and incubated for 10 minutes. The grids were
1135 washed again with PBS for 5 minutes and then blocked for 30 minutes using Pierce™
1136 Protein-Free T20 (TBS) Blocking buffer (Thermo Scientific). Next, the samples were
1137 incubated for 2 hours with the primary antibody at a concentration of 1:100, which was
1138 diluted in blocking buffer. Specific primary antibodies against TasA and CalY were used
1139 to detect TasA and CalY, respectively. For the immunolocation of BC1280-His and N-
1140 domain-His, we employed the primary antibody anti-6-His produced in rabbits (Merck).
1141 Subsequently, the grids were washed twice with TBS-T, each time with a 5-minute
1142 incubation. The samples were incubated for 1 hour with a goat anti-rabbit 20-nm gold-
1143 conjugated secondary antibody (Ted Pella, Redding, CA, USA) at a dilution of 1:100 in
1144 blocking buffer. Next, the grids were subjected to two additional washes with TBS-T, each

1145 lasting 5 minutes. The samples were fixed with 2.5% glutaraldehyde for 10 minutes. The
1146 samples were washed with Milli-Q water for 5 minutes and subsequently negatively
1147 stained with a 1% uranyl acetate solution. Finally, the grids were rinsed with a single drop
1148 of water and dried under dark conditions.

1149 As a negative control to check the specificity of the 20-nm gold secondary
1150 antibody, we performed the same protocol, but instead of adding the primary antibody,
1151 we added blocking buffer at that step.

1152 The samples were examined using a Thermo Fisher Scientific Tecnai G^2 20 TWIN
1153 transmission electron microscope at an accelerating voltage of 120 kV. The final images
1154 were taken using an Olympus Veleta side-mounted CCD with a resolution of 2k × 2k Mpx.
1155

1156 **Cellular fractionation**

1157 For protein immunodetection by Western blotting, the samples were fractionated into
1158 three components: the extracellular medium, cell wall-associated proteins, and cellular
1159 contents (membrane and cytosol). Preliminary studies of BC1280 and the N-domain were
1160 conducted through cellular fractionation using the *Δbc1280* strain harboring the plasmid
1161 pUTE657-*bc1280*-6xHis or pUTE657-N-*domain*-6xHis, respectively. The strains were
1162 subsequently grown in TY media supplemented with agitation at 28°C. When the cultures
1163 reached an OD_{600nm} of 0.4, 10 μM IPTG was added to each sample, and the cultures
1164 were incubated for 4 hours.

1165 All the samples were subjected to the following protocol. The samples were
1166 centrifuged at 6000 × g for 5 minutes, after which the supernatant was passed through a
1167 0.45 μm PES filter and retained as the extracellular medium fraction. The pellets were

1168 resuspended in 10 ml of PBS containing 100 μg/ml lysozyme and incubated for 2 hours
1169 at 37°C. Subsequently, the samples were centrifuged at 9000 × g for 20 min at 4°C, after
1170 which the resulting supernatant and the pellet were retained as the cell wall and cellular
1171 fractions, respectively. The cellular fraction was then resuspended in a small volume of
1172 PBS. Finally, all the samples were precipitated with trichloroacetic acid (TCA) to a final
1173 concentration of 10%. The mixture was then incubated on ice for 1 hour. Next, the
1174 precipitated proteins were centrifuged at 13000 × g for 20 minutes at 4°C, and the
1175 resulting pellets were washed twice with 1 ml of acetone. Finally, the samples were dried
1176 at 37°C for 5 minutes, resuspended in 1x Laemmli buffer (Bio-Rad), and loaded onto a
1177 12% SDS–PAGE gel for the immunodetection assay.
1178
1179 **Polyacrylamide gels and western blotting**
1180 The samples were diluted in 1x Laemmli buffer (Bio-Rad) and then heated at 100 °C for
1181 5 minutes. The proteins were separated by 12% SDS–PAGE with the Spectra™
1182 molecular weight marker (Thermo Fisher Scientific) and subsequently transferred to a
1183 polyvinylidene fluoride (PVDF) membrane (Bio-Rad) using the Trans-Blot Turbo Transfer
1184 System (Bio-Rad) at 25 V for 30 minutes. Next, the membrane was blocked for 1 hour
1185 using 5% nonfat milk diluted in 50 mM Tris-HCl (150 mM NaCl, pH 7.5) containing 0.1%
1186 Tween-20 (TBS-T). The membrane was then incubated with the primary antibody in a
1187 solution of 3% nonfat milk in TBS-T. The membrane was washed three times with TBS-
1188 T, with 10 minutes of incubation between each wash. Next, the membrane was incubated
1189 with the secondary antibody against rabbit IgG conjugated to horseradish peroxidase
1190 (Bio-Rad) for 2 hours at a concentration of 1:3000 and diluted in TBS-T. Following the

1191 incubation, the membrane was washed twice with TBS-T and once with TBS. Finally, for
1192 immunodetection, the membranes were exposed to Pierce ECL Western blotting
1193 Substrate (Thermo Fisher Scientific).

1194 For the immunodetection of BC1280-His or N-domain-His, we used an anti-6-His
1195 (Merck) primary antibody produced in rabbits at a dilution of 1:2500.

1196 To study the molecular weight of the oligomers formed by BC1280 and N-domain$_{39-}$
1197 $_{190}$ under native conditions, the proteins were diluted in native sample buffer (Bio-Rad),
1198 loaded onto polyacrylamide gels (Any kD™ Mini-PROTEAN® TGX™ Precast Protein
1199 Gels, Bio-Rad), and analyzed using the NativeMark™ (Invitrogen) marker as a reference.
1200 The native gels were run in buffer containing 25 mM Tris and 192 mM glycine at a
1201 constant voltage of 200 V. To visualize the proteins, the gels were stained with Coomassie
1202 brilliant blue.

1203

1204 **Immunolocalization studies by fluorescence microscopy**

1205 For the immunofluorescence experiments, the samples were grown as previously
1206 described for transmission electron microscopy studies. Then, 150 µl of each sample was
1207 added to well slides coated with 0.1% poly-L-lysine (Sigma–Aldrich) and incubated for 2
1208 hours to facilitate bacterial adhesion to the substrate. The sample was then removed, and
1209 the bacteria that had attached to the slide surface were fixed with fixation buffer (3%
1210 paraformaldehyde and 0.1% glutaraldehyde diluted in PBS) for 10 minutes. Briefly, the
1211 wells were washed twice with PBS, and the samples were incubated for 1 hour with
1212 blocking buffer (3% m/v bovine serum albumin (BSA) and 0.2% v/v Triton X-100 in PBS).
1213 Then, the buffer was removed, and the wells were incubated for 3 hours with the primary

1214 antibody at a concentration of 1:100 diluted in blocking buffer. The primary antibodies
1215 used for visualization by TEM were also used for this experiment. The wells were then
1216 rinsed three times with washing buffer (0.2% m/v BSA and 0.05% v/v Triton X-100 in
1217 PBS), and each incubation lasted for 5 minutes. Next, the wells were incubated for 2
1218 hours with the secondary antibody goat anti-rabbit IgG-Atto488 (manually labeled) at a
1219 dilution of 1:400 in blocking buffer. The samples were washed once with washing buffer
1220 and twice with PBS, for an incubation period of 5 minutes each. The immunostainings
1221 were fixed for 5 minutes using fixation buffer, and the wells were subsequently rinsed with
1222 PBS three times. To detect the cell wall, the slide was incubated with wheat germ
1223 agglutinin labeled with Alexa Fluor 647 (WGA, Thermo Fisher Scientific) for 20 minutes
1224 at a dilution of 1:100 in PBS. Finally, the bacterial DNA was stained for 20 minutes using
1225 Hoechst at a dilution of 1:1000. As a negative control, immunostaining was performed
1226 without incubation with the primary antibody.

1227 The immunostaining was visualized using confocal laser scanning microscopy. For
1228 the fluorescence corresponding to Atto-488 visualization, we employed an excitation
1229 wavelength of 488 nm and detected the emission between 497 and 572 nm. The signal
1230 corresponding to Alexa Fluor 647 was visualized with an excitation wavelength of 561
1231 nm, and emission was detected between 576 and 686 nm. The Hoescht signal was
1232 visualized independently using a specific dichroic filter.

1233 For colocalization studies of BC1280-CalY/TasA, the same immunostaining
1234 protocol was followed with a few modifications. First, BC1280-6xHis was labeled using
1235 the anti-6-His antibody produced in rabbits (Merck), followed by goat anti-rabbit antibody
1236 conjugated to Atto-488 as the primary and secondary antibody, respectively. After three

1237 washes were performed with washing buffer, the samples were treated with quenching
1238 solution (0.1 M glycine in PBS) for 10 minutes. Next, TasA or CalY were detected using
1239 the following specific primary antibodies: anti-TasA and anti-CalY, which were produced
1240 in rabbit, followed by anti-rabbit conjugated to Alexa-Fluor 647 produced in goat as the
1241 secondary antibody. Following the final steps, the samples were fixed, stained with
1242 Hoechst, and finally visualized via CLSM, as described previously.

1243 The colocalization images were analyzed using the software ImageJ and the
1244 colocalization threshold plugin. The region of interest (ROI) was defined for the region
1245 corresponding to Alexa Fluor 647, and both Pearson´s and Mander´s coefficients were
1246 calculated.[83] As a negative control, the band corresponding to Alexa Fluor 647
1247 overlapped with itself, resulting in 100% colocalization. For each sample, a minimum of 5
1248 different images were analyzed, and statistical significance was determined for Pearson´s
1249 coefficient using an unpaired t test with Welch´s correlation, with a p value < 0.005.

1250

1251 **Protein expression and purification**

1252 BC1280 and N-domain$_{39\text{-}190}$ were purified under native conditions. For the heterologous
1253 expression of BC1280, the plasmid pET24b-*BC1280* was transformed into *E. coli* BL21
1254 (DE3) (Merck, Kenilworth, NJ, USA), and colonies were selected with 50 µg/ml
1255 kanamycin. The plasmid pET24-*N-domain$_{39\text{-}190}$* was subsequently transformed into *E. coli*
1256 Lemo21 (DE3)pLyss (New England Biolabs, USA) for N-domain$_{39\text{-}190}$ expression, after
1257 which colonies were selected with 50 µg/ml kanamycin and 5 µg/ml chloramphenicol. The
1258 following steps were performed for both proteins. A single fresh colony was picked and
1259 cultured in 10 ml of LB supplemented with the corresponding antibiotics mentioned

1260 previously. The preculture was incubated under shaking conditions for 6 hours at 37°C.
1261 A total of 10% (v/v) of the precultures were inoculated into 1 L of LB supplemented with
1262 the corresponding antibiotics. The culture was incubated at 37°C with agitation until the
1263 optical density reached 0.6. For BC1280, the culture was induced with 400 μM IPTG and
1264 incubated for 16 hours at 28°C. Otherwise, the culture corresponding to the
1265 overexpression of N-domain$_{39-190}$ was induced with 400 μM IPTG and 100 μM rhamnose
1266 and incubated for 3 hours at 28°C. Then, the cells were harvested (6000 × g, 30 minutes
1267 at 4°C, JLA 8.1 rotor, Beckman Coulter, Brea, CA, USA), and the pellets were frozen at -
1268 80°C until use.

1269 The pellets were resuspended in 18 ml of buffer A (20 mM Na_2PO_4, 500 mM NaCl
1270 and 20 mM imidazole, pH 8) with cOmplete™ EDTA-free Protease Inhibitor Cocktail
1271 (Roche) and CelLytic™ B-Cell Lysis reagent (Merck). The lysates were incubated for 30
1272 min with agitation at room temperature. The samples were sonicated using a Branson
1273 450 digital sonifier on ice with 3 pulses for 45 seconds at 40% amplitude and then
1274 centrifuged at 15000 × g for 60 minutes at 4°C (F34-6-38 rotor; Eppendorf, Hamburg,
1275 Germany). The supernatant was loaded onto a HisTrap HP 5 ml column (GE Healthcare)
1276 and purified using an AKTA Start FPLC system (GE Healthcare). The column was
1277 preequilibrated with buffer A, and after sample loading, it was washed with the same
1278 buffer. Next, the proteins were eluted using a linear gradient of buffer B (20 mM Na_2PO_4,
1279 500 mM NaCl and 500 mM imidazole, pH 8). Since the elution mixture was not completely
1280 pure, the sample was loaded onto a HiLoad 16/600 Superdex 75 pg column (Cytiva), and
1281 the pure protein was subsequently eluted in buffer containing 20 mM Tris and 50 mM
1282 NaCl, pH 7.5. The elution fractions were loaded onto a 12% Tris-Tricine SDS–PAGE gel,

1283 and the samples containing the purified protein were concentrated using Amicon Ultra15
1284 centrifugal filter units (Millipore) with a 3 kDa cutoff. Subsequently, the purified fractions
1285 were separated via 12% SDS–PAGE to assess their purity, which was determined
1286 through Coomassie blue staining. Finally, the gel bands were analyzed using tandem
1287 mass spectrometry.

