Molecular characterization of the N-terminal half of TasA during functional amyloid assembly and its contribution to *Bacillus subtilis* biofilm formation

Jesús Cámara-Almirón[1,#a], Laura Domínguez-García[1], Nadia El Mammeri[2#b], Alons Lends[2#c], Birgit Habenstein[2], Antonio de Vicente[1], Antoine Loquet[2], Diego Romero[1]

[1] Departamento de Microbiología, Instituto de Hortofruticultura Subtropical y Mediterránea "La Mayora", Universidad de Málaga, (Campus Universitario de Teatinos), Málaga, Spain

[2] CNRS, Chemistry and Biology of Membranes and Nanoobjects (CBMN), Institut Europeen de Chimie et Biologie (IECB), University of Bordeaux, Pessac, France

[#a]Current Address: Department of Fundamental Microbiology, Faculty of Biology and Medicine, University of Lausanne, Biophore Building, Lausanne, Switzerland

[#b]Current address: Department of Chemistry, Massachusetts Institute of Technology, 170 Albany Street, Cambridge, MA 02139

[#c]Current address: Latvian Institute of Organic Synthesis, Aizkraukles 21, Riga LV-1006, Latvia

* Corresponding author

E-mail: diego_romero@uma.es

Abstract

Biofilms are bacterial communities that result from a cell differentiation process that leads to the secretion of an extracellular matrix (ECM) by part of the bacterial population. In *Bacillus subtilis,* the main protein component of the ECM is the functional amyloid TasA, which forms a fiber-based scaffold that confers structure to the ECM. The N-terminal half of TasA is strongly conserved among *Bacillus* species and contains a protein domain, the amyloid core (AcTasA), which is critical for the formation of the amyloid architecture. In this study, we demonstrate that recombinantly purified AcTasA *in vitro* retains biochemical properties previously observed for the entire protein. Further analysis of the AcTasA amino acid sequence revealed two amyloidogenic stretches and a region of imperfect amino acid repeats, which are known to contribute to functional amyloid assembly. Biochemical characterization of these amyloidogenic stretches found in AcTasA revealed their amyloid capacity *in vitro*, contributing to the amyloid nature of AcTasA. Moreover, the study of the imperfect amino acid repeats revealed the critical role of residues D64, K68 and D69 in the structural function of TasA. *In vivo* and *in vitro* experiments with versions of TasA carrying the substitutions D64A, K68A, and D69A demonstrated a partial loss of function of the protein either in the assembly of the ECM or in the stability of the core and amyloid polymerization. Taken together, our findings allow us to better understand the polymerization process of TasA during biofilm formation and provide knowledge into the sequence determinants that promote the molecular behavior of functional amyloids.

41 # Introduction

42 Amyloids are proteins that have the ability to transition from monomers into insoluble
43 fibers that share a common quaternary structure enriched in beta-sheets stacked in the so-called
44 cross-beta architecture [1]. These proteins have traditionally been studied in the context of human
45 diseases, as several amyloid proteins and peptides are considered the causative agents of many
46 protein misfolding disorders that lead to known neurodegenerative pathologies [2,3]. However, the
47 beneficial role of amyloid proteins in many biological functions required for homeostasis is
48 currently well established, which is why part of this broad family of proteins has been subclassified
49 as functional amyloids [3]. Alternative functions reported for these proteins in bacteria are
50 antimicrobial, promotion of bacterial virulence in plants or providing a structural scaffold during
51 biofilm formation (see [4-6] for extended reviews on the subject).

52 Biofilm formation is an intrinsic ability of bacterial species that allows their association in
53 communities in which bacterial cells are closely linked to each other by the secretion of a self-
54 produced extracellular matrix (ECM) [7]. The ECM is mostly comprised of extracellular
55 polysaccharides, proteins and other molecules that allow for a more efficient interaction of the
56 microbial community with the environment and act as a shield that protects cells against external
57 stressors [7,8]. Functional amyloids and other proteinaceous filaments play a key role in the assembly
58 of these complex multicellular communities, providing mainly structural support. In species from
59 the *Bacillus* genus, the main protein component of biofilms is TasA [9], which differs between the
60 two phylogenetically distinct groups of *Bacillus*, the *subtilis* group and the *cereus* group [10]. TasA,

61 in its monomeric form, is a stable soluble globular protein that is able to transition during biofilm
62 formation into filaments that exhibit typical amyloid features (resistance to protease degradation
63 or enrichment in the beta-sheet secondary structure) [10-12]. TasA fibers have been shown to form *in*
64 *vivo*, by direct observation in biofilm samples [12], and *in vitro,* in protein extracted from biofilm
65 samples in a pre-aggregated conformation, where the aggregation properties and morphology of
66 the native filaments have been studied [13-16], or in heterologously expressed recombinant protein
67 purified from *E. coli*, where the structural properties of the protein and fibers have been
68 characterized [10,11]. The assembly of TasA is assisted by TapA, a two-domain partially disordered
69 accessory protein [17] that is also essential for biofilm formation, which promotes TasA
70 polymerization and anchors the fibers to the cell surface [10,18]. The main biological function of these
71 fibers in biofilms is structural, creating a scaffold that provides integrity to the ECM [12]. However,
72 the protein plays other biologically relevant roles, contributing to membrane stability and
73 dynamics or acting as a developmental signal that maintains a motile subpopulation within
74 biofilms [19,20].

75 The formation of the amyloid fold is not a trait exclusively restricted to amyloid proteins
76 but is more a universal property of polypeptide chains when certain physicochemical conditions
77 are achieved [1]. In amyloids, this specific folding pattern is favored, which is partly due to the amino
78 acid context where certain regions, the presence of different amino acid patterns (such as imperfect
79 repeats, especially in bacterial functional amyloids [21,22]) or the appearance of specific amino acids
80 in certain conformations determine the amyloid propensity of a given protein under physiological

81 conditions [23]. These regions have been studied extensively in amyloid proteins, as they form what
82 is known as the amyloid core of the protein, the part of the protein that contributes more to the
83 cross-beta structure and is more tightly folded [24], conferring the fibers their unique
84 physicochemical properties.

85 In previous works [10], we studied the structural properties of TasA in its fibrillar form, where
86 an amyloid core, comprising mostly the N-terminal half of the protein, was delimited. In the
87 present work, we have revealed some of the sequence determinants of this amyloid core that
88 promote the biochemical behavior of TasA. By combining bioinformatics, mutational analysis and
89 structural, biochemical, and biophysical studies, we defined specific residues critical for the
90 structural function of TasA in the ECM. All these results allow for a better comprehension of the
91 structural function of TasA during ECM production and biofilm formation and thus improve our
92 understanding of the role of functional amyloids and other filamentous proteins in the persistence
93 and interaction of bacteria in and with the environment.

94

95 # Results

96

97 ## TasA contains an N-terminal amyloid core that retains amyloid
98 properties

99 Solid-state nuclear magnetic resonance spectroscopy analysis (SSNMR) on assembled
100 recombinant TasA fibers allowed us to delimitate a rigid amyloid core (AcTasA) of 110 amino

101 acids that extends from residues K35 to K144 [10] (Fig 1 residues labeled in red). AcTasA is located
102 in the N-terminal half of the protein, which is the region of TasA that is highly conserved across
103 different *Bacillus* species (Fig 1). Based on these findings, we hypothesized a relevant contribution
104 of this region to the biochemical and biophysical properties of the protein and to the final
105 architecture of the TasA fiber and functionality. We analyzed the properties of the AcTasA
106 sequence using two complementary bioinformatic approaches. First, we searched for amino acid
107 repeats within the sequence, which have been demonstrated to be involved in the polymerization
108 process of other amyloid proteins, including bacterial functional amyloids [21,22,25,26]; second, we
109 looked for aggregation-prone regions, also known as amyloidogenic stretches, in AcTasA. These
110 tools revealed the presence of i) two amyloidogenic regions that extend from L78 to G90 (called
111 here the segment LG-13) and from D104 to G117 (called DG-14) (Fig 1, labeled, respectively, in
112 blue and green) and ii) a sequence of imperfect amino acid repeats in which the sequence
113 KDxxFxxxxxxLxxKExxxxxNxxxxKxxxGxxxx is repeated twice and extends from K35 to S101
114 (Fig 1 underlined in cyan).
115
116 **Fig 1. The N-terminal Domain of TasA Revealed Amyloidogenic Stretches and Imperfect**
117 **Amino Acid Repeats.** Sequence Comparison of the N-terminal Domain of TasA between
118 Different Species of *Bacillus*. The different features detected by bioinformatic tools are labeled
119 over the sequence. The gray colors indicate different degrees of sequence conservation. The region
120 corresponding to the amyloid core of TasA is labeled in red (from K35 to K144). The two

121 amyloidogenic stretches are labelled in blue (L78 to G90, LG-13) and green (D104, G117, DG-
122 14). The sequence underlined in cyan corresponds to the imperfect amino acid repeat in which the
123 sequence KDxxFxxxxxxLxxKExxxxxNxxxxKxxxGxxxx is repeated twice in the core. Arrows
124 indicate residues that have been analyzed by site-directed mutagenesis.
125
126 To experimentally corroborate the predicted amyloid properties of this region, AcTasA,
127 heterologously expressed in *E. coli*, was purified to homogeneity and maintained soluble in
128 monomeric form in a 1% acetic solution [10] prior to studying the polymerization on buffer at
129 physiological pH (20 mM Tris, 50 mM NaCl at pH 7). Transmission electron microscopy (TEM)
130 analysis of AcTasA in buffer after 1 week of incubation revealed filaments with a bundle-like
131 supramolecular organization that resembled those formed by the entire TasA protein (Fig 2A, top
132 and bottom micrographs, respectively). The bundle widths varied from ~5 to ~25 nm with a
133 tendency to stack themselves to form larger objects. The smallest molecular entity observable,
134 presumably protofibrils, averaged ~3.5 nm wide. Given the ability of AcTasA to aggregate and
135 form filament-like structures, we further characterized the kinetics of the aggregation process over
136 time using dynamic light scattering (DLS), a noninvasive technique that allows us to monitor the
137 size distribution of the aggregates found in the sample based on the signal intensity associated with
138 each hydrodynamic radius (R_H). The absence of major changes in the size distribution of the
139 population during the time studied and the mean R_H of ~10 nm of the population immediately after
140 buffer exchange (time 0 h), which is far above what is attributed to WT TasA [10,13], were two

141 complementary pieces of evidence suggesting that polymerization of AcTasA occurs rapidly upon
142 buffer exchange to neutral pH (Fig 2B). Two populations of similar abundance values and mean
143 R_H values of ~12 nm and ~22 nm were observed at 72 h, indicative of a transition into larger
144 aggregates. According to this dynamic, a more homogeneous size distribution of aggregates with
145 a mean R_H of ~25 nm dominated at 144 h (Fig 2B, middle and bottom graphs). Overall, this set of
146 experiments confirms that AcTasA rapidly transitions at pH 7 from monomers to aggregates that
147 exhibit an unequivocal fibrillar form. Next, we wondered whether these AcTasA assemblies
148 exhibited amyloid biochemical features similar to those of full-length WT TasA. Amyloid proteins
149 exhibit unique tinctorial properties as they bind to specific dyes. One such dye is thioflavin-T
150 (ThT), which upon binding to the beta-sheets naturally present in amyloid samples, undergoes a
151 shift in the fluorescence emission maximum that can be used to estimate the gain of beta-sheet
152 secondary structure; this occurs concomitantly with the transition from the monomeric to the
153 fibrillar form [27]. ThT binding assays of purified AcTasA showed a progressive increase in the
154 fluorescence intensity of ThT at the amyloid-specific wavelength over time in a concentration-
155 dependent manner (Fig 2C). Moreover, this ThT binding kinetics was shown in a polymerization
156 curve similar to that observed for WT TasA *in vitro* and other amyloid proteins [10,28-30].
157
158 **Fig 2. The TasA Amyloid Core Region Exhibits Amyloid Properties *In Vitro*.**
159 **(**A) Transmission electron micrographs of negatively stained AcTasA samples show dense bundles
160 of fibrillar assemblies. The scale bars are 500 nm (top) and 100 nm (bottom). (B) Aggregation

161 kinetics of AcTasA as measured by DLS at different time points. (C) Kinetics of fluorescence
162 emission at 480 nm of AcTasA samples indicates ThT binding properties. Experiments were
163 performed at different protein concentrations. Average values of three independent experiments
164 are represented. Error bars indicate the SEM. (D) Coomassie stained SDS–PAGE gel of the
165 AcTasA samples digested with proteinase K at different time points. The first lane indicates the
166 molecular marker. Each lane contains a sample corresponding to a specific digestion time.
167
168 As mentioned above, the amyloid core is the part of the protein most intensely folded in
169 the final tridimensional structure of the fiber, and therefore, it exhibits remarkable
170 physicochemical robustness, including resistance to proteases [31]. In our previous work, AcTasA
171 was detected after TasA fibrils were subjected to partial proteinase K digestion [10]. Thus, we
172 wondered if the aggregation properties of isolated AcTasA preserved the structural fold that
173 confers TasA fibers their resistance against proteolytic activity. In a similar protease digestion
174 experiment, polymerized AcTasA resisted treatment with proteinase K even 1 h after incubation,
175 as shown by the ~13 kDa band resolved in the comassie-stained SDS–PAGE gel of the digestion
176 samples compared to the untreated control (Fig 2D, last lane on the right and lane labeled with a
177 "C", respectively).
178

179 Amyloidogenic features of the N-terminal half of TasA contribute to
180 the molecular behavior of AcTasA

181 Our previous findings demonstrated that AcTasA exhibits biochemical features similar to
182 those observed for the whole TasA protein and proved that these properties are retained in the N-
183 terminal half. Thus, we asked if the features previously predicted in our *in silico* analysis (Fig 1)
184 were relevant to defining the structure or functionality of TasA. Our bioinformatic predictions of
185 amyloidogenic stretches consistently showed two regions of AcTasA, designated LG-13 and DG-
186 14, with a high tendency for aggregation (Fig 3A) (see Fig 1, labeled in blue and green,
187 respectively). To study whether these two regions contribute to biofilm formation and have any
188 impact on the structural function of TasA, we performed a mutagenesis analysis in which either of
189 these two stretches was deleted from the protein sequence. As described in a previous work [19], to
190 avoid undesirable effects when manipulating the *tasA* gene in the endogenous operon, these
191 experiments were performed using a strain lacking the whole *tapA-sipW-tasA* operon and
192 complemented with a version of the operon containing the mutated versions of *tasA* and integrated
193 in the neutral *lacA* locus. Complementation of the Δ*(tapA-sipW-tasA)* strain with an operon
194 containing a native version of TasA rescued the wrinkled phenotype associated with the WT strain
195 in solid and liquid cultures (Fig 3B, left images). However, complementation with the operon
196 containing a TasA version lacking amino acids from E82 to N88 (Δ82-88, LG-13 region) or from
197 Q108 to V116 (Δ108-116, DG-14 region) failed to restore the biofilm formation phenotype (Fig
198 3B, middle and right images). To check if these mutated versions of TasA were stable, we
199 performed a biofilm fractionation assay followed by protein precipitation and Western blot
200 analysis using anti-TasA antibodies. A strong anti-TasA reacting band was observed in samples

