1 Diehl et al.

2 Structural changes of TasA in biofilm formation of *Bacillus*

3 *subtilis*

4

5

6

7 **Supporting Information including additional Material &**

8 **Methods, Data, Figures and Tables**

9

A   Supporting Material & Methods

Constructs, expression and purification

For PCR amplification of the gene for mature TasA (aa 28-261, Uniprot P54507) chromosomal DNA of *B. subtilis* 168 and primer pair 1 (Table S4) were used. The PCR product was cloned into a modified pET24a vector (kanamycin resistance), pCA528 (1), to generate pCA528_His_Sumo_TasA28-261. For the truncated version (aa 28-239), pCA528_His_Sumo_TasA28-239, a stop codon was introduced at residue 240 via modified QuikChange site-directed mutagenesis, using forward and reverse primer 2 (Table S4). The different constructs were transformed into T7Express pRARE2. For this purpose, T7Express (New England Biolabs) was modified with pRARE2 (plasmid for rare tRNAs carrying chloramphenicol resistance, Novagen/Merck Millipore).

For unlabelled recombinant protein expression, transformants were grown on LB-medium at 37 °C to an OD_{600nm} of 0.8, then the temperature was lowered to 22 °C and the culture was induced by 0.5 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) for overnight cultivation. For ^{15}N- and/or ^{13}C-labelling and/or deuteration (D_2O 99,85% Eurisotop) minimal medium with 4 g glucose ($^{13}C_6$ glucose or d_7 $^{13}C_6$ glucose - Sigma) and 1 g $^{15}NH_4Cl$ (Eurisotop) per litre in M9 medium with twofold concentration of M9 salts, additional 0.1 x BME vitamins (Sigma) and trace elements was used (2). Selenomethionine- (SeMet) and $^{15}N,^{13}C$-valine-labelled $TasA_{239}$ was produced by supplementing 1 l M9 medium as above with 2 g glucose only, but also 0.5 g NH_4Cl, 100 mg of L-lysine, L-threonine, L-phenylalanine and SeMet, 50 mg of L-leucine, L-isoleucine, and $^{13}C,^{15}N$ L-valine. After 8 h of cultivation another 2 g glucose, 0.5 g NH_4Cl and 100

40   mg SeMet were supplied for overnight growth. The wet cells were collected by

41   centrifugation and suspended in buffer A (20 mM Tris/HCl pH 7.5 500 mM NaCl

42   5 mM imidazol) and a protease inhibitor cocktail (cOmplete EDTA free - Roche),

43   Benzonase (Merck) as well as 1 mM $MgCl_2$ were added. Cells were disintegrated

44   under high pressure (LM10 Microfluidizer) and a clear supernatant was

45   obtained after centrifugation at 60,000 g for 30 min. This supernatant was then

46   filtrated using a 0.45 μm device and applied to a metal chelating (MC) column

47   equilibrated with buffer A at a Workstation Vision (Applied Biosystems). His-tag

48   bound proteins were eluted by an imidazol gradient with a maximum of 500 mM.

49   Fractions with target protein were pooled, Sumo-protease (the clone was kindly

50   supplied by P. Loll, Drexel University, US) added, and the solution was dialysed

51   against buffer B (20 mM Tris/HCl pH 7.5 200 mM NaCl) overnight at 10 °C using

52   membranes with 8 kDa cut-off. The reaction was stopped the following day by a

53   protease inhibitor cocktail as above and MC-chromatography was repeated to

54   remove protease, His-tag and intact fusion protein. The flow-through was

55   concentrated by ultrafiltration (using an Amicon stirring device, regenerated

56   cellulose membrane with 10 kDa cut-off, or alternatively a Jumbosep

57   polyethersulfon centrifugal device from PALL, causing less precipitation of TasA)

58   to 4 ml and applied to a 320 ml Superdex 200 column equilibrated with buffer C

59   (20 mM sodium phosphate pH 7.0. 50 mM NaCl). Purification success and protein

60   stability were checked by SDS-PAGE (3).

61   Final products were concentrated up to 500 μM in buffer C, sterile-filtrated and

62   0.02% sodium azide as well as cOmplete (protease inhibitor cocktail) were

63   added.

64   *Preparation of $^2H,^{15}N,^{13}C$-TasA$_{261}$ biofilm*

65 Cells of *B. subtilis* Δ*tasA* (DK 1042 *tasA*::*kan*) (4) were grown overnight and used

66 to inoculate (at a dilution of 1:100) 1.4 ml medium (MOLP) (5) at pH 7 in 24-

67 well plates. After 1-2 h of incubation 100-500 µg of ^{2}H,^{15}N,^{13}C-labelled TasA were

68 added and the culture further incubated for 2-3 days at 30 °C (6, 7). A sufficient

69 amount of biofilm (containing appr. 270 µg TasA) was collected, washed with a

70 mixture of 70% buffer C and 30% D_2O, and transferred into a 1.9 mm rotor.

71 *Crystallization and structure determination.*

72 TasA$_{239}$ was crystallized employing the sitting-drop vapour-diffusion method.

73 Experiments were assisted by a Gryphon pipetting robot (Matrix Technologies

74 Co.) and a Rock Imager 1000 storage system (Formulatrix). Crystals appeared

75 within 2-6 days and were directly flash-frozen in liquid nitrogen without

76 addition of a cryo-protectant. Crystallization was successful by mixing 200 nl of

77 purified SeMet-labelled TasA$_{239}$ (residues 28-239) at a concentration of 19.5

78 mg/ml with an equal volume of reservoir solution containing 34% PEG

79 2000MME, 0.1 M ammonium sulfate, 0.2 M lithium salicylate, 0.1 M sodium

80 acetate (pH 4.6). Wild type TasA$_{239}$ was crystallized as described above for

81 SeMet-incorporated TasA$_{239}$, but at a concentration of 20 mg/ml and without

82 lithium salicylate in the reservoir solution. All diffraction data were recorded at

83 BL14.1 at BESSY II (Helmholtz-Zentrum Berlin, HZB), processed and scaled using

84 the XDSapp (8). The crystallographic phase problem for the SeMet derivate of

85 TasA$_{239}$ was solved by using HKL2MAP (SHELX suite)(9). The structure of wild

86 type TasA$_{239}$ was solved by molecular replacement with Phaser (10) using the

87 SeMet-TasA$_{239}$ structure as search model. Both protein structures were manually

88 built using COOT (11). The SeMet-TasA$_{239}$ structure was iteratively refined using

89 Refmac (12), and the wild type TasA$_{239}$ structure using Phenix (13). The SeMet-

90   TasA$_{239}$ structure consists of amino acids 30-117 and 125-239, and the wild type

91   TasA$_{239}$ structure without bound salicylate comprises residues 39-74, 77-117

92   and 125-238. Residues 28-29, 118-124 for SeMet-TasA$_{239}$ and residues 28-38,

93   75-76, 118-124 and residue 239 for the wild type TasA$_{239}$ were disordered and

94   therefore not visible in the electron density.

95   97.4% of the residues in the SeMet-TasA$_{239}$ structure were in the allowed regions

96   of the Ramachandran map and 98.4% for the non-derivatized TasA$_{239}$ structure.

97   The Ramachandran statistics of both structures were analyzed using Molprobity

98   (14). Figures and structure superimpositions were generated with PyMol

99   (http://www.pymol.org). For the recognition of PPII helices the ASSP (A

100   Program for assigning Secondary Structures in proteins) server was helpful (Fig.

101   S15) (15).

102   *Thioflavin T (ThT) assay*

103   5 µl samples of TasA solutions with a concentration of 200 µM and a given pH

104   and temperature (see Results) were diluted with 95 µl buffer C while transferred

105   into a 96-well plate (Costar 3615) and mixed 1:1 with 40 µM ThT solution in the

106   same buffer. Fluorescence (Ex 438nm/ Em 495nm) was acquired using a Tecan

107   Reader Safire (16).

108   *Fibrillation by extrusion*

109   A protein sample of about 10 mg/ml was pumped into 70% ethanol at a rate of 9

110   µl/min via a syringe equipped with PEEK tubing (inner diameter of 0.125 mm)

111   according to Teule et al. 2009 (17).

112   *Solution NMR*

113   All samples used for NMR spectroscopy contained TasA (labelled with ^{15}N,^{13}C or

114   ^{2}H,^{15}N,^{13}C for assignment, or ^{15}N for studies of interactions or stability) in buffer

C. All NMR spectra were recorded at 300 K either on AV-III-600 or AV-III-750

NMR spectrometers (600 or 750 MHz ^{1}H frequency, respectively; Bruker Biospin,

Karlsruhe, Germany) equipped with cryoprobes of either TXI or TCI type, both

equipped with one-axis self-shielded gradients.

Two-dimensional, [15]N-[1]H-correlations were performed as HSQC experiments

(18), or in case of triple labelled samples as TROSY (19) experiments. In

preparation of BEST experiments, SOFAST-HMQC (20) spectra were also

recorded. Independent of the type of technique, 512(^{1}H)*128(^{15}N) complex

points were acquired, with tHmax = 51.2 ms, tNmax = 42.5 ms and either 8, 16 or

32 scans, depending on the concentration of the samples.

BEST triple resonance experiments (21) were performed with a recovery delay

of 600 ms and using the following parameters: data size

512(^{1}H)*50(^{15}N)*50(^{13}C) complex points, tHmax = 51.2 ms, tNmax = 16.6 ms,

tCmax = 10.0 or 5.0 ms; carrier frequencies were placed at 4.7 ppm (^{1}H), 119

ppm (^{15}N) and 55 ppm or 45 ppm (^{13}C); HNCA/HN(CO)CA 16 scans,

HNCACB/HN(CO)CACB 32 scans (nuclei listed in brackets are used for transfer

but chemical shifts are not evolved).

TROSY triple resonance experiments (22) were performed using a recovery

delay of 1.5 s. For HNCACB and HN(CO)CACB, 8 scans each were performed

using a data size of 512(^{1}H)*50(^{15}N)*64(^{13}C) complex points, tHmax = 51.2 ms,

tNmax = 16.6 ms, tCmax = 6.4 ms. Carrier frequencies were placed at 4.7 ppm

(^{1}H), 119 ppm (^{15}N) and 45 ppm (^{13}C). For HNCO and HN(CA)CO experiments, 8

and 32 scans were acquired, respectively, producing a data matrix with

512(^{1}H)*48(^{15}N)*40(^{13}C) complex points and tHmax = 51.2 ms, tNmax = 15.9 ms,

139   tCmax = 16.0 ms. The carrier was set to 4.7 ppm (^{1}H), 119 ppm (^{15}N) and 175

140   ppm (^{13}C).

141   In all cases, NMR data were processed and spectra viewed using TOPSPIN 3.2 or

142   earlier versions (Bruker Biospin, Karlsruhe, Germany). The processed data were

143   converted to UCSF format (23) and subsequently transferred to CCPN (24) for

144   evaluation. For H/D exchange experiments, ^{13}C,^{15}N-labelled TasA$_{239}$ was

145   lyophilised and re-dissolved in deuterated buffer. ^{15}N HSQC spectra were

146   recorded after 11 min, 1 h, 6 h, and 5 days.

147   *Solid state MAS NMR experiments*

148   All samples used for ^{1}H-detected MAS NMR spectroscopy contained ^{2}H,^{13}C, ^{15}N,-

149   labelled TasA. A sample labelled only with ^{13}C and ^{15}N was made to count signal

150   sets of the various amino acids in carbon-carbon correlations. The proton

151   content of the backbone amide groups was adjusted to 70% by using appropriate

152   buffers for equilibration. Samples were packed into a 1.9 mm rotor by

153   centrifugation at 120,000 x g, 8 °C for 1 h using an Optima L-90K ultracentrifuge

154   with a SW40Ti rotor (Beckman Coulter) and a device similar to that described by

155   Böckmann et al. (25).

156   Proton-detected MAS NMR experiments were recorded on a standard-bore 900

157   MHz spectrometer equipped with a 1.9 mm quadruple-resonance probe (Bruker

158   BioSpin, Karlsruhe, Germany). The MAS frequency was set to 40 kHz and the VT

159   gas flow to a nominal value of 230 K. According to an external calibration with

160   DSS and H$_2$O, the achieved sample temperature was around 288 K (25).

161   2D (H)NH type experiments using cross-polarization (CP) and MISSISSIPPI

162   solvent suppression were used for all samples (10). Typical π/2-pulse lengths

163   were 2.5 μs for ^{1}H, 5 μs for ^{13}C and 7 μs for ^{15}N. For the ^{1}H/^{15}N CP, a contact time

164 of 1500 μs was chosen. On the proton channel, a square pulse with RF strength of

165 $3\omega_r/2$ was employed, whereas the ^{15}N spins were locked with a tangent ramp

166 and a mean RF strength of $\omega_r/2$ (ω_r corresponds to the MAS frequency). For the

167 back transfer from ^{15}N to ^{1}H, a short CP contact time of 800 μs was applied to

168 ensure selectivity. Water suppression was achieved using the MISSISSIPI

169 sequence (with RF field strength of around 5 kHz) without homospoil gradients.

170 Low power WALTZ-16 decoupling for ^{1}H (5 kHz) or ^{15}N (3.12 kHz) was applied

171 during the ^{15}N evolution period or the ^{1}H acquisition period, respectively. For the

172 2D (H)NH experiment, the effective acquisition time in the indirect dimension

173 was set to 25.6 ms while using a recovery delay of 1.5 s. For the sedimented

174 sample 96 scans were used, for the samples containing fibrils 64 scans and for

175 the biofilm sample 512 scans.