1288 CalY and TasA were purified via heterologous expression in *E. coli* BL21(DE3)
1289 pLyss and from inclusion bodies under denaturing conditions following the protocol
1290 established by El Mammeri et al. [32]

1291

1292 **Tandem mass spectrometry analysis of protein bands**

1293 The SDS–PAGE gel bands that corresponded to the protein purification and proteinase
1294 K digestion experiments were subjected to analysis via tandem mass spectrometry using
1295 a nanoion trap system [HPLC–electrospray ionization–tandem mass spectrometry (MS)].
1296 The bands were cut and destained using a mixture of 50% acetonitrile (ACN) and 25 mM
1297 ammonium bicarbonate. Then, the samples were dehydrated and dried using ACN.
1298 Disulfide bridges were reduced with 10 mM dithiothreitol (DTT) diluted in 50 mM
1299 ammonium bicarbonate, followed by incubation at 56°C for 30 minutes. Subsequently,
1300 the excess DTT was removed, and the cysteine residues were carbamidomethylated
1301 using 55 mM iodoacetamide diluted in 50 mM ammonium bicarbonate for 20 minutes at
1302 room temperature in the dark. The gel bands were then dehydrated again, and the
1303 proteins were digested using 10 ng/µl trypsin (Promega) at 30°C overnight. To extract the
1304 peptides from the gel, the samples were incubated with a solution of 0.1% ACN/formic
1305 acid (FA) for 30 minutes at room temperature. To eliminate ACN and ammonium

1306 bicarbonate, the samples were dried using a SpeedVacTM. Next, the samples were
1307 resuspended in a solution containing 0.1% FA, ultrasonicated for 3 minutes, and
1308 centrifuged at 13,000 × g for 5 minutes. Subsequently, the samples were purified and
1309 concentrated using C18 ZipTip® (Merck) following the instructions supplied by the
1310 manufacturer. Finally, the samples were injected into an Easy nLC 1200 UHPLC system,
1311 which was coupled to a Q Exactive HF-X Hybrid Quadrupole-Orbitrap mass spectrometer
1312 (Thermo Fisher Scientific). The software versions used for acquisition and analysis were
1313 Tune 2.9 and Xcalibur 4.1.31.9. The mobile phases used in the HPLC consisted of i)
1314 buffer A, which contained 0.1% FA dissolved in water, and ii) buffer B, which contained
1315 0.1% FA dissolved in 80% ACN. Peptides were loaded onto a precolumn (Acclaim
1316 PepMap 100, 75 µm × 2 cm; C18, 3 µm; 100 A; Thermo Fisher Scientific) at a flow rate
1317 of 20 µl/min, while elution was performed using a 50 cm analytical column (PepMap RSLC
1318 C18, 2 µm; 100 A, 75 µm × 50 cm; Thermo Fisher Scientific). Elution was carried out
1319 through a gradient concentration over 60 minutes, transitioning from 5% to 20% buffer B.
1320 This was followed by a 5-minute gradient from 20% to 32% buffer B, resulting in a 10-
1321 minute elution with 95% buffer B. Then, the column was equilibrated with 5% buffer B
1322 using a constant flow rate of 300 µl/min. Before sample analysis, an external calibration
1323 was performed using LTQ Velos ESI Positive Ion Calibration Solution (Pierce, IL, USA),
1324 along with internal calibration using the polysiloxane ion signal at m/z 445.120024
1325 obtained from ambient air. The MS1 scans were conducted within a m/z range of 375-
1326 1.600 at a resolution of 120.000. Employing a data-dependent acquisition strategy, the
1327 15 most intense precursor ions with a charge ranging from +2 to +5 within a window of
1328 1.2 m/z were selected from fragmentation to generate the corresponding MS2 spectra.

1329 The fragmentation ions were generated through high-energy collision-induced
1330 dissociation with an initial mass set at 110 m/z and subsequently detected using a mass-
1331 analyzed Orbitrap at a resolution of 30.000. The dynamic exclusion for the selected ions
1332 was set to 30 seconds, and the maximum accumulation times for MS1 and MS2 were 50
1333 ms and 70 ms, respectively. Finally, for protein sequence identification, the raw data were
1334 analyzed using Proteome Discoverer 2.4 (Thermo Fisher Scientific) with the Sequest HT
1335 search tool with mass tolerance parameters of 10 ppm and 0.02 Da for precursor and
1336 fragment ions, respectively.
1337

1338 **Dynamic light scattering experiments**

1339 The samples corresponding to the elution peaks obtained by size exclusion
1340 chromatography for BC1280 and N-domain$_{39-190}$ were filtered through a 0.46-μm syringe
1341 filter. The size measurements were performed using a Malvern Zetasizer Nano ZS
1342 (Malvern Panalytical, Malvern, United Kingdom) with a laser wavelength of 632.8 nm as
1343 the excitation source and a 1 cm pathway polystyrene cuvette as the sample holder. A
1344 total volume of 1 ml of the protein solution was transferred to the cuvette, and the dynamic
1345 light scattering signal was measured with a count rate ranging from 100 to 230 kcps,
1346 depending on the sample. Data collection and subsequent analysis were conducted using
1347 Zetasizer software v.6.34 (Malvern Panalytical).

1348 **Transmission electron microscopy**

1349 To study the morphology of BC1280, N-domain$_{39-191}$ and C-domain$_{191-279}$ purified *in vitro*,
1350 the samples were subjected to 2-fold serial dilutions and incubated on copper grids for 2

1351 minutes. Then, the samples were stained with a 2% uranyl acetate solution for 1 minute
1352 and dried under dark conditions.
1353 The samples were examined using a Thermo Fisher Scientific Tecnai G^2 20 TWIN
1354 transmission electron microscope at an accelerating voltage of 120 kV. The final images
1355 were captured using an Olympus Veleta side-mounted CCD with a resolution of 2k × 2k
1356 Mpx.
1357

1358 **Thioflavin T (ThT) assays**

1359 The polymerization kinetics were studied using thioflavin T staining. The experiment was
1360 conducted in a 96-well plate, and 40 μM ThT was added to each well at a specific
1361 concentration, after which the mixture was diluted in buffer containing 20 mM Tris and 50
1362 mM NaCl at pH 7.4. The polymerization kinetics of CalY were assayed at 40 μM using
1363 low concentrations of BC1280 and N-domain$_{39\text{-}190}$ (2 and 10 μM).
1364 The plates were incubated at 37°C in a plate reader and shaken at 100 rpm before
1365 each measurement (Tecan Infinite M1000 Pro; Tecan, Männedorf, Switzerland). The
1366 fluorescence intensity was monitored every 30 minutes through the bottom of the plate
1367 using an excitation wavelength of 440 nm and an emission wavelength of 480 nm. All
1368 measurements were recorded in triplicate.
1369

1370 **Sequence analysis by different bioinformatics tools**

1371 The primary sequence of BC1280 was obtained from UniProt (https://www.uniprot.org/)
1372 with the accession number Q81GC7. SignalP5.0 was used to predict the cleavage site
1373 for the signal peptidase.[84] The hydrophobicity was analyzed using the ProtScale tool

1374 provided by the ExPASy server, and the prediction of disordered regions was conducted
1375 using FoldUnfold with default parameters.[45,47] Putative amyloidogenic regions were
1376 predicted employing the following bioinformatics tools: AmylPred2, FoldAmyloid,
1377 MetAmyl, PASTA2.0 and TANGO.[40–44]
1378

1379 **AlphaFold prediction**

1380 The secondary structure model of BC1280 was predicted by AlphaFold using the
1381 sequence available in UniProt with the accession number Q81GC7 and default
1382 parameters.[48,49] The pdb model obtained was subsequently used for predicting chemical
1383 shifts via SPARTA.[53] Finally, the predicted correlations overlapped with the experimental
1384 data obtained by SSNMR.
1385

1386 ***In vitro* assembly of oligomers**

1387 BC1280 and N-domain$_{39\text{-}190}$ were diluted to a final concentration of 50 μM in buffer
1388 containing 20 mM Tris and 50 mM NaCl at pH 7.5. The oligomers self-assembled under
1389 agitation for 2 weeks at 37°C and a speed of 200 rpm. The aggregates were then
1390 ultracentrifuged (40k rpm speed, 5 hours, 18°C, TLA 120.1 rotor; Beckman Coulter), and
1391 the resulting pellets were analyzed using X-ray diffraction, ATR-FTIR, and SSNMR.
1392

1393 **X-ray diffraction measurements**

1394 Filament diffraction patterns were obtained at 4°C using a Rigaku FR-X rotating anode X-
1395 ray generator (Rigaku, Tokyo, Japan) equipped with an EIGER 1 M hybrid pixel detector
1396 (Dectris, Baden, Switzerland) at the copper wavelength. The concentrated hydrated

1397 samples were mounted in MicroLoops from Mitegen (Ithaca, NY, USA) on a goniometer
1398 head under cold nitrogen flow. Each diffraction pattern represents a 360° rotation along
1399 the φ axis, with an exposure time of 720 s. No correction, such as smooth filtering or
1400 baseline correction, was applied to the data. WinPLOTR (https://cdifx.univ-
1401 rennes1.fr/winplotr/winplotr.htm) was used for converting q values to reticular distance (d)
1402 values before plotting.
1403

1404 **Proteinase K (PK) digestion**

1405 A 22.5 µg solution of aggregated N-domain$_{39\text{-}190}$ was treated for 45 minutes at 37°C with
1406 3 µg/ml proteinase K in 20 mM Tris and 50 mM NaCl at pH 7.5. The proteinase activity
1407 was stopped using 5 mM PMSF. The samples were mixed with an equal volume of
1408 Laemmli buffer (Bio-Rad) and heated at 100°C for 5 minutes. Subsequently, the samples
1409 were separated using 12% SDS–PAGE, and the bands were visualized by Coomassie
1410 blue staining. The bands of interest were excised and subjected to analysis by tandem
1411 mass spectrometry.
1412

1413 **Attenuated total reflection FTIR spectroscopy**

1414 The samples were treated under the same conditions as those used for the X-ray studies.
1415 Infrared spectra for BC1280 and N-domain$_{39\text{-}190}$ were recorded with an FT-IR
1416 spectrophotometer (model Vertex70; Bruker) in attenuated total reflectance (ATR) mode
1417 utilizing a Golden Gate Single Reflection Diamond ATR System accessory (Specac). No
1418 sample preparation was needed, and spectra were recorded with an air background and
1419 64 scans for both the background and sample measurements. The spectral resolution

1420 was set at 4 cm-1". The raw data within the spectral range of the amide I region were
1421 used for estimating the total amount of secondary structure. To increase the resolution of
1422 the minimal absorbance peaks, a second derivative was calculated. Next, a deconvolution
1423 approach was employed using PeakFit software to determine the contribution percentage
1424 for each type of secondary structure: parallel β-sheet (1634 cm^{-1}), antiparallel β-sheet
1425 (1620 cm^{-1} and 1690 cm^{-1}), α-helix (1658 cm^{-1}), turns (1675 cm^{-1}) and random coil (1646
1426 cm^{-1}), in accordance with the well-established assignments.[58]
1427

1428 **Proteomic analysis by the isobaric tag for relative and absolute quantification**
1429 **(iTRAQ) method**

1430 To estimate the relative levels of TasA, BC1280 and CalY during biofilm formation in the
1431 wild-type strain *B. cereus* ATCC14579, we conducted sample preprocessing and
1432 subsequent analysis following the protocols described by Caro-Astorga et al.,in 2020.[26]
1433

1434 **NMR spectroscopy**

1435 Solid-state NMR experiments were performed on a 600 MHz spectrometer (Bruker
1436 Biospin, Germany) equipped with a triple resonance 4mm probe. Experiments were
1437 performed at 11 kHz MAS frequency at a sample temperature of 283K. Chemical shifts
1438 were referenced with DSS. For 2D ^{13}C-^{13}C PDSD spectra (mixing time of 50 and 200 ms),
1439 an initial cross-polarization step of 0.5 ms was used and high-power SPINAL-64
1440 decoupling was applied during acquisition times. 2D ^{13}C-^{13}C PDSD (50 ms) and ^{1}H-^{13}C
1441 INEPT spectra were recorded for 1-3 days, and 6 days for the ^{13}C-^{13}C PDSD spectrum
1442 (200 ms).

1443 Solution NMR $^{1}H,^{15}N$ SOFAST-HMQC experiment was carried out at 298 K on a protein
1444 sample containing 200 μM BC1280, 20mM Tris, 50mM NaCl, pH7.5, on a 700 MHz
1445 spectrometer (Bruker Biospin, Germany) equipped with a 5mm TXI $^{1}H/^{13}C/^{15}N/^{2}H$ probe.
1446 The NMR data were processed using the TOPSPIN 4.0.6 software and analyzed by
1447 CCPNMR.[85]

1448 **Circular dichroism**

1449 CalY (40 μM) alone or in combination with BC1280 (at 2 or 10 μM) was incubated in 10
1450 mM sodium phosphate buffer at pH 8 for 16 hours at 37ºC under agitation. Circular
1451 dichroism spectra were recorded using a JASCO J-815 spectrometer, using a 0.1 cm
1452 path-length cuvette in the range of 180-260 nm, with a 0.5 nm step and a 1 s collection
1453 time per step. The scan rate was 50 nm/min. Final spectra were obtained as the average
1454 of six scans after a blank correction. The spectrometer was continuously purged with dry
1455 N2 gas.CD spectra were buffer subtracted, and the results were expressed as mean
1456 residue ellipticity.

REFERENCES

1. Elhariry, H.M. (2011). Attachment strength and biofilm forming ability of Bacillus cereus on green-leafy vegetables: Cabbage and lettuce. Food Microbiology *28*, 1266–1274. https://doi.org/10.1016/j.fm.2011.05.004.

2. Hu, H.-J., Chen, Y.-L., Wang, Y.-F., Tang, Y.-Y., Chen, S.-L., and Yan, S.-Z. (2017). Endophytic *Bacillus cereus* Effectively Controls *Meloidogyne incognita* on Tomato Plants Through Rapid Rhizosphere Occupation and Repellent Action. Plant Disease *101*, 448–455. https://doi.org/10.1094/PDIS-06-16-0871-RE.

3. Cormontagne, D., Rigourd, V., Vidic, J., Rizzotto, F., Bille, E., and Ramarao, N. (2021). Bacillus cereus Induces Severe Infections in Preterm Neonates: Implication at the Hospital and Human Milk Bank Level. Toxins *13*, 123. https://doi.org/10.3390/toxins13020123.

4. Bottone, E.J. (2010). *Bacillus cereus* , a Volatile Human Pathogen. Clin Microbiol Rev *23*, 382–398. https://doi.org/10.1128/CMR.00073-09.

5. Thorsen, L., Azokpota, P., Munk Hansen, B., Rønsbo, M.H., Nielsen, K.F., Hounhouigan, D.J., and Jakobsen, M. (2011). Formation of cereulide and enterotoxins by Bacillus cereus in fermented African locust beans. Food Microbiology *28*, 1441–1447. https://doi.org/10.1016/j.fm.2011.07.003.

6. McKillip, J.L. (2000). Prevalence and expression of enterotoxins in Bacillus cereus and other Bacillus spp., a literature review. Antonie van Leeuwenhoek *77*, 393–399. https://doi.org/10.1023/A:1002706906154.

7. Dragoš, A., and Kovács, Á.T. (2017). The Peculiar Functions of the Bacterial Extracellular Matrix. Trends in Microbiology *25*, 257–266. https://doi.org/10.1016/j.tim.2016.12.010.

8. Jamal, M., Ahmad, W., Andleeb, S., Jalil, F., Imran, M., Nawaz, M.A., Hussain, T., Ali, M., Rafiq, M., and Kamil, M.A. (2018). Bacterial biofilm and associated infections. Journal of the Chinese Medical Association *81*, 7–11. https://doi.org/10.1016/j.jcma.2017.07.012.

9. Limoli, D.H., Jones, C.J., and Wozniak, D.J. (2015). Bacterial Extracellular Polysaccharides in Biofilm Formation and Function. Microbiol Spectr *3*, 3.3.29. https://doi.org/10.1128/microbiolspec.MB-0011-2014.

10. Lopez, D., Vlamakis, H., and Kolter, R. (2010). Biofilms. Cold Spring Harbor Perspectives in Biology *2*, a000398–a000398. https://doi.org/10.1101/cshperspect.a000398.