201 from WT or the control strains, but the signal was absent in samples of the strains carrying the
202 mutated alleles (Fig 3C). This result, along with the fact that the two strains bearing the mutated
203 versions of TasA showed the same phenotype, suggested that the deletion of these regions renders
204 the protein unstable, which most likely leads to the degradation of the protein by quality control
205 proteases.
206
207 **Fig 3. Amyloidogenic Regions are Important for Protein Folding and Contribute to the**
208 **Amyloid Properties of AcTasA.**
209 **(**A) The hexapeptide amyloidogenicity profile of AcTasA as predicted by MetAmyl reveals two
210 amyloidogenic stretches found in this sequence. (B) Mutagenesis analysis of AcTasA in which the
211 predicted amyloid stretches were deleted from the TasA sequence and the mutated alleles were
212 introduced in a Δ*(tapA-sipW-tasA)* background. The images show the phenotypes of *B. subtilis*
213 biofilms grown in solid and liquid biofilm-inducing media of the strains carrying either the native
214 allele (unmodified) or the alleles carrying the deletions. Scale bars = 1 cm. (C) Western blot using
215 an α-TasA antibody of protein extracts corresponding to the different biofilm fractions (cells,
216 medium and matrix) of the strains carrying the abovementioned alleles. Immunoblot images have
217 been cropped and spliced for illustrative purposes. Black lines over the blot images delineate
218 boundaries of immunoblot splicing. In all cases, all the slices shown were derived from a single
219 blot. (D) Top. Kinetics of ThT fluorescence emission at 480 nm of the synthetic peptides
220 corresponding to the amyloid stretch predicted. The average of three independent experiments is

221 shown. Error bars indicate the SEM. Bottom. Transmission electron microscopy micrographs of
222 negatively stained LG-13 and DG-14 synthetic peptides at the final experimental time-point. Scale
223 bar = 200 nm. (E) ^{1}D ^{13}C cross-polarization spectra of TasA fibrils (dark blue) and the assemblies
224 formed by peptides LG-13 (red) and DG14 (light blue). (F) Zoomed region of the 1D 13C cross-
225 polarization spectra shown in E between 160 and 190 ppm.
226
227 Complementarily, we studied the amyloidogenic behavior of these two amino acids
228 stretches *in vitro* using synthetic peptides. None of the peptides showed ThT binding activity (β-
229 sheet formation) during the first 20 hours of incubation. However, after this latent phase, the LG-
230 13 peptide exhibited an exponential increase in the fluorescence signal, reaching saturation a few
231 hours later (Fig 3D, blue). Transmission electron microscopy analysis at the end of the ThT
232 experiment revealed molecular entities of small size without any significant organization for the
233 DG-14 peptide (Fig 3D, left image). Abundant fibrillar assemblies made of bundles comprised of
234 several protofilaments (~5 nm diameter each) ranging between 6 and 14 nm in diameter, which is
235 a molecular organization very similar to the full-length mature TasA assemblies observed *in vitro*,
236 characterized samples of the LG-13 peptide. We further studied the rigid core of the self-assembled
237 fibrils made of LG-13 or the molecular entities of DG-14 TasA peptides by solid-state NMR
238 spectroscopy. Using magic-angle spinning NMR, we measured 1D dipolar-based ^{13}C cross-
239 polarization (CP) spectra to assess the secondary structure conformation of TasA subunits along
240 the fibril axis (Fig 3E). Both peptides showed strong and sharp ^{13}C NMR signals (Fig 3E),

241 indicating that the structural core of these fibrils is made of well-ordered and rigid subunits. We
242 compared those 1D ^{13}C spectra to that of full-length TasA fibrillated *in vitro* [10]. Overall, the
243 spectral envelopes of the three samples are quite similar, indicating a very similar secondary
244 conformation. In particular, the backbone carbonyl ^{13}C atoms accurately probe the conformation
245 of amino acids using NMR chemical shifts. Indeed, β-sheet carbonyl chemical shifts are ~173
246 ppm, while α-helical residues are found at ~178 ppm (using a DSS-based NMR scale) (Fig 3F).
247 Similar to full-length TasA, both LG-13 and DG-14 peptides form β-rich fibrils. Overall, SSNMR
248 analysis suggests that LG-13 and DG-14 are amyloidogenic segments of TasA that maintain the
249 same secondary structure conformation, and they also have the ability to self-assemble into fibrils
250 formed by well-ordered subunits associated with low structural polymorphism.
251

252 AcTasA contains an imperfect amino acid repeat with specific
253 amino acids critical for protein functionality

254 Repeated regions within amyloid proteins have been demonstrated to contribute to the
255 assembly of amyloid filaments [21,22,32], and an imperfect amino acid repeat,
256 KDxxFxxxxxxLxxKExxxxxNxxxxKxxxGxxxx, which is repeated twice, was predicted within the
257 AcTasA region (Fig 1, cyan line). To explore the contribution of this region to the structure and
258 functionality of TasA, we performed a mutational analysis by alanine scanning on selected amino
259 acids K35, D36, D64, K68, D69, F72 and G96, which showed different levels of conservation (Fig
260 1, arrows). The strain expressing the native *tasA* allele was unaffected in its ability to form a WT-

261 like biofilm in solid or liquid medium (Fig 4A). Strains expressing some of the mutated alleles
262 (K35A, D36A; F72A or G96A) also displayed the characteristic wrinkled morphology of *B.*
263 *subtilis* biofilms (Fig 4A). However, the strains that expressed the D64A, K68A, and D69A alleles
264 failed to fully restore the WT biofilm formation phenotype (Fig 4A).
265
266 **Fig 4. Mutagenesis Analysis of the Imperfect Amino Acid Repeat Reveals Amino Acids**
267 **Essential for Biofilm Formation.**
268 (A) Alanine scanning site-directed mutagenesis analysis of AcTasA in which specific amino acids
269 from the imperfect repeat region of AcTasA were substituted from the TasA sequence and the
270 mutated alleles were introduced in a Δ*(tapA-sipW-tasA)* background. The images show the
271 phenotypes of *B. subtilis* biofilms grown in solid and liquid biofilm-inducing media of the strains
272 carrying either the native allele (unmodified) or the alleles carrying the substitutions. Scale bars =
273 1 cm. (B) Western blot using an α-TasA antibody of protein extracts corresponding to the different
274 biofilm fractions (cells, medium and matrix) of the strains carrying the abovementioned alleles
275 that were affected in colony morphology or pellicle formation. Immunoblot images have been
276 cropped for illustrative purposes.
277
278 Apart from its structural role in fiber assembly and biofilm formation, the presence of TasA
279 in the cell membrane plays an important role in bacterial physiology, maintaining cell membrane
280 stability and preventing bacterial cell death [19]. Thus, we wondered if the phenotypes that we

281 observed in these two strains carrying the TasA variants could somehow be related to this
282 physiological function of TasA in the cell membrane. Confocal microscopy analysis of strains
283 expressing the *tasA* alleles that were fused to mCherry and staining with membrane dye showed a
284 homogenous distribution of the TasA-related fluorescence signal across the cell membrane except
285 in the poles, where it formed two small foci. The same distribution of TasA was observed in all
286 samples, regardless of the mutations introduced (S1 Fig); thus, we discarded any effect in the
287 observed phenotype that could be attributed to physiological alterations caused by a
288 mislocalization of TasA.

289 We then speculated that this partial loss of phenotype could be related to the instability of
290 the mutated TasA alleles due to the alanine substitutions. Biofilm fractionation assay from pellicles
291 followed by immunoblot analysis using the anti-TasA antibody (Fig 4B) showed the presence of
292 an anti-TasA reacting band in all the fractions from the WT strain, the strain carrying the native
293 TasA allele or the strain carrying the TasA D64A allele. However, a weaker signal was observed
294 in fractions of cells expressing the TasA K68A and D69A versions. In addition, two well-defined
295 reacting bands appeared in the medium and ECM fractions of samples from this strain, which
296 suggests interference of the amino acid substitution with protein processing. Nonetheless, we
297 concluded from this experiment that the TasA version of the mutants with an altered biofilm
298 formation phenotype is sufficiently stable to be detected in protein extracts; thus, the mutant
299 phenotype should not be explained by the instability and degradation of TasA. TEM analysis of
300 48 h pellicles of these two mutated strains showed a network of branched fibers decorating the

301 surfaces of cells expressing the native TasA or the TasA D64A alleles (S2 Fig, top and second
302 from the top micrograph panels). In contrast, straight fibrillar structures were scarcely observed in
303 cell preparations of the strain expressing the TasA K68, D69A allele (S2 Fig).

304 Thus, it appeared that the structural function of TasA in the formation of the fibers and
305 required for biofilm formation was compromised in this strain. We performed an *in silico* analysis
306 in which the mutated proteins were modeled and compared to the WT protein (Fig 5). The TasA
307 WT structure had to be predicted and used as a reference when comparing the variant proteins,
308 although the structure of the TasA monomer has already been experimentally determined and
309 deposited in the Protein Data Bank [11]. Differences between the modeled WT TasA predicted by
310 AlphaFold and the crystallographic model were found in the loop between K118 and Y124, which
311 was not modeled in the crystal structure, or in some regions that are modeled differently in the
312 predicted WT structure, such as the first amino acids or the absence of beta-sheet β7 (S3 Fig). We
313 represented the different protein models using the crystal structure as a reference, using the same
314 schematic representation of the different structural features of the protein that was used in another
315 study [11] (S3 Fig). When the models of the mutated versions of TasA and WT TasA were compared
316 (Fig 5A and B), some noticeable structural differences were found: i) TasA D64A showed the
317 disruption of the β6 beta-sheet by a random coil conformation adopted by residues from K126 to
318 D130 and ii) the appearance of a new alpha-helix, labeled α5', from residues G175 to T177 (Fig
319 5A). The disruption of the alpha-helix α5 was the only difference between the model of TasA
320 K68A, D69A and the TasA WT model (Fig 5B). To further analyze these structural differences,

321 we additionally modeled the possible hydrogen bonds within each protein structure. The D64
322 residue present in the loop between β2 and β3 in the WT protein forms five hydrogen bonds with
323 neighboring amino acids: two with K61 and one with F200 (Fig 5C top, left image). The
324 substitution of this residue seems to interfere with one of the hydrogen bonds formed with K61
325 (Fig 5C top). This single change seems to perturb the structure of the protein and might explain
326 the structural differences when compared to the TasA WT model. In the case of TasA K68A and
327 D69A, the K68 residue is bound to M196 from β8 through two hydrogen bonds (Fig 5C bottom).
328 The substitution of these amino acids preserves the hydrogen bonds connecting the two beta-sheets
329 by A68, but an additional bond is formed between A69 and G42. This change in the hydrogen
330 bond configuration might result in the loss of α5, whose residues are modeled as random coils in
331 the predicted structure of this TasA variant.
332
333 **Fig 5. Structure Predictions and Comparison of the WT TasA and Variant Proteins.**
334 **(**A) Structure comparison between the TasA model and the D64A allele. The numbering in the
335 scheme follows the representation described for the crystal structure of TasA in its monomeric
336 form published by Diehl et al. [11]. The coloring scheme indicates the position of the residue in the
337 sequence of the protein, where blue indicates the N-terminus and red indicates the C-terminus.
338 Warmer colors indicate proximity to the C-terminal end. (B) Structure comparison between the
339 TasA model and the K68A and D69A alleles. (C) Superimposition of the WT – D64A models

340 (top) and WT – K68A and D69A models (bottom) and a detailed view of the hydrogen bond
341 analysis of the mutated proteins.
342

343 Molecular characterization of the TasA variants revealed two 344 different alterations in the core of the mutated proteins

345 Three residues of the AcTasA core, D64, K68 and D69, are relevant for the functionality
346 of TasA, and the predicted structural changes of TasA associated with their substitution explain
347 their high degree of conservation within the AcTasA sequence (see Fig 1). Thus, we reasoned that
348 these amino acids participate in the polymerization of TasA. The two versions of TasA were
349 heterologously expressed and purified, and their molecular properties were studied *in vitro* using
350 the protocols previously optimized for TasA, TapA or AcTasA [10]. The structural changes observed
351 for these two versions of TasA were suggestive of alterations at the level of structural stability
352 compared to the WT TasA. We used circular dichroism (CD) to measure the mean residue
353 ellipticity (MRE) at 222 nm, a representative wavelength of the protein secondary structure, as a
354 function of temperature (S4 Fig) to follow the unfolding of the proteins and investigate differences
355 associated with the specific amino acid substitutions. Interestingly, the three proteins showed a
356 high degree of thermostability. The thermal scan of WT TasA showed gain of secondary structure
357 upon heating (S3 Fig), a thermal behavior that is distinct compared to that of globular proteins,
358 especially at high temperatures, where heating normally induces protein unfolding and, therefore,
359 an increase in the MRE. However, similar to what has been reported for other amyloid proteins [33],

360 a decrease in the MRE as a function of temperature was recorded instead of reaching a completely
361 unfolded state at high temperatures. TasA D64A, despite the structural alterations, exhibited a
362 similar thermal behavior as that observed for TasA WT, showing enrichment in the secondary
363 structure with increasing temperature and a higher signal at 222 nm, suggesting a lower proportion
364 of secondary structure, especially in the range of 60 °C, compared to the WT protein. Major
365 differences were evident between TasA K68A and D69A and TasA WT. A slight gain of secondary
366 structure upon heating was observed in comparison to the other versions of TasA; however, the
367 signal variations were very low compared to the other versions of TasA, indicating that the protein
368 does not suffer profound changes in secondary structure at this wavelength with increasing
369 temperature; therefore, changes in the thermal stability can be deduced.