176 Data were processed with TOPSPIN 3.2 or 4 (Bruker BioSpin, Karlsruhe,

177 Germany), applying shifted-sinebell (in t_1) and Lorentzian-to-Gaussian

178 transformation (in t_2) functions prior to Fourier transformation. Zero filling was

179 applied to 4096 (t_1) x 1024 (t_2) points. For better visualisation and comparison,

180 the processed data were first converted to UCSF format (23) and subsequently

181 transferred to the CCPN software (24).

182 *CD measurements*

183 CD spectra (260 nm-185 nm, average of 5 scans) of 10 μM protein samples (20

184 mM phosphate buffer, 50 mM NaF, pH 7.0 or 4.0) were recorded with a step size

185 of 1 nm at 25 °C on an Applied Photophysics CD spectrometer (Chirascan) in 1

186 mm cuvettes. Melting curves were obtained in the range of 20 to 90 °C with a

187 ramp of 1 °C per min and back. The results were analysed with DiChroweb (26,

188 27).

189 *Isothermal titration calorimetry (ITC)*

190 ITC experiments were performed using a VP-ITC titration microcalorimeter (GE

191 Healthcare, Freiburg, Germany). All titrations were performed at 20 °C and pH

192 7.0. 9 mM $MnCl_2$ or 4 mM Lithium salicylate (in the syringe) were titrated in 8μl

193 steps into a 80 μM TasA solution (in the cell). The protein as well as the titration

194 components were dissolved in buffer containing 20 mM Tris/HCl and 50 mM

195 NaCl, pH 7.0. Raw data (incremental heat per molecule of added ligand) were

196 fitted by nonlinear least squares with the ORIGIN7 software using a one-site

197 binding model.

198 *Protease activity*

199 Protease activity was determined by monitoring cleavage of Azocasein (Sigma A-

200 2765) via UV/VIS spectroscopy. Proteinase K (Invitrogen) was used as a positive

201 control. Briefly, 1 ml samples of 1.25% substrate and 840 μg $TasA_{261}$ or 6.7 μg

202 proteinase K dissolved in buffer C or 20 mM sodium acetate buffer, 50 mM NaCl,

203 pH 4.0, were incubated at 37 °C. At distinct time points 0.1 ml samples were

204 drawn, and undigested substrate was precipitated by 0.4 ml of 5% TCA. After

205 centrifugation at 10,000 g for 10 min, 0.2 ml of the supernatant was neutralized

206 with 0.8 ml 0.5 M NaOH. Absorbance of the final solutions was measured at 440

207 nm. Different TasA preparations (without protease inhibitor cocktail cOmplete)

208 and the influence of 10 mM EDTA as well as 1-5 mM $MnCl_2$ and $ZnCl_2$ were

209 tested.

210 *Analytical ultracentrifugation*

211 Analytical ultracentrifugation (AUC) was performed with an Optima XL-1

212 (Beckman) centrifuge. Sedimentation velocity experiments of $TasA_{261}$ stored at -

213 20 °C and pH 7 were performed at 20 °C and 35 krpm with 70 μM protein in

214   buffer C as well as samples of 40 μM $TasA_{261}$ in buffer C that was, however,

215   dialysed to pH 3.5 or 3.0. The respective data are shown in Fig. 4a.

216   Furthermore, freshly prepared $TasA_{261}$ samples (about 200 μM) were allowed to

217   form fibres over 2 weeks at 40 °C in buffer C pH 7.0 and pH 3.0. Samples were

218   diluted to 58 μM and sedimentation velocity monitored at 15 krpm. Prior to

219   these experiments, time point 0 was recorded at a rotor speed of 25 krpm. The

220   respective data are shown in Fig. S6a and b.

221   Scans were recorded by interference and absorbance optics in 5 min intervals.

222   Sedimentation coefficient distributions c(s) were calculated from absorbance

223   data with the program Sedfit. The protein partial specific volume and the buffer

224   physical constants were calculated from amino acid and buffer composition,

225   respectively, using SEDNTERP(28). Plots were created with GUSSI (available at

226   http://biophysics.swmed.edu/MBR/software.html) (29).

227   *Electron Microscopy*

228   For electron microscopy, formvar-carbon film coated grids were glow-

229   discharged, a drop of protein sample was applied on the grids, blotted via filter

230   paper and negatively stained by 3% aqueous uranyl acetate. Dry grids were

231   imaged at Tecnai G2 200 kV or Zeiss 900 80kV transmission electron

232   microscopes. Images were taken at x50,000 magnification.

233   *Mass spectrometry*

234   For the analysis of TasA gel bands, appropriate sections of the gel were excised,

235   washed first with water, then with 25 mM ammonium bicarbonate in

236   acetonitrile/water (1:1), then with 50 mM ammonium bicarbonate, shrunk by

237   dehydration in acetonitrile and subsequently most of the liquid removed in a

238   speed-vacuum centrifuge. The gel pieces were then incubated in 20 μl of 50 mM

239   ammonium bicarbonate containing 50 ng trypsin (sequencing grade modified,

240   Promega) at 37 °C overnight. The enzymatic reaction was terminated by addition

241   of 20 µl of 0.5% trifluoroacetic acid in acetonitrile. The liquid was separated and

242   solvent evaporated under vacuum. Tryptic peptides were re-dissolved in 6 µl

243   0.1% trifluoroacetic acid, 5% acetonitrile in water. LC-MS/MS analyses were

244   performed using a capillary liquid chromatography system (Ultimate 3000

245   nanoLC) connected to an Orbitrap Elite mass spectrometer (Thermo Scientific,

246   Germany). Reversed-phase LC separations were performed on a capillary

247   column (Acclaim PepMap100, C18, 2 µm, 100 Å, 75 µm i.d. × 25 cm, Thermo

248   Scientific) at an eluent flow rate of 200 nl/min using a linear gradient of 3–50%

249   of phase B in 60 min. Mobile phase A was 0.1% formic acid in water and mobile

250   phase B was 0.1% formic acid in acetonitrile. Mass spectra were acquired in a

251   data-dependent mode with one MS survey scan with a nominal resolution of

252   60,000 (Orbitrap) and MS/MS scans of the 15 most intense precursor ions in the

253   linear trap quadrupole.

254   For limited proteolysis of soluble and fibrillar TasA**,** samples of 1.6 µg protein

255   were incubated in 6 µl 50 mM ammonium bicarbonate. Subsequently, 2 µl of 50

256   mM ammonium bicarbonate containing 16 ng trypsin (sequencing grade

257   modified, Promega) were added (E:S=1:100) and the solution was incubated at

258   37 °C. Aliquots (0.5 µl of proteolysis reaction) were withdrawn from the reaction

259   mixture after 15 min, 30 min, and 17 h, and proteolysis was stopped by mixing

260   with 0.5 µl of MALDI matrix (α-cyano-hydroxycinnamic acid) on the MALDI

261   target plate. MS and MS/MS measurements were performed using a MALDI-TOF-

262   TOF instrument (AB SCIEX TOF/TOF 5800; Applied Biosystems, Framingham,

263   MA, USA) equipped with a neodymium-doped yttrium-lithium-fluoride laser

264 (Nd:YLF, 349 nm). MS spectra were acquired in positive ion reflector mode, and

265 each spectrum obtained was a mean of 4000 laser shots. For MS/MS, a maximum

266 of 30 precursor ions were selected automatically. GPS Explorer (version 3.6,

267 Applied Biosystems) was used to process the spectra.

268 MS and MS/MS data were searched in-house against a self-made database which

269 contains the SwissProt database (version 2014_12; 547,085 sequences) and the

270 recombinant TasA sequence using a MASCOT server. For MALDI-MS, the mass

271 tolerance of precursor and sequence ions was set to 100 ppm and 0.35 Da,

272 respectively. For LC-MS/MS, the mass tolerance of precursor and sequence ions

273 was set to 10 ppm and 0.35 Da, respectively. A maximum of two missing peptides

274 due to insufficient cleavage was allowed. Methionine oxidation and the

275 acrylamide modification of cysteine were used as variable modifications. Peptide

276 identifications were accepted if their ions scores indicate identity or extensive

277 homology ($p < 0.05$).

278 *Computational protein modeling*

279 Modeller (30) was used to build homology models of CalY1 and CalY2 from both

280 *B. cereus* and *B. anthracis* using the SeMet-TasA$_{239}$ crystal structure as a

281 template. Ten models were produced, with those having the lowest DOPE scores

282 (31) being shown in Fig. S5 and used for metal binding site prediction.

283 *Metal binding site prediction*

284 The presence of potential Zn^{2+} binding sites in the SeMet-TasA$_{239}$ crystal

285 structure and camelysin homology models was assessed using two prediction

286 servers: i) Metal Ion-Binding Site Prediction and Docking Server (MIB) (32) and

287 ii) IonCom (33). In both instances, standard settings were used.

288

B Supporting Data

Binding of divalent cations and structural relationship to camelysins

Based on sequence alignments and making use of our structural model, we generated 3D models (Fig. S5a) based on a hand-curated multiple sequence alignment of four *Bacillus cereus* and *anthracis* camelysins (Fig. S5b. Compared to TasA, helix fragments are deleted and loops connecting the conserved helical segments are much shorter in these modelled proteins, in particular α2 of TasA appears clipped and α3 is no longer present. All other secondary structure elements of TasA are clearly built in the modelled camelysin structures. A potential motif for binding of divalent cations was previously predicted for camelysin using sequence alignments (34). The respective residues are also conserved in the TasA sequence (D31, D64, N74, D104, Q214 and Q218, see arrows in Fig. S5b). However, in our 3D structure of TasA$_{239}$ these residues do not appear to form clusters with a structural arrangement that is suitable for metal binding (green residues in Fig. 2c) as required for protease activity, although four of them are located on top of the structure. In the modelled camelysin structures, the respective residues were also not suitably clustered.

Intriguingly, *B. subtilis* biofilm formation takes place in minimal media solely when bivalent metal ions like Mg^{2+} or Mn^{2+} (35, 36) are present. We investigated therefore binding of divalent cations by isothermal titration calorimetry (ITC), NMR and crystallization experiments. ITC measurements indicate very weak manganese binding to TasA$_{239}$ with an apparent affinity of ~340 μM, indicating weak electrostatic interactions (Fig. S4e,f). Along this line, crystallization of TasA$_{239}$ in presence of manganese and zinc ions did not yield structures with any bound cation. In particular, protein was co-crystallized with 6 mM MnCl$_2$ and a 1.6 Å dataset was recorded at the Mn absorption edge. No anomalous signal could be detected.

313   Titration of zinc ions into a TasA$_{261}$ solution did not lead to chemical shift changes in the ^{15}N-

314   ^{1}H spectra, however, the strongly paramagnetic manganese ions caused a loss of cross-peak

315   intensity at comparably high concentrations (2.7-fold molar excess compared to TasA).

316   Affected residues are located within β-sheet 2 which comprises strands β3, β5, β6, β8 and the

317   loop preceding β8 (Fig. 2e, Fig S5b), and form negatively charged surface areas (Fig. 2d).

318   Protease activity was not detected by an unspecific assay that employs azocasein as substrate.

319   These results suggest that TasA does not bind metals such as Mn^{2+} or Zn^{2+} with sufficient

320   affinity to induce metalloprotease activity as described for camelysin (34, 37).

321

322   **Structural differences of fibrillar and monomeric forms investigated by mass**

323   **spectrometry and NMR**

324   The monomeric form of TasA was also characterized by H/D exchange monitored by

325   NMR spectroscopy and together with the fibrillar form by protease digestion combined

326   with mass spectrometry. In those experiments, residues in areas protected from H/D

327   exchange or proteolytic digestion are detected. In essence, several stretches with highly

328   protected amides are observed in NMR experiments on soluble TasA$_{239}$, (Fig. S5b),

329   indicating negligible exchange over several hours or days. As expected, those slowly

330   exchanging amide groups are located in conserved secondary structure elements (Fig.

331   S5b).

332   When applying tryptic digestion and mass spectrometry (resolving peptides between 1

333   and 4 kDa) to soluble TasA$_{261}$, we observe cleavage in case of fibrils, but no cleavage at

334   lysines 136 (at the beginning of α2) and 201 (directly following β8) in case of the folded

335   monomer. Residues in the respective peptides ($_{127}$NIILDDANLKDLYLM$_{141}$ and

336     $_{194}$VQMEIQFK$_{201}$) were identified as H/D exchange-resistant stretches in the soluble

337   form by NMR. This suggests that K136 and K201, which are protected against tryptic

digestion in soluble TasA261, become exposed in the fibrillar form due to a conformational change.

In summary, TasA undergoes vast structural transitions on its way from the cytosol to natural biofilms. There are two types of high-molecular weight forms, gel-like and fibrillar, of which one particular fibrillar form takes part in biofilm formation.

C Supporting References

1. Andréasson C, Fiaux J, Rampelt H, Mayer MP, Bukau B (2008) Hsp110 is a nucleotide-activated exchange factor for Hsp70. *J Biol Chem* 283(14):8877–8884.

2. Studier FW (2005) Protein production by auto-induction in high density shaking cultures. *Protein Expr Purif* 41(1):207–234.

3. Laemmli UK (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* 227:680–685.

4. Konkol MA, Blair KM, Kearns DB (2013) Plasmid-encoded ComI inhibits competence in the ancestral 3610 strain of Bacillus subtilis. *J Bacteriol* 195(18):4085–4093.