11. Lemon, K.P., Earl, A.M., Vlamakis, H.C., Aguilar, C., and Kolter, R. (2008). Biofilm Development with an Emphasis on Bacillus subtilis. In Bacterial Biofilms Current Topics in Microbiology and Immunology., T. Romeo, ed. (Springer Berlin Heidelberg), pp. 1–16. https://doi.org/10.1007/978-3-540-75418-3_1.

12. Bridier, A., Le Coq, D., Dubois-Brissonnet, F., Thomas, V., Aymerich, S., and Briandet, R. (2011). The Spatial Architecture of Bacillus subtilis Biofilms Deciphered Using a Surface-

Associated Model and In Situ Imaging. PLoS ONE *6*, e16177. https://doi.org/10.1371/journal.pone.0016177.

13. Pham, C.L.L., Kwan, A.H., and Sunde, M. (2014). Functional amyloid: widespread in Nature, diverse in purpose. Essays in Biochemistry *56*, 207–219. https://doi.org/10.1042/bse0560207.

14. Álvarez-Mena, A., Cámara-Almirón, J., De Vicente, A., and Romero, D. (2020). Multifunctional Amyloids in the Biology of Gram-Positive Bacteria. Microorganisms *8*, 2020. https://doi.org/10.3390/microorganisms8122020.

15. Blanco, L.P., Evans, M.L., Smith, D.R., Badtke, M.P., and Chapman, M.R. (2012). Diversity, biogenesis and function of microbial amyloids. Trends in Microbiology *20*, 66–73. https://doi.org/10.1016/j.tim.2011.11.005.

16. Fowler, D.M., Koulov, A.V., Balch, W.E., and Kelly, J.W. (2007). Functional amyloid – from bacteria to humans. Trends in Biochemical Sciences *32*, 217–224. https://doi.org/10.1016/j.tibs.2007.03.003.

17. Dragoš, A., Kovács, Á.T., and Claessen, D. (2017). The Role of Functional Amyloids in Multicellular Growth and Development of Gram-Positive Bacteria. Biomolecules *7*, 60. https://doi.org/10.3390/biom7030060.

18. Romero, D., and Kolter, R. (2014). Functional amyloids in bacteria. International Microbiology, 65–73. https://doi.org/10.2436/20.1501.01.208.

19. Majed, R., Faille, C., Kallassy, M., and Gohar, M. (2016). Bacillus cereus Biofilms—Same, Only Different. Front. Microbiol. *7*. https://doi.org/10.3389/fmicb.2016.01054.

20. Lin, Y., Briandet, R., and Kovács, Á.T. (2022). Bacillus cereus sensu lato biofilm formation and its ecological importance. Biofilm *4*, 100070. https://doi.org/10.1016/j.bioflm.2022.100070.

21. Houry, A., Briandet, R., Aymerich, S., and Gohar, M. (2010). Involvement of motility and flagella in Bacillus cereus biofilm formation. Microbiology *156*, 1009–1018. https://doi.org/10.1099/mic.0.034827-0.

22. Candela, T., Fagerlund, A., Buisson, C., Gilois, N., Kolstø, A., Økstad, O.A., Aymerich, S., Nielsen-Leroux, C., Lereclus, D., and Gohar, M. (2019). CalY is a major virulence factor and a biofilm matrix protein. Molecular Microbiology *111*, 1416–1429. https://doi.org/10.1111/mmi.14184.

23. Vilain, S., Pretorius, J.M., Theron, J., and Brözel, V.S. (2009). DNA as an Adhesin: *Bacillus cereus* Requires Extracellular DNA To Form Biofilms. Appl Environ Microbiol *75*, 2861–2868. https://doi.org/10.1128/AEM.01317-08.

24. Caro-Astorga, J., Álvarez-Mena, A., Hierrezuelo, J., Guadix, J.A., Heredia-Ponce, Z., Arboleda-Estudillo, Y., González-Munoz, E., De Vicente, A., and Romero, D. (2020). Two

1529 genomic regions encoding exopolysaccharide production systems have complementary
1530 functions in B. cereus multicellularity and host interaction. Sci Rep *10*, 1000.
1531 https://doi.org/10.1038/s41598-020-57970-3.

1532 25. Caro-Astorga, J., PÃ©rez-GarcÃ-a, A., De Vicente, A., and Romero, D. (2015). A genomic
1533 region involved in the formation of adhesin fibers in Bacillus cereus biofilms. Front.
1534 Microbiol. *5*. https://doi.org/10.3389/fmicb.2014.00745.

1535 26. Caro-Astorga, J., Frenzel, E., Perkins, J.R., Álvarez-Mena, A., De Vicente, A., Ranea,
1536 J.A.G., Kuipers, O.P., and Romero, D. (2020). Biofilm formation displays intrinsic offensive
1537 and defensive features of Bacillus cereus. npj Biofilms Microbiomes *6*, 3.
1538 https://doi.org/10.1038/s41522-019-0112-7.

1539 27. Stöver, A.G., and Driks, A. (1999). Secretion, Localization, and Antibacterial Activity of
1540 TasA, a *Bacillus subtilis* Spore-Associated Protein. J Bacteriol *181*, 1664–1672.
1541 https://doi.org/10.1128/JB.181.5.1664-1672.1999.

1542 28. Tjalsma, H., Bolhuis, A., Van Roosmalen, M.L., Wiegert, T., Schumann, W., Broekhuizen,
1543 C.P., Quax, W.J., Venema, G., Bron, S., and Van Dijl, J.M. (1998). Functional analysis of
1544 the secretory precursor processing machinery of *Bacillus subtilis* : identification of a
1545 eubacterial homolog of archaeal and eukaryotic signal peptidases. Genes Dev. *12*, 2318–
1546 2331. https://doi.org/10.1101/gad.12.15.2318.

1547 29. Romero, D., Aguilar, C., Losick, R., and Kolter, R. (2010). Amyloid fibers provide structural
1548 integrity to *Bacillus subtilis* biofilms. Proc. Natl. Acad. Sci. U.S.A. *107*, 2230–2234.
1549 https://doi.org/10.1073/pnas.0910560107.

1550 30. Diehl, A., Roske, Y., Ball, L., Chowdhury, A., Hiller, M., Molière, N., Kramer, R., Stöppler,
1551 D., Worth, C.L., Schlegel, B., et al. (2018). Structural changes of TasA in biofilm formation
1552 of *Bacillus subtilis*. Proc. Natl. Acad. Sci. U.S.A. *115*, 3237–3242.
1553 https://doi.org/10.1073/pnas.1718102115.

1554 31. Erskine, E., Morris, R.J., Schor, M., Earl, C., Gillespie, R.M.C., Bromley, K.M., Sukhodub,
1555 T., Clark, L., Fyfe, P.K., Serpell, L.C., et al. (2018). Formation of functional, non-
1556 amyloidogenic fibres by recombinant *Bacillus subtilis* TasA. Molecular Microbiology *110*,
1557 897–913. https://doi.org/10.1111/mmi.13985.

1558 32. Mammeri, N.E., Hierrezuelo, J., Tolchard, J., Cámara-Almirón, J., Caro-Astorga, J.,
1559 Álvarez-Mena, A., Dutour, A., Berbon, M., Shenoy, J., Morvan, E., et al. (2019). Molecular
1560 architecture of bacterial amyloids in *Bacillus* biofilms. FASEB j. *33*, 12146–12163.
1561 https://doi.org/10.1096/fj.201900831R.

1562 33. Böhning, J., Ghrayeb, M., Pedebos, C., Abbas, D.K., Khalid, S., Chai, L., and Bharat,
1563 T.A.M. (2022). Donor-strand exchange drives assembly of the TasA scaffold in Bacillus
1564 subtilis biofilms. Nat Commun *13*, 7082. https://doi.org/10.1038/s41467-022-34700-z.

1565 34. Fagerlund, A., Dubois, T., Økstad, O.-A., Verplaetse, E., Gilois, N., Bennaceur, I., Perchat,
1566 S., Gominet, M., Aymerich, S., Kolstø, A.-B., et al. (2014). SinR Controls Enterotoxin

1567 Expression in Bacillus thuringiensis Biofilms. PLoS ONE *9*, e87532.
1568 https://doi.org/10.1371/journal.pone.0087532.

1569 35. Jorda, J., Xue, B., Uversky, V.N., and Kajava, A.V. (2010). Protein tandem repeats - the
1570 more perfect, the less structured: Structural state of perfect protein repeats. FEBS Journal
1571 *277*, 2673–2682. https://doi.org/10.1111/j.1742-464X.2010.07684.x.

1572 36. Helmann, J.D. (2016). Bacillus subtilis extracytoplasmic function (ECF) sigma factors and
1573 defense of the cell envelope. Current Opinion in Microbiology *30*, 122–132.
1574 https://doi.org/10.1016/j.mib.2016.02.002.

1575 37. Mascher, T. (2023). Past, Present, and Future of Extracytoplasmic Function σ Factors:
1576 Distribution and Regulatory Diversity of the Third Pillar of Bacterial Signal Transduction.
1577 Annu. Rev. Microbiol. *77*, 625–644. https://doi.org/10.1146/annurev-micro-032221-024032.

1578 38. Gade Malmos, K., Blancas-Mejia, L.M., Weber, B., Buchner, J., Ramirez-Alvarado, M.,
1579 Naiki, H., and Otzen, D. (2017). ThT 101: a primer on the use of thioflavin T to investigate
1580 amyloid formation. Amyloid *24*, 1–16. https://doi.org/10.1080/13506129.2017.1304905.

1581 39. Sunde, M., Serpell, L.C., Bartlam, M., Fraser, P.E., Pepys, M.B., and Blake, C.C.F. (1997).
1582 Common core structure of amyloid fibrils by synchrotron X-ray diffraction 1 1Edited by F.
1583 E. Cohen. Journal of Molecular Biology *273*, 729–739.
1584 https://doi.org/10.1006/jmbi.1997.1348.

1585 40. Tsolis, A.C., Papandreou, N.C., Iconomidou, V.A., and Hamodrakas, S.J. (2013). A
1586 Consensus Method for the Prediction of 'Aggregation-Prone' Peptides in Globular Proteins.
1587 PLoS ONE *8*, e54175. https://doi.org/10.1371/journal.pone.0054175.

1588 41. Garbuzynskiy, S.O., Lobanov, M.Yu., and Galzitskaya, O.V. (2010). FoldAmyloid: a
1589 method of prediction of amyloidogenic regions from protein sequence. Bioinformatics *26*,
1590 326–332. https://doi.org/10.1093/bioinformatics/btp691.

1591 42. Emily, M., Talvas, A., and Delamarche, C. (2013). MetAmyl: A METa-Predictor for
1592 AMYLoid Proteins. PLoS ONE *8*, e79722. https://doi.org/10.1371/journal.pone.0079722.

1593 43. Walsh, I., Seno, F., Tosatto, S.C.E., and Trovato, A. (2014). PASTA 2.0: an improved server
1594 for protein aggregation prediction. Nucleic Acids Research *42*, W301–W307.
1595 https://doi.org/10.1093/nar/gku399.

1596 44. Fernandez-Escamilla, A.-M., Rousseau, F., Schymkowitz, J., and Serrano, L. (2004).
1597 Prediction of sequence-dependent and mutational effects on the aggregation of peptides and
1598 proteins. Nat Biotechnol *22*, 1302–1306. https://doi.org/10.1038/nbt1012.

1599 45. Galzitskaya, O.V., Garbuzynskiy, S.O., and Lobanov, M.Yu. (2006). FoldUnfold: web server
1600 for the prediction of disordered regions in protein chain. Bioinformatics *22*, 2948–2949.
1601 https://doi.org/10.1093/bioinformatics/btl504.

1602 46. Romero, P., Obradovic, Z., Li, X., Garner, E.C., Brown, C.J., and Dunker, A.K. (2001).
1603 Sequence complexity of disordered protein. Proteins *42*, 38–48.
1604 https://doi.org/10.1002/1097-0134(20010101)42:1<38::AID-PROT50>3.0.CO;2-3.

1605 47. Wilkins, M.R., Gasteiger, E., Bairoch, A., Sanchez, J.-C., Williams, K.L., Appel, R.D., and
1606 Hochstrasser, D.F. (1998). Protein Identification and Analysis Tools in the ExPASy Server.
1607 In 2-D Proteome Analysis Protocols (Humana Press), pp. 531–552.
1608 https://doi.org/10.1385/1-59259-584-7:531.

1609 48. Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O.,
1610 Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A., et al. (2021). Highly accurate
1611 protein structure prediction with AlphaFold. Nature *596*, 583–589.
1612 https://doi.org/10.1038/s41586-021-03819-2.

1613 49. Varadi, M., Anyango, S., Deshpande, M., Nair, S., Natassia, C., Yordanova, G., Yuan, D.,
1614 Stroe, O., Wood, G., Laydon, A., et al. (2022). AlphaFold Protein Structure Database:
1615 massively expanding the structural coverage of protein-sequence space with high-accuracy
1616 models. Nucleic Acids Research *50*, D439–D444. https://doi.org/10.1093/nar/gkab1061.

1617 50. Meier, B.H., Riek, R., and Böckmann, A. (2017). Emerging Structural Understanding of
1618 Amyloid Fibrils by Solid-State NMR. Trends in Biochemical Sciences *42*, 777–787.
1619 https://doi.org/10.1016/j.tibs.2017.08.001.

1620 51. Loquet, A., El Mammeri, N., Stanek, J., Berbon, M., Bardiaux, B., Pintacuda, G., and
1621 Habenstein, B. (2018). 3D structure determination of amyloid fibrils using solid-state NMR
1622 spectroscopy. Methods *138–139*, 26–38. https://doi.org/10.1016/j.ymeth.2018.03.014.

1623 52. Van Der Wel, P.C.A. (2017). Insights into protein misfolding and aggregation enabled by
1624 solid-state NMR spectroscopy. Solid State Nuclear Magnetic Resonance *88*, 1–14.
1625 https://doi.org/10.1016/j.ssnmr.2017.10.001.

1626 53. Shen, Y., and Bax, A. (2010). SPARTA+: a modest improvement in empirical NMR
1627 chemical shift prediction by means of an artificial neural network. J Biomol NMR *48*, 13–22.
1628 https://doi.org/10.1007/s10858-010-9433-9.

1629 54. Wang, Y., and Jardetzky, O. (2002). Probability-based protein secondary structure
1630 identification using combined NMR chemical-shift data. Protein Science *11*, 852–861.
1631 https://doi.org/10.1110/ps.3180102.

1632 55. Ulamec, S.M., Brockwell, D.J., and Radford, S.E. (2020). Looking Beyond the Core: The
1633 Role of Flanking Regions in the Aggregation of Amyloidogenic Peptides and Proteins. Front.
1634 Neurosci. *14*, 611285. https://doi.org/10.3389/fnins.2020.611285.