370 To test whether these differences in structure and stability are translated into differences in
371 the amyloid behavior of protein and fiber formation, we analyzed the aggregation kinetics of the
372 two alleles using DLS. TasA D64A showed a mean R_H of ~1.6 nm at the beginning of the
373 experiment, which was slightly smaller than that reported for TasA WT [10]. However, the most
374 representative population in the TasA D64A sample reached a mean R_H between ~7.8 and ~9 nm
375 after 72 h of incubation and finally reached a mean size between ~9 and ~10 nm after 144 h (Fig
376 6A left). The kinetics of aggregation of TasA K68A and D69A were, however, substantially
377 different, given that the mean size distribution of the particles present in the sample was ~2 nm at
378 t = 0 h and remained nearly invariable over time (Fig 6A right). This finding suggested that the
379 aggregation kinetics of this protein are slower than those observed for the WT protein or the TasA

380 D64A allele, indicating a defect of TasA K68A and D69A in aggregation. We next measured the
381 characteristic beta-sheet enrichment of amyloid proteins in these molecular assemblies by ThT
382 binding experiments. The dynamics of fluorescence emission of TasA D64A (Fig 6B blue) at the
383 amyloid-specific wavelength were similar to those of TasA WT (Fig 6B gray), indicating a similar
384 beta-sheet content. However, the fluorescent signal of TasA K68A and D69A saturated earlier than
385 that of the TasA WT or TasA K68A and D69A proteins, even at the higher protein concentration
386 used in the experiment, and the maximum intensity of the signal was, overall, nearly three times
387 lower than that observed for the TasA WT or TasA D64A proteins (Fig 6B red), suggesting limited
388 beta-sheet enrichment. Morphological characterization by TEM of the molecular assemblies
389 formed by the two mutated variants showed a tendency to form large aggregates compared to TasA
390 WT (Fig 6C). The aggregates of TasA D64A were composed of associated bundles of filaments
391 of different sizes, from ~5 nm to hundreds of nanometers. Some individual fibrillar entities could
392 also be observed within these bundles, with an average size between ~1.5 and ~2.5 nm wide and
393 with variable length (Fig 6C center panel). TasA K68A and D69A, however, formed large and
394 densely packed aggregates with individual fibrillar entities that were hardly visible (Fig 6C right
395 panel). The partial proteinase-K digestion experiment showed that the assemblies formed by both
396 mutated proteins were resistant to protease activity, as indicated by the bands that were present
397 after digestion in the Coomassie-stained polyacrylamide gels (S5 Fig). Differences were, however,
398 noticeable after 20 minutes of digestion. Treatment with TasA D64A rendered resistance bands
399 between 15 and 25 kDa and between 10 and 15 kDa, which progressively disappeared over time.

400 A similar pattern was obtained for TasA K68A and D69A; however, the band between 10 and 15
401 kDa remained after 60 minutes of digestion, indicating that this variant is likely more resistant to
402 digestion than TasA D64A.
403
404 **Fig 6. *In Vitro* Analysis of the TasA Variants Reveals Two Types of Molecular Behaviors.**
405 **(**A) Aggregation kinetics of the D69A allele (left) and K68A, D69A allele (right) as measured by
406 DLS at different time points. (B) Kinetics of ThT fluorescence emission at 480 nm of the WT,
407 D64A or K68A, D69A allele and superimposition of the plots corresponding to the 30 μM protein
408 concentration. Experiments were performed at different protein concentrations. Average values of
409 three independent experiments are represented. Error bars indicate the SEM. (C) Transmission
410 electron microscopy micrographs of negatively stained samples of the WT and TasA variants.
411 White squares indicate areas of the pictures that have been zoomed in. Scale bars = 500 and 100
412 nm from top to bottom.
413
414 All these results indicate that TasA D64A is a stable protein in cells with similar
415 biochemical stability and behavior compared to TasA WT and is, in principle, not affected in its
416 ability to form aggregates enriched in beta-sheets that are resistant to proteinase degradation.
417 However, TasA K68A and D69A appear to be less stable than TasA WT and exhibit a stronger
418 tendency to form aggregates with limited amyloid features. These two distinct molecular behaviors

419 reflect the importance of these residues in the core's contribution to TasA functionality and explain
420 the two different phenotypes derived from these amino acid substitutions.

421 To further characterize the molecular conformation of TasA D64A, TasA K68A, and D69A
422 amyloid fibrils, we employed solid-state NMR spectroscopy. Recombinant D64A or K68A, D69A
423 mutant proteins were overexpressed in *E. coli* and ^{13}C-labeled following protocols previously
424 optimized [10]. *In vitro* self-assembled filaments of the mutant proteins were investigated using
425 magic-angle spinning SSNMR. We recorded a two-dimensional ^{13}C-^{13}C correlation experiment
426 using a first cross-polarization ^{1}H-^{13}C step to reveal the conformation of rigid residues involved in
427 the amyloid core. Comparison of ^{13}C-^{13}C spectral fingerprints (Fig 7A in red and green) showed
428 that D64A, K68A, D69A assemblies exhibit broader NMR signals. ^{13}C line-widths for the two
429 mutants are ~200-400 Hz (measured with full-width at half-height), suggesting a high propensity
430 of the 2 mutants to adopt polymorphic amyloid structures. For WT assemblies (Fig 7A, in black),
431 narrow peaks are observed with ^{13}C line-widths of ~100-200 Hz, indicating a low amount of
432 structural polymorphism at the local level. These line-width values are not surprising for WT
433 fibrils, since functional amyloids often lead to the aggregation of protein subunits into a unique
434 polymorphic structure characterized by well-resolved SSNMR signals [34]. The observation of much
435 broader signals for the two mutants is nevertheless not unexpected for amyloid fibril samples, since
436 pathological amyloids involved in misfolding diseases very often lead to broad NMR spectra due
437 to the presence of severe local structural polymorphism [35]. Next, we scrutinized the chemical shift
438 values for the alanine Cα-Cβ spectral region. This spectral region is helpful for probing the

439 secondary structure. Following the procedure described for WT TasA fibrils [10], we compared the
440 two mutants to WT fibrils (Fig 7B). We observed two trends for the chemical shift values of the
441 mutant fibril NMR signals: (i) several signals are conserved in their chemical shift values between
442 the two mutants and the WT sample, suggesting that these residues have a conserved local
443 conformation between the three samples. Nevertheless, (ii) numerous signals observed in the WT
444 sample disappeared for the two mutants, indicating a partial loss of the structural integrity for the
445 mutants. This observation was also made for other spectral regions of the two-dimensional
446 experiments. Altogether, this chemical shift analysis suggests that the well-ordered amyloid core
447 observed for WT TasA fibrils is only partially conserved for the two mutants, and additionally, a
448 significant number of residues in the amyloid core of D64A, K68A, D69A mutants have lost their
449 structural stability. Although the chemical shift analysis did not reveal a significant difference
450 between D64A and K68A, D69A mutants, the comparison of cross-polarization (CP) and J-based
451 (INEPT) polarization transfer showed a different behavior. While CP transfer reveals rigid residues
452 of the structural core, INEPT signals probe the highly mobile residues of the assembly. We derived
453 the CP/INEPT signal ratio from ^{1}D ^{13}C experiments (Fig 7C), and we observed that the ratio was
454 comparable for WT and D64A samples (ratio ~2) but decreased for K68A and D69A samples
455 (ratio ~1). The relative loss of CP signals for the K68A and D69A samples suggests overall more
456 dynamic assemblies, in line with the TEM observations.
457

458 **Fig 7. Recombinant WT TasA, TasA D64A or TasA K68A, D69A show Similar Structural**
459 **Fingerprints.**
460 (A) 2D [^{13}C]-[^{13}C] SSNMR experiments of recombinant TasA WT (black), TasA D64A (green)
461 and TasA K68A, D69A (red) filaments revealing the rigid residues. (B) Overlay of 2D SSNMR
462 [^{13}C]-[^{13}C] spectra of TasA WT (black), TasA D64A (green) and TasA K68A and D69A (red)
463 fibrils in the alanine Cα-Cβ spectral region. C) CP/INEPT signal ratio from ^{1}D ^{13}C experiments in
464 samples of the WT TasA, TasA D64A and TasA K68A and D69A fibrils.
465

466 *In vivo* characterization of the TasA variants demonstrates a
467 complete loss of structural function in the K68A and D69A alleles

468 Previous experiments have demonstrated that the two amino acid substitutions introduced
469 in the imperfect repeat region of AcTasA have an effect on morphology and biofilm formation and
470 have revealed that i) the D64A allele shows a molecular behavior similar to that of the WT protein,
471 which is able to self-assemble with a tendency to form dense aggregates, consisting of large
472 bundles of fibers that exhibit amyloid properties and ii) the K68A, D69A protein is unstable,
473 exhibits limited amyloid properties and a tendency toward the formation of aggregates instead of
474 fibers.
475 Despite the loss of functionality of both TasA variants compared to the WT protein, the
476 strains expressing these two versions of the protein still formed a pellicle, although it was different
477 from that of the WT strain or the strain expressing the native protein. One possible explanation for

478 this contradictory finding is the compensatory structural activity mediated by EPS, the other major
479 component of the ECM in *B. subtilis* biofilms. To test this hypothesis, the EPS operon was deleted
480 in all the strains expressing the different versions of TasA. The biofilm formation phenotype of
481 the strains carrying the native or D64A variant proteins resembled that of the Δ*epsA-O* strain alone,
482 in which a weak and fragmented pellicle floated in the liquid medium (Fig 8A). However, the
483 phenotype of the strain expressing $TasA_{K68A,\ D69A}$ mirrored that of the double Δ*tasA*, Δ*epsA-O*
484 deletion strain, showing a complete absence of the pellicle (Fig 8A right image). According to
485 these biofilm phenotypes, cells of the Δ*epsA-O*, Δ*eps-A-O* $TasA_{native}$, and Δ*epsA-O* $TasA_{D64A}$
486 strains were uniformly decorated with a network of fibers that was morphologically
487 indistinguishable between these strains (Fig 8B). The identical growth dynamics of all the strains
488 (Fig 8C) and the estimated cell density (Fig 8D) in MSgg led us to discard any growth defect of
489 the strain expressing $TasA_{K68A,\ D69A}$.

490 **Fig 8. K68A and D69A Completely Abrogated Biofilm Formation.**

491 **(**A) The images show the biofilm formation phenotypes of the Δ*epsA-O*, Δ*tasA*Δ*epsA-O*, and
492 Δ*epsA-O* strains carrying the different TasA variant alleles. (B) Transmission electron microscopy
493 micrographs of negatively stained samples from the Δ*epsA-O* or Δ*epsA-O* strain carrying the native
494 TasA allele or the D64A allele grown under biofilm-inducing conditions. White squares indicate
495 areas of the images that have been zoomed in. Scale bars = 1 μm (top) and 200 nm (bottom). (C)
496 Growth curves of the WT, Δ*epsA-O*, and Δ*epsA-O* strains carrying the different TasA variant
497 alleles in liquid MSgg medium. (D) Colony counts at different time points of the growth curve of

498 the strains WT, Δ*epsA-O*, and Δ*epsA-O* carrying the different TasA variant alleles in liquid MSgg
499 medium.

500 As previously reported [18], the mixture of Δ*tasA* and Δ*eps* strains complement themselves
501 when coinoculated in liquid medium, providing each other externally with their missing ECM
502 components and restoring the wrinkled phenotype characteristic of *B. subtilis* biofilms. The same
503 phenomenon takes place in solid medium, in which the coinoculation of both strains in a 1:1
504 proportion results in a phenotype very different from the other two mutant strains, leading to the
505 formation of wrinkles (S6 Fig, top images). The three Δ*epsA-O* strains expressing the different
506 versions of TasA showed a phenotype that resembled that of the Δ*eps* strain in solid medium (S6
507 Fig middle row images). Restoration of the wrinkly phenotype was only achieved when Δ*tasA* was
508 cocultured with Δ*epsA*-O carrying the $TasA_{native}$ or $TasA_{D64A}$ alleles (S6 Fig, bottom images),
509 demonstrating that the biofilm-defective phenotype of the Δ*epsA-O* $TasA_{K68A, D69A}$ strain is mostly
510 due to a lack of structural functionality of AcTasA and the inability of this strain to assemble a
511 functional ECM.

512

513 # Discussion

514 The N-terminal domain of TasA contains two amyloid stretches and one sequence of
515 imperfect amino acid repeats. Our study has proven that at least one of the predicted amyloidogenic
516 stretches found in AcTasA, which exhibits amyloid properties, has the potential to contribute to
517 the process of polymerization of TasA and that some highly conserved amino acids within this

518 imperfect repeat are potential determinants of the amyloid nature and the related structural
519 functionality of TasA. Unfortunately, we were unable to characterize the TasA variants lacking
520 these amyloidogenic regions, as their deletion resulted in the degradation of the protein. These
521 regions are known to be of utmost importance to explain the molecular behavior of amyloid
522 proteins [36,37], and given the ability of amyloids to self-assemble through protein–protein
523 interactions, these stretches serve novel approaches in synthetic biology and biotechnology, for
524 example, to capture free amyloid peptides or proteins or even bacteria themselves when exhibiting
525 these sequences on their surface [38]. The fact that these mutated proteins are eliminated to
526 undetectable levels and the different phenotypes of strains expressing their alleles compared to the
527 Δ*tasA* strain or any other ECM-defective strain suggest a relevant contribution to the native folding
528 features of TasA.

529 The unique amino acid substitution D64A in AcTasA was sufficient to render a mutated
530 protein that retained amyloid features but was incapable of contributing to biofilm formation. The
531 fact that this variant is able to complement biofilm formation, at least partially, when it is produced
532 in a ΔEPS background and coinoculated with a Δ*tasA* strain, suggests that when both components
533 are produced in the same cell, they are not completely functional. Indeed, this is reversed when
534 they are expressed in different cell populations. We speculate that the fibers formed by $TasA_{D64A}$,
535 although functional from a structural and biochemical point of view, modify the surface of the
536 cells, altering the deposition of the EPS in the ECM or disturbing any other putative interaction
537 between the amyloid-like filaments and the EPS. Such a phenomenon would not be unprecedented,

538 given that similar interactions between functional amyloids and polysaccharides have been
539 reported in other bacterial species. In biofilms of *Streptomyces coelicolor,* the fimbriae that
540 mediate adhesion to surfaces are composed of bundles of amyloid fibers formed by chaplins,
541 functional amyloids involved in reducing the water surface tension and promoting the development
542 of aerial structures [39]. These fimbriae are maintained attached to the cell surface via their
543 interaction with cellulose, which seems to serve as a template for the polymerization of the
544 fimbriae [40]. Moreover, a more recent work has proposed that coiled-coil domains in intermediate
545 filament-like proteins have an intrinsic affinity for cellulose in different bacterial species [41],
546 demonstrating a connection between specific domains in filamentous proteins and interactions
547 with polysaccharides.