5. Ahimou F, Jacques P, Deleu M (2000) Surfactin and iturin A effects on Bacillus subtilis surface hydrophobicity. *Enzyme Microb Technol* 27(10):749–754.

6. Branda SS, Gonzalez-Pastor JE, Ben-Yehuda S, Losick R, Kolter R (2001) Fruiting body formation by Bacillus subtilis. *Proc Natl Acad Sci U A* 98(20):11621–6.

7. Romero D, Aguilar C, Losick R, Kolter R (2010) Amyloid fibers provide structural integrity to Bacillus subtilis biofilms. *Proc Natl Acad Sci U A* 107(5):2230–4.

8. Krug M, Weiss MS, Heinemann U, Mueller U (2012) XDSAPP: a graphical user interface for the convenient processing of diffraction data using XDS. *J Appl Crystallogr* 45(3):568–572.

9. Pape T, Schneider TR (2004) HKL2MAP: a graphical user interface for macromolecular phasing with SHELX programs. *J Appl Crystallogr* 37(5):843–844.

10. McCoy AJ, et al. (2007) Phaser crystallographic software. *J Appl Crystallogr* 40(4):658–674.

371   11.   Emsley P, Lohkamp B, Scott WG, Cowtan K (2010) Features and
372   development of Coot. *Acta Crystallogr D Biol Crystallogr* 66(4):486–501.

373   12.   Murshudov GN, Vagin AA, Dodson EJ (1997) Refinement of macromolecular
374   structures by the maximum-likelihood method. *Acta Crystallogr D Biol*
375   *Crystallogr* 53(3):240–255.

376   13.   Adams PD, et al. (2010) PHENIX: a comprehensive Python-based system for
377   macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr*
378   66(2):213–221.

379   14.   Chen VB, et al. (2010) MolProbity: all-atom structure validation for
380   macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr*
381   66(1):12–21.

382   15.   Gill AC, et al. (2000) Post-translational hydroxylation at the N-terminus of
383   the prion protein reveals presence of PPII structure in vivo. *EMBO J*
384   19(20):5324–5331.

385   16.   Biancalana M, Koide S (2010) Molecular mechanism of Thioflavin-T binding
386   to amyloid fibrils. *Biochim Biophys Acta* 1804(7):1405–1412.

387   17.   Teulé F, et al. (2009) A protocol for the production of recombinant spider
388   silk-like proteins for artificial fiber spinning. *Nat Protoc* 4(3):341–355.

389   18.   Bodenhausen G, Ruben DJ (1980) Natural abundance nitrogen-15 NMR by
390   enhanced heteronuclear spectroscopy. *Chem Phys Lett* 69(1):185–189.

391   19.   Pervushin K, Riek R, Wider G, Wüthrich K (1997) Attenuated T2 relaxation
392   by mutual cancellation of dipole–dipole coupling and chemical shift
393   anisotropy indicates an avenue to NMR structures of very large biological
394   macromolecules in solution. *Proc Natl Acad Sci* 94(23):12366–12371.

395   20.   Schanda P, Kupče Ē, Brutscher B (2005) SOFAST-HMQC experiments for
396   recording two-dimensional deteronuclear correlation spectra of proteins
397   within a few seconds. *J Biomol NMR* 33(4):199–211.

398   21.   Lescop E, Schanda P, Brutscher B (2007) A set of BEST triple-resonance
399   experiments for time-optimized protein resonance assignment. *J Magn*
400   *Reson* 187(1):163–169.

401   22.   Salzmann M, Pervushin K, Wider G, Senn H, Wüthrich K (1998) TROSY in
402   triple-resonance experiments: new perspectives for sequential NMR
403   assignment of large proteins. *Proc Natl Acad Sci* 95(23):13585–13590.

404   23.   Goddard TD, Kneller DG *SPARKY 3* (University of California, San Francisco).

405   24.   Vranken WF, et al. (2005) The CCPN data model for NMR spectroscopy:
406   development of a software pipeline. *Proteins Struct Funct Bioinforma*
407   59(4):687–696.

408   25. Böckmann A, et al. (2009) Characterization of different water pools in solid-
409   state NMR protein samples. *J Biomol NMR* 45(3):319.

410   26. Whitmore L, Wallace BA (2008) Protein secondary structure analyses from
411   circular dichroism spectroscopy: methods and reference databases.
412   *Biopolymers* 89(5):392–400.

413   27. Whitmore L, Wallace BA (2004) DICHROWEB, an online server for protein
414   secondary structure analyses from circular dichroism spectroscopic data.
415   *Nucleic Acids Res* 32(Web Server issue):W668-673.

416   28. Laue T, Shah B, Ridgeway T, Pelletier S Computeraided interpretation of
417   analytical sedimentation data for proteins.(1992) Analytical
418   ultrancentrifugation in biochemistry and polymer science. *Roy Soc*
419   *Chem*:90–125.

420   29. Schuck P (2000) Size-distribution analysis of macromolecules by
421   sedimentation velocity ultracentrifugation and lamm equation modeling.
422   *Biophys J* 78(3):1606–1619.

423   30. Sali A, Blundell TL (1993) Comparative protein modelling by satisfaction of
424   spatial restraints. *J Mol Biol* 234(3):779–815.

425   31. Shen M-Y, Sali A (2006) Statistical potential for assessment and prediction
426   of protein structures. *Protein Sci Publ Protein Soc* 15(11):2507–2524.

427   32. Lin Y-F, et al. (2016) MIB: Metal Ion-Binding Site Prediction and Docking
428   Server. *J Chem Inf Model* 56(12):2287–2291.

429   33. Hu X, Dong Q, Yang J, Zhang Y (2016) Recognizing metal and acid radical
430   ion-binding sites by integrating ab initio modeling with template-based
431   transferals. *Bioinforma Oxf Engl* 32(21):3260–3269.

432   34. Grass G, et al. (2004) Camelysin is a novel surface metalloproteinase from
433   Bacillus cereus. *Infect Immun* 72(1):219–228.

434   35. Chu F, Kearns DB, Branda SS, Kolter R, Losick R (2006) Targets of the
435   master regulator of biofilm formation in Bacillus subtilis. *Mol Microbiol*
436   59(4):1216–28.

437   36. Dogsa I, Brloznik M, Stopar D, Mandic-Mulec I (2013) Exopolymer diversity
438   and the role of levan in Bacillus subtilis biofilms. *PloS One* 8(4):e62044.

439   37. Fricke B, et al. (2001) The cell envelope-bound metalloprotease (camelysin)
440   from Bacillus cereus is a possible pathogenic factor. *Biochim Biophys Acta*
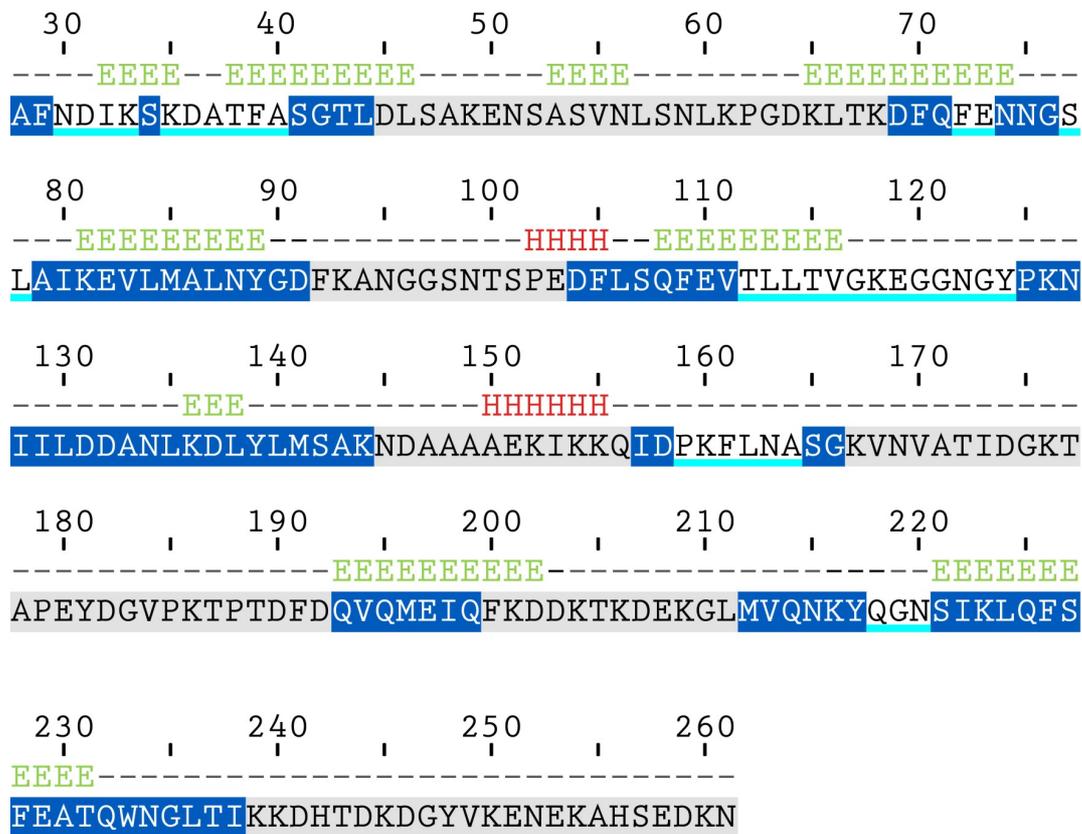441   1537(2):132–146.

442

443

D   Supplementary Data, Figures

Fig. S1

446

```
          30            40            50            60            70
          |    |    |    |    |    |    |    |    |    |    |    |
       ----|-EEEE--|EEEEEEEEE|------|--EEEE---|-----|--EEEEEEEEEE|---
       AFNDIKSKDATFASGTLDLSAKENSASVNLSNLKPGDKLTKDFQFENNGS

          80            90           100           110           120
          |    |    |    |    |    |    |    |    |    |    |    |
       ---|EEEEEEEEE|----------|--HHHH--|EEEEEEEE|----------|
       LAIKEVLMALNYGDFKANGGSNTSPEDFLSQFEVTLLTVGKEGGNGYPKN

         130           140           150           160           170
          |    |    |    |    |    |    |    |    |    |    |    |
       --------|----EEE|----------|--HHHHH--|----------|----------
       IILDDANLKDLYLMSAKNDAAAAEKIKKQIDPKFLNASGKVNVATIDGKT

         180           190           200           210           220
          |    |    |    |    |    |    |    |    |    |    |    |
       ---------------|--EEEEEEEEEE|------------|----------|-EEEEEEE
       APEYDGVPKTPTDFDQVQMEIQFKDDKTKDEKGLMVQNKYQGNSIKLQFS

         230           240           250           260
          |    |    |    |    |    |    |    |    |
       EEEE|----------|----------|----------|-----
       FEATQWNGLTIKKDHTDKDGYVKENEKAHSEDKN
```

EEE	beta-strand (Jpred3)	AAA	folded (FoldUnfold)
HHH	helix (Jpred3)	AAA	treated as folded (FoldUnfold)
		AAA	unfolded (FoldUnfold)

447
448
449

Fig. S1. Bioinformatic analysis of TasA.
Jpred3 (http://www.compbio.dundee.ac.uk/jpred/) and FoldUnfold (http://bioinfo.protres.ru/ogu/) were applied to predict secondary structure and disordered regions of TasA. The analysis enabled the design of a truncated construct TasA$_{239}$ via changing the triplet coding for K240 into a stop codon by mutagenesis. In contrast to TasA$_{261}$, the shortened version TasA$_{239}$ lead to well-diffracting crystals (2.6 Å) readily.

457

461

462

463
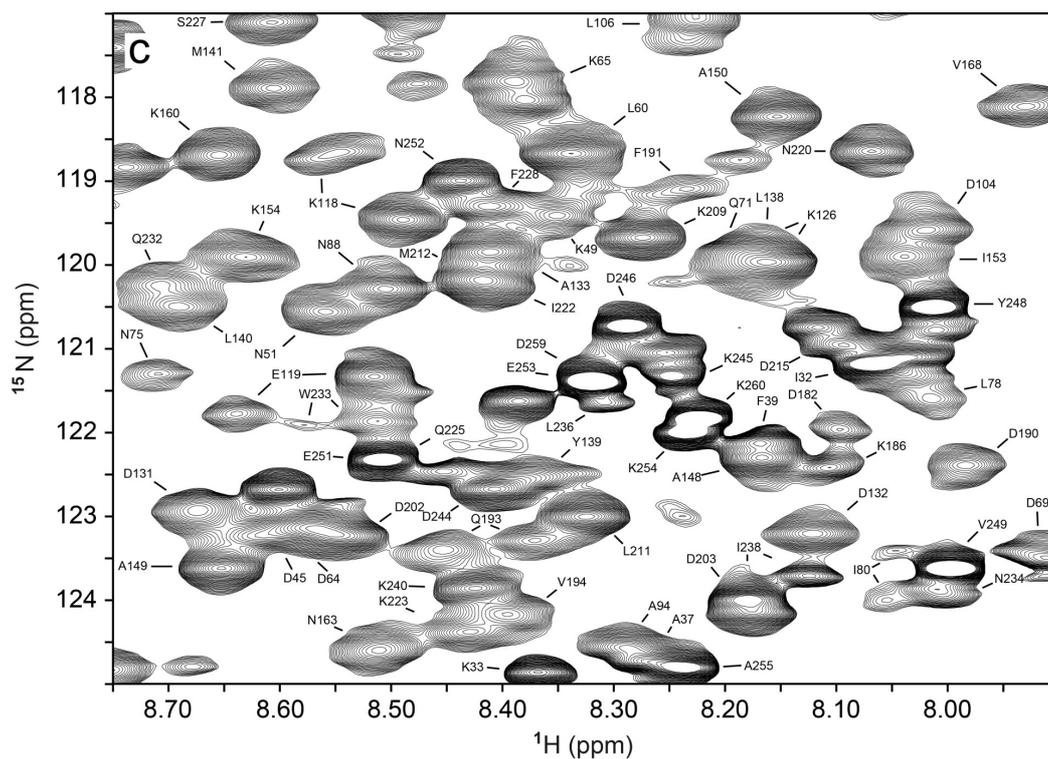
Fig. S2. Comparison of ^{15}N-^{1}H correlations of TasA$_{239}$ and TasA$_{261}$, and assignment of TasA$_{261}$. a) Superposition of solution NMR ^{15}N-^{1}H-HSQC spectra, recorded from full-length, ^{15}N-labelled recombinant TasA$_{261}$ (blue) and the truncated TasA$_{239}$ (red). b) Assignment of solution NMR spectra of TasA$_{261}$. An expansion of the crowded region (^{1}H: 7.9 - 8.7 ppm; ^{15}N: 117 - 125 ppm) is shown in c).