1635 56. Bhopatkar, A.A., and Kayed, R. (2023). Flanking regions, amyloid cores, and
1636 polymorphism: the potential interplay underlying structural diversity. Journal of Biological
1637 Chemistry *299*, 105122. https://doi.org/10.1016/j.jbc.2023.105122.

1638 57. Matlahov, I., and Van Der Wel, P.C.A. (2018). Hidden motions and motion-induced
1639 invisibility: Dynamics-based spectral editing in solid-state NMR. Methods *148*, 123–135.
1640 https://doi.org/10.1016/j.ymeth.2018.04.015.

1641 58. Goormaghtigh, E., Cabiaux, V., and Ruysschaert, J.-M. (1994). Determination of Soluble
1642 and Membrane Protein Structure by Fourier Transform Infrared Spectroscopy. In
1643 Physicochemical Methods in the Study of Biomembranes Subcellular Biochemistry., H. J.
1644 Hilderson and G. B. Ralston, eds. (Springer US), pp. 405–450. https://doi.org/10.1007/978-
1645 1-4615-1863-1_10.

1646 59. Romero, D., Vlamakis, H., Losick, R., and Kolter, R. (2011). An accessory protein required
1647 for anchoring and assembly of amyloid fibres in *B. subtilis* biofilms. Molecular
1648 Microbiology *80*, 1155–1168. https://doi.org/10.1111/j.1365-2958.2011.07653.x.

1649 60. Roske, Y., Lindemann, F., Diehl, A., Cremer, N., Higman, V.A., Schlegel, B., Leidert, M.,
1650 Driller, K., Turgay, K., Schmieder, P., et al. (2023). TapA acts as specific chaperone in TasA
1651 filament formation by strand complementation. Proc. Natl. Acad. Sci. U.S.A. *120*,
1652 e2217070120. https://doi.org/10.1073/pnas.2217070120.

1653 61. Abbasi, R., Mousa, R., Dekel, N., Amartely, H., Danieli, T., Lebendiker, M., Levi-Kalisman,
1654 Y., Shalev, D.E., Metanis, N., and Chai, L. (2019). The Bacterial Extracellular Matrix
1655 Protein TapA Is a Two-Domain Partially Disordered Protein. ChemBioChem *20*, 355–359.
1656 https://doi.org/10.1002/cbic.201800634.

1657 62. Lukaszczyk, M., Pradhan, B., and Remaut, H. (2019). The Biosynthesis and Structures of
1658 Bacterial Pili. In Bacterial Cell Walls and Membranes Subcellular Biochemistry., A. Kuhn,
1659 ed. (Springer International Publishing), pp. 369–413. https://doi.org/10.1007/978-3-030-
1660 18768-2_12.

1661 63. Shanmugasundarasamy, T., Karaiyagowder Govindarajan, D., and Kandaswamy, K. (2022).
1662 A review on pilus assembly mechanisms in Gram-positive and Gram-negative bacteria. The
1663 Cell Surface *8*, 100077. https://doi.org/10.1016/j.tcsw.2022.100077.

1664 64. Pradhan, B., Liedtke, J., Sleutel, M., Lindbäck, T., Zegeye, E.D., O´Sullivan, K., Llarena, A.,
1665 Brynildsrud, O., Aspholm, M., and Remaut, H. (2021). Endospore Appendages: a novel pilus
1666 superfamily from the endospores of pathogenic Bacilli. The EMBO Journal *40*, e106887.
1667 https://doi.org/10.15252/embj.2020106887.

1668 65. Hendrickx, A.P.A., Poor, C.B., Jureller, J.E., Budzik, J.M., He, C., and Schneewind, O.
1669 (2012). Isopeptide bonds of the major pilin protein BcpA influence pilus structure and
1670 bundle formation on the surface of *Bacillus cereus*. Molecular Microbiology *85*, 152–163.
1671 https://doi.org/10.1111/j.1365-2958.2012.08098.x.

1672 66. Zegeye, E.D., Pradhan, B., Llarena, A.-K., and Aspholm, M. (2021). Enigmatic Pilus-Like
1673 Endospore Appendages of Bacillus cereus Group Species. IJMS *22*, 12367.
1674 https://doi.org/10.3390/ijms222212367.

1675 67. Wilson, M.R., Yerbury, J.J., and Poon, S. (2008). Potential roles of abundant extracellular
1676 chaperones in the control of amyloid formation and toxicity. Mol. BioSyst. *4*, 42–52.
1677 https://doi.org/10.1039/B712728F.

1678 68. Kovacs, D., and Tompa, P. (2012). Diverse functional manifestations of intrinsic structural
1679 disorder in molecular chaperones. Biochemical Society Transactions *40*, 963–968.
1680 https://doi.org/10.1042/BST20120108.

1681 69. Foit, L., George, J.S., Zhang, B.W., Brooks, C.L., and Bardwell, J.C.A. (2013). Chaperone
1682 activation by unfolding. Proc. Natl. Acad. Sci. U.S.A. *110*.
1683 https://doi.org/10.1073/pnas.1222458110.

1684 70. Díaz-Villanueva, J., Díaz-Molina, R., and García-González, V. (2015). Protein Folding and
1685 Mechanisms of Proteostasis. IJMS *16*, 17193–17230.
1686 https://doi.org/10.3390/ijms160817193.

1687 71. Arnaud, M., Chastanet, A., and Débarbouillé, M. (2004). New Vector for Efficient Allelic
1688 Replacement in Naturally Nontransformable, Low-GC-Content, Gram-Positive Bacteria.
1689 Appl Environ Microbiol *70*, 6887–6891. https://doi.org/10.1128/AEM.70.11.6887-
1690 6891.2004.

1691 72. Pflughoeft, K.J., Sumby, P., and Koehler, T.M. (2011). *Bacillus anthracis sin* Locus and
1692 Regulation of Secreted Proteases. J Bacteriol *193*, 631–639.
1693 https://doi.org/10.1128/JB.01083-10.

1694 73. Madeira, F., Pearce, M., Tivey, A.R.N., Basutkar, P., Lee, J., Edbali, O., Madhusoodanan,
1695 N., Kolesnikov, A., and Lopez, R. (2022). Search and sequence analysis tools services from
1696 EMBL-EBI in 2022. Nucleic Acids Research *50*, W276–W279.
1697 https://doi.org/10.1093/nar/gkac240.

1698 74. Untergasser, A., Cutcutache, I., Koressaar, T., Ye, J., Faircloth, B.C., Remm, M., and Rozen,
1699 S.G. (2012). Primer3—new capabilities and interfaces. Nucleic Acids Research *40*, e115–
1700 e115. https://doi.org/10.1093/nar/gks596.

1701 75. Vargas, P., Felipe, A., Michán, C., and Gallegos, M.-T. (2011). Induction of *Pseudomonas*
1702 *syringae* pv. *tomato* DC3000 MexAB-OprM Multidrug Efflux Pump by Flavonoids Is
1703 Mediated by the Repressor PmeR. MPMI *24*, 1207–1219. https://doi.org/10.1094/MPMI-03-
1704 11-0077.

1705 76. Livak, K.J., and Schmittgen, T.D. (2001). Analysis of Relative Gene Expression Data Using
1706 Real-Time Quantitative PCR and the 2−ΔΔCT Method. Methods *25*, 402–408.
1707 https://doi.org/10.1006/meth.2001.1262.

1708 77. O'Toole, G.A., Pratt, L.A., Watnick, P.I., Newman, D.K., Weaver, V.B., and Kolter, R.
1709 (1999). [6] Genetic approaches to study of biofilms. In Methods in Enzymology (Elsevier),
1710 pp. 91–109. https://doi.org/10.1016/S0076-6879(99)10008-9.

1711 78. Falgueras, J., Lara, A.J., Fernández-Pozo, N., Cantón, F.R., Pérez-Trabado, G., and Claros,
1712 M.G. (2010). SeqTrim: a high-throughput pipeline for pre-processing any type of sequence
1713 read. BMC Bioinformatics *11*, 38. https://doi.org/10.1186/1471-2105-11-38.

1714 79. Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. Nat
1715 Methods *9*, 357–359. https://doi.org/10.1038/nmeth.1923.

1716 80. Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis,
1717 G., Durbin, R., and 1000 Genome Project Data Processing Subgroup (2009). The Sequence
1718 Alignment/Map format and SAMtools. Bioinformatics *25*, 2078–2079.
1719 https://doi.org/10.1093/bioinformatics/btp352.

1720 81. Robinson, M.D., McCarthy, D.J., and Smyth, G.K. (2010). edgeR : a Bioconductor package
1721 for differential expression analysis of digital gene expression data. Bioinformatics *26*, 139–
1722 140. https://doi.org/10.1093/bioinformatics/btp616.

1723 82. Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and
1724 dispersion for RNA-seq data with DESeq2. Genome Biol *15*, 550.
1725 https://doi.org/10.1186/s13059-014-0550-8.

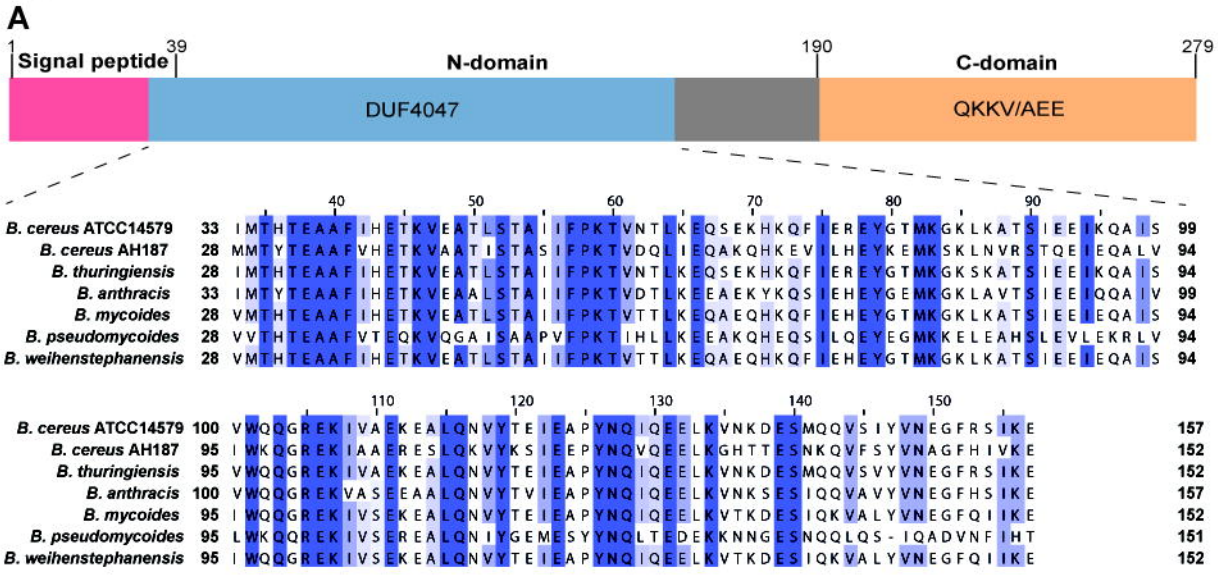
1726 83. Dunn, K.W., Kamocka, M.M., and McDonald, J.H. (2011). A practical guide to evaluating
1727 colocalization in biological microscopy. American Journal of Physiology-Cell Physiology
1728 *300*, C723–C742. https://doi.org/10.1152/ajpcell.00462.2010.

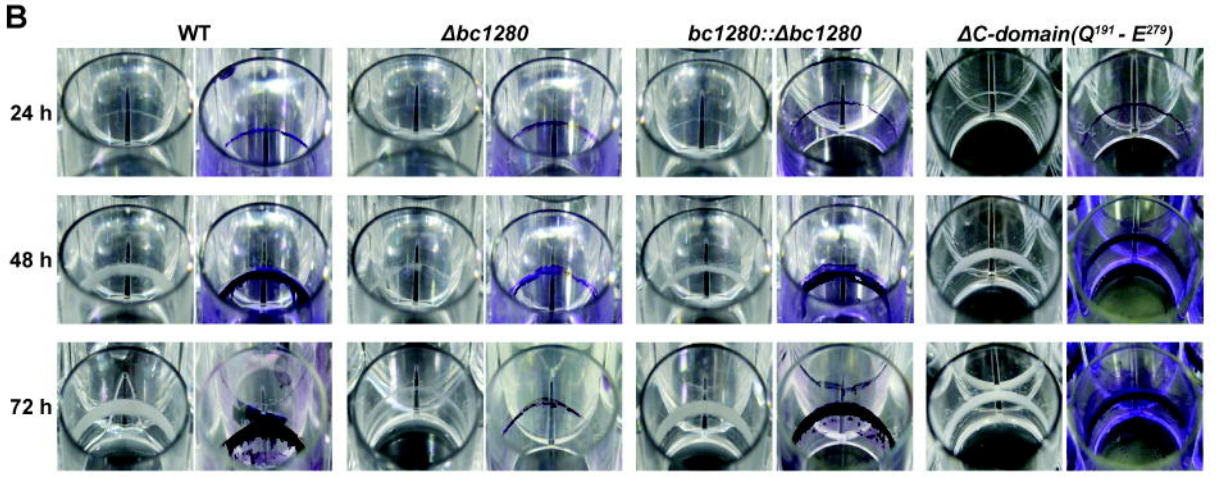
1729 84. Almagro Armenteros, J.J., Tsirigos, K.D., Sønderby, C.K., Petersen, T.N., Winther, O.,
1730 Brunak, S., Von Heijne, G., and Nielsen, H. (2019). SignalP 5.0 improves signal peptide
1731 predictions using deep neural networks. Nat Biotechnol *37*, 420–423.
1732 https://doi.org/10.1038/s41587-019-0036-z.

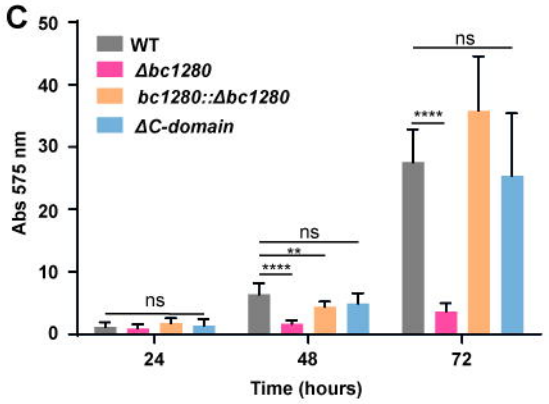
1733 85. Skinner, S.P., Fogh, R.H., Boucher, W., Ragan, T.J., Mureddu, L.G., and Vuister, G.W.
1734 (2016). CcpNmr AnalysisAssign: a flexible platform for integrated NMR analysis. J Biomol
1735 NMR *66*, 111–124. https://doi.org/10.1007/s10858-016-0060-y.