548 The substitutions K68A and D69A, which have been characterized in this work, were
549 previously studied in which an additional function for TasA in cell physiology via its localization
550 to the cell membrane was demonstrated [19]. A strain expressing this version of TasA has a defective
551 biofilm formation phenotype, which is disconnected from the cell membrane stability required for
552 the proper physiological function of TasA [19]. These amino acids are well conserved in the AcTasA
553 region of imperfect repeats, and contrary to the other TasA variant investigated, their substitution
554 affects the amyloid nature and stability of the entire protein, compromising its structural function
555 in biofilm formation. Imperfect amino acid repeats are a common theme in many functional and
556 nonfunctional amyloids [21,22,25,32,42,43], a reason for their use for the localization of important

557 residues for protein fibrillation. Indeed, the alteration of these regions can render altered amyloid
558 behavior, either inhibiting or improving the process of amyloidogenesis.

559 A recent work has provided insight into the molecular structure of native TasA filaments
560 purified from *B. subtilis* biofilms with atomic resolution [16]. It was proposed that, as a monomer,
561 the unstructured N-terminal end of the protein (amino acids from A28 to S41) along with the β1
562 sheet undergoes a conformational change that results in these residues extending away from the
563 monomer folding into another beta-sheet that complements the next subunit in a process known as
564 donor strand complementation and the refolding of β3. Based on this finding, the authors propose
565 that the structural features of the native TasA filaments fit better in the category of amyloid-like
566 proteins, as these fibers are neither a linear arrangement of globular subunits nor is their structure
567 representative of the cross-β architecture typical of amyloid proteins. Nonetheless, it has been
568 previously demonstrated that recombinant TasA *in vitro* shows the typical cross-β architecture in
569 X-ray diffraction experiments, clearing any doubts regarding this question, at least concerning
570 recombinant TasA. In addition, native filaments purified directly from *B. subtilis* biofilms, even if
571 purified to homogeneity, may be affected by the presence of the accessory protein TapA in the
572 cells [18]; therefore, caution must be exercised. With the data that are available currently, the core
573 region of TasA fits rather well in the proposed model in which the N-terminal domain of TasA is
574 involved in donating the strand required for the assembly of the subunits and for the interactions
575 that stabilize this strand. It is tempting to speculate that the studied amino acid substitutions in
576 D64, K68, D69 participate in this process of stabilization of the donor strand; therefore, their

577 variation, especially in K68A, D69A, interferes with this process, affecting the assembly of the
578 filaments.

579 Taken together, our findings, along with those from previously published works, lead us to
580 a better understanding of the polymerization process of TasA during biofilm formation and provide
581 knowledge into the sequence determinants that promote the molecular behavior of functional
582 amyloids.

583

584 # Materials and methods

585

586 ## Bacterial strains and culture conditions

587 The bacterial strains used in this study are listed in S1 Table. Bacterial cultures were grown
588 at 37 °C from frozen stocks on Luria-Bertani (LB: 1% tryptone (Oxoid), 0.5% yeast extract
589 (Oxoid) and 0.5% NaCl) plates. Isolated bacteria were inoculated in appropriate medium. Biofilm
590 assays were performed on MSgg medium: 100 mM morpholinepropane sulfonic acid (MOPS) (pH
591 7), 0.5% glycerol, 0.5% glutamate, 5 mM potassium phosphate (pH 7), 50 μg/ml tryptophan, 50
592 μg/ml phenylalanine, 50 μg/ml threonine, 2 mM $MgCl_2$, 700 μM $CaCl_2$, 50 μM $FeCl_3$, 50 μM
593 $MnCl_2$, 2 μM thiamine, and 1 μM $ZnCl_2$. For *in vitro* cell complementation assays with cell-free
594 supernatants, the medium optimized for lipopeptide production (MOLP) was used and prepared as
595 previously described. For cloning and plasmid replication, *Escherichia coli* DH5α was used. For
596 protein purification, *Escherichia coli* BL21(DE3) was used. *Bacillus subtilis* 168 is a naturally

597 competent domesticated strain used as a first step to transform the different constructs into *Bacillus*
598 *subtilis* NCIB3610 by SPP1 phage-mediated generalized transduction. The final antibiotic
599 concentrations for *B. subtilis* were MLS (1 μg/ml erythromycin, 25 μg/ml lincomycin),
600 spectinomycin (100 μg/ml), tetracycline (10 μg/ml), chloramphenicol (5 μg/ml), and kanamycin
601 (10 μg/ml). For the selection of plasmids in *E. coli*, ampicillin at 100 μg/ml was used.
602

603 Plasmid and strain construction

604 TasA was purified using the pDFR6 (pET22b-*tasA*) plasmid containing the ORF of TasA
605 without the signal peptide or the stop codon, which was constructed as previously described. To
606 purify $TasA_{D64A}$ and $TasA_{K68A, D69A}$, the ORF of the mutated alleles excluding the signal peptide
607 were amplified from strains JC78 and JC81 using the primers TasA_Exp_C_NdeI_F and
608 TasA_Exp_C_XhoI_R. The resulting PCR products were then cloned into pET22b by digestion
609 with NdeI and XhoI, and the final plasmids were maintained in *E. coli* (strains JC104 and JC106,
610 respectively).
611 The region corresponding to the amyloid core of TasA was amplified from the *tasA* ORF
612 using primers TasA_Exp_C_NdeI_F and TasA_Exp_C_XhoI_R. The resulting PCR product was
613 digested using NdeI and XhoI and cloned into pET22b. The final plasmid was maintained in *E.*
614 *coli* (strain JC118).
615 Strains JC72, JC75, JC76, JC80 and JC82 were generated by site-directed mutagenesis
616 using the primer pairs del82-88, del82-88-antisense, del108-116, del108-116-antisense

617 KD_AA_35-36, KD_AA_35-36_antisense, F_A_72, F_A_72_antisense, E_A_82,
618 E_A_82_antisense, G_A_96, G_A_96_antisense and D_A_64, D_A_64_antisense with a
619 QuickChange Lightning Site Directed Mutagenesis Kit (Agilent Technologies), following the
620 manufacturer's instructions.

621 Strains JC209, JC222 and JC224 containing the mCherry fusions of the native protein or
622 the $TasA_{K68A,\ D69A}$ or $TasA_{D64A}$ variants, respectively, were generated using the oligos op-tasA_fw,
623 op-tasA_rv, mCherry_fw, mCherry_rv, tasA_downstream_fw, and tasA_downstream_rv and the
624 NEBuilder HiFi DNA Assembly Cloning Kit (New England Biolabs).

625 All of the *B. subtilis* strains generated were constructed by transforming *B. subtilis* 168 via
626 its natural competence and then using the positive clones as donors for transferring the constructs
627 into *B. subtilis* NCIB3610 via generalized SPP1 phage transduction [44]

628

629 Bioinformatic analysis of the TasA sequence

630 For the prediction of amyloidogenic regions within the N-terminal region of the TasA
631 protein sequence, we used different sequence-based tools that are freely available online and
632 applied different computational methods [45]. MetAmyl [46] and AmylPred 2.0 [47] were used as
633 consensus predictors, as both applications integrate the results provided by different online tools
634 that predict protein aggregation or amyloidogenicity based on sequence analysis. APPNN [48] and
635 FISH amyloid [49] were used as predictors based on machine-learning approaches. The consensus
636 regions obtained from the different methods were selected for further experimental analyses.

637 The analysis of imperfect amino acid repeats present within the TasA sequence was
638 performed using the webtool RADAR [50], available at the EMBL-EBI website.
639 The structure prediction of WT TasA, $TasA_{D64A}$ and $TasA_{K68A, D69A}$ was performed using
640 the crystal structure of the TasA monomer as a template. Structures were predicted using
641 AlphaFold [51] in a Google Colab notebook (ColabFold) [52]. Protein models were visualized and
642 analyzed using UCSF ChimeraX (v1.3) [53].
643

644 **Biofilm formation assays**

645 To analyze biofilm formation and the colony architecture on plates and pellicle formation
646 in liquid medium, the different strains were grown in LB plates at 37 °C overnight. Then,
647 suspensions of the corresponding strains were made in PBS, and these were adjusted to an OD_{600}
648 of 1. To analyze biofilm formation on the plates, 2 µl drops of the bacterial suspensions were
649 spotted onto the surface of MSgg plates that were incubated at 30 °C. Pictures were taken at 24,
650 48 or 72 hours to observe the colony architecture at different time points. To analyze the formation
651 of pellicles in the air liquid interface of liquid media, 10 µl of each bacterial suspension was added
652 to 1 ml of liquid MSgg medium in 24-well plates. Then, the plates were incubated at 30 °C without
653 agitation or aeration for 48 hours.
654

655 **Biofilm fractionation**

656 The presence of TasA was analyzed in the studied strains by fractioning the biofilms in
657 medium, cells or ECM as previously described [9]. To have sufficient biological material, one 24-
658 well plate was used for each bacterial strain, of which only the 8 central wells were used. Briefly,
659 pellicles were carefully lifted from the wells and placed in a separate tube containing 8 ml of MS
660 medium (MSgg without glycerol and glutamate). The remaining spent medium of the well was
661 filtered sterilized and kept at 4 °C for further analysis. Then, to separate the cells from the ECM,
662 the pellicles were subjected to mild sonication (approximately 10 pulses of 30 s in a Branson 450
663 digital sonifier, with 30 s pauses in ice), and the resulting suspensions were centrifuged at 9000 x
664 g to separate the ECM from the cells. The supernatant (ECM fraction) was filtered, sterilized and
665 kept at 4 °C. The cell fraction was resuspended in 8 ml of MS medium and kept at 4 °C until further
666 processing.
667

668 **Synthetic peptides**

669 Peptides corresponding to the amyloidogenic regions detected within the TasA amyloid
670 core were synthetically produced and purified by Proteogenix (Schiltigheim, France). The
671 sequence of the peptides was: for peptide LG-13 NH_2-LAIKEVMALNYG-COOH and for peptide
672 DG-14 NH_2-DFLSQFEVTLLTVG-COOH. A stock solution of each peptide was prepared in
673 DMSO. Then, the peptides were diluted in Tris 20 mM and NaCl 50 mM to prepare the working
674 solution for the different experiments at the appropriate concentration.
675

676 ## Protein precipitation

677 To precipitate the proteins of each biofilm fraction, 2 ml of the medium and ECM fractions
678 were directly treated with 10% trichloroacetic acid (TCA) and incubated on ice for 1 h. For the
679 cell fractions, 2 ml of the suspension was treated with 0.1 mg/ml lysozyme and incubated at 37 °C
680 for 30 min. Then, the lysozyme-treated suspension was subjected to TCA treatment as described
681 above. After incubation, proteins were collected by centrifugation at 13000 x g for 20 min at 4 °C.
682 The pellets were washed twice with cold acetone and dried by air. The proteins were resolubilized
683 in 8 M urea, 100 mM Tris, and 150 mM NaCl by incubating at 100 °C for 5 mins.

684

685 ## Protein expression and purification

686 Protein was expressed and purified as previously described [10]. Briefly, *E. coli* BL21(DE3)
687 cells were freshly transformed with the pET22b plasmid containing either WT TasA (pDFR6
688 plasmid), the different mutated versions of the protein (D64A, and K68A, D69A) or the amyloid
689 core region. Colonies were selected from the plates, resuspended in 10 mL of LB with 100 µg/mL
690 ampicillin and incubated overnight at 37 °C with shaking. This preinoculum was then used to
691 inoculate 500 mL of LB + ampicillin, and the culture was incubated at 37 °C until an OD600 of
692 0.7–0.8 was reached. Next, the culture was induced with 1 mM isopropyl β-D-1-
693 thiogalactopyranoside (IPTG) and incubated O/N at 30 °C with shaking to induce the formation of
694 inclusion bodies. After that, cells were harvested by centrifugation (5000 × g, 15 min, 4 °C),
695 resuspended in buffer A (Tris 50 mM, 150 mM NaCl, pH 8), and then centrifuged again. These

696 pellets were stored frozen at −80 °C until use. After thawing, cells were resuspended in buffer A
697 and broken down by sonication on ice using a Branson 450 digital sonifier (3 × 45 s, 60%
698 amplitude). After sonication, the lysates were centrifuged (15,000 × g, 60 min, 4 °C), and the
699 supernatant was discarded, as proteins were mainly expressed in inclusion bodies. The
700 proteinaceous pellet was resuspended in buffer A supplemented with 2% Triton X-100, incubated
701 at 37 °C with shaking for 20 min to further eliminate any remaining cell debris, and centrifuged
702 (15,000 × g, 10 min, 4 °C). The pellet was then extensively washed with buffer A (37 °C, 2 h),
703 centrifuged (15,000 × g for 10 min, 4 °C), resuspended in denaturing buffer (Tris 50 mM NaCl
704 500 mM, 6 M GuHCl), and incubated at 60 °C overnight to completely solubilize the inclusion
705 bodies. Lysates were clarified via sonication on ice (3 × 45 s, 60% amplitude) and centrifugation
706 (15,000 × g, 1 h, 16 °C) and were then passed through a 0.45-µm filter prior to affinity
707 chromatography. Proteins were purified using an AKTA Start FPLC system (GE Healthcare). The
708 lysates were loaded into a HisTrap HP 5 mL column (GE Healthcare) previously equilibrated with
709 binding buffer (50 mM Tris, 0.5 M NaCl, 20 mM imidazole, 8 M urea, pH 8). Protein was eluted
710 from the column with elution buffer (50 mM Tris, 0.5 M NaCl, 500 mM imidazole, 8 M urea, pH
711 8). After the affinity chromatography step, the buffer was exchanged with 1% acetic acid at pH 3
712 and 0.02% sodium azide by using a HiPrep 26/10 desalting column (GE Healthcare). This ensured
713 that the proteins were maintained in their monomeric form. The purified proteins were stored under
714 these conditions at 4 °C (maximum 1 month) until further use.
715

716 ## Assembly of TasA filaments

717 To assemble the TasA filaments *in vitro* from the corresponding monomeric purified
718 proteins, first, the acidic buffer was removed and exchanged with Tris 20 mM NaCl 50 mM using
719 FPLC and a HiPrep 26/10 desalting column as described above. Filaments were assembled with a
720 protein concentration above 1 mg/ml and incubated at 30 °C in standing conditions for one week.