469

471

472

473
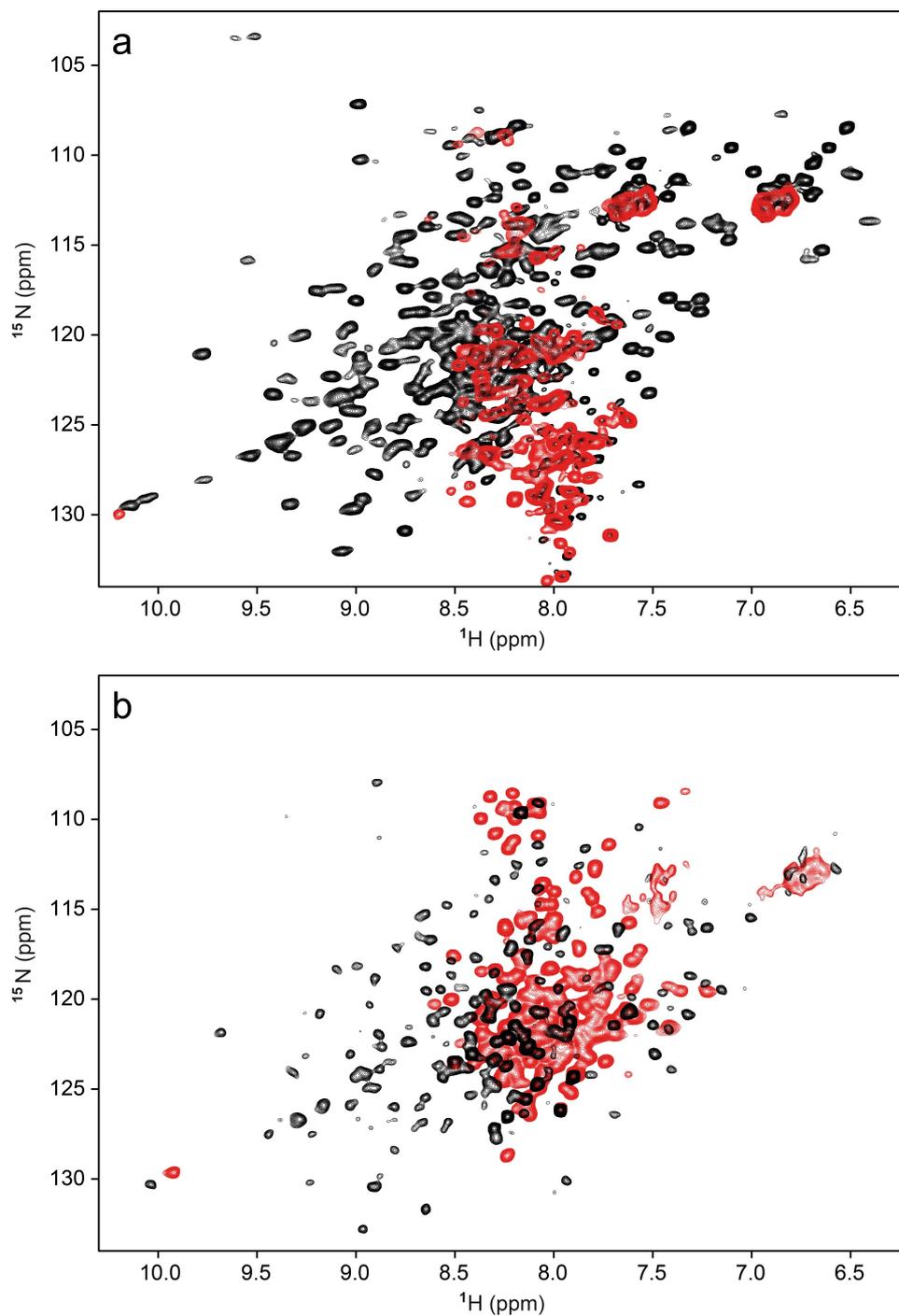
Fig. S3. Stability of TasA$_{261}$ analysed by NMR. a) Solution NMR ^{15}N-^{1}H-HSQC spectrum of ^{15}N-labelled recombinant TasA$_{261}$ stored for four weeks at room temperature and stabilized by a protease inhibitor cocktail (cOmplete, Roche) as well as 0.02% NaN$_3$ (black), superimposed with a spectrum recorded on a sample without additives (red). b) Superposition of a solution NMR ^{15}N-^{1}H-HSQC spectrum of TasA$_{261}$ (black) with a spectrum of TasA$_{261}$ in presence of 1% SDS (red).

480

Fig. S4

482
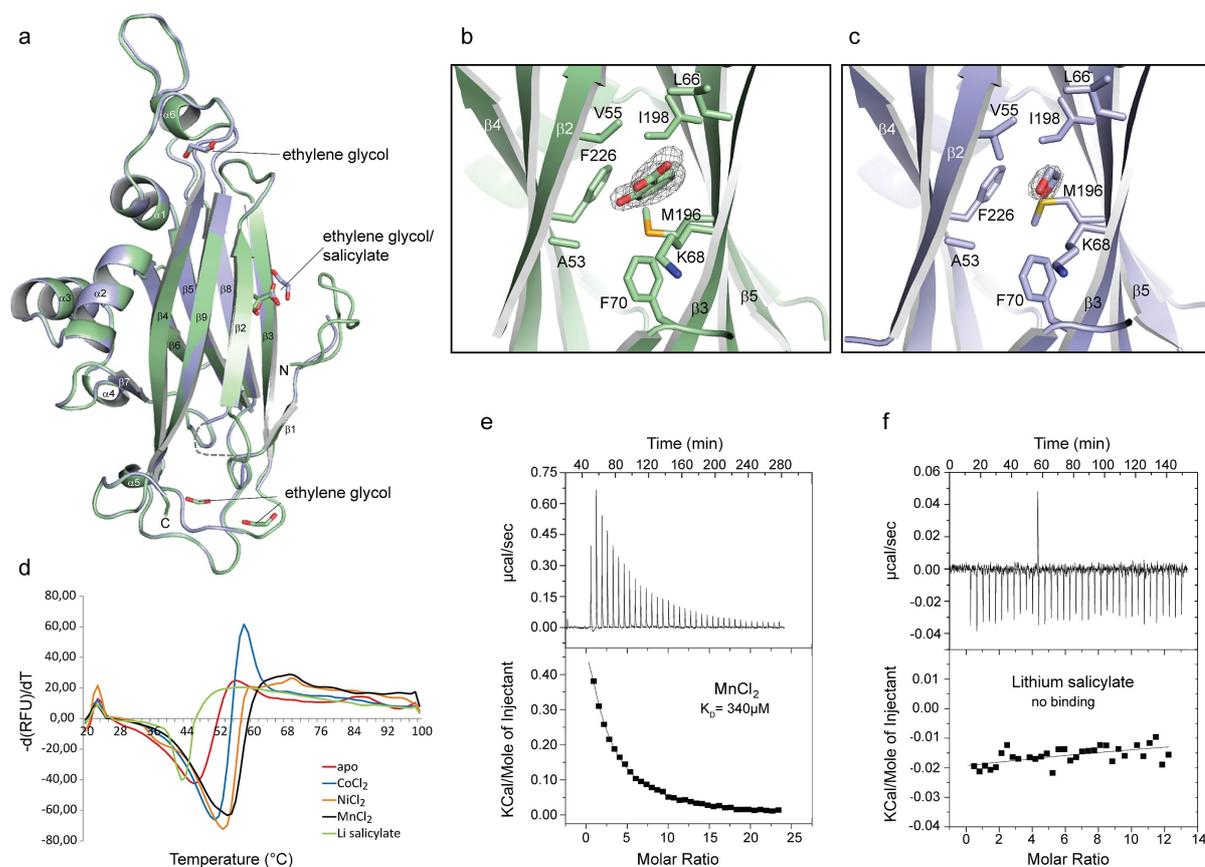
Fig. S4. Small molecule interaction, thermal shift assay and isothermal titration calorimetry (ITC) of TasA$_{239}$. a) Superimposed overall structures of SeMet TasA$_{239}$ crystallized in presence and absence of salicylate, shown in light green and light blue, respectively. Bound small molecules - ethylene glycol and salicylate - are shown in stick representation according to the colour of the TasA structures. b,c) Magnified view into the hydrophobic pocket of TasA, with b) detailed view of bound salicylate and c) ethylene glycol enclosed in their electron density (shown as grey mesh). The residues forming the hydrophobic pocket are depicted in stick representation. d) Thermal shift experiments with TasA$_{239}$ in absence (apo) and in the presence of divalent cations and lithium salicylate. The melting temperatures were determined from the first derivative of the melting curves, shown in different colours according the legend in the graph; e,f) Isothermal titration calorimetry e) of TasA$_{261}$ titrated with 9 mM MnCl$_2$ and f) TasA$_{239}$ with 4 mM lithium salicylate. The MnCl$_2$ data were fitted to a one-set-of-site binding model showing an apparent affinity of 340.1 ± 17.6 µM with N=1.05 ± 0.20 sites, ΔH=2309 ± 478.2 cal/mol, ΔS=23.7 cal/mol/deg. Lithium salicylate did not show any specific binding to TasA.

501

502   *Fig. S5*

CalY1 *B. cereus*   CalY1 *B. anthracis*   CalY2 *B. cereus*   CalY2 *B. anthracis*
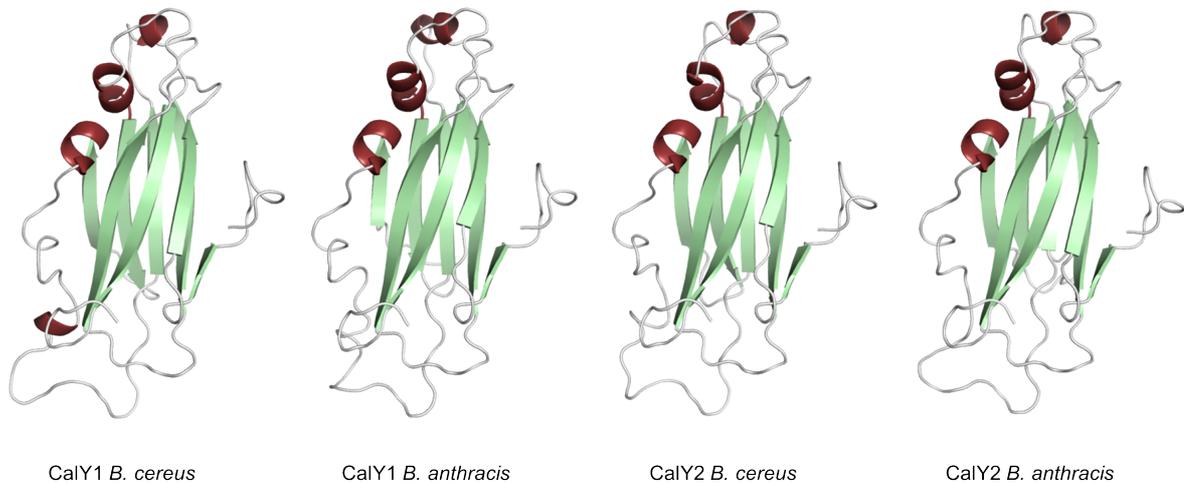
503

504

505   **Fig. S5a. Camelysin models.** The *tasA* operons in the pathogenic strains *B. cereus*
506   or *B. anthracis* contain a SipW homologue together with *tasA*-like genes, named
507   *calY1* and *calY2*. We generated four camelysin 3D models based on a hand-curated
508   multiple sequence alignment of the four *Bacillus cereus* sequences (Fig. S5B). TasA
509   belongs to the MEROPS family M73. The special feature of metalloproteases of the
510   type M73 along with M63 members is that metal ions are not coordinated via
511   histidines. Predicted alternative binding sites (see main text) are not clustered in the
512   TasA structure nor in the camelysin models. Application of prediction tools for metal
513   binding   sites   like   MIB   (http://bioinfo.cmu.edu.tw/MIB)   or   IonCom
514   (http://zhanglab.ccmb.med.umich.edu/IonCom/) did not yield convincing results for
515   TasA. However, not all members of the M73 family are known to be active
516   peptidases.