1736

Figure 1


A
1
39
190
279
Signal peptide
N-domain
C-domain
DUF4047
QKKV/AEE
B. cereus ATCC14579
B. cereus AH187
B. thuringiensis
B. anthracis
B. mycoides
B. pseudomycoides
B. weihenstephanensis


B
WT
Δbc1280
bc1280::Δbc1280
ΔC-domain(Q191 - E279)
24 h
48 h
72 h


C
WT
Δbc1280
bc1280::Δbc1280
ΔC-domain
Abs 575 nm
Time (hours)
ns
**
****

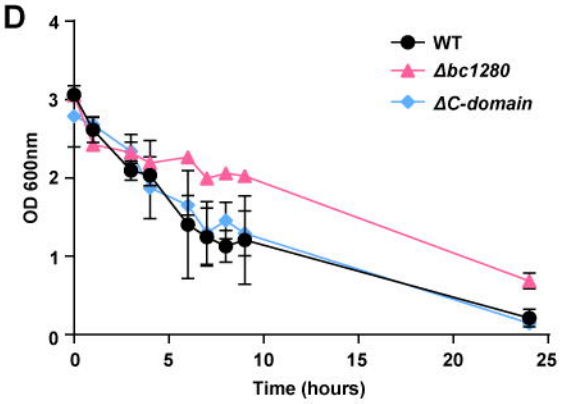

D
WT
Δbc1280
ΔC-domain
OD 600nm
Time (hours)

Figure 2

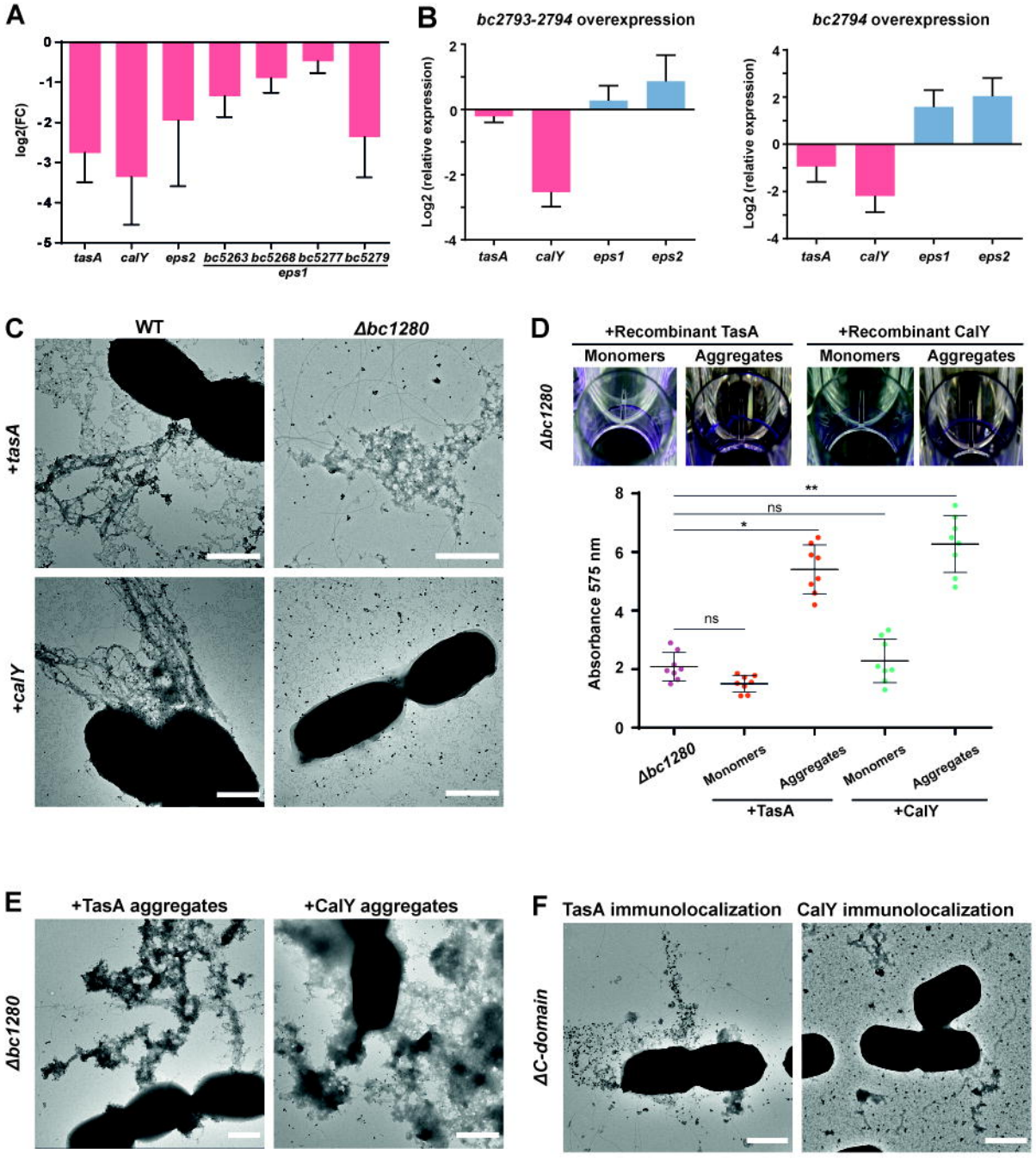

A
log2(FC)
0
-1
-2
-3
-4
-5
tasA
calY
eps2
bc5263
bc5268
bc5277
bc5279
eps1
B
bc2793-2794 overexpression
Log2 (relative expression)
2
0
-2
-4
tasA
calY
eps1
eps2
bc2794 overexpression
Log2 (relative expression)
4
2
0
-2
-4
tasA
calY
eps1
eps2
C
WT
Δbc1280
+tasA
+calY
D
+Recombinant TasA
+Recombinant CalY
Monomers
Aggregates
Monomers
Aggregates
Δbc1280
Absorbance 575 nm
8
6
4
2
0
**
ns
*
ns
Δbc1280
Monomers
Aggregates
Monomers
Aggregates
+TasA
+CalY
E
+TasA aggregates
+CalY aggregates
Δbc1280
F
TasA immunolocalization
CalY immunolocalization
ΔC-domain

Figure 3

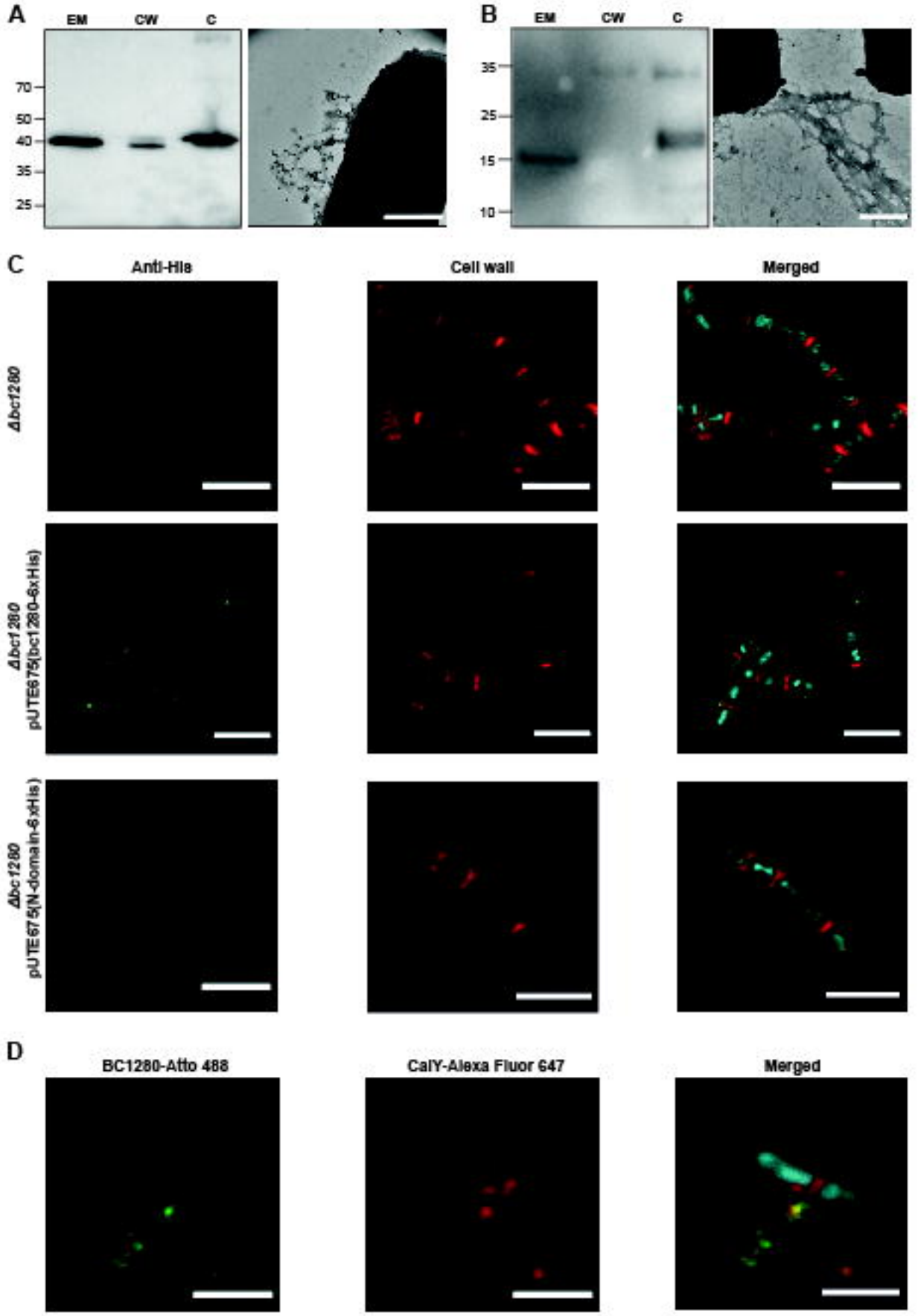

A
EM
CW
C
70
50
40
35
25
B
EM
CW
C
35
25
15
10
C
Anti-His
Cell wall
Merged
Δbc1280
Δbc1280
pUTE675(bc1280-6xHis)
Δbc1280
pUTE675(N-domain-6xHis)
D
BC1280-Atto 488
CalY-Alexa Fluor 647
Merged

Figure 4

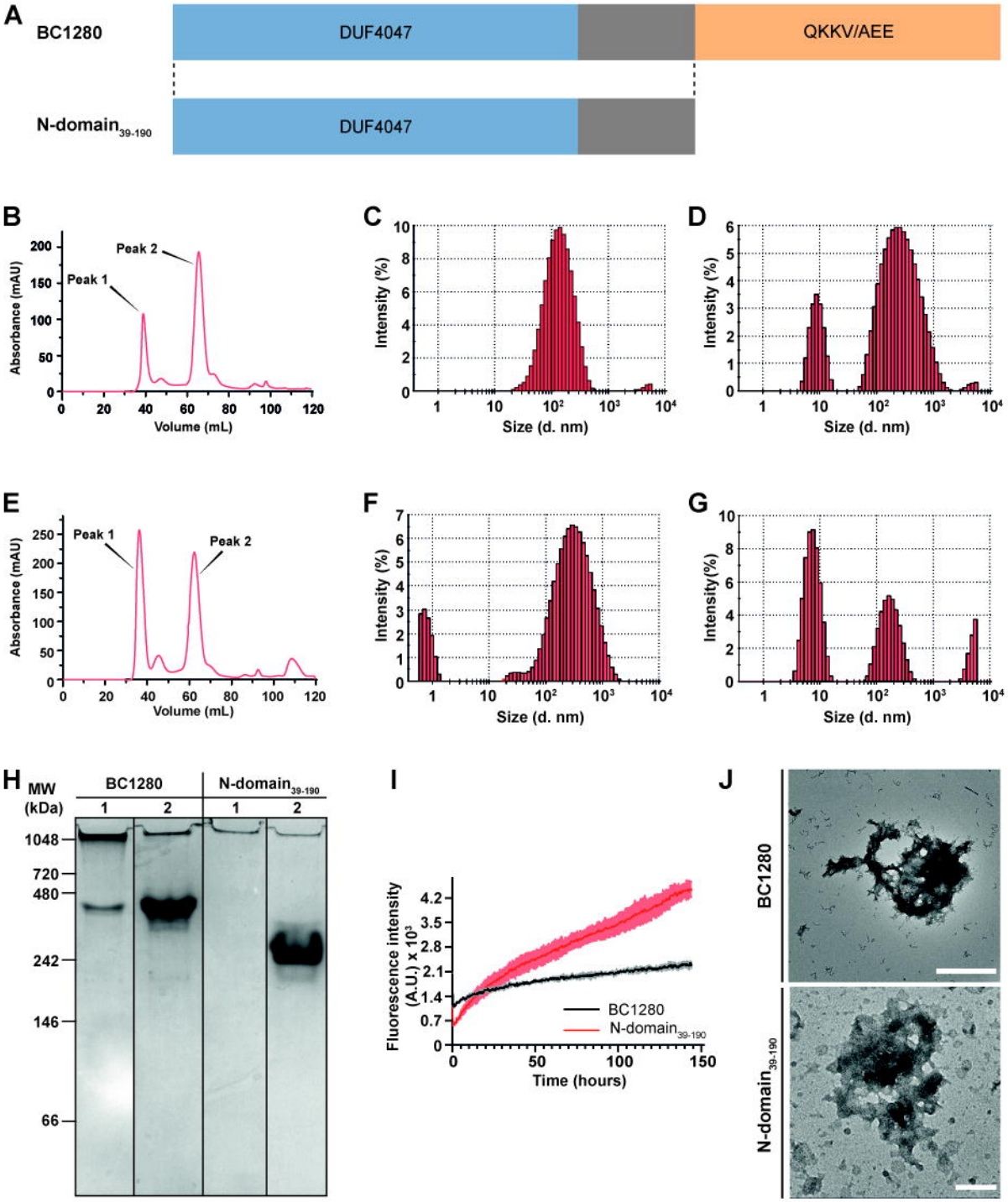
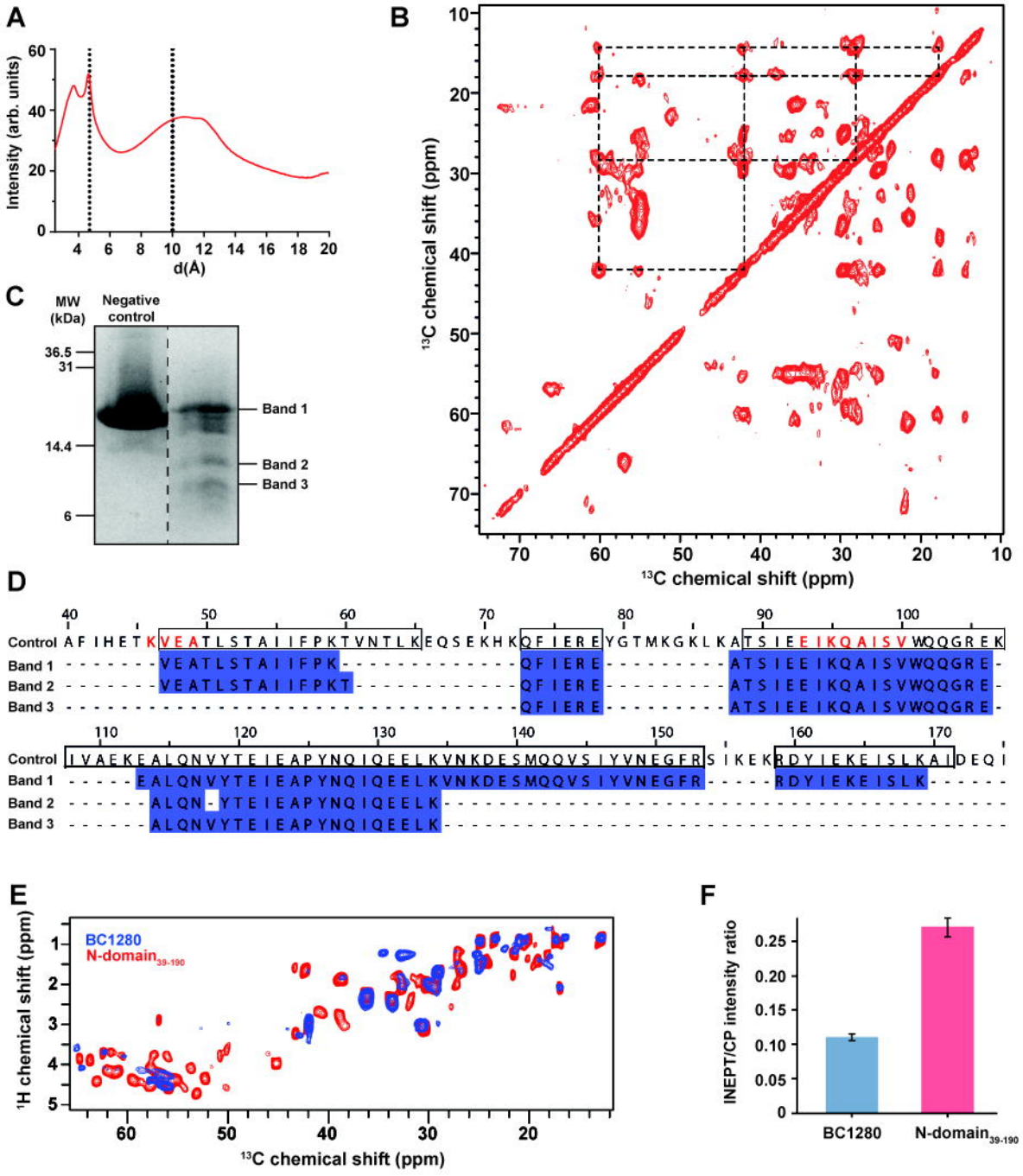