721

722 ## SDS–PAGE and Western blot

723 Protein samples from the different assays were diluted in 2x Laemmli sample buffer (Bio-
724 Rad) and heated at 100 °C for 5 min. Proteins were separated by SDS–PAGE in 12% acrylamide
725 gels and then transferred onto PVDF membranes using a Trans-Blot Turbo Transfer System (Bio-
726 Rad) and PVDF transfer packs (Bio-Rad). For immunodetection of TasA, the membranes were
727 incubated with an anti-TasA antibody (rabbit) at a 1:20,000 dilution in Pierce Protein-Free (TBS)
728 blocking buffer (Thermo Fisher). A secondary anti-rabbit IgG antibody conjugated to horseradish
729 peroxidase (Bio-Rad) was used at a 1:3000 dilution in the same buffer. The membranes were
730 developed using Pierce ECL Western blotting Substrate (Thermo Fisher).

731

732 ## Transmission electron microscopy

733 To visualize the fibers formed by the different TasA variants, the assembled protein
734 samples were deposited over carbon-coated copper grids and incubated for 1 hour. Next, the excess
735 sample was bloated off the grid, and the samples were negatively stained by floating the grids in a

736 2% uranyl acetate solution for 30 s and then washed by floating the grid in distilled water for 30
737 seconds.

738 To visualize the filaments formed by the synthetic peptides corresponding to the
739 amyloidogenic regions of the amyloid core of TasA, the samples were taken from the microplates
740 used in the thioflavin-T binding experiments after saturation of the signal (36 h of incubation) and
741 were deposited in grids and negatively stained as described above.

742 To study the morphology of fibers on the surface of cells, pellicles of the different strains
743 were grown as described above. After 48 h of incubation, carbon-coated copper grids were
744 deposited into the wells over the fully formed pellicles and incubated for 2 h at 30 °C in standing
745 conditions. After incubation, the grids were washed twice with distilled water for 30 s, negatively
746 stained with 1% uranyl acetate for 20 s and washed again once with distilled water for 30 s.

747 All the samples were visualized in an FEI Tecnai G2 20 TWIN transmission electron
748 microscope at an accelerating voltage of 80 kV. The images were taken using a side-mounted CCD
749 Olympus Veleta with 2k x 2k Mp.

750

751 **Confocal laser scanning microscopy**

752 The localization of TasA was studied using a TasA-mCherry translational fusion in a Δ*tasA*
753 genetic background (see S1 Table). Strains JC209, JC222 and JC224 containing the WT protein
754 fused to mCherry or the fusion of the different variants, respectively, were grown in solid medium
755 under biofilm-inducing conditions, as described above, and the different colonies were taken at 48

756 h and resuspended in distilled water using a 25 $^{5/8}$ G needle. Bacterial suspensions were treated
757 with CellBrite Fix 488 Membrane Stain (Biotium, stock solution at 1000X) at a final concentration
758 of 1X. Images of the stained bacteria were acquired sequentially to obtain images from the
759 membrane stain and mCherry fluorescence. The CellBrite image was acquired by exciting the dye
760 at 488 nm and recording the emissions from 498 to 553 nm, followed by a second acquisition
761 corresponding to the mCherry fluorescence with excitation at 561 nm and recording of the
762 emissions from 572 to 665 nm.

763 All images were obtained by visualizing the samples using an inverted Leica SP5 system
764 with a 63x NA 1.4 HCX PL APO oil-immersion objective. For each experiment, the laser settings,
765 scan speed, PMT or HyD detector gain, and pinhole aperture were kept constant for all of the
766 acquired images.

767

768 Thioflavin-T binding assays

769 To ensure that the assay is performed when TasA is in monomeric form, the purified
770 proteins or the amyloid core were buffer exchanged and immediately used afterward in 20 mM
771 Tris and 50 mM NaCl. The assay was performed in 96-well microplates. TasA or the different
772 variants were mixed at different final concentrations of 7.5, 15 or 30 μM with 30 μM thioflavin-T
773 in a final volume of 200 μl. In the experiments with the amyloid core region, the purified core was
774 mixed at a final concentration of 15, 30 or 60 μM with thioflavin-T under the same conditions.

775 Measurements were performed at 30 °C every 30 min in a Fluostar Omega plate reader (BMG
776 Labtech) equipped with 450 nm and 480 nm filters for excitation and emission, respectively.
777 For the experiments performed with peptide samples, working solutions of both peptides
778 were prepared in 20 mM Tris and 50 mM NaCl, and the peptides were used at a final concentration
779 of 200 μM.
780

781 **Dynamic light scattering experiments**

782 Freshly purified TasA, TasA amyloid core or the different TasA variants were buffer
783 exchanged to 50 mM Tris and 20 mM NaCl containing 0.02% sodium azide immediately prior to
784 the experiment. The final concentration was adjusted to 30 μM for all the samples, and these were
785 incubated at 30 °C in standing conditions. Samples were taken at different time points to measure
786 particle size in a Zetasizer Nano ZS (Malvern) equipped with a 632.8 nm laser as the excitation
787 source and using 1 cm path length polystyrene cuvettes. Zetasizer software version 8.01 (Malvern)
788 was used for data analysis.
789

790 **Circular dichroism measurements**

791 WT TasA or the different variants were freshly purified and buffer exchanged to 20 mM
792 phosphate buffer at pH 7 containing 0.02% sodium azide. For the thermal denaturation
793 experiments, changes in ellipticity at 222 nm were recorded as a function of the increasing
794 temperature. The thermal denaturation spectra were recorded by measuring the spectra of all the

795 samples in 1 mm quartz cuvettes between 210 and 230 nm, with temperature steps of 10 °C and
796 incubation times of 5 minutes between each step. Measurements were performed with the
797 following parameters: resolution of 1 nm, bandwidth of 1 nm, sensitivity of 100 mdeg and 5
798 accumulations.
799

800 **SSNMR spectroscopy**

801 All experiments were recorded at 600 MHz on a Bruker spectrometer equipped with a triple
802 resonance 3.2 mm probe using a magic-angle spinning frequency of 11 kHz. The sample
803 temperature was set to 278 K, and chemical shifts were calibrated using DSS as an internal
804 reference. A decoupling strength of 90 kHz was used with small phase incremental alternation
805 with 64 steps (SPINAL-64). 1D ^{13}C experiments on LG-13 and DG-14 were recorded using a
806 cross-polarization contact time of 1 ms and 20k scans. 2D ^{13}C-^{13}C experiments were recorded
807 using a proton-driven spin diffusion mixing time of 50 ms.
808

809 **Partial proteinase-K digestion experiments**

810 Experiments were performed as previously described [10,54]. Briefly, 50 µg of assembled WT
811 TasA, the different variants or the amyloid core were incubated for 1 h at 37 °C with 0.02 mg/ml
812 proteinase K in 20 mM Tris pH 8 and 50 mM NaCl, and samples were taken at different time
813 points for further analysis (1, 5, 15, 20, 30, 45 or 60 min). Reactions were stopped in each sample

814 by the addition of 1 vol. Laemmli buffer and heated at 100 °C for 5 min. Then, samples were
815 separated by SDS–PAGE and stained with Coomassie Brilliant Blue.
816

817 Acknowledgments

818 We thank Juan Félix López Téllez from the Nanoimaging Unit of the Plataforma Bionand,
819 Instituto de Investigación Biomédica de Málaga (IBIMA), for his technical assistance in the
820 transmission electron microscopy imaging and analysis. We thank José Luis Zafra Paredes and
821 José María Montenegro Martos from the Vibrational Spectroscopy unit and the Electronic
822 Spectroscopy unit, respectively, of the Servicios Centrales de Apoyo a la Investigación (SCAI)
823 from the University of Málaga for their technical assistance with the circular dichroism and
824 dynamic light scattering experiments. We thank David Vela Corcía for his help with the protein
825 modelling and hydrogen bond analysis. We also thank Estelle Morvan, Axelle Grélard and the
826 Biophysical and Structural Chemistry Platform at IECB, CNRS UAR 3033, INSERM US001,
827 University of Bordeaux.
828

829 Author contributions

830 D. Romero conceived the study. D. Romero, and J. Cámara-Almirón. designed the
831 experiments. J. Cámara-Almirón performed the main experimental work. L. Domínguez-García.
832 performed the experiments with the AcTasA and provided assistance with the experimental work.
833 A. Loquet and B. Habenstein designed the SSNMR spectroscopy experiments. N. El Mammeri

and A. Lends carried out the SSNMR spectroscopy experiments. J. Cámara-Almirón, D. Romero, A. Loquet and N. El Mammeri wrote the manuscript. L. Domínguez-García, N. El Mammeri, A. Loquet, B. Habenstein, A. Lends, and A. de Vicente contributed critically to writing the final version of the manuscript.

838 # References

839 1. Gallardo, R., Ranson, N.A., and Radford, S.E. (2020). Amyloid structures: much more than
840 just a cross-beta fold. Curr Opin Struct Biol *60*, 7-16. 10.1016/j.sbi.2019.09.001.

841 2. Ke, P.C., Zhou, R., Serpell, L.C., Riek, R., Knowles, T.P.J., Lashuel, H.A., Gazit, E.,
842 Hamley, I.W., Davis, T.P., Fandrich, M., et al. (2020). Half a century of amyloids: past,
843 present and future. Chem Soc Rev *49*, 5473-5509. 10.1039/c9cs00199a.

844 3. Hassan, M.N., Nabi, F., Khan, A.N., Hussain, M., Siddiqui, W.A., Uversky, V.N., and
845 Khan, R.H. (2022). The amyloid state of proteins: A boon or bane? Int J Biol Macromol
846 *200*, 593-617. 10.1016/j.ijbiomac.2022.01.115.

847 4. Alvarez-Mena, A., Camara-Almiron, J., de Vicente, A., and Romero, D. (2020).
848 Multifunctional amyloids in the biology of gram-positive bacteria. Microorganisms *8*,
849 2020. 10.3390/microorganisms8122020.

850 5. Levkovich, S.A., Gazit, E., and Bar-Yosef, D.L. (2021). Two decades of studying
851 functional amyloids in microorganisms. Trends Microbiol *29*, 251-265.
852 10.1016/j.tim.2020.09.005.

853 6. Salinas, N., Povolotsky, T.L., Landau, M., and Kolodkin-Gal, I. (2020). Emerging roles of
854 functional bacterial amyloids in gene regulation, toxicity, and immunomodulation.
855 Microbiol Mol Biol Rev *85*, e00062-00020. 10.1128/MMBR.00062-20.

856 7. Yin, W., Wang, Y., Liu, L., and He, J. (2019). Biofilms: the microbial "protective clothing"
857 in extreme environments. Int J Mol Sci *20*, 3423. 10.3390/ijms20143423.

858 8. Yan, J., and Bassler, B.L. (2019). Surviving as a Community: Antibiotic Tolerance and
859 Persistence in Bacterial Biofilms. Cell Host Microbe *26*, 15-21.
860 10.1016/j.chom.2019.06.002.
861 9. Branda, S.S., Chu, F., Kearns, D.B., Losick, R., and Kolter, R. (2006). A major protein
862 component of the *Bacillus subtilis* biofilm matrix. Mol Microbiol *59*, 1229-1238.
863 10.1111/j.1365-2958.2005.05020.x.
864 10. El Mammeri, N., Hierrezuelo, J., Tolchard, J., Camara-Almiron, J., Caro-Astorga, J.,
865 Alvarez-Mena, A., Dutour, A., Berbon, M., Shenoy, J., Morvan, E., et al. (2019). Molecular
866 architecture of bacterial amyloids in Bacillus biofilms. FASEB J *33*, 12146-12163.
867 10.1096/fj.201900831R.
868 11. Diehl, A., Roske, Y., Ball, L., Chowdhury, A., Hiller, M., Moliere, N., Kramer, R.,
869 Stoppler, D., Worth, C.L., Schlegel, B., et al. (2018). Structural changes of TasA in biofilm
870 formation of Bacillus subtilis. Proc Natl Acad Sci U S A *115*, 3237-3242.
871 10.1073/pnas.1718102115.
872 12. Romero, D., Aguilar, C., Losick, R., and Kolter, R. (2010). Amyloid fibers provide
873 structural integrity to Bacillus subtilis biofilms. Proc Natl Acad Sci U S A *107*, 2230-2234.
874 10.1073/pnas.0910560107.
875 13. Chai, L., Romero, D., Kayatekin, C., Akabayov, B., Vlamakis, H., Losick, R., and Kolter,
876 R. (2013). Isolation, characterization, and aggregation of a structured bacterial matrix
877 precursor. J Biol Chem *288*, 17559-17568. 10.1074/jbc.M113.453605.

878 14. Azulay, D.N., Ghrayeb, M., Ktorza, I.B., Nir, I., Nasser, R., Harel, Y.S., and Chai, L.
879 (2020). Colloidal-like aggregation of a functional amyloid protein. Phys Chem Chem Phys
880 *22*, 23286-23294. 10.1039/d0cp03265d.
881 15. Ghrayeb, M., Hayet, S., Lester-Zer, N., Levi-Kalisman, Y., and Chai, L. (2021). Fibrilar
882 polymorphism of the bacterial extracellular matrix protein TasA. Microorganisms *9*, 529.
883 10.3390/microorganisms9030529.
884 16. Bohning, J., Ghrayeb, M., Pedebos, C., Abbas, D.K., Khalid, S., Chai, L., and Bharat,
885 T.A.M. (2022). Donor-strand exchange drives assembly of the TasA scaffold in *Bacillus*
886 *subtilis* biofilms. Nat Commun *13*, 7082. 10.1038/s41467-022-34700-z.
887 17. Abbasi, R., Mousa, R., Dekel, N., Amartely, H., Danieli, T., Lebendiker, M., Levi-
888 Kalisman, Y., Shalev, D.E., Metanis, N., and Chai, L. (2019). The Bacterial Extracellular
889 Matrix Protein TapA Is a Two-Domain Partially Disordered Protein. Chembiochem *20*,
890 355-359. 10.1002/cbic.201800634.
891 18. Romero, D., Vlamakis, H., Losick, R., and Kolter, R. (2011). An accessory protein required
892 for anchoring and assembly of amyloid fibres in *B. subtilis* biofilms. Mol Microbiol *80*,
893 1155-1168. 10.1111/j.1365-2958.2011.07653.x.
894 19. Camara-Almiron, J., Navarro, Y., Diaz-Martinez, L., Magno-Perez-Bryan, M.C., Molina-
895 Santiago, C., Pearson, J.R., de Vicente, A., Perez-Garcia, A., and Romero, D. (2020). Dual
896 functionality of the amyloid protein TasA in Bacillus physiology and fitness on the
897 phylloplane. Nat Commun *11*, 1859. 10.1038/s41467-020-15758-z.