517

518

519
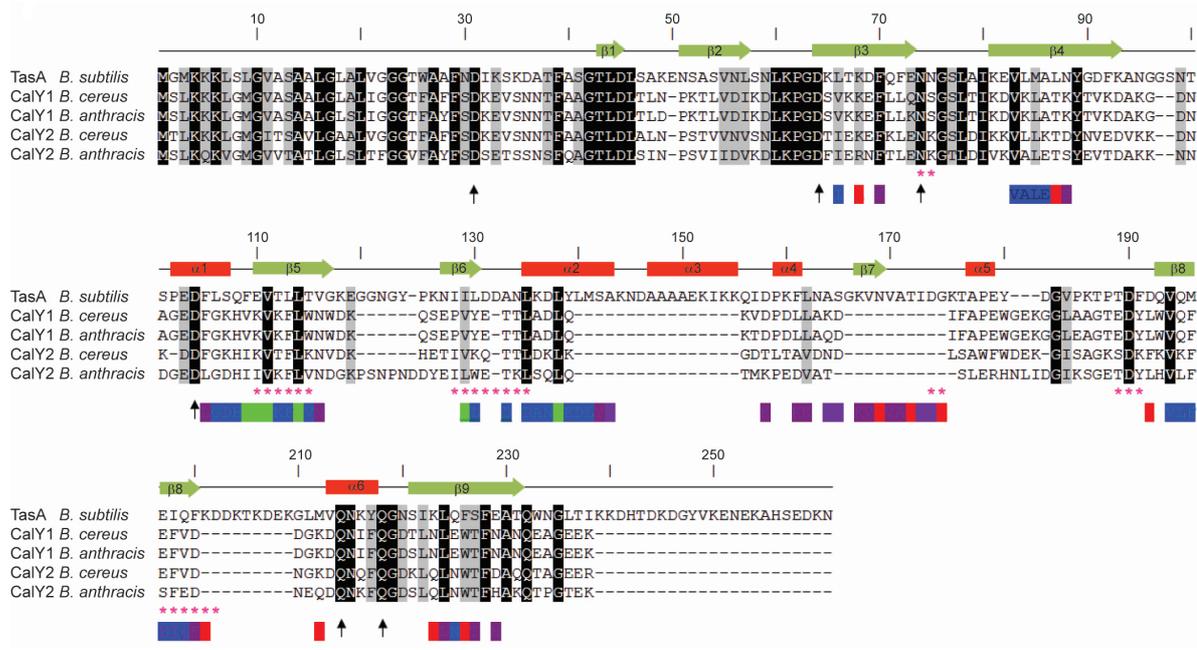
Fig. S5b. Alignment of TasA and camelysins from *B. cereus* and *B. anthracis*. The secondary structure of TasA is shown on top of the alignment. Residues whose signals experienced line broadening upon titration with Mn^{2+} are indicated by red asterisks below the sequences. Black arrows indicate residues proposed for metal coordination. Protected areas of $TasA_{239}$ in solution determined by H/D exchange and NMR analysis are shown below (amide proton signal still visible after red = 11 min; purple = 1 hour; blue = 6 hours; green = 5 days).

Fig. S6

550
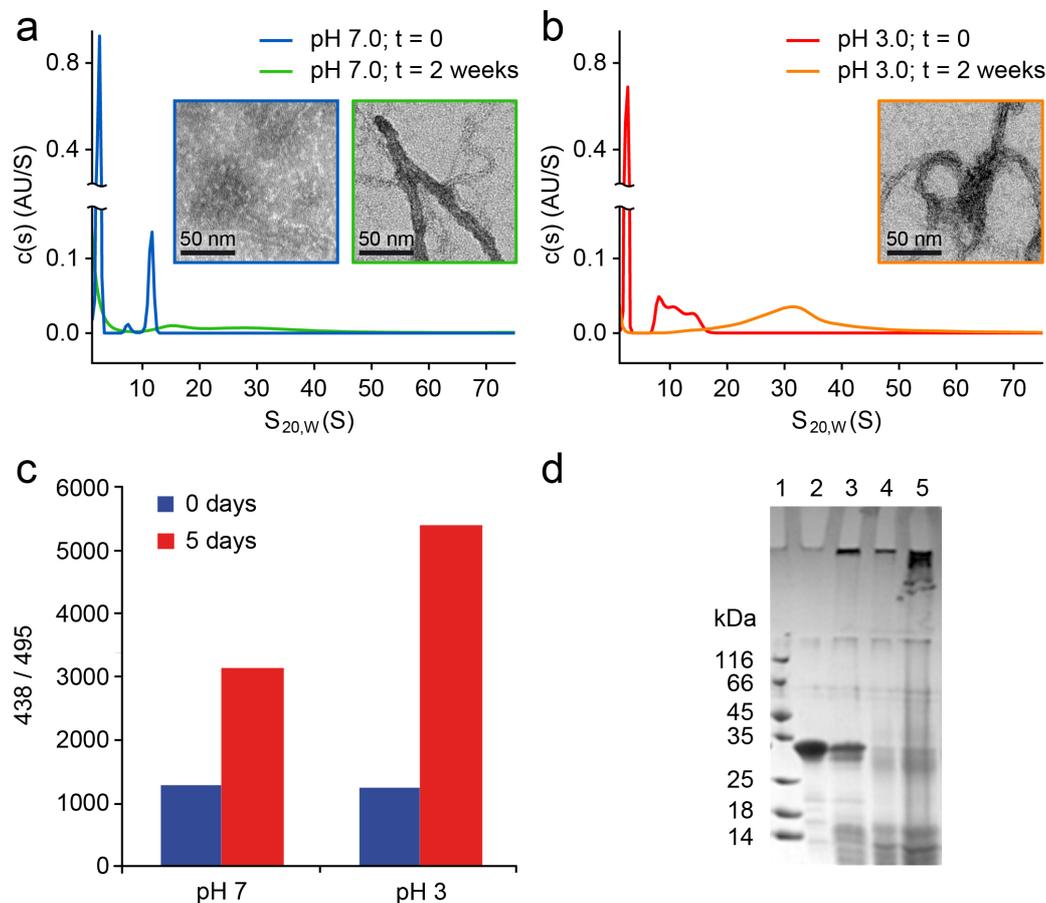
Fig. S6. TasA$_{261}$ oligomerisation/fibrillation at 40 °C over two weeks. a) Analytical ultracentrifugation (AUC) at pH 7 of freshly prepared TasA$_{261}$ (blue curve) and incubated for two weeks at 40 °C (green curve). At *t* = 0 (directly after gel filtration and sample concentration, prior to incubation at 40°C) mainly monomer and distinct oligomers are present. After two weeks, the monomer population is much smaller and a broad distribution of oligomers is formed. Small, initial fibrillar structures are already observed in electron micrographs at time point zero (blue frame) and solid, large fibers after incubation for 2 weeks at 40°C (green frame). b) AUC of freshly prepared TasA$_{261}$ (red curve) at pH 3 and incubated for two weeks at 40 °C (orange curve). At *t* = 0, (sample treated as in a) with pH additionally adjusted to 3) mainly monomer and several oligomers are present; two weeks later, the monomer is almost gone and a wide variety of high molecular weight species appear. The large fibrils that are visible in our electron micrographs do not

566 appear in the AUC profiles because they are too large and thus out of range
567 with respect to the x-axis in a) and b), and pelleted within the run-up time of
568 the centrifuge. c) Thioflavin-T (ThT) assay of a fresh preparation of TasA$_{261}$
569 (blue bars) and after incubation for 5 days at 40 °C (red bars), both at pH 7
570 (left) and pH 3 (right). The ThT response after 5 days is significant at both pH
571 settings. d) SDS-PAGE (Lane 1: marker proteins; Lane 2: control sample
572 taken before start of incubation; Lane 3: TasA$_{261}$ incubated at pH 7 and 40 °C;
573 Lanes 4 and 5: TasA$_{261}$ incubated at pH 3 at 40 °C. Lane 4 shows 40%
574 protein as in lane 5). In this case the monomer band is strongly reduced and
575 high molecular weight species are accumulated on top of the stacking gel.
576 Mass spectrometric analysis of these bands demonstrated intact TasA in
577 these aggregates/fibres. The corresponding EM picture for pH 7 is shown in
578 Fig. 4c.
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599