A
BC1280
DUF4047
QKKV/AEE
N-domain39-190
DUF4047
B
Peak 1
Peak 2
Absorbance (mAU)
Volume (mL)
C
Intensity (%)
Size (d. nm)
D
Intensity (%)
Size (d. nm)
E
Peak 1
Peak 2
Absorbance (mAU)
Volume (mL)
F
Intensity (%)
Size (d. nm)
G
Intensity(%)
Size (d. nm)
H
MW (kDa)
BC1280
N-domain39-190
1
2
1
2
1048
720
480
242
146
66
I
Fluorescence intensity (A.U.) x 10³
Time (hours)
BC1280
N-domain39-190
J
BC1280
N-domain39-190

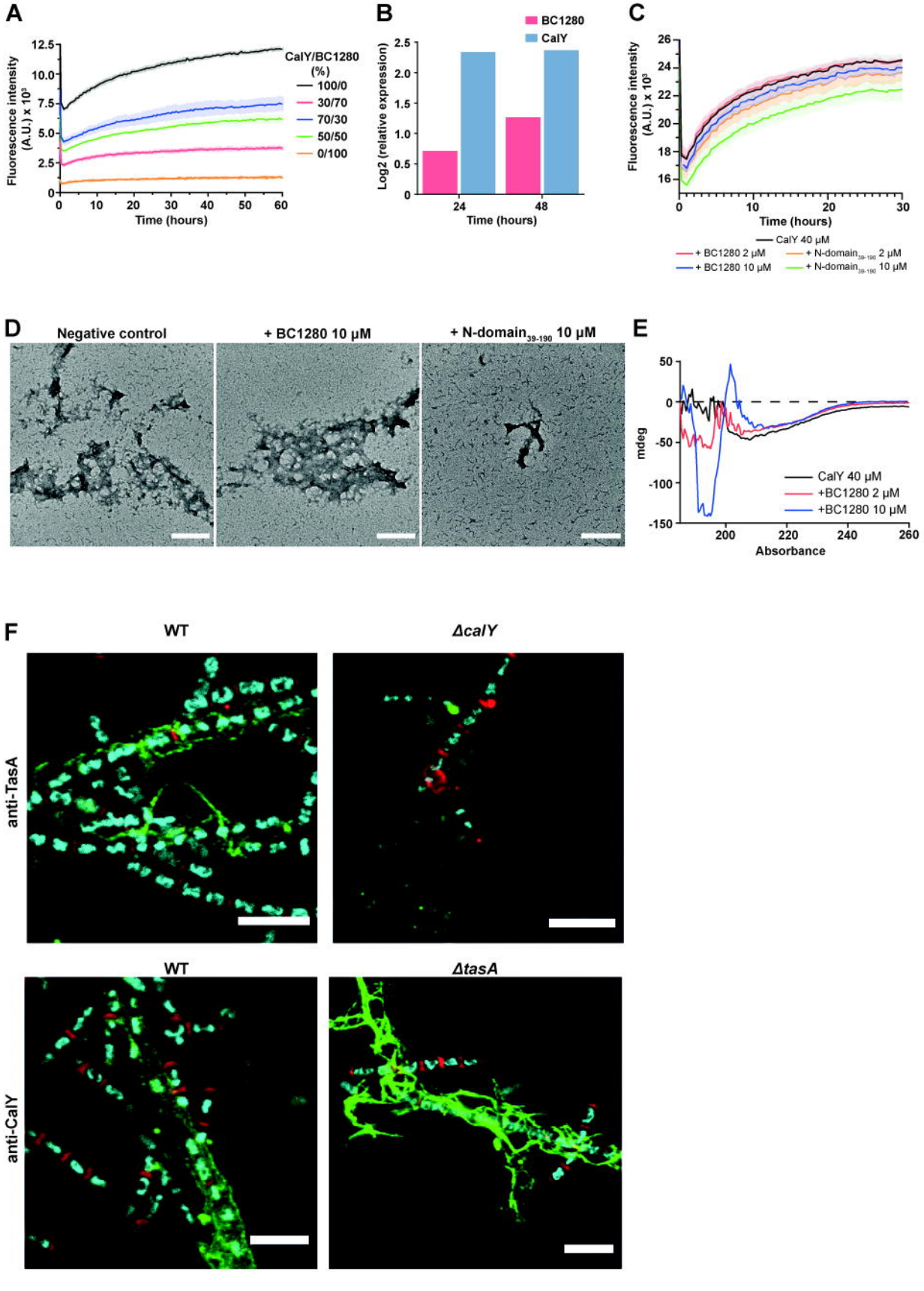
Figure 5



G

		Percentage of structural elements (%)	
Secondary structure element	**Wavenumbers (cm^{-1})**	**BC1280**	**N-domain$_{39-190}$**
β-sheet antiparallel	1620,1690	25%	33%
β-sheet parallel	1634	22%	19%
Random coil	1646	24%	24%
α-helix	1658	15%	12%
Turn	1675	14%	12%

Figure 6


A
Fluorescence intensity (A.U.) x 10³
Time (hours)
CalY/BC1280 (%)
100/0
30/70
70/30
50/50
0/100
B
BC1280
CalY
Log2 (relative expression)
Time (hours)
C
Fluorescence intensity (A.U.) x 10³
Time (hours)
CalY 40 µM
+ BC1280 2 µM
+ BC1280 10 µM
+ N-domain39-190 2 µM
+ N-domain39-190 10 µM
D
Negative control
+ BC1280 10 µM
+ N-domain39-190 10 µM
E
mdeg
Absorbance
CalY 40 µM
+BC1280 2 µM
+BC1280 10 µM
F
WT
ΔcalY
anti-TasA
WT
ΔtasA
anti-CalY

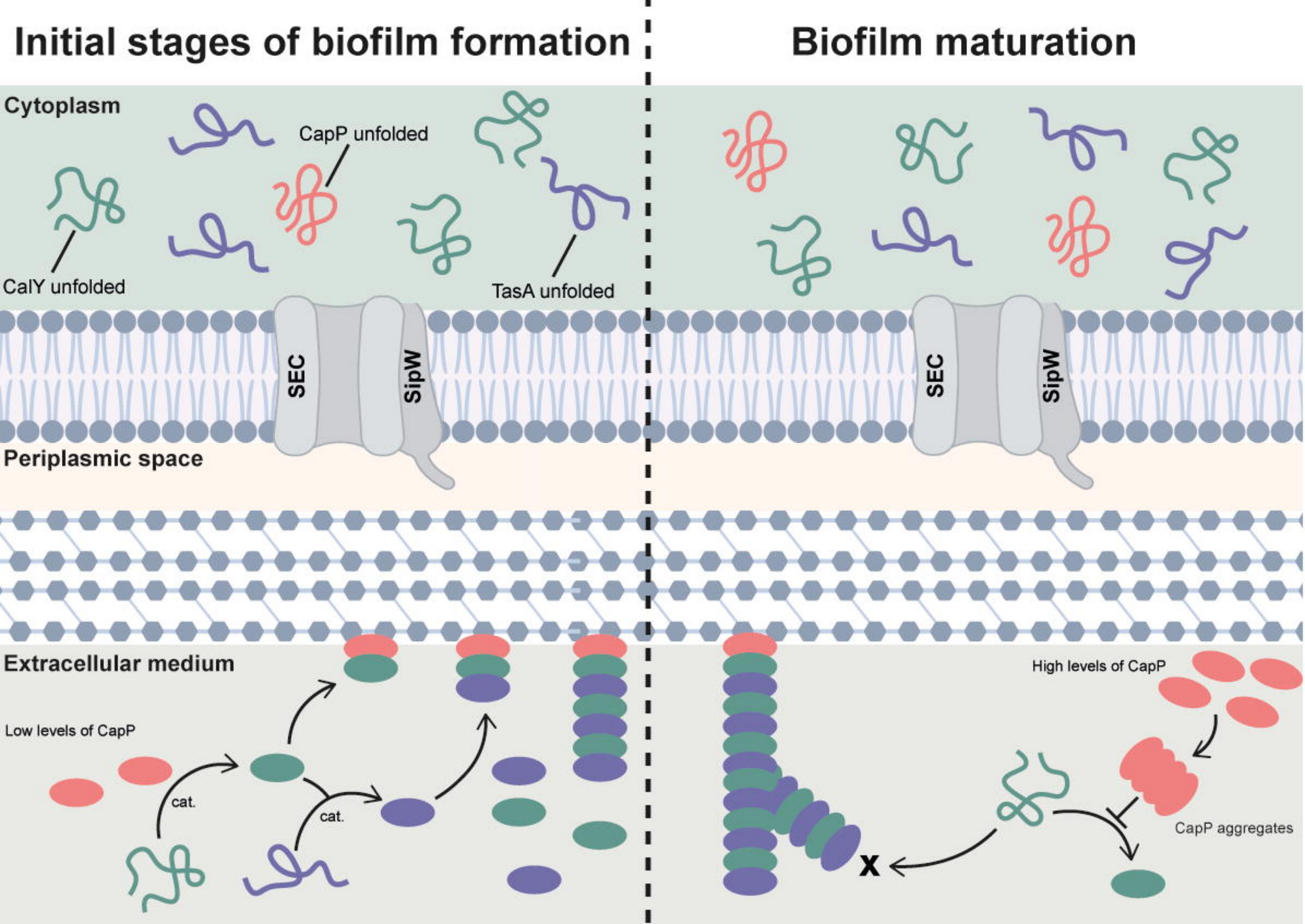

Initial stages of biofilm formation
Biofilm maturation
Cytoplasm
CapP unfolded
CalY unfolded
TasA unfolded
SEC
SipW
SEC
SipW
Periplasmic space
Extracellular medium
Low levels of CapP
cat.
cat.
High levels of CapP
CapP aggregates
X

