1 **A new twist on bacterial motility – two distinct type IV pili revealed by cryoEM**

2

3 Alexander Neuhaus[1,2], Muniyandi Selvaraj[3,-], Ralf Salzer[4,X], Julian D. Langer[5,6], Kerstin Kruse[4],

4 Kelly Sanders[1,2], Bertram Daum[1,2], Beate Averhoff[4] and Vicki A. M. Gold[1,2,#],

5

6 [1] Living Systems Institute, University of Exeter, Stocker Road, EX4 4QD, United Kingdom

7 [2] College of Life and Environmental Sciences, Geoffrey Pope, University of Exeter, Stocker

8 Road, Exeter, EX4 4QD, United Kingdom

9 [3] Department of Structural Biology, Max Planck Institute of Biophysics, Max-von-Laue Str. 3,

10 60438 Frankfurt am Main, Germany

11 - Current address: Laboratory of Structural Biology, Helsinki Institute of Life Science, 00014

12 University of Helsinki, Finland

13 [4] Molecular Microbiology and Bioenergetics, Institute of Molecular Biosciences, Goethe

14 University Frankfurt, Max-von-Laue Str. 9, 60438 Frankfurt am Main, Germany

15 [X] Current address: Structural Studies Division, Medical Research Council – Laboratory of

16 Molecular Biology, Cambridge Biomedical Campus, Francis Crick Ave, Cambridge CB2 0QH,

17 United Kingdom

18 [5] Department of Molecular Membrane Biology, Max Planck Institute of Biophysics, Max-von-

19 Laue Str. 3, 60438 Frankfurt am Main, Germany

20 [6] Proteomics, Max Planck Institute for Brain Research, Max-von-Laue Str. 4, 60438 Frankfurt

21 am Main, Germany

22

23 [#]corresponding author

24

Summary

Many bacteria express flexible protein filaments on their surface that enable a variety of important cellular functions. Type IV pili are examples of such filaments and are comprised of a helical assembly of repeating pilin subunits. Type IV pili are involved in motility (twitching), surface adhesion, biofilm formation and DNA uptake (natural transformation). They are therefore powerful structures that enable bacterial proliferation and genetic adaptation, potentially leading to the development of pathogenicity and antibiotic resistance. They are also targets for drug development.

By a complement of experimental approaches, we show that the bacterium *Thermus thermophilus* produces two different forms of type IV pilus. We have determined the structures of both and built atomic models. The structures answer key unresolved questions regarding the molecular architecture of type IV pili and identify a new type of pilin. We also delineate the roles of the two filaments in promoting twitching and natural transformation.

41 **Main**

42   Type IV pili (T4P) are flexible extracellular protein filaments found on many bacteria. They

43   form multifunctional fibres involved in twitching motility, adhesion, immune evasion,

44   bacteriophage infection, virulence and colony formation. T4P have also been linked to DNA

45   uptake, called natural transformation, which is a powerful mechanism that enables genetic

46   adaptation [1–3]. The filaments are homopolymers composed of thousands of pilin subunits,

47   which form helical arrays measuring several micrometres in length. How T4P are involved in

48   seemingly unrelated functions such as motility and natural transformation is so far unclear.

49

50   Depending on the bacterial species, pilins range from 90 to 250 amino acids in length. They

51   are produced as prepilins with a typical class III signal peptide [4,5]. The preprotein is

52   translocated via the Sec pathway into the cell membrane where the signal peptide is cleaved

53   by prepilin peptidase, priming the pilin for incorporation into the growing pilus. Filament

54   assembly is ATP-dependent and occurs at an inner membrane platform which contains PilC,

55   PilM, PilN and PilO [6]. In *Thermus thermophilus,* assembly of pilins into a T4P filament depends

56   on the assembly ATPase PilF, which interacts with the inner membrane platform via PilM [7].

57   Two retraction ATPases, PilT1 and PilT2, are essential for T4P depolymerisation [8,9]. T4P are

58   extruded by the outer membrane secretin PilQ [10–13]. Recently, it has been suggested that

59   expression of the *T. thermophilus* major pilin PilA4 is temperature dependent, leading to

60   hyperpiliation at suboptimal growth temperatures [14]. The first two *in situ* structures of T4P

61   assembly machineries were solved only recently in both open (pilus assembled) and closed

62   (pilus retracted) states [11,15], yet detailed information regarding the molecular interactions

63   governing filament assembly was lacking.

64

65   Crystal structures of full length pilins or head domains from various bacteria are available

66   in different oligomeric states [6,16–22]. Pilins have a conserved N-terminal α-helix, with a 4-5

67   stranded anti-parallel β-sheet at the C-terminus. The α-helix forms the core of the filament,

68   while the globular β-sheet head domain is solvent exposed and subject to post-translational

69   modifications [16,17]. To date, five low-resolution electron cryo-microscopy (cryoEM) structures

70   of isolated T4P have been reported. The first, a 12.5 Å structure from *Neisseria gonorrhoeae*,

71   was sufficient to place crystal structures of pilins into the data but not to resolve their

72   structure within the map [17]. Four subsequent structures from *Pseudomonas aeruginosa*, two

73   *Neisseria* species and enterohemorrhagic *Escherichia coli* have been determined in the 5-8 Å

74   resolution range [23–25].

75

76   In this study, we combine different modes of cryoEM