Fig. S7

602
603

604
605
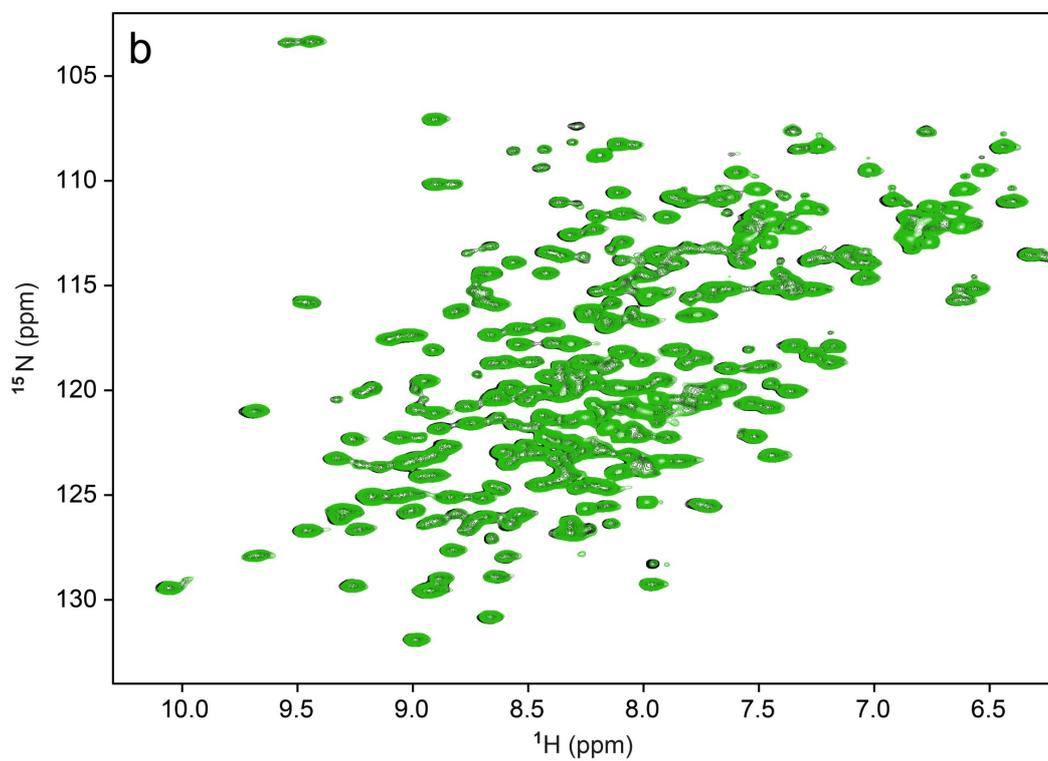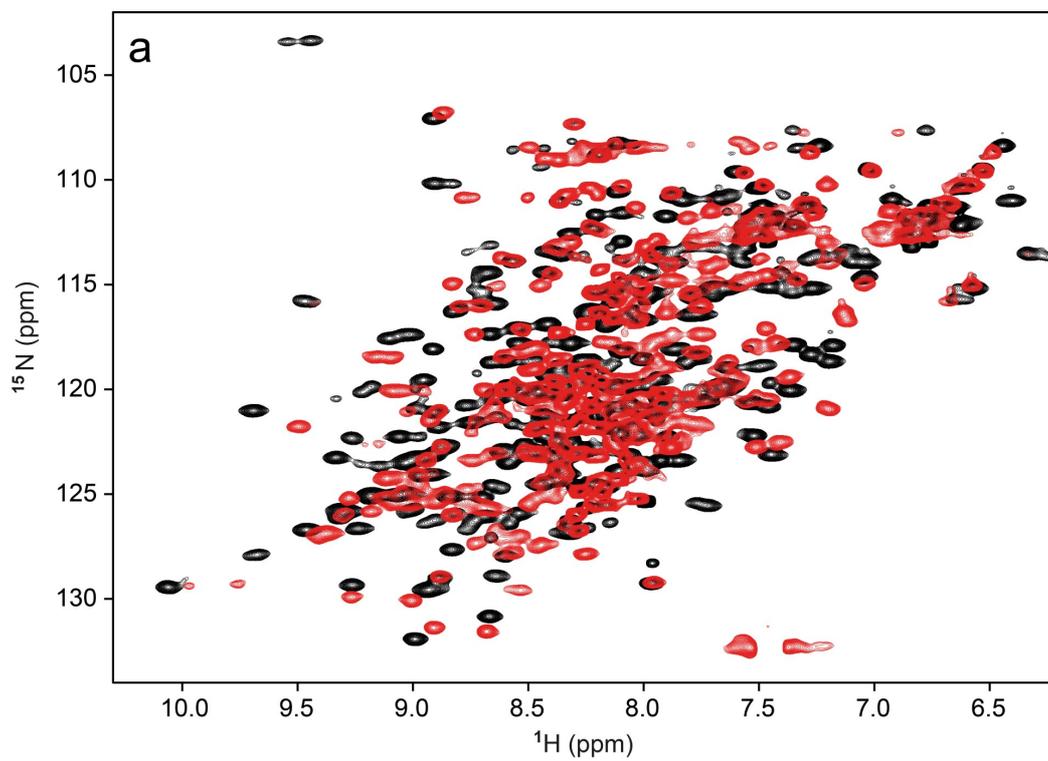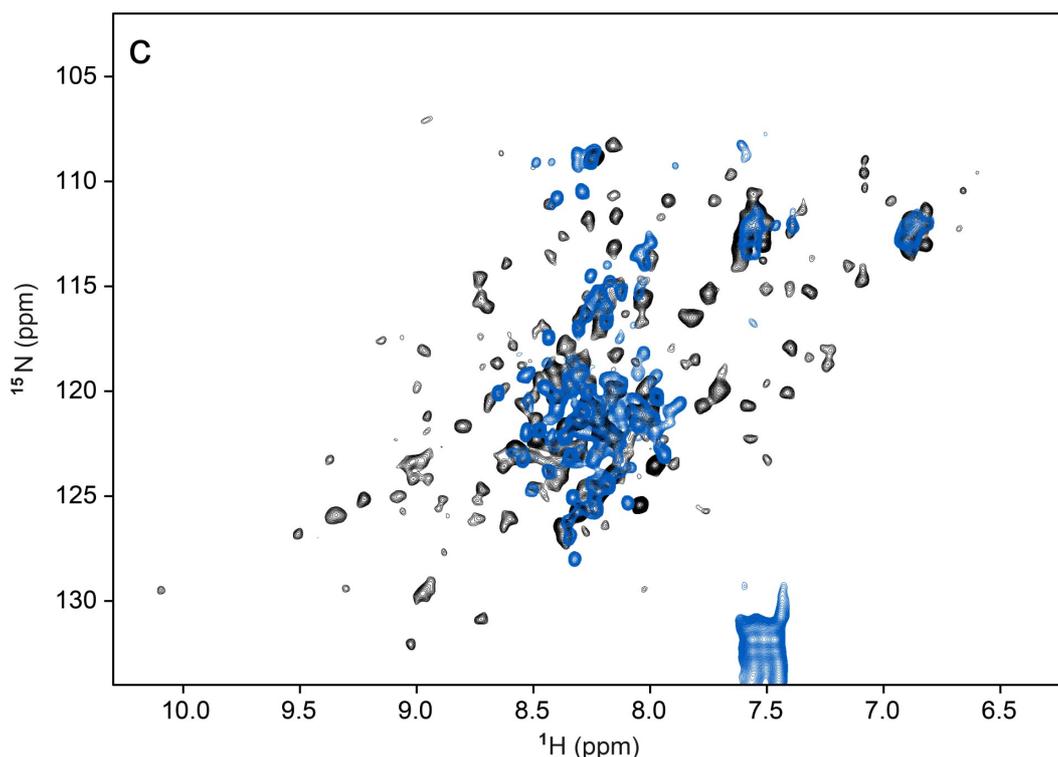
Fig. S7. pH-dependency of TasA$_{261}$. a) Superposition of solution NMR ^{15}N-^{1}H-
HSQC spectra of ^{15}N-labelled recombinant TasA$_{261}$ at pH 7 (black) and after dialysis
against pH 3 (red). b) HSQC spectra of TasA$_{261}$ at pH 7 (black) and of a TasA$_{261}$
preparation back-dialysed from pH 3 to pH 7 (green). c) HSQC spectra of TasA$_{261}$ at
pH 7 (black) and of a TasA$_{261}$ preparation rapidly adjusted to pH 4 (blue). The
solution NMR data of TasA were obtained at pH 7, but well diffracting crystals grew
only around pH 5 with truncated TasA$_{239}$. To assess if the conformation in the crystal
corresponds to the solution conformation and at which pH the protein conformation is
changed, TasA$_{261}$ was stepwise dialysed against buffer with decreasing pH. This mild
procedure is comparable with the gentle equilibration in crystallisation setups. Due to
the similarity of the spectra overlaid in (a) and (b), the conformation at pH 3 seems to
be very well comparable with that observed at pH 7. To our surprise, the
conformation is not compromised at pH 3 if the protein is dialysed (a), and a dialysis
back to pH 7 corroborates the presence of an intact TasA structure (b). In contrast,
fast adjustment to acidic pH by HCl addition causes an increase in disorder (c, blue
spectrum). This process is difficult to observe by CD spectroscopy that indicates 58%
unstructured, 7% helical (5% distorted helices) and 35% β-strand (13% βII) for
TasA$_{261}$ at pH 7 and 20 °C with only slightly different numbers at pH 7 and 90 °C.
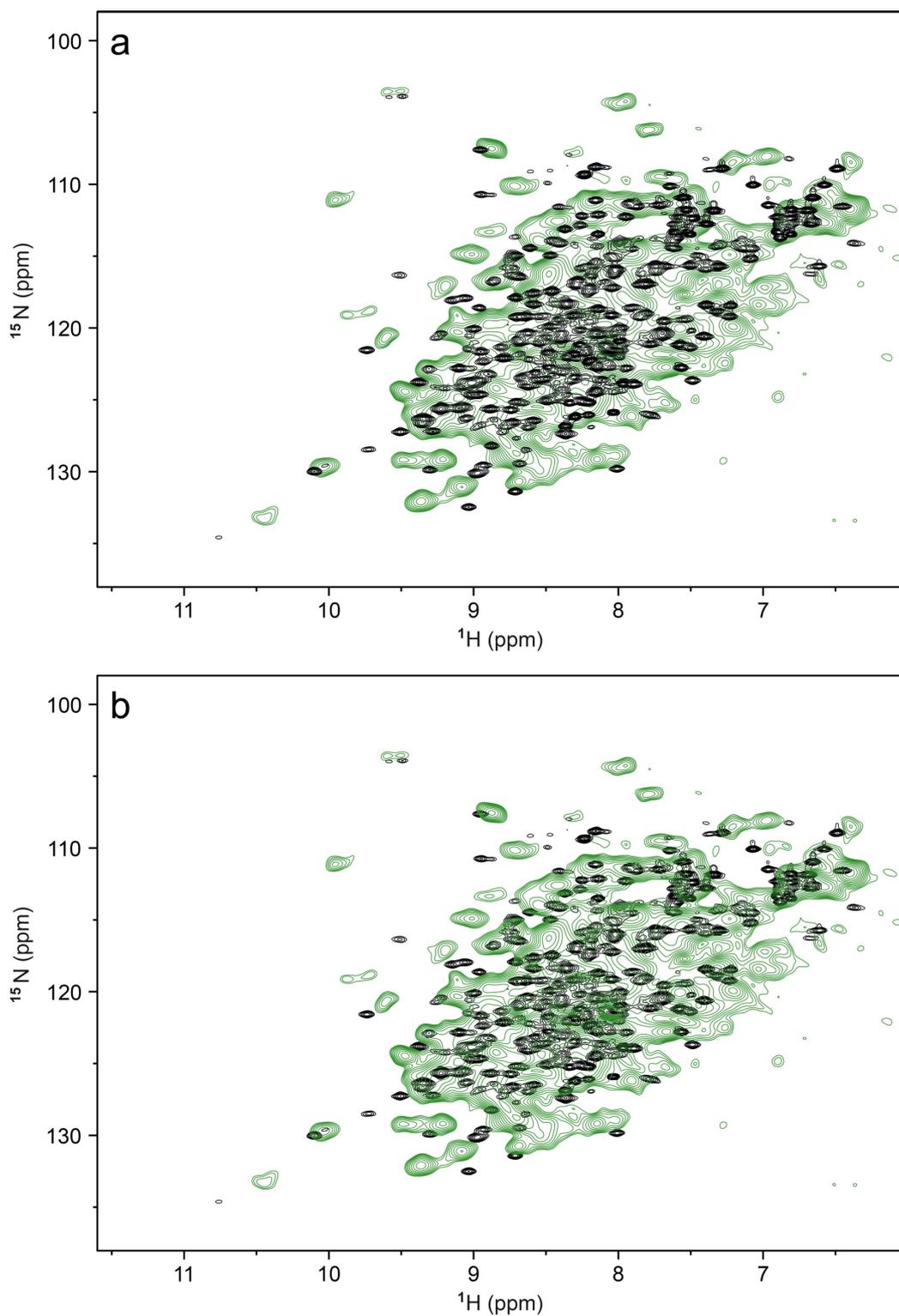(Fig. S14a and Table S3).