Figure S1

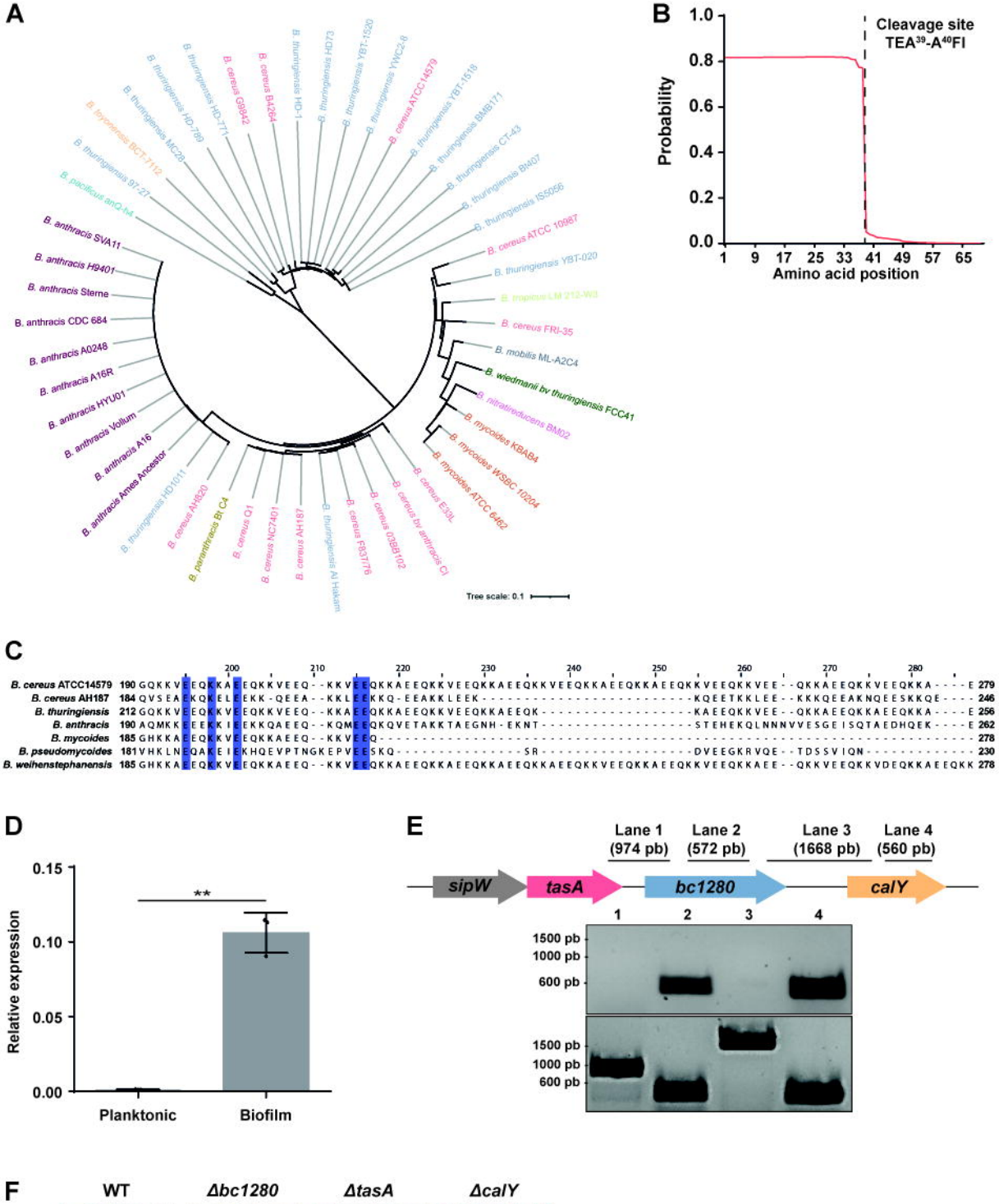
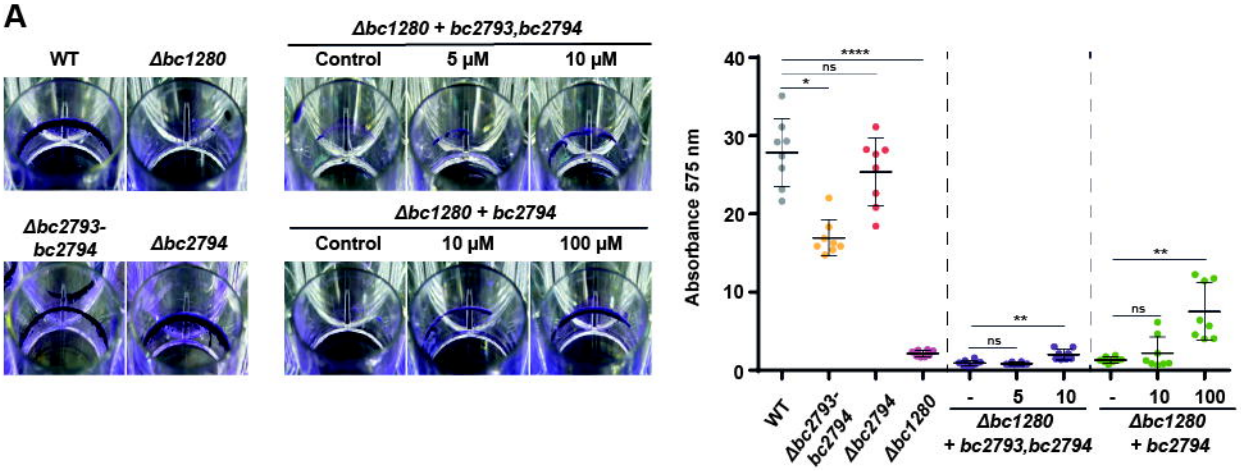
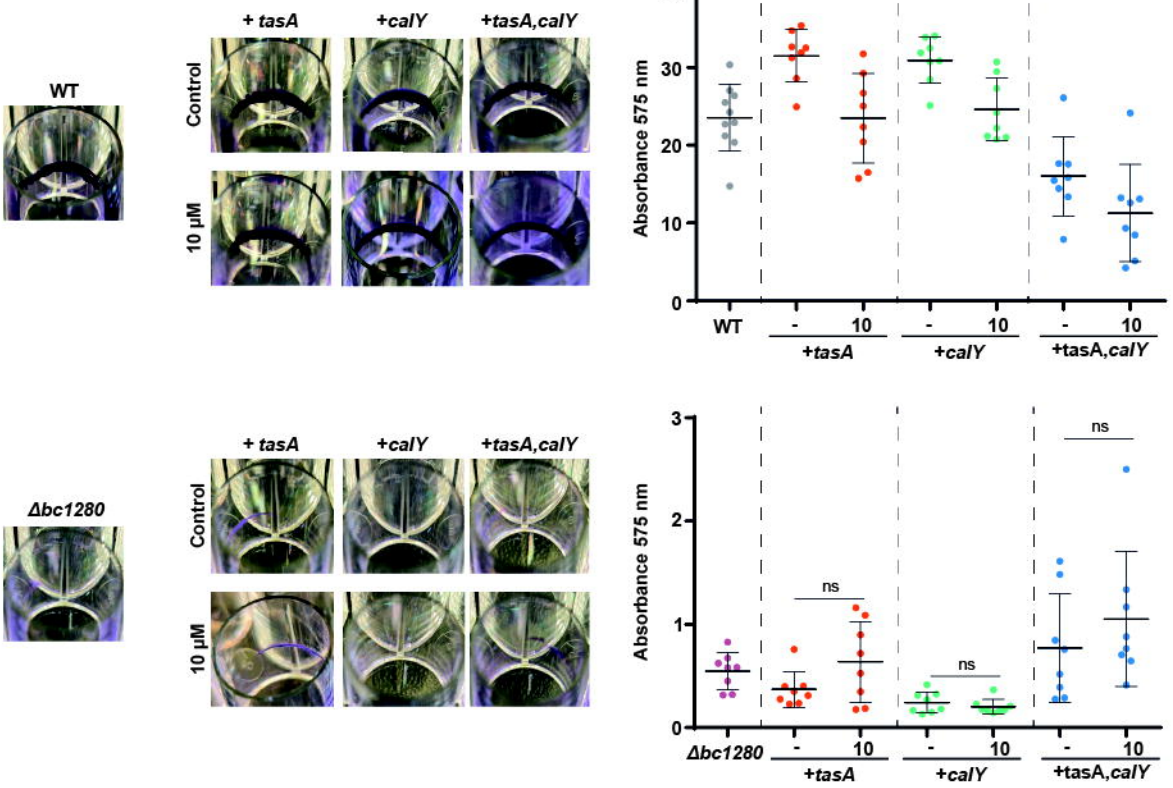

A
B. thuringiensis HD-1
B. thuringiensis HD73
B. thuringiensis YBT-1520
B. thuringiensis YWC2-8
B. cereus ATCC14579
B. thuringiensis YBT-1518
B. thuringiensis BMB171
B. thuringiensis CT-43
B. thuringiensis BM07
B. thuringiensis IS5056
B. cereus ATCC 10987
B. thuringiensis YBT-020
B. tropicus LM 212-W3
B. cereus FRI-35
B. mobilis ML-A2C4
B. wiedmanii bv thuringiensis FCC41
B. nitratireducens BM02
B. mycoides KBAB4
B. mycoides WSBC 10204
B. mycoides ATCC 6462
B. cereus E30L
B. cereus bv anthracis Cl
B. cereus 03BB102
B. cereus FB3775
B. thuringiensis Al Hakam
B. cereus AH187
B. cereus NC7401
B. cereus Q1
B. parasanthracis Bt C4
B. cereus AH820
B. thuringiensis HD1011
B. anthracis Ames Ancestor
B. anthracis A16
B. anthracis Vollum
B. anthracis HYU01
B. anthracis A16R
B. anthracis A0248
B. anthracis CDC 684
B. anthracis Sterne
B. anthracis H9401
B. anthracis SVA11
B. pacificus anQ-14
B. thuringiensis 97-27
B. tropicus BCT-7112
B. thuringiensis MC29
B. thuringiensis HD-789
B. thuringiensis HD-771
B. cereus G9842
B. cereus BA264
Tree scale: 0.1
B
Cleavage site
TEA39-A40FI
Probability
1.0
0.8
0.6
0.4
0.2
0.0
1 9 17 25 33 41 49 57 65
Amino acid position
C
200 210 220 230 240 250 260 270 280
B. cereus ATCC14579 190 GQKKVEEQKKAEEQKKVEEQ--KKVEEQKKAEEQKKVEEQKKAEEQKKAEEQKKVEEQKKVEE--QKKAEEQKKVEEQKKA----E 279
B. cereus AH187 184 QVSEAEKQKELEEKK-QEEA--KKLEEKKQ-EEAKKLEEK----------KQEETKKLEE--KKKQEEAKNQEESKKQE---- 246
B. thuringiensis 212 GQKKVEEQKKVEEQKKVEEQ--KKAEEQKKAEEQKKVEEQKKAEEQK----------KAEEQKKVEE--QKKAEEQKKAEEQKKA----E 256
B. anthracis 190 AQMKKEEEKKIEEKKQAEEQ--KQMEEQKQVETAKKTAEGNH-EKNT----------STEHEKQLNNNVVESGEISQTAEDHQEK----E 262
B. mycoides 185 GHKKAEEQKKVEEQKKAEEQ--KKVEEQ---------- 278
B. pseudomycoides 181 VHKLNEQAKEIEKHQEVPTNGKEPVEESKQ----------SR----------DVEEGKRVQE--TDSSVIQN---------- 230
B. weihenstephanensis 185 GHKKAEEQKKVEEQKKAEEQ--KKVEEQKKAEEQKKAEEQKKAEEQKKAEEQKKVEEQKKAEEQKKVEEQKKAEE--QKKVEEQKKVDEQKKAEEQKK 278
D
Relative expression
0.15
0.10
0.05
0.00
**
Planktonic
Biofilm
E
Lane 1 (974 pb)
Lane 2 (572 pb)
Lane 3 (1668 pb)
Lane 4 (560 pb)
sipW
tasA
bc1280
calY
1 2 3 4
1500 pb
1000 pb
600 pb
F
WT
Δbc1280
ΔtasA
ΔcalY
24 h
48 h
72 h

Figure S2

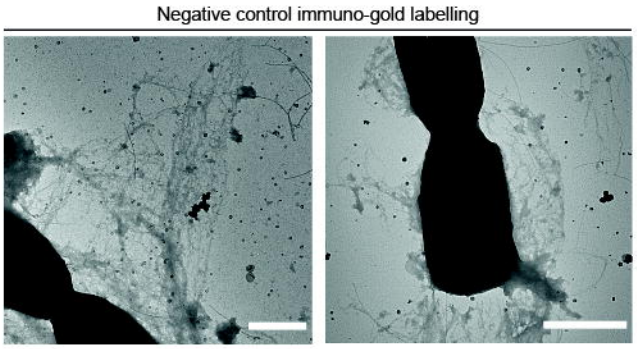
A



B



C



D

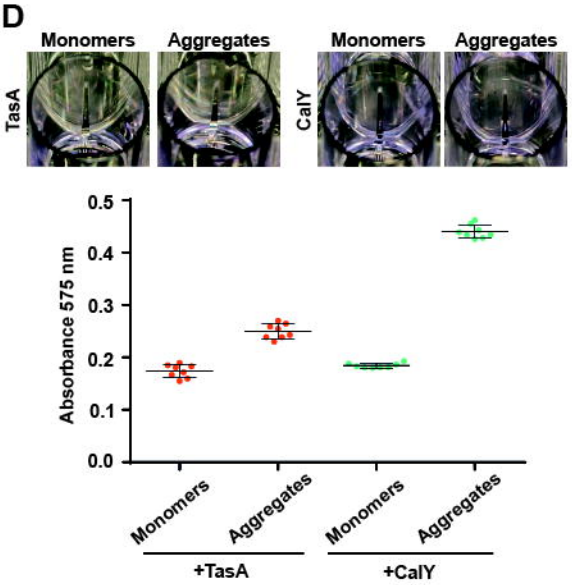


Figure S3

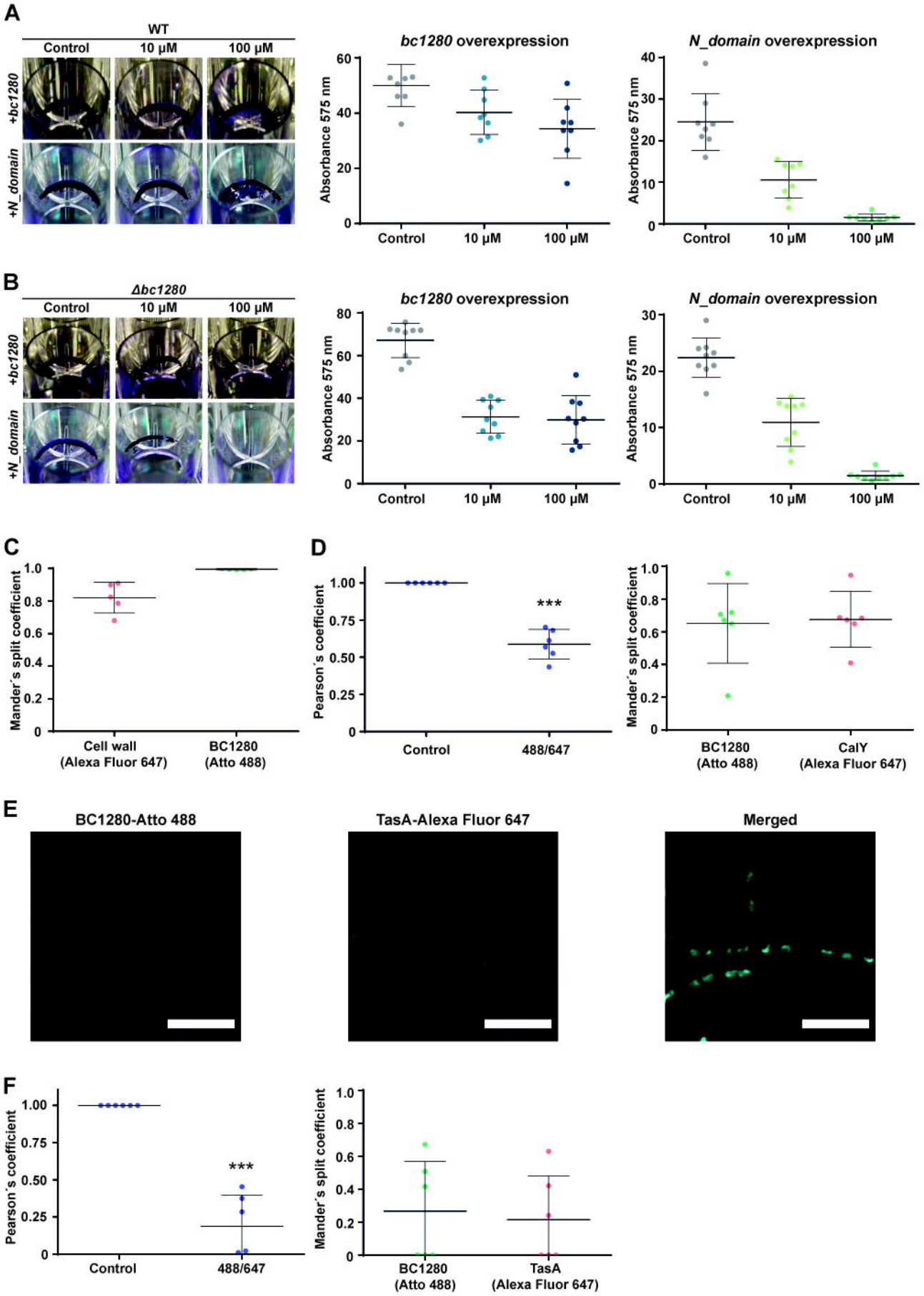

A
WT
Control
10 µM
100 µM
+bc1280
+N_domain
bc1280 overexpression
N_domain overexpression
Absorbance 575 nm
Control
10 µM
100 µM
B
Δbc1280
C
Mander´s split coefficient
Cell wall
(Alexa Fluor 647)
BC1280
(Atto 488)
D
Pearson´s coefficient
***
488/647
CalY
(Alexa Fluor 647)
E
BC1280-Atto 488
TasA-Alexa Fluor 647
Merged
F
TasA
(Alexa Fluor 647)