898 20. Steinberg, N., Keren-Paz, A., Hou, Q., Doron, S., Yanuka-Golub, K., Olender, T., Hadar,
899 R., Rosenberg, G., Jain, R., Camara-Almiron, J., et al. (2020). The extracellular matrix
900 protein TasA is a developmental cue that maintains a motile subpopulation within *Bacillus*
901 *subtilis* biofilms. Sci Signal *13*, eaaw8905. 10.1126/scisignal.aaw8905.
902 21. Szulc, N., Gasior-Glogowska, M., Wojciechowski, J.W., Szefczyk, M., Zak, A.M.,
903 Burdukiewicz, M., and Kotulska, M. (2021). Variability of amyloid propensity in imperfect
904 repeats of CsgA protein of *Salmonella enterica* and *Escherichia coli*. Int J Mol Sci *22*,
905 5127. 10.3390/ijms22105127.
906 22. Rasmussen, C.B., Christiansen, G., Vad, B.S., Lynggaard, C., Enghild, J.J., Andreasen, M.,
907 and Otzen, D. (2019). Imperfect repeats in the functional amyloid protein FapC reduce the
908 tendency to fragment during fibrillation. Protein Sci *28*, 633-642. 10.1002/pro.3566.
909 23. Knowles, T.P., Vendruscolo, M., and Dobson, C.M. (2014). The amyloid state and its
910 association with protein misfolding diseases. Nat Rev Mol Cell Biol *15*, 384-396.
911 10.1038/nrm3810.
912 24. Ulamec, S.M., Brockwell, D.J., and Radford, S.E. (2020). Looking Beyond the Core: The
913 Role of Flanking Regions in the Aggregation of Amyloidogenic Peptides and Proteins.
914 Front Neurosci *14*, 611285. 10.3389/fnins.2020.611285.
915 25. Harada, R., Kobayashi, N., Kim, J., Nakamura, C., Han, S.W., Ikebukuro, K., and Sode,
916 K. (2009). The effect of amino acid substitution in the imperfect repeat sequences of alpha-

917 synuclein on fibrillation. Biochim Biophys Acta *1792*, 998-1003.
918 10.1016/j.bbadis.2009.06.010.
919 26. Wang, X., and Chapman, M.R. (2008). Sequence determinants of bacterial amyloid
920 formation. J Mol Biol *380*, 570-580. 10.1016/j.jmb.2008.05.019.
921 27. Xue, C., Lin, T.Y., Chang, D., and Guo, Z. (2017). Thioflavin T as an amyloid dye: fibril
922 quantification, optimal concentration and effect on aggregation. R Soc Open Sci *4*, 160696.
923 10.1098/rsos.160696.
924 28. Horvath, I., Weise, C.F., Andersson, E.K., Chorell, E., Sellstedt, M., Bengtsson, C.,
925 Olofsson, A., Hultgren, S.J., Chapman, M., Wolf-Watz, M., et al. (2012). Mechanisms of
926 protein oligomerization: inhibitor of functional amyloids templates alpha-synuclein
927 fibrillation. J Am Chem Soc *134*, 3439-3444. 10.1021/ja209829m.
928 29. Perov, S., Lidor, O., Salinas, N., Golan, N., Tayeb-Fligelman, E., Deshmukh, M., Willbold,
929 D., and Landau, M. (2019). Structural Insights into Curli CsgA Cross-beta Fibril
930 Architecture Inspire Repurposing of Anti-amyloid Compounds as Anti-biofilm Agents.
931 PLoS Pathog *15*, e1007978. 10.1371/journal.ppat.1007978.
932 30. Yang, W., Willemse, J., Sawyer, E.B., Lou, F., Gong, W., Zhang, H., Gras, S.L., Claessen,
933 D., and Perrett, S. (2017). The propensity of the bacterial rodlin protein RdlB to form
934 amyloid fibrils determines its function in Streptomyces coelicolor. Sci Rep *7*, 42867.
935 10.1038/srep42867.

936 31. Kushnirov, V.V., Dergalev, A.A., and Alexandrov, A.I. (2020). Proteinase K resistant
937 cores of prions and amyloids. Prion *14*, 11-19. 10.1080/19336896.2019.1704612.
938 32. Cherny, I., Rockah, L., Levy-Nissenbaum, O., Gophna, U., Ron, E.Z., and Gazit, E. (2005).
939 The formation of *Escherichia coli* curli amyloid fibrils is mediated by prion-like peptide
940 repeats. J Mol Biol *352*, 245-252. 10.1016/j.jmb.2005.07.028.
941 33. Fontana, N.A., Fonseca-Maldonado, R., Mendes, L.F.S., Meleiro, L.P., and Costa-Filho,
942 A.J. (2018). The yeast GRASP Grh1 displays a high polypeptide backbone mobility along
943 with an amyloidogenic behavior. Sci Rep *8*, 15690. 10.1038/s41598-018-33955-1.
944 34. Daskalov, A., El Mammeri, N., Lends, A., Shenoy, J., Lamon, G., Fichou, Y., Saad, A.,
945 Martinez, D., Morvan, E., Berbon, M., et al. (2021). Structures of Pathological and
946 Functional Amyloids and Prions, a Solid-State NMR Perspective. Front Mol Neurosci *14*,
947 670513. 10.3389/fnmol.2021.670513.
948 35. Tycko, R. (2014). Physical and structural basis for polymorphism in amyloid fibrils.
949 Protein Sci *23*, 1528-1539. 10.1002/pro.2544.
950 36. De La Paz, M.L., and Serrano, L. (2004). Sequence determinants of amyloid fibril
951 formation. Proc Natl Acad Sci U S A *101*, 87-92. 10.1073/pnas.2634884100.
952 37. Ventura, S., Zurdo, J., Narayanan, S., Parreno, M., Mangues, R., Reif, B., Chiti, F.,
953 Giannoni, E., Dobson, C.M., Aviles, F.X., and Serrano, L. (2004). Short amino acid
954 stretches can mediate amyloid formation in globular proteins: the Src homology 3 (SH3)
955 case. Proc Natl Acad Sci U S A *101*, 7258-7263. 10.1073/pnas.0308249101.

956 38. Vendrell-Fernandez, S., Lozano-Picazo, P., Cuadros-Sanchez, P., Tejero-Ojeda, M.M., and
957 Giraldo, R. (2022). Conversion of the OmpF Porin into a Device to Gather Amyloids on
958 the E. coli Outer Membrane. ACS Synth Biol *11*, 655-667. 10.1021/acssynbio.1c00347.
959 39. Claessen, D., Rink, R., de Jong, W., Siebring, J., de Vreugd, P., Boersma, F.G., Dijkhuizen,
960 L., and Wosten, H.A. (2003). A novel class of secreted hydrophobic proteins is involved
961 in aerial hyphae formation in *Streptomyces coelicolor* by forming amyloid-like fibrils.
962 Genes Dev *17*, 1714-1726. 10.1101/gad.264303.
963 40. de Jong, W., Wosten, H.A., Dijkhuizen, L., and Claessen, D. (2009). Attachment of
964 *Streptomyces coelicolor* is mediated by amyloidal fimbriae that are anchored to the cell
965 surface via cellulose. Mol Microbiol *73*, 1128-1140. 10.1111/j.1365-2958.2009.06838.x.
966 41. Soderholm, N., Javadi, A., Flores, I.S., Flardh, K., and Sandblad, L. (2018). Affinity to
967 cellulose is a shared property among coiled-coil domains of intermediate filaments and
968 prokaryotic intermediate filament-like proteins. Sci Rep *8*, 16524. 10.1038/s41598-018-
969 34886-7.
970 42. McGlinchey, R.P., Shewmaker, F., McPhie, P., Monterroso, B., Thurber, K., and Wickner,
971 R.B. (2009). The repeat domain of the melanosome fibril protein Pmel17 forms the
972 amyloid core promoting melanin synthesis. Proc Natl Acad Sci U S A *106*, 13731-13736.
973 10.1073/pnas.0906509106.

974 43. Daskalov, A., Dyrka, W., and Saupe, S.J. (2015). Theme and variations: evolutionary
975 diversification of the HET-s functional amyloid motif. Sci Rep *5*, 12494.
976 10.1038/srep12494.
977 44. Yasbin, R.E., and Young, F.E. (1974). Transduction in Bacillus subtilis by bacteriophage
978 SPP1. J Virol *14*, 1343-1348. 10.1128/JVI.14.6.1343-1348.1974.
979 45. Santos, J., Pujols, J., Pallares, I., Iglesias, V., and Ventura, S. (2020). Computational
980 prediction of protein aggregation: Advances in proteomics, conformation-specific
981 algorithms and biotechnological applications. Comput Struct Biotechnol J *18*, 1403-1413.
982 10.1016/j.csbj.2020.05.026.
983 46. Emily, M., Talvas, A., and Delamarche, C. (2013). MetAmyl: a METa-predictor for
984 AMYLoid proteins. PLoS One *8*, e79722. 10.1371/journal.pone.0079722.
985 47. Tsolis, A.C., Papandreou, N.C., Iconomidou, V.A., and Hamodrakas, S.J. (2013). A
986 consensus method for the prediction of 'aggregation-prone' peptides in globular proteins.
987 PLoS One *8*, e54175. 10.1371/journal.pone.0054175.
988 48. Familia, C., Dennison, S.R., Quintas, A., and Phoenix, D.A. (2015). Prediction of Peptide
989 and Protein Propensity for Amyloid Formation. PLoS One *10*, e0134679.
990 10.1371/journal.pone.0134679.
991 49. Gasior, P., and Kotulska, M. (2014). FISH Amyloid - a new method for finding
992 amyloidogenic segments in proteins based on site specific co-occurrence of aminoacids.
993 BMC Bioinformatics *15*, 54. 10.1186/1471-2105-15-54.

994 50. Heger, A., and Holm, L. (2000). Rapid automatic detection and alignment of repeats in
995 protein sequences. Proteins *41*, 224-237. 10.1002/1097-0134(20001101)41:2<224::aid-
996 prot70>3.0.co;2-z.
997 51. Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O.,
998 Tunyasuvunakool, K., Bates, R., Zidek, A., Potapenko, A., et al. (2021). Highly accurate
999 protein structure prediction with AlphaFold. Nature *596*, 583-589. 10.1038/s41586-021-
1000 03819-2.
1001 52. Mirdita, M., Schütze, K., Moriwaki, Y., Heo, L., Ovchinnikov, S., and Steinegger, M.
1002 (2022). ColabFold: making protein folding accessible to all. Nat Methods *19*, 679-682.
1003 53. Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C.,
1004 and Ferrin, T.E. (2004). UCSF Chimera--a visualization system for exploratory research
1005 and analysis. J Comput Chem *25*, 1605-1612. 10.1002/jcc.20084.
1006 54. Dos Reis, S., Coulary-Salin, B., Forge, V., Lascu, I., Begueret, J., and Saupe, S.J. (2002).
1007 The HET-s prion protein of the filamentous fungus *Podospora anserina* aggregates *in vitro*
1008 into amyloid-like fibrils. J Biol Chem *277*, 5703-5706. 10.1074/jbc.M110183200.

1009

1010 # Supporting information

1011 **S1 Fig. All the TasA Variants Exhibit Normal Localization in the Cell Membrane.** Confocal
1012 microscopy micrographs of cells carrying the different TasA alleles fused to a mCherry reporter
1013 and stained with a membrane dye (CellBrite). Scale bars = 5 µm.

1014 **S2 Fig. Transmission Electron Microscopy Images of Cells Carrying the Native Proteins or**
1015 **the Protein Variants.** Transmission electron microscopy images of negatively stained samples
1016 from cells carrying the native or the D64A or K69 and D69A alleles or Δ*tasA* cells growing under
1017 biofilm inducing conditions in MSgg medium. White squares indicate areas of the images that
1018 have been zoomed in. Scale bars = 500 nm (left), 200 nm (middle) and 100 nm (right).

1019 **S3 Fig. Structure Comparison between the Crystal Structure of the TasA Monomer and the**
1020 **Predicted Model.** Structure comparison between the crystal structure of TasA and the model
1021 predicted by AlphaFold. The numbering in the scheme follows the representation described for the
1022 crystal structure of TasA in its monomeric form published by Diehl et al. [11]. The image on the right
1023 shows the superimposition of the two structures. The coloring scheme indicates the position of the
1024 residue in the sequence of the protein, where blue indicates the N-terminus and red indicates the
1025 C-terminus. Warmer colors indicate proximity to the C-terminal end.

1026 **S4 Fig. The TasA Variants Show Differences in Thermal Stability Compared to the WT**
1027 **Protein.** Mean residue ellipticity at 222 nm measured as a function of temperature in the WT or
1028 the different TasA variants.

1029 **S5 Fig. Comassie Stained SDS–PAGE Gel of the D64A (left) or K68A and D69A Samples**
1030 **Digested with Proteinase K at Different Time Points.** The M lane indicates the molecular
1031 marker. Each lane contains a sample corresponding to a specific digestion time. SDS–PAGE gel
1032 images have been cropped and spliced for illustrative purposes. The dashed lines over the gel
1033 images indicate the boundaries of the image splicing. The two slices in the two different gel images
1034 are derived from a single gel.

1035 **S6 Fig. The TasA D64A Allele, but not the K68A or D69A Allele, is Able to Perform**
1036 **Extracellular Complementation When Mixed with a Δ*tasA* Strain.** Colony morphology
1037 phenotypes of Δ*tasA* and Δ*epsA-O* strains, of the coinoculation between Δ*tasA* and Δ*epsA-O*, of
1038 the Δ*epsA-O* strains carrying the different TasA alleles and of the coinoculation of these strains
1039 with the Δ*tasA* strain. Scale bars = 1 cm.