625
626
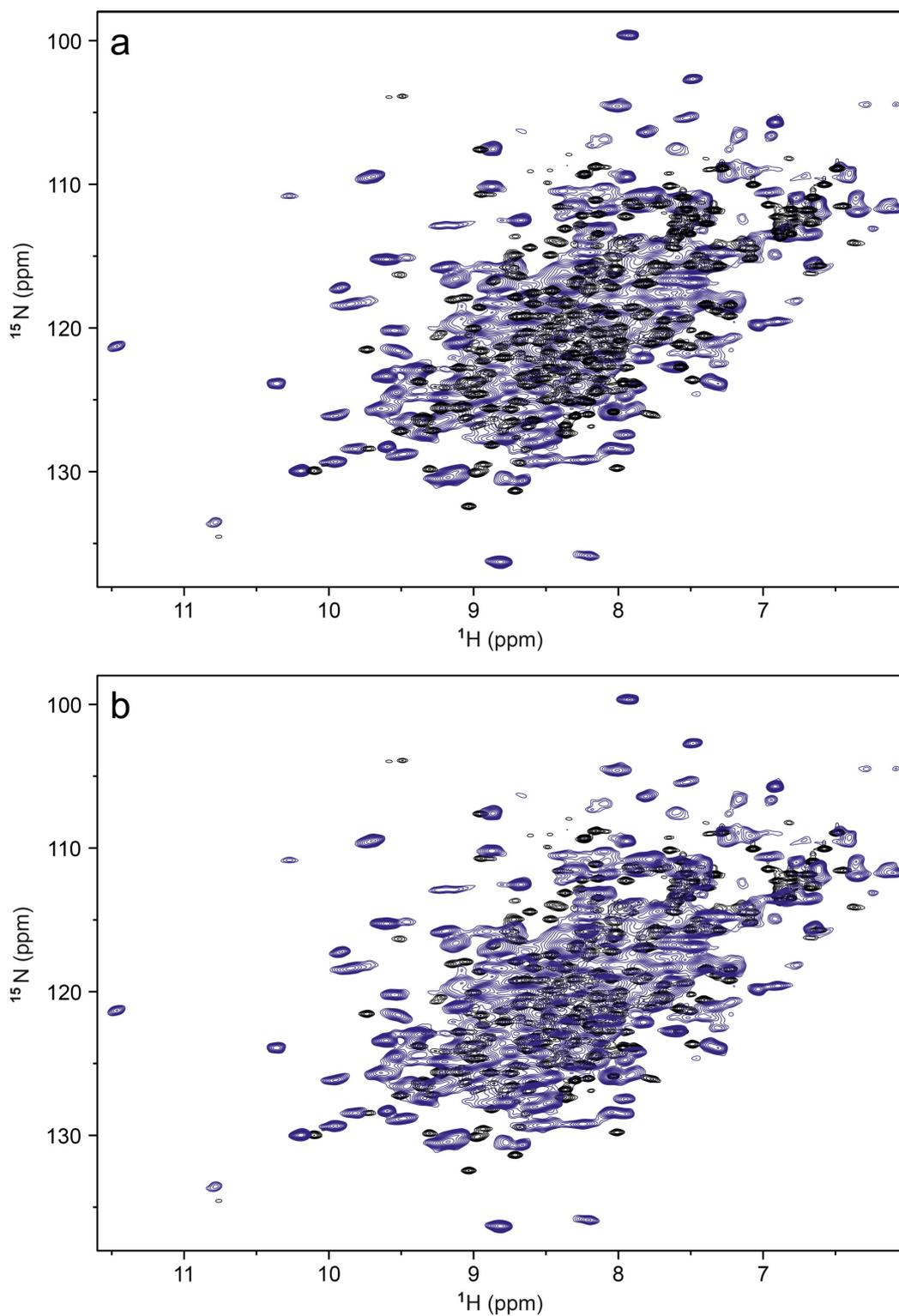
Fig. S8

Fig. S8. NMR data of soluble and precipitated TasA$_{261}$. Superposition of a solution NMR ^{15}N-^{1}H-HSQC spectrum of ^{2}H,^{15}N,^{13}C-labelled recombinant TasA$_{261}$ (black) and a MAS NMR ^{15}N-^{1}H correlation of ^{2}H,^{15}N,^{13}C-labelled precipitated protein (blue). a) Black contours on top; b) blue contours on top.
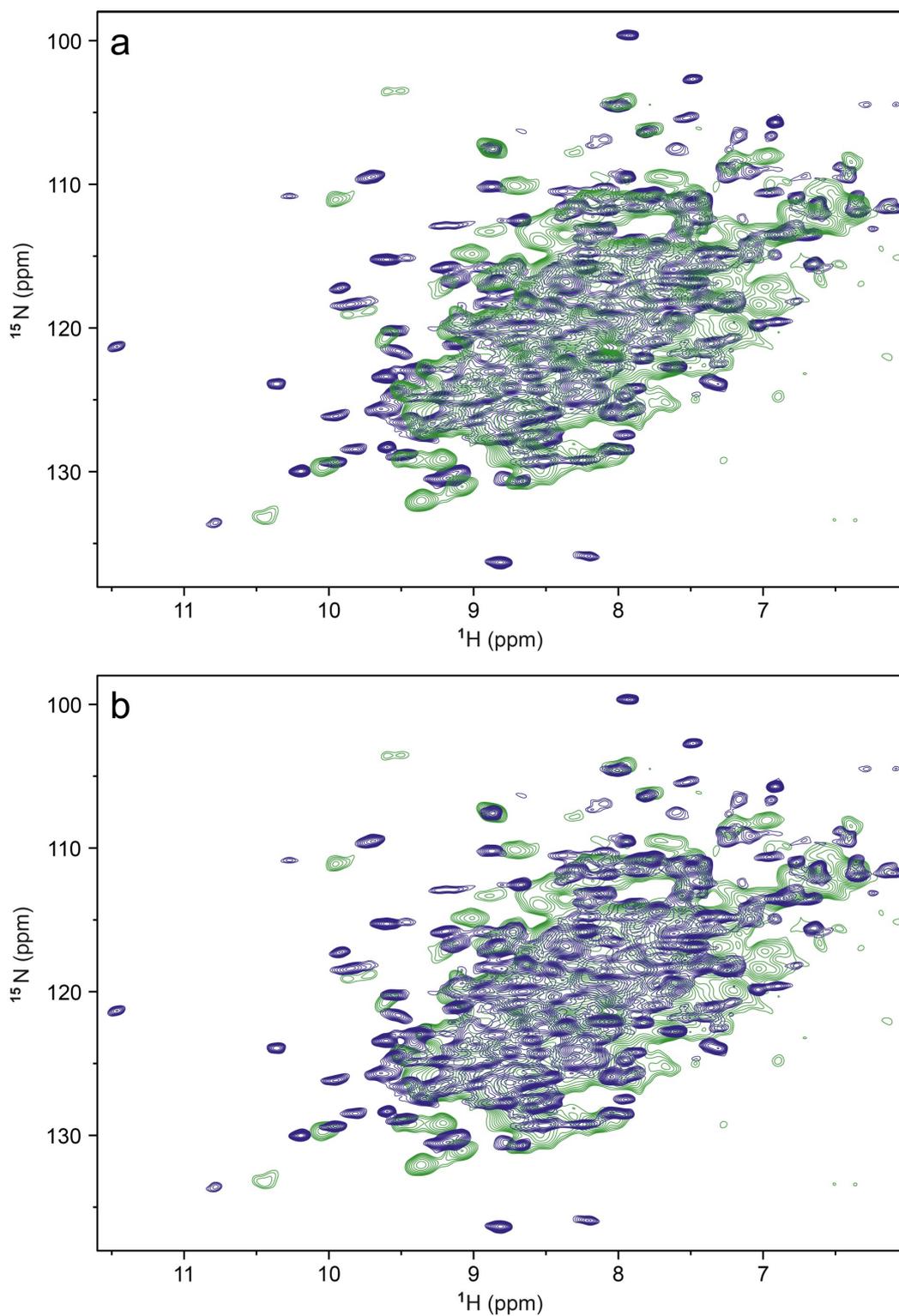
636   *Fig. S9*

637

638

639   **Fig. S9. NMR data of soluble TasA$_{261}$ and a gel-like TasA$_{239}$ preparation.**
640   Superposition of a solution NMR ^{15}N-^{1}H-HSQC spectrum of ^{2}H,^{15}N,^{13}C-labelled
641   recombinant TasA$_{261}$ (black) and a MAS NMR ^{15}N-^{1}H correlation of ^{2}H,^{15}N,^{13}C-
642   labelled, gel-like protein preparation (green). a) Black contours on top; b) green
643   contours on top.

Fig. S10

Fig. S10. NMR data of gel-like TasA$_{239}$ and precipitated TasA$_{261}$. Superposition of a MAS NMR ^{15}N-^{1}H correlation of ^{2}H,^{15}N,^{13}C-labelled, gel-like TasA$_{239}$ preparation (green) and a MAS NMR ^{15}N-^{1}H correlation of ^{2}H,^{15}N,^{13}C-labelled precipitated protein (blue). a) Green contours on top; b) blue contours on top.
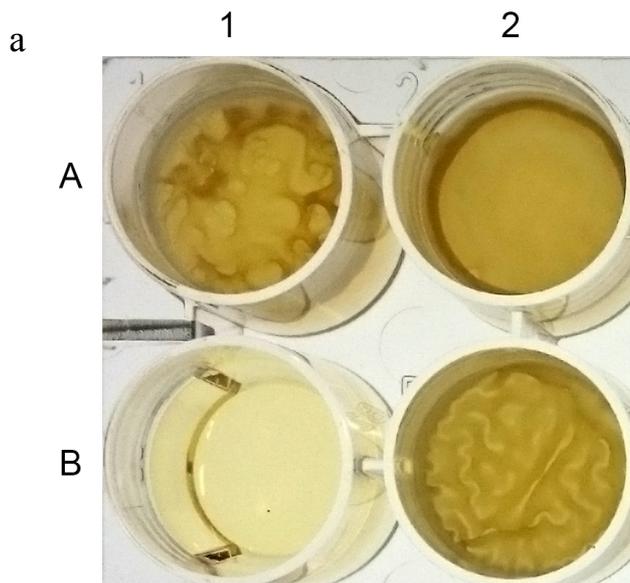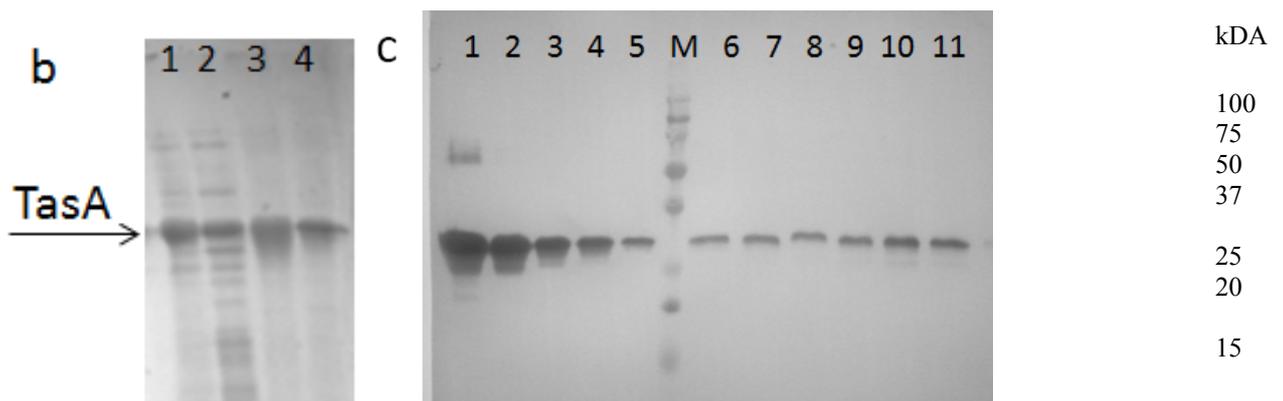
Fig. S11. Biofilm experiments with TasA$_{261}$. a) *ΔtasA* mutant biofilm rescued by ^{2}H, ^{13}C, ^{15}N-labelled TasA$_{261}$ in MOLP medium after 72 h at 30 °C. A1: Wild type *B. subtilis*; A2: *ΔtasA* mutant; B1: control, B2: *ΔtasA* mutant; rescued by recombinant TasA$_{261}$ (200 μg/1.4 ml).   b) Stability test of TasA$_{261}$ in spent cell free *B. subtilis* medium and analysis by SDS PAGE. Lane 1: TasA at t_0; lane 2: TasA after 6 days at room temperature; lane 3: control TasA at t_0; lane 4: control, TasA 6 days at room temperature. c) Immunoblot with Anti-TasA antibody of spent medium to check residual amount of TasA after biofilm formation by *B. subtilis* ΔtasA. Lanes 1-5 TasA calibration 500; 250; 80;40; 20 μg / 1,4 ml respectively; lanes 6-11 samples from set ups with 100 (lane 6) 250 (lanes 7 and 8) and 500 (lanes 9-11) μg TasA /1,4 ml. Up to 30 μg/ 1,4 ml are left in the spent medium from the originally provided 500 μg/1,4 ml.   Thus, more then 90 % of TasA became part of the biofilm as TasA is not degraded in the *B.subtilis* culture by proteases as shown in b).

671   *Fig. S12*
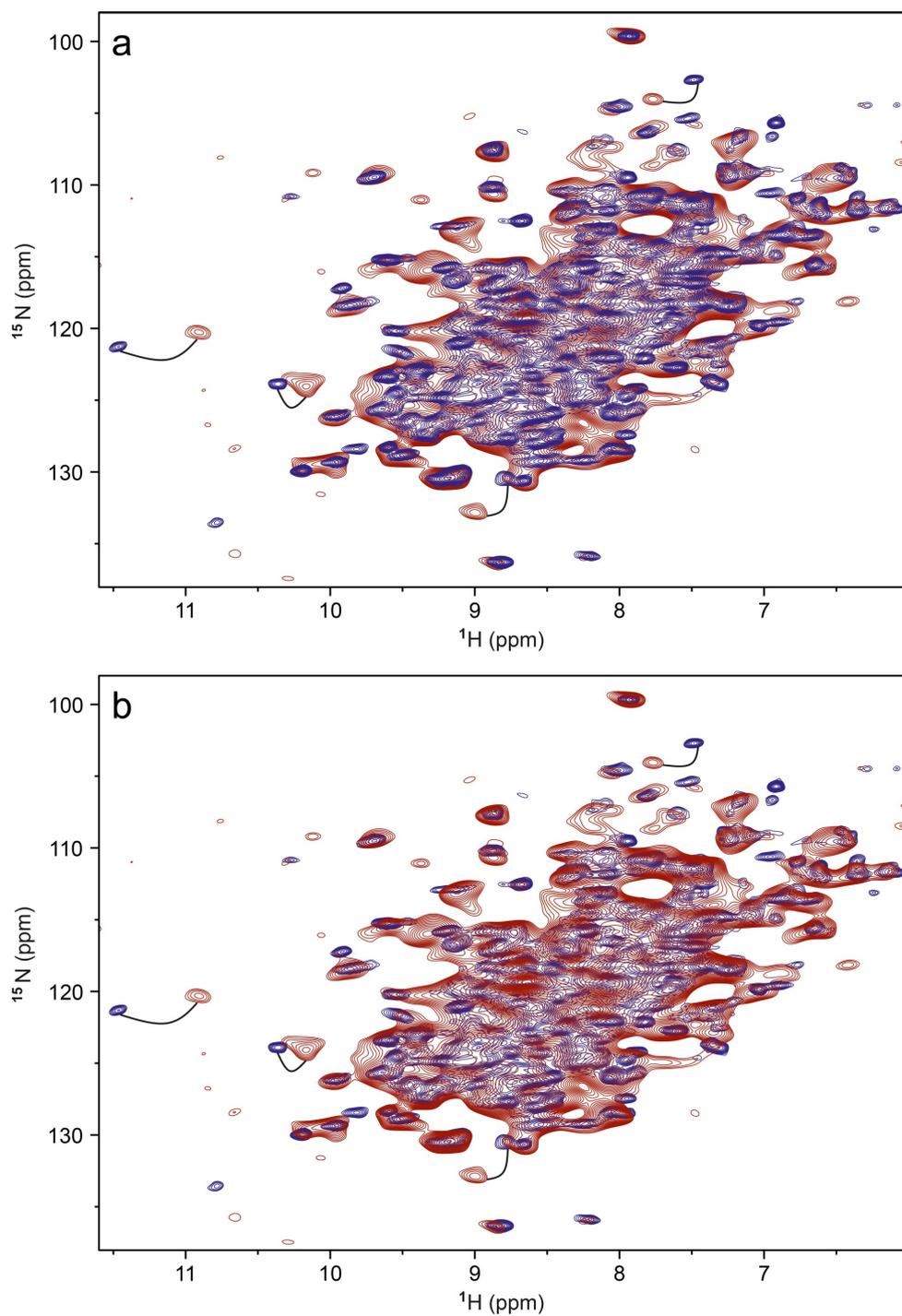
672

673

674

675   **Fig. S12. NMR data of precipitated TasA$_{261}$ and biofilm.** Superposition of a MAS
676   NMR ^{15}N-^{1}H-correlation of ^{2}H,^{15}N,^{13}C-labelled precipitated, fibrillar TasA$_{261}$ (blue) and
677   a MAS NMR ^{15}N-^{1}H-correlation of a biofilm formed by the *B. subtilis ΔtasA* strain after
678   supply of recombinant, ^{2}H,^{15}N,^{13}C-labelled TasA$_{261}$ (red). a) Blue contours on top; b)
679   red contours on top.