Figure S4

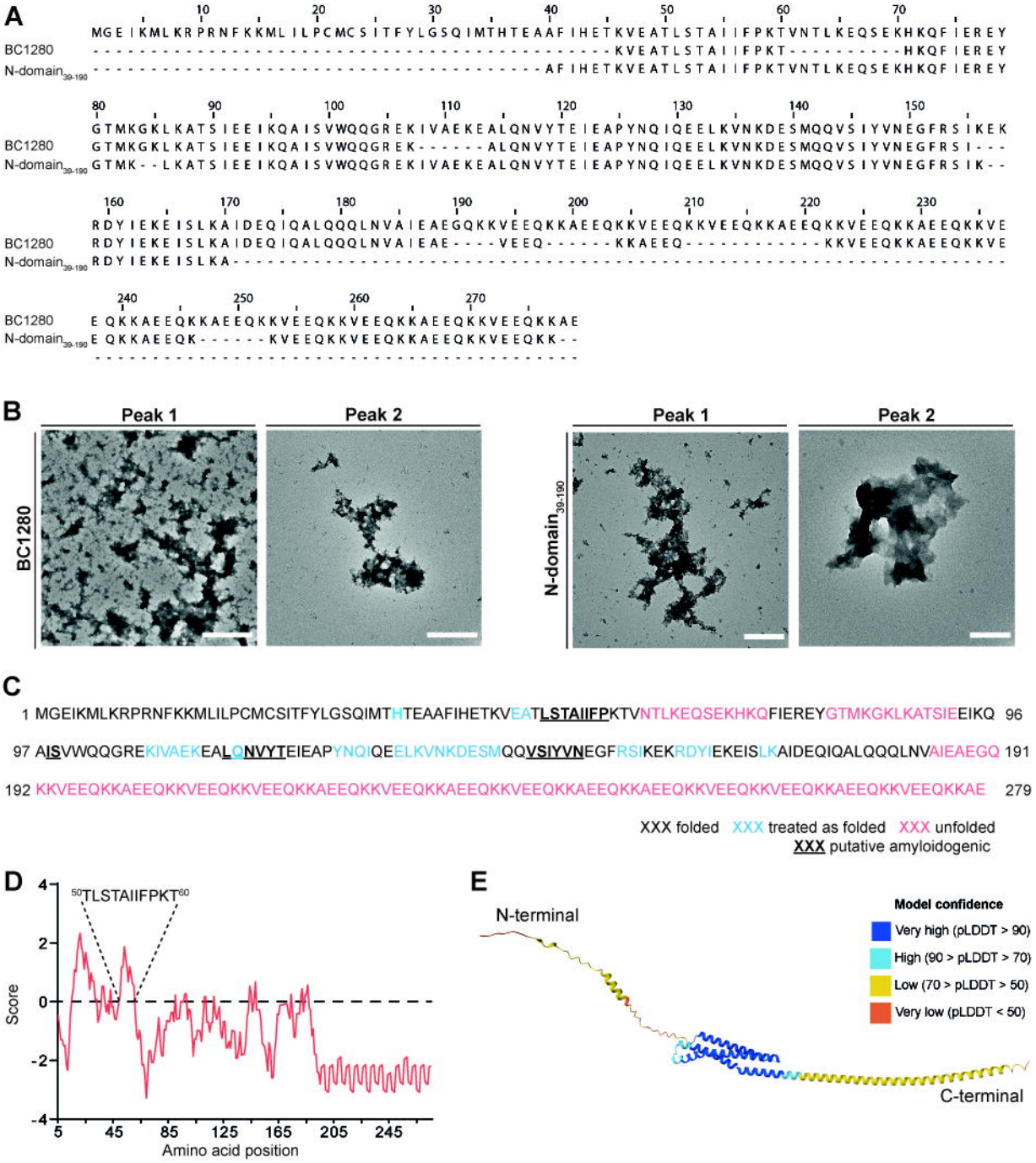

A
BC1280
N-domain39-190
B
Peak 1
Peak 2
BC1280
N-domain39-190
C
1 MGEIKMLKRPRNFKKMLILPCMCSITFYLGSQIMTHTEAAFIHETKVEATLSTAIIFPKTVNTLKEQSEKHKQFIEREYGTMKGKLKATSIEEIKQ 96
97 AISVWQQGREKIVAEKEALQNVYTEIEAPYNQIQEELKVNKDESMQQVSIYVNEGFRSIKEKRDYIEKEISLKAIDEQIQALQQQLNVAIEAEGQ 191
192 KKVEEQKKAEEQKKVEEQKKVEEQKKAEEQKKVEEQKKAEEQKKVEEQKKAEEQKKAEEQKKVEEQKKVEEQKKAEEQKKVEEQKKAE 279
XXX folded XXX treated as folded XXX unfolded
XXX putative amyloidogenic
D
50TLSTAIIFPKT60
Score
Amino acid position
E
N-terminal
C-terminal
Model confidence
Very high (pLDDT > 90)
High (90 > pLDDT > 70)
Low (70 > pLDDT > 50)
Very low (pLDDT < 50)

Figure S5

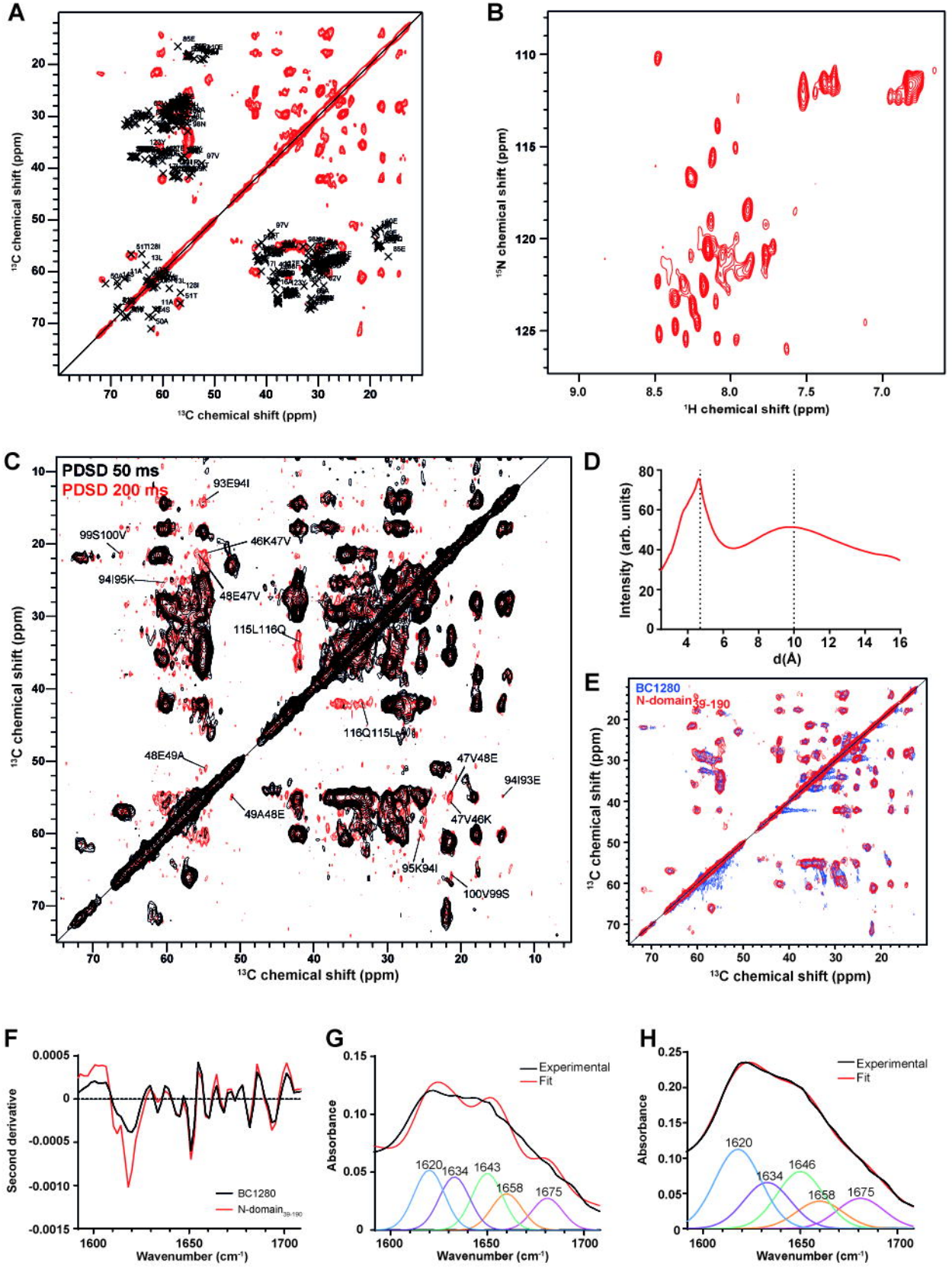
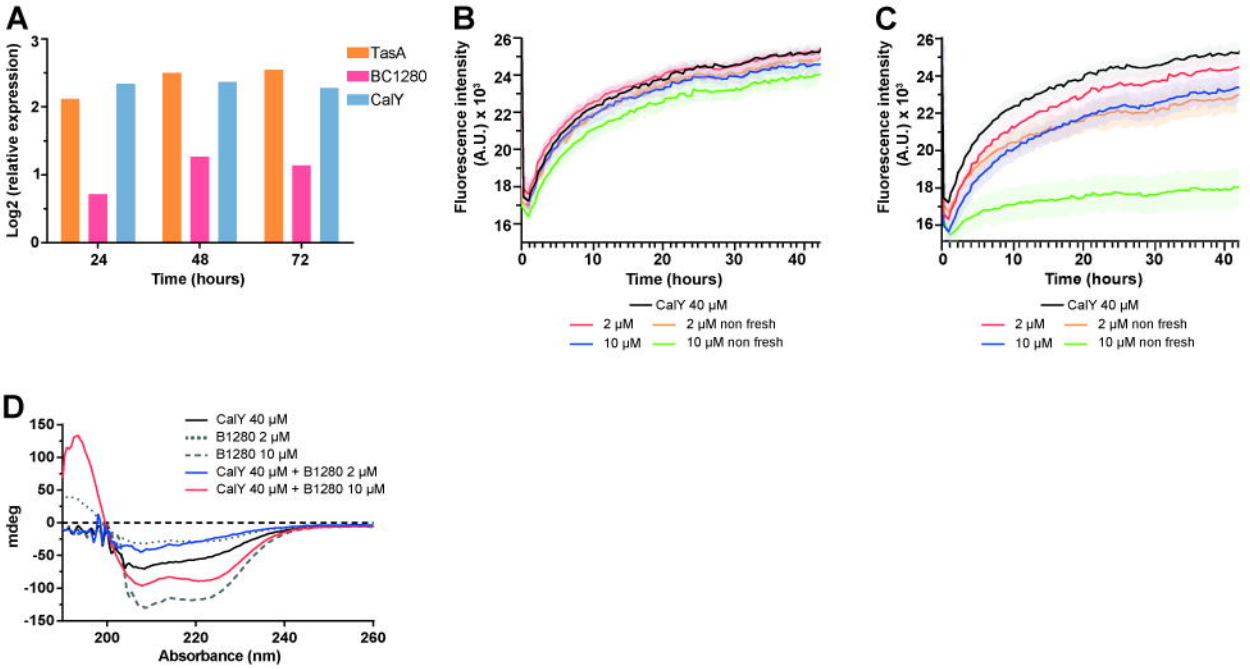

A
^{13}C chemical shift (ppm)
B
^{15}N chemical shift (ppm)
^{1}H chemical shift (ppm)
C
PDSD 50 ms
PDSD 200 ms
93E94I
99S100V
46K47V
94I95K
48E47V
115L116Q
116Q115L
48E49A
47V48E
94I93E
49A48E
47V46K
95K94I
100V99S
D
Intensity (arb. units)
d(Å)
E
BC1280
N-domain$_{39-190}$
F
Second derivative
BC1280
N-domain$_{39-190}$
Wavenumber (cm^{-1})
G
Absorbance
Experimental
Fit
1620
1634
1643
1658
1675
H
1620
1634
1646
1658
1675

Figure S6


A
Log2 (relative expression)
TasA
BC1280
CalY
Time (hours)
B
Fluorescence intensity (A.U.) x 10³
Time (hours)
CalY 40 µM
2 µM
2 µM non fresh
10 µM
10 µM non fresh
C
Fluorescence intensity (A.U.) x 10³
Time (hours)
CalY 40 µM
2 µM
2 µM non fresh
10 µM
10 µM non fresh
D
CalY 40 µM
B1280 2 µM
B1280 10 µM
CalY 40 µM + B1280 2 µM
CalY 40 µM + B1280 10 µM
mdeg
Absorbance (nm)