Figure 1

```
                                                
```

Figure 2

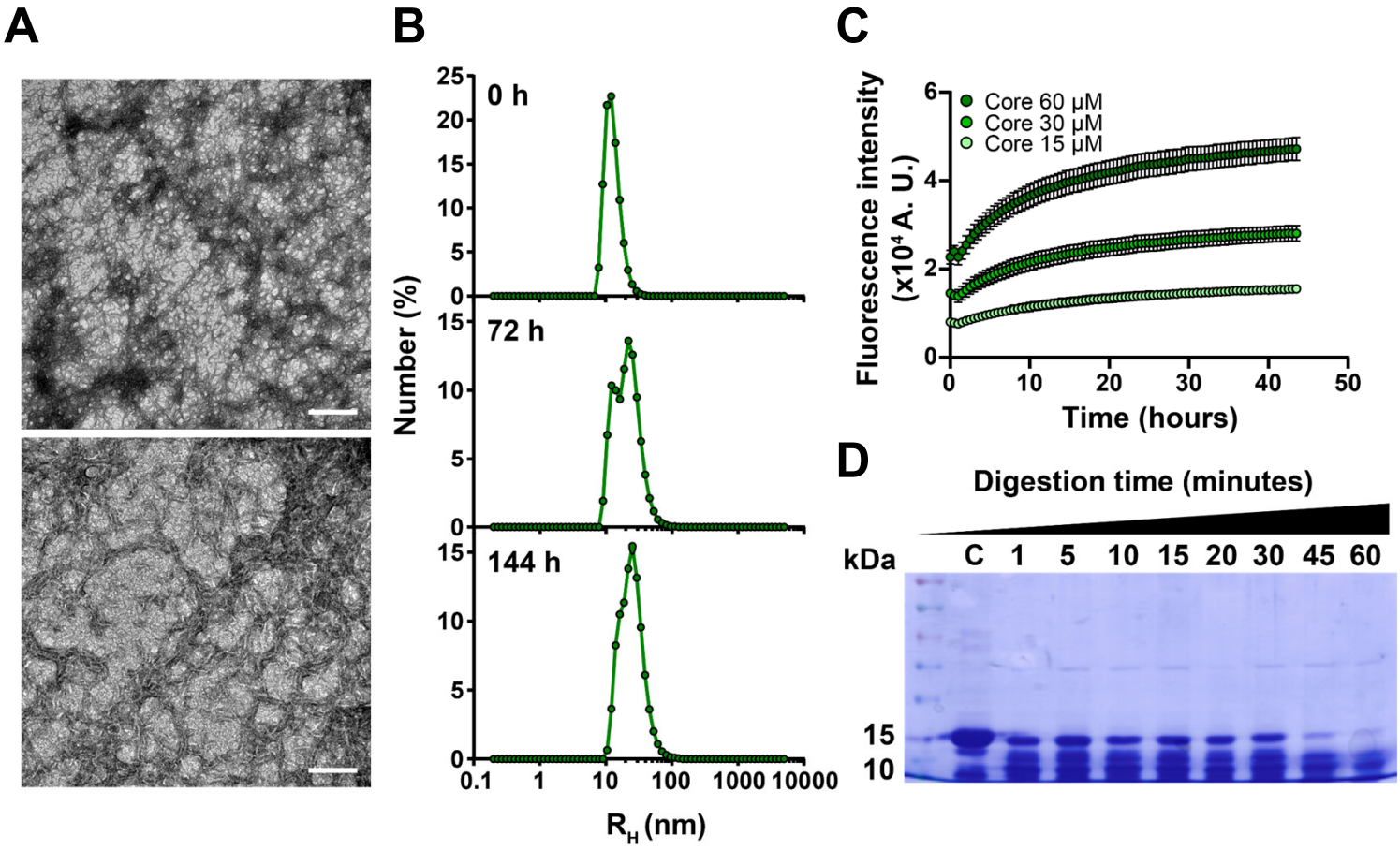

A
B
C
D
0 h
72 h
144 h
Number (%)
R_H (nm)
Core 60 μM
Core 30 μM
Core 15 μM
Fluorescence intensity
(x10⁴ A. U.)
Time (hours)
Digestion time (minutes)
kDa
C 1 5 10 15 20 30 45 60
15
10

Figure 3

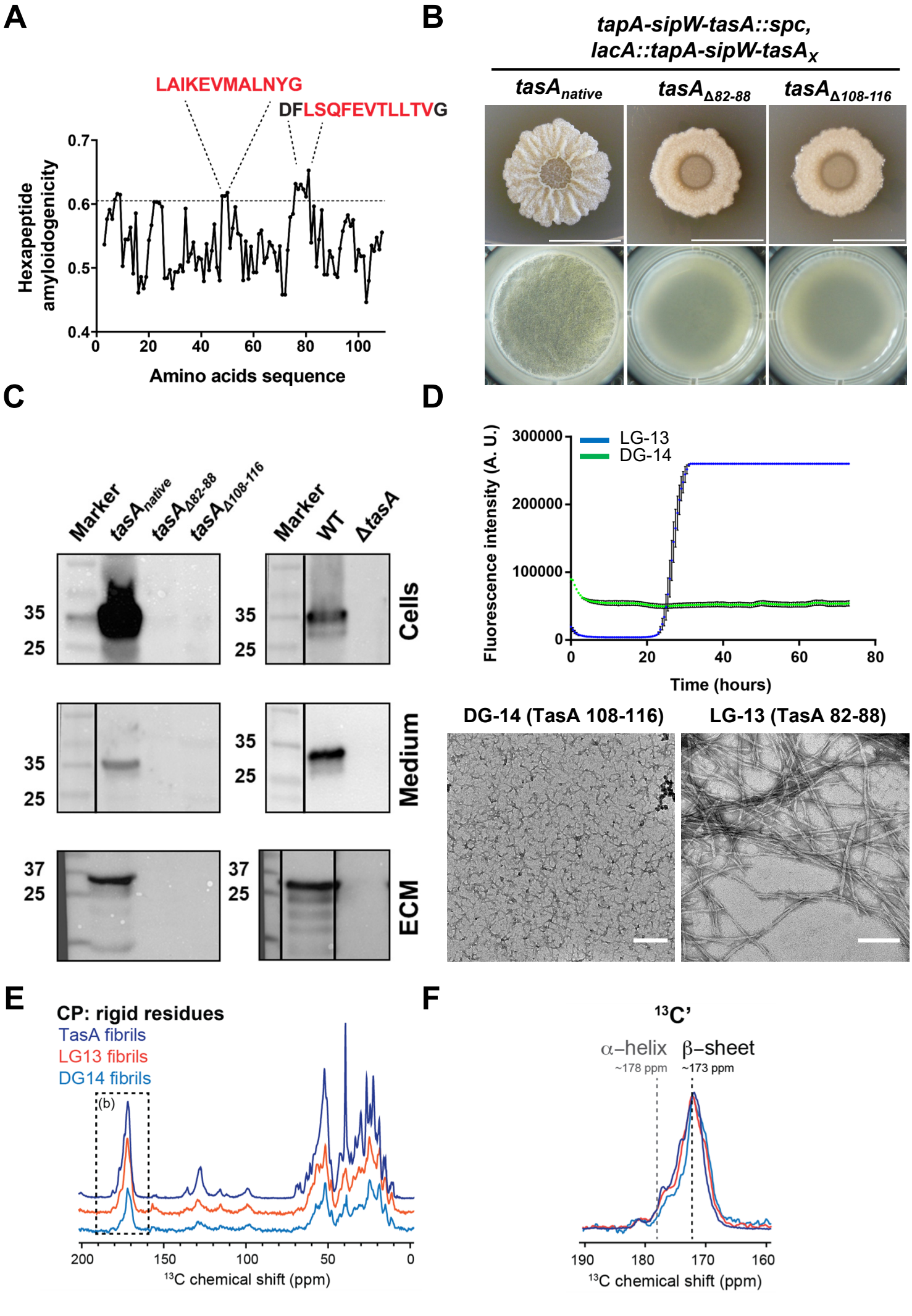

A
LAIKEVMALNYG
DFLSQFEVTLLTVG
Hexapeptide amyloidogenicity
0.7
0.6
0.5
0.4
0
20
40
60
80
100
Amino acids sequence
B
tapA-sipW-tasA::spc,
lacA::tapA-sipW-tasAX
tasAnative
tasAΔ82-88
tasAΔ108-116
C
Marker
tasAnative
tasAΔ82-88
tasAΔ108-116
Marker
WT
ΔtasA
35
25
Cells
Medium
37
ECM
D
LG-13
DG-14
Fluorescence intensity (A. U.)
300000
200000
100000
0
Time (hours)
DG-14 (TasA 108-116)
LG-13 (TasA 82-88)
E
CP: rigid residues
TasA fibrils
LG13 fibrils
DG14 fibrils
(b)
200
150
100
50
0
13C chemical shift (ppm)
F
13C'
α–helix
β–sheet
~178 ppm
~173 ppm
190
180
170
160
13C chemical shift (ppm)

Figure 4

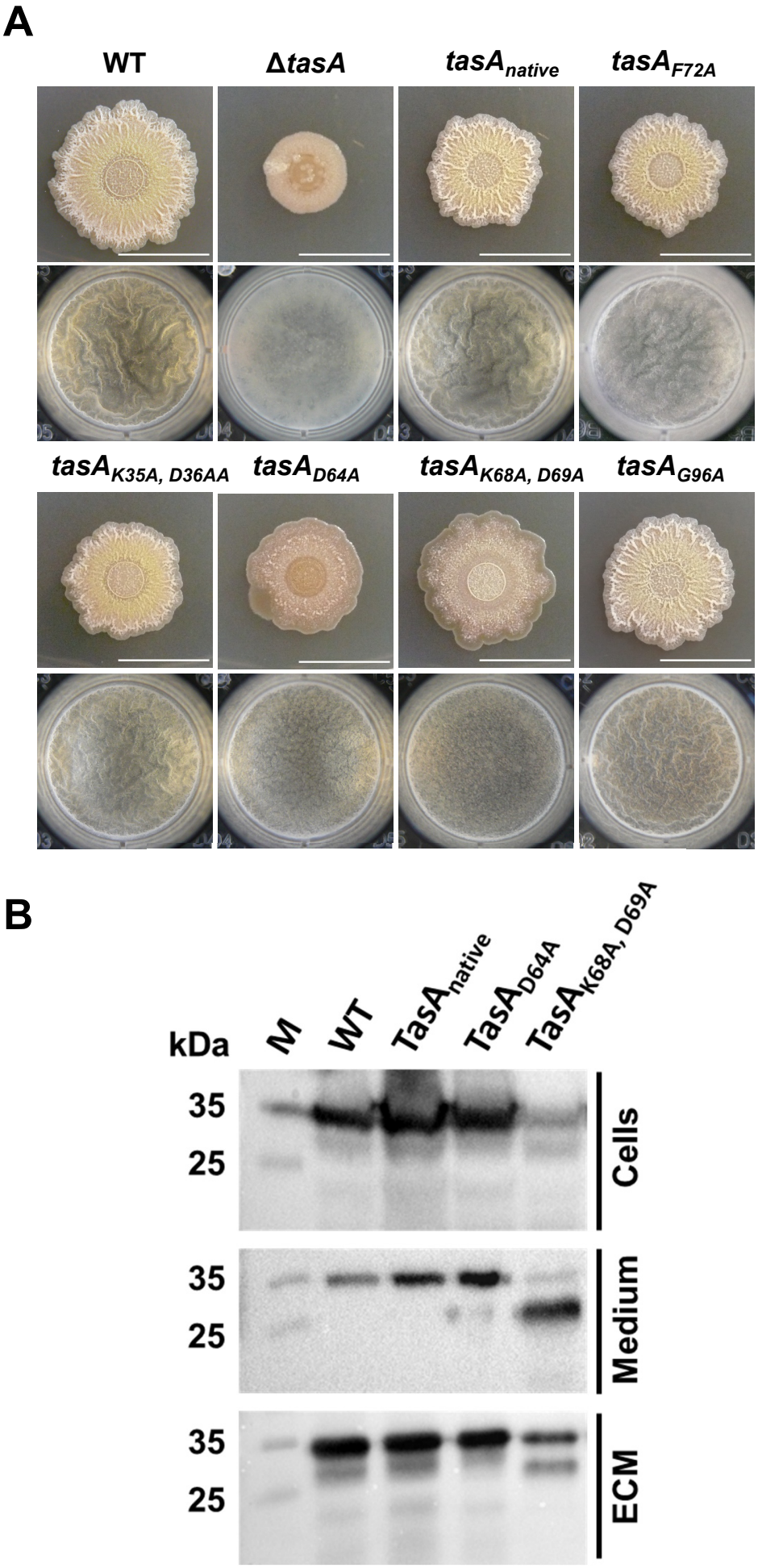

A
WT
ΔtasA
tasA native
tasA F72A
tasA K35A, D36AA
tasA D64A
tasA K68A, D69A
tasA G96A
B
kDa
M
WT
TasA native
TasA D64A
TasA K68A, D69A
35
25
Cells
35
25
Medium
35
25
ECM

Figure 5

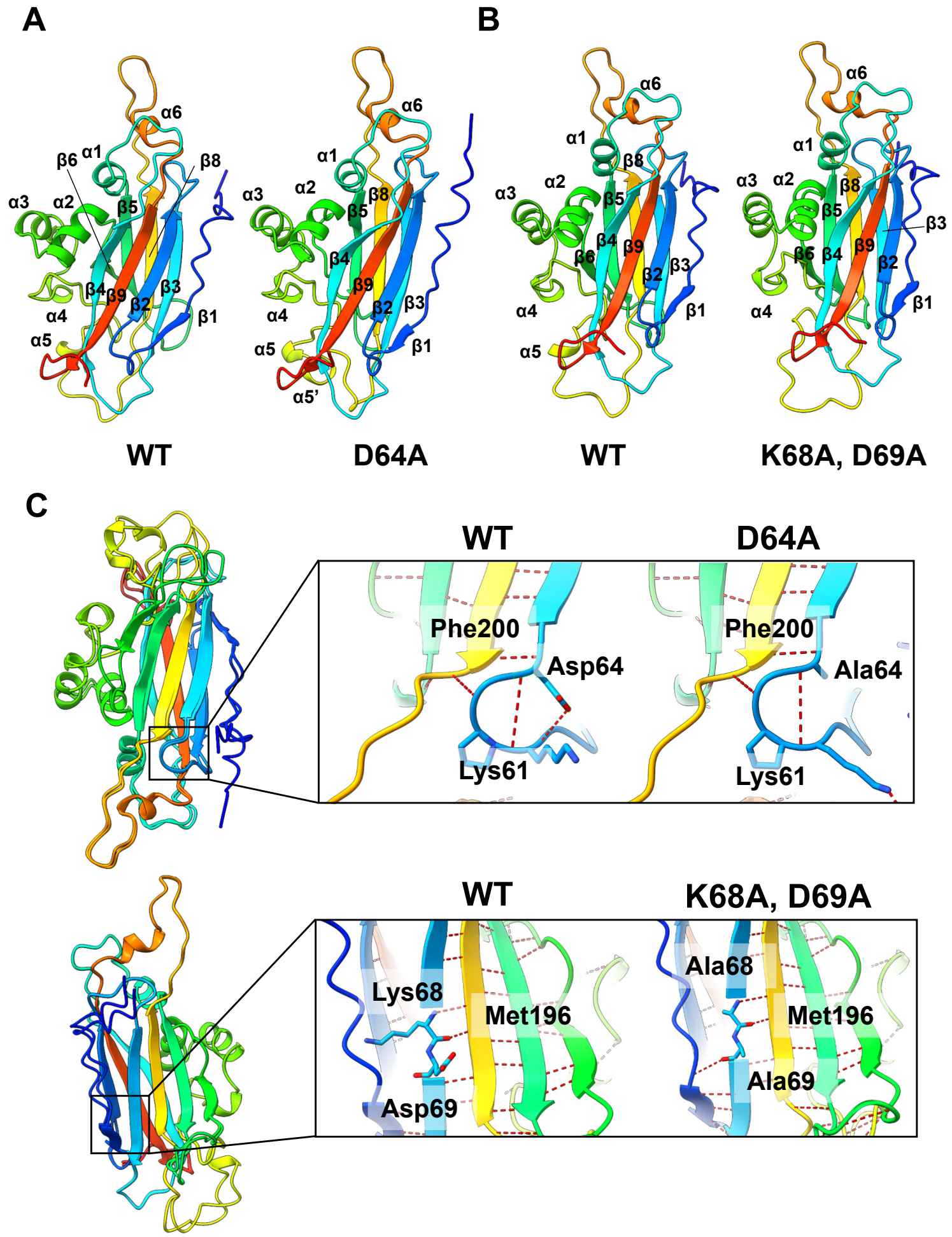

A
B
α6
α1
β6
β8
α3
α2
β5
β4
β9
β2
β3
α4
β1
α5
α5'
WT
D64A
WT
K68A, D69A
C
WT
D64A
Phe200
Asp64
Lys61
Phe200
Ala64
Lys61
WT
K68A, D69A
Lys68
Met196
Asp69
Ala68
Met196
Ala69