680

Fig. S13

682

683
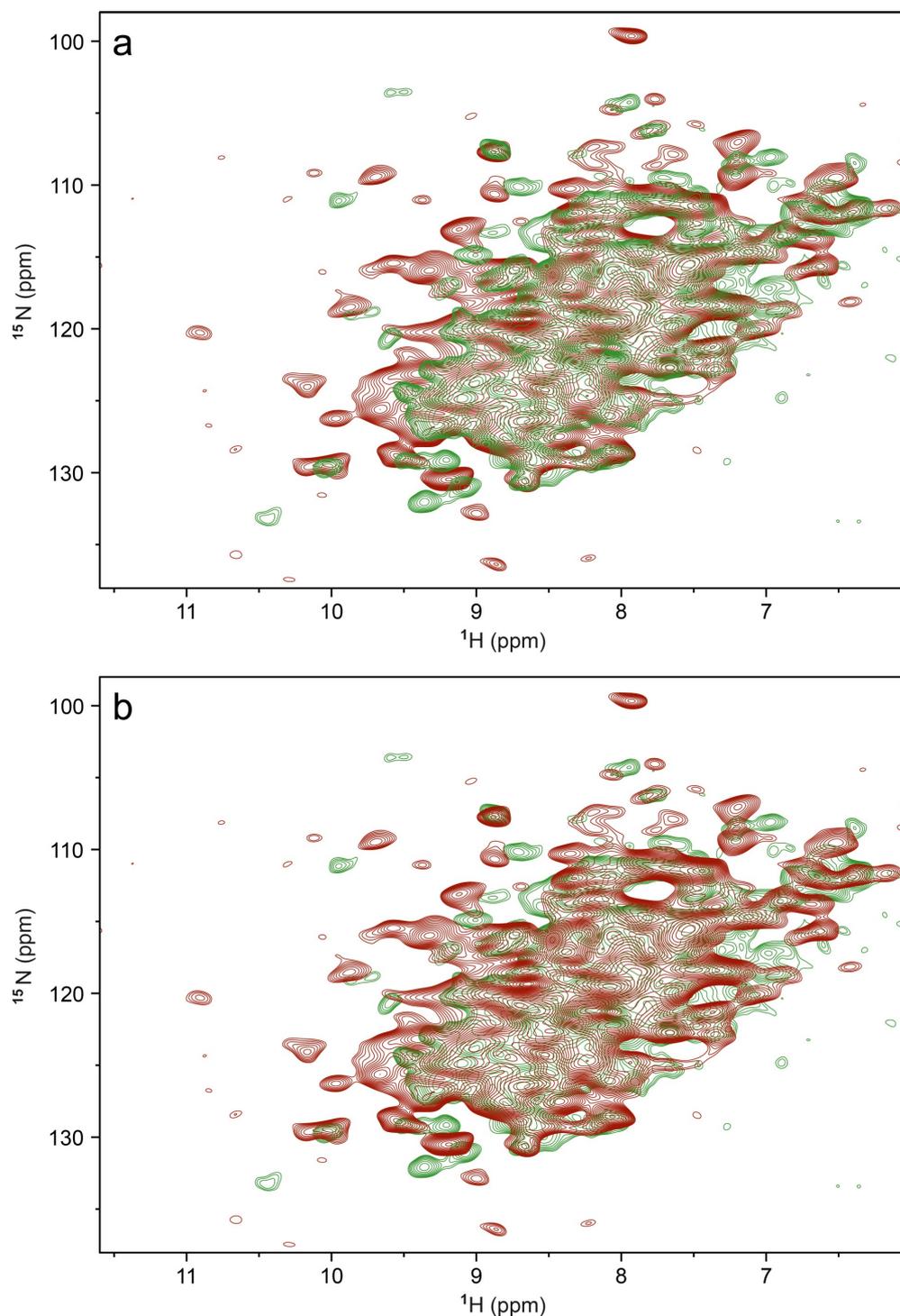
Fig. S13. NMR data of gel-like TasA$_{239}$ and biofilm. Superposition of a MAS NMR ^{15}N-^{1}H-correlation of ^{2}H,^{15}N,^{13}C-labelled gel-like TasA$_{239}$ (green) and a MAS NMR ^{15}N-^{1}H-correlation of a biofilm, formed by a *B. subtilis ΔtasA* strain after supply of recombinant, ^{2}H,^{15}N,^{13}C-labelled TasA$_{261}$ (red). a) Green contours on top; b) red contours on top.

689
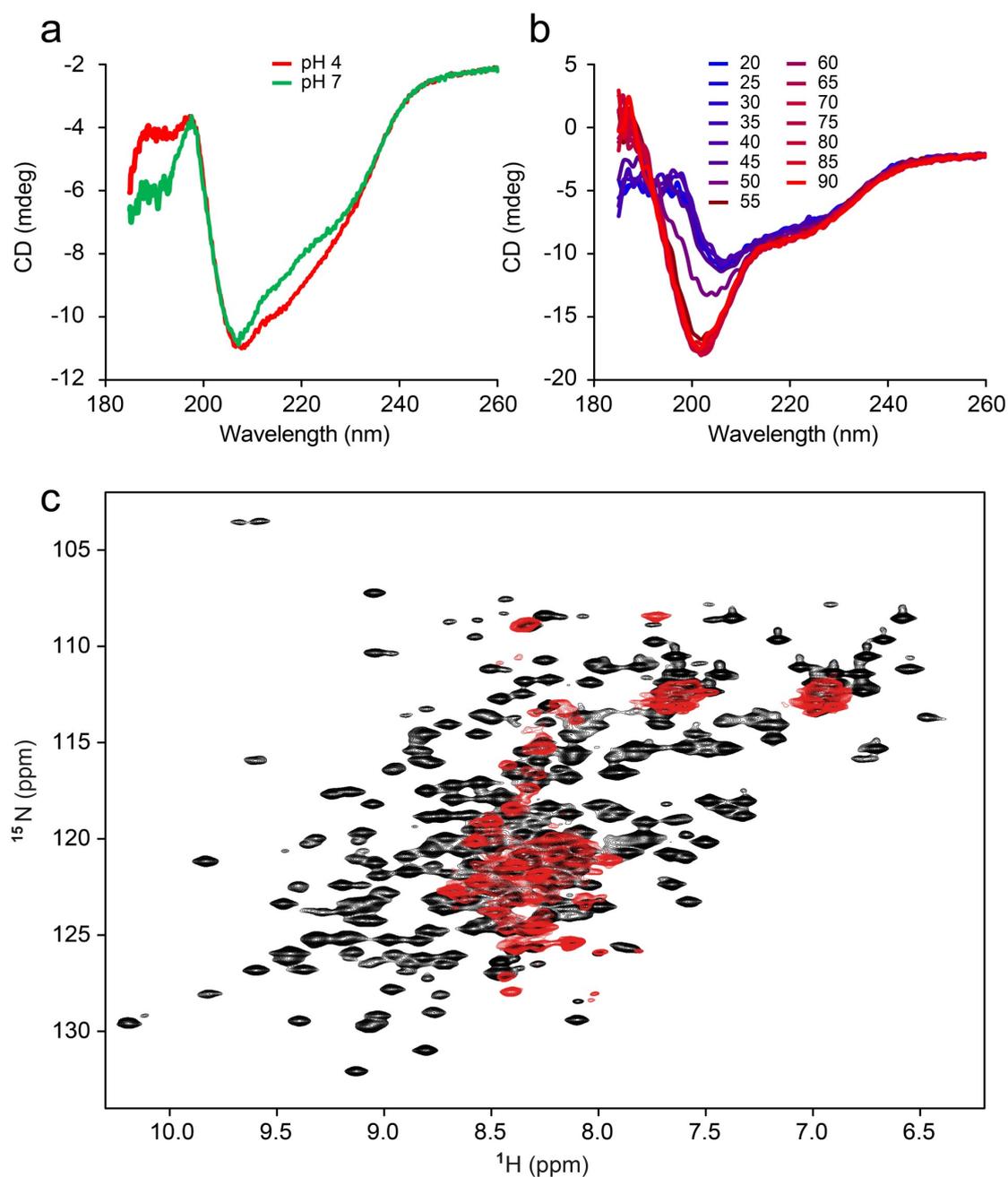
Fig. S14. pH- and temperature stability of TasA$_{261}$. a) CD spectrum of 5 µM TasA$_{261}$ in 20 mM phosphate buffer, 50 mM NaF at pH 7 (green) and adjusted to pH 4 (red). b) CD spectra of TasA$_{261}$ recorded during increasing the temperature from 20 to 90 °C in 5 °C steps. The purple curve (50 °C) indicates the melting point. c) Solution NMR ^{15}N-^{1}H-HSQC spectra of TasA$_{261}$ (black) and after elevating the temperature to 60 °C and back in steps of 1 °C per minute (red).

701   *Fig. S15*

702
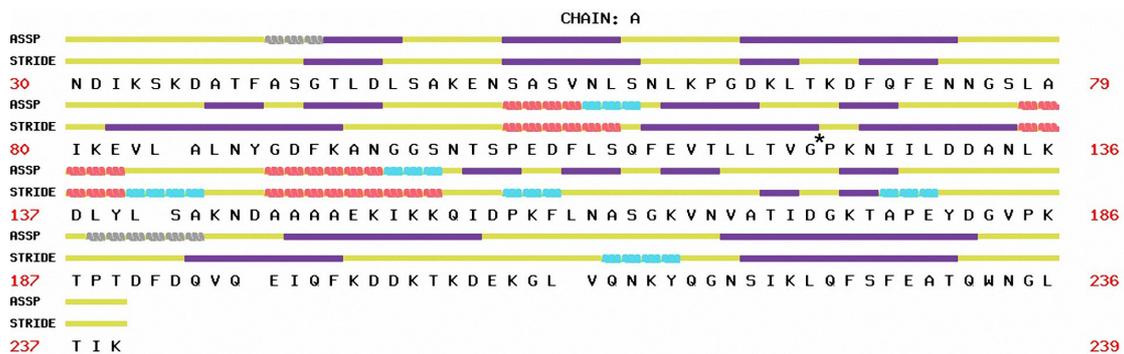
703

704

705   **Fig. S15. ASSP/STRIDE analysis of the TasA$_{239}$ X-ray data.** Helices: red; β-
706   strands: purple; loops: yellow. Atypical structural elements are colour coded as
707   following: 3$_{10}$ helices: light blue; polyproline-II helices (PPII): grey. Gaps in the amino
708   acid sequence stand for SeMet. The asterisk marks the position of 7 missing
709   residues due to low electron density in the crystal.

710

711

712

713

714

715

E Supplementary Data, Tables

Table S1. X-ray diffraction data collection and refinement statistics.

	TasA (SeMet) + Salicylate	TasA (apo) - Salicylate
Data collection		
Beamline	BESSY 14.1	BESSY 14.1
Wavelength (Å)	0.97973	0.91841
Space group	P1	P2$_1$
Cell dimensions		
a, b, c (Å)	40.73, 43.15, 51.74	41.66, 43.04, 51.69
α, β,γ (°)	89.92, 91.86, 90.01	90.00, 90.17, 90.00
Resolution (Å)*	43.15 - 1.56 (1.65 - 1.56)	43.04 - 1.86 (1.97- 1.86)
R$_{meas}$*	6.5 (43.4)	9.4 (77.9)
< I / σ(I) >*	9.51 (2.07)	13.83 (2.09)
Completeness (%)*	94.2 (91.9)	99.6 (98.3)
Redundancy*	2.1	4.4
Refinement		
Resolution (Å)	1.56	1.86
No. reflections	47368	14834
R$_{work}$ / R$_{free}$ (%)	15.22 / 17.52	17.26 / 21.18
No. atoms		
Protein	3409	1472
Ligands/ions	32	176
Water	331	142
Mean B factor (Å^2)	25.10	28.60
R.m.s deviations		
Bond lengths (Å)	0.016	0.007
Bond angles (°)	1.85	0.99
Mols/AU	2	1

718 *Data in highest resolution shell are indicated in parentheses.

720   *Table S2. Analytical ultracentrifugation results.*

Sample		Monomer			Total Oligomeric Fraction		
history of sample	f/f_0	fraction	S_{20},w	MW (kDa)	fraction	S_{20},w	MW (kDa)
TasA$_{261}$ pH 7 *fresh*	1.4	79%	2.3	~27	21%	11.5 6.5	~300 ~140
TasA$_{261}$ pH 7 *2 weeks, 40 °C*	1,6	20%	2.3	~27	50% 25%	20-60 6-20	~2000-9000 ~350-2000
TasA$_{261}$ pH 3 *adjusted, fresh*	1,4	66%	2.3	~27	33%	7-15	~150-500
TasA$_{261}$ pH 3 *adjusted, 2 weeks 40 °C*	2.6	~2%	2.3	~27	5% 92%	50-75 10-50	>7000 ~600-7000
TasA$_{239}$ *fresh, pH 7*	1.4	~99%	2.1	~24	~1%		
TasA$_{261}$ *Stored -20 °C, pH 7*	1.3	40%	2.7	~32	60%	12	~300
TasA$_{261}$ *as above, dialysed pH 3.5*	1.4	92%	2.6	~32	5%	5	~100
TasA$_{261}$ *as above, dialysed pH 3*	1.4	60%	2.6	~32	22% 18%	11-18 5-11	~300-650 ~100-300

(Figure S6a & b — upper block; Figure 4a — lower block)

721

722 *Table S3. Comparison of structural elements of TasA according to Jpred*
723 *prediction, CD - data deconvolution by DiChroWeb (and determined by X-ray*
724 *crystallography.*

Proportion in %	Jpred	CD pH 7 20°C	CD pH 7 90°C	X-ray
Helix	3	7 (5)	6 (5)	18
Strand	32	35 (13)	33 (13)	32
Turn		14	14	
disordered		44	46	
others	66			50 (incl PPII)

725 In parenthesis values for distorted α helices and β strands (βII)

726

727

728

729 *Table S4. Primer pairs for TasA cloning and modification.*

TasA$_{261}$	**underlined - TasA specific**	Ddel	Sall
forward 1	5'CCAGTGGGTCTCAGGTGGT<u>GCATTTAACG ACATTAAATC AAAGG</u>3'		
reverse 1	5'ATTATAGTCGACTTA<u>ATTTTTTATCCTCGCTATGCGC</u>3'		

730

TasA$_{239}$	red – mutation for generation of a stop codon	
forward 2	5'CAATCAAA**T**AGGACCATACTGATAAAGACGGTTATGTGAAAGAAAATG3'	
reverse 2	5'AGTATGGTCC T**A**TTTGATTGTCAAGCCGTT CCACTGTGTAG3'	

731 Synthesized by: 1 Thermo Fisher Scientific (Schwerte, Germany) 2 BioTez (Berlin, Germany)

732

733

734