Figure 6

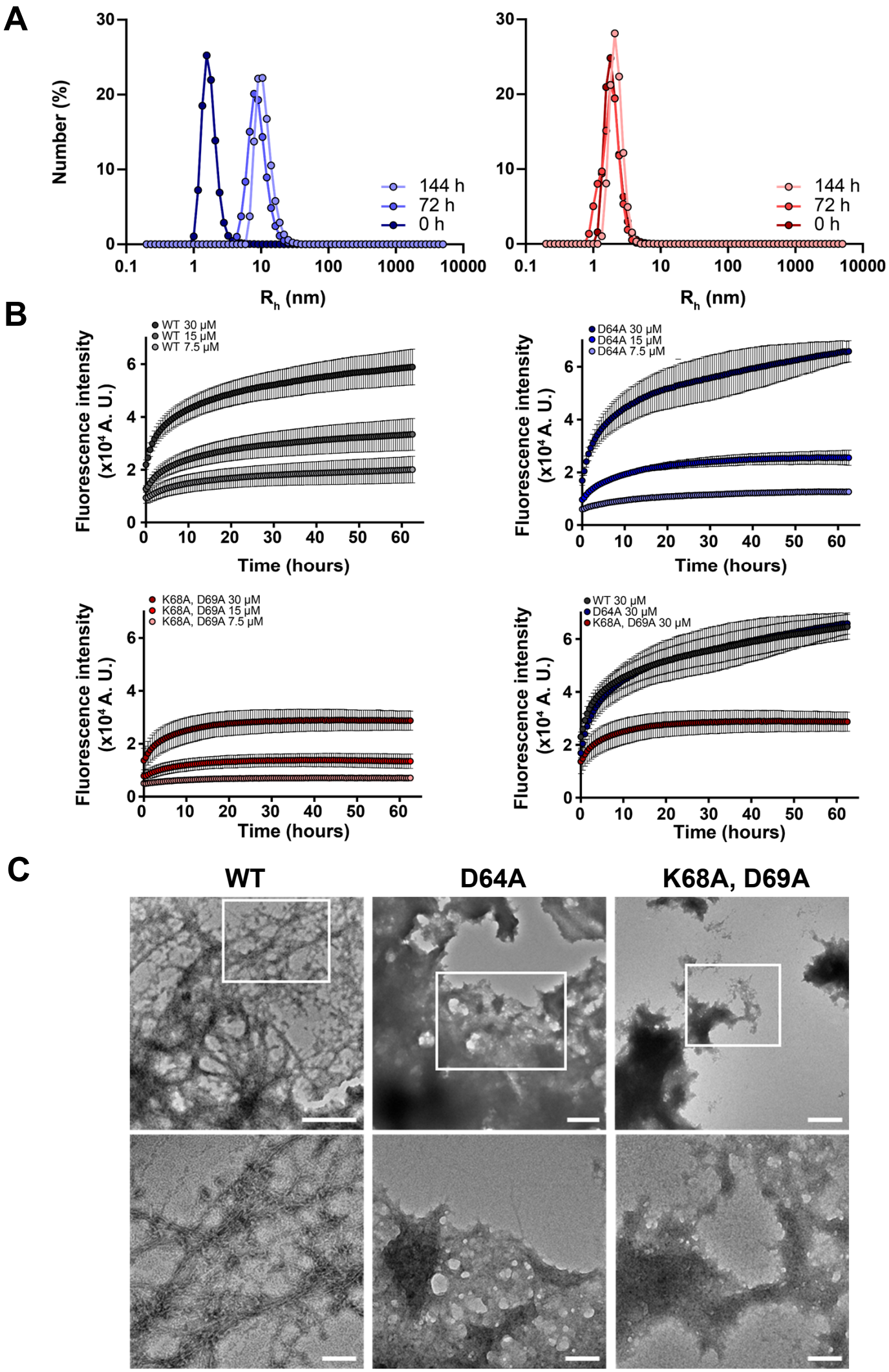

A
Number (%)
144 h
72 h
0 h
R_h (nm)
B
WT 30 μM
WT 15 μM
WT 7.5 μM
D64A 30 μM
D64A 15 μM
D64A 7.5 μM
K68A, D69A 30 μM
K68A, D69A 15 μM
K68A, D69A 7.5 μM
WT 30 μM
D64A 30 μM
K68A, D69A 30 μM
Fluorescence intensity (x10⁴ A. U.)
Time (hours)
C
WT
D64A
K68A, D69A

Figure 7

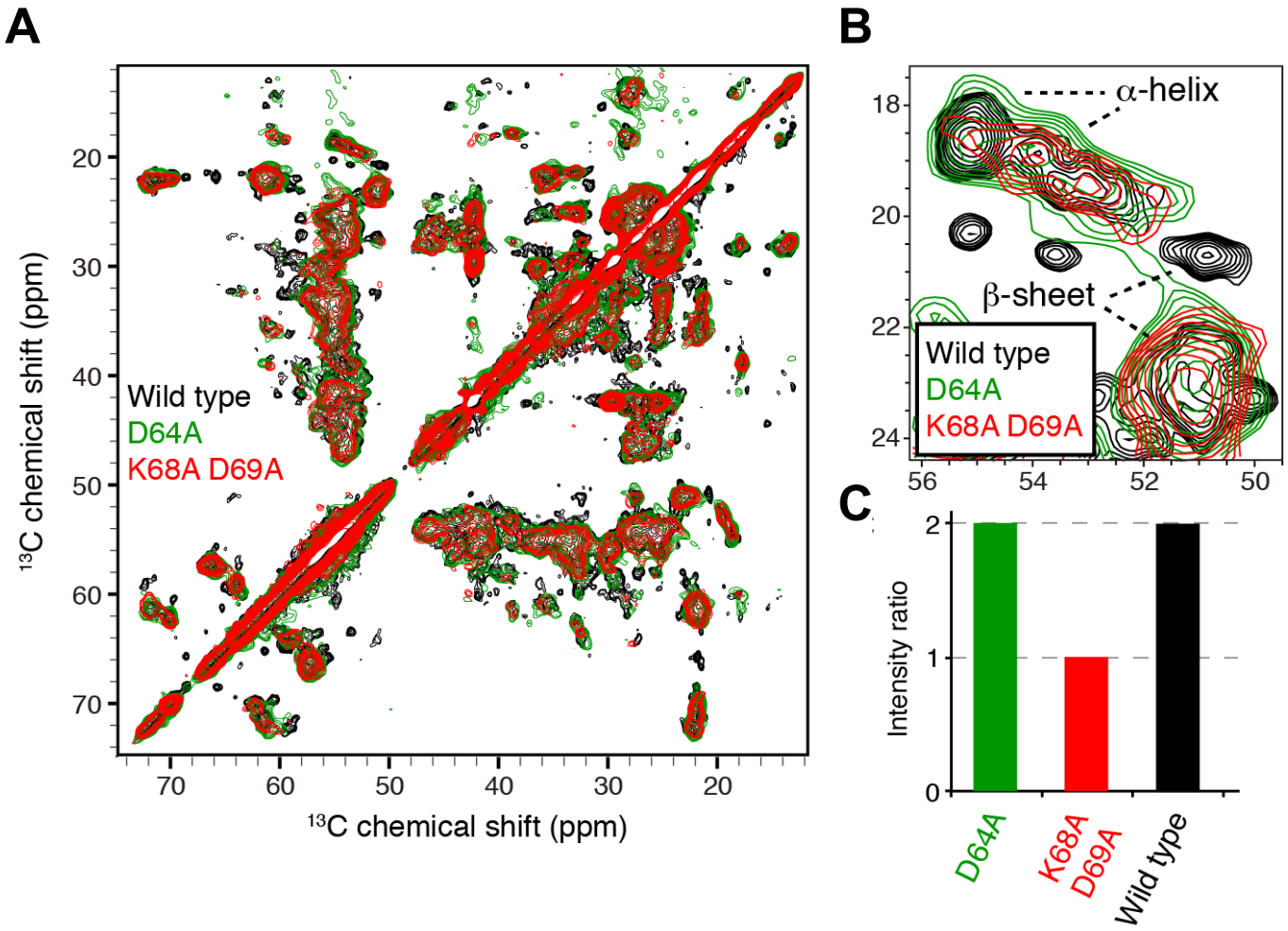
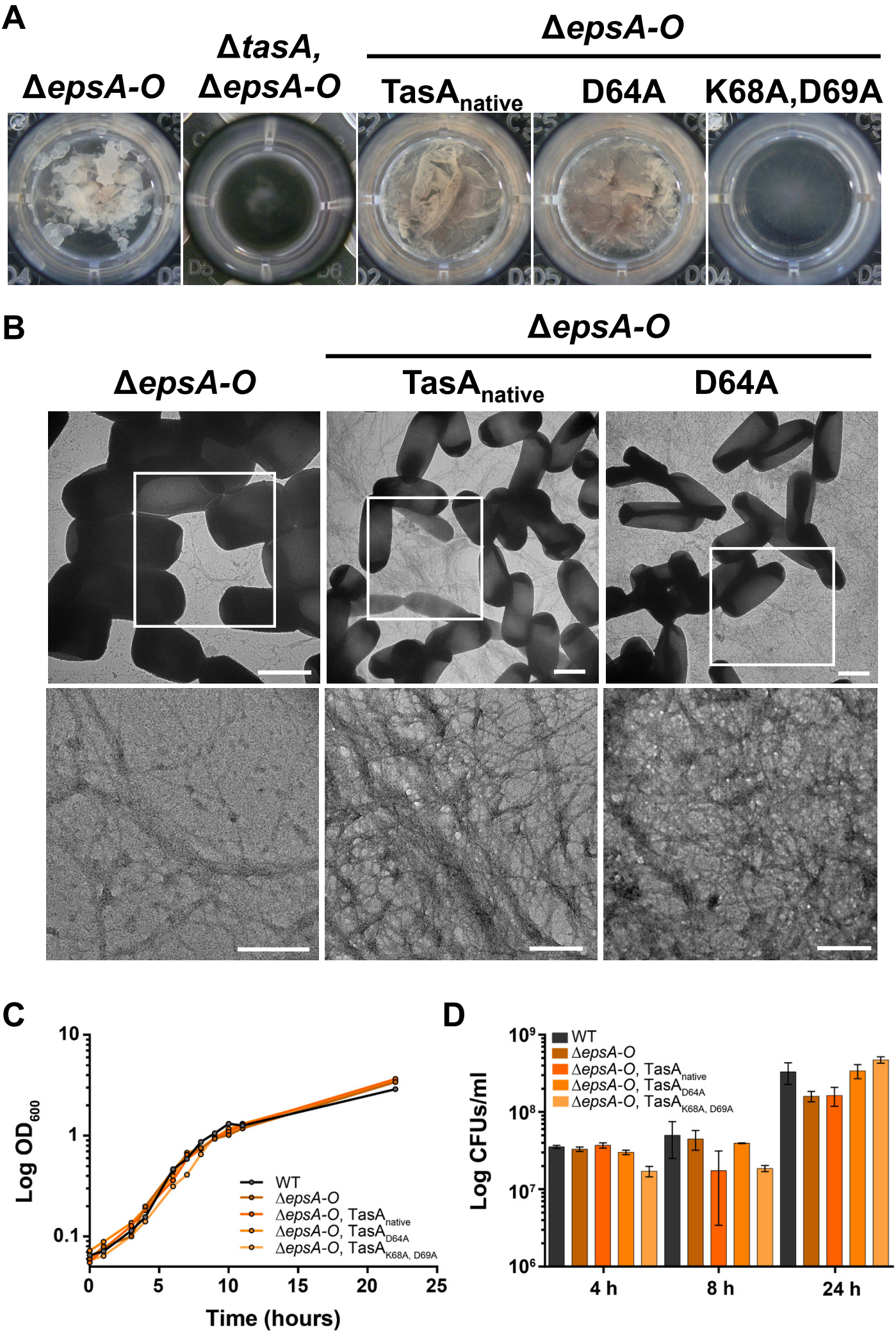

A
B
C
Wild type
D64A
K68A D69A
^{13}C chemical shift (ppm)
α-helix
β-sheet
Intensity ratio
D64A
K68A D69A
Wild type

Figure 8


A
ΔepsA-O
ΔtasA, ΔepsA-O
ΔepsA-O
TasAnative
D64A
K68A,D69A
B
ΔepsA-O
ΔepsA-O
TasAnative
D64A
C
Log OD600
Time (hours)
WT
ΔepsA-O
ΔepsA-O, TasAnative
ΔepsA-O, TasAD64A
ΔepsA-O, TasAK68A, D69A
D
Log CFUs/ml
4 h
8 h
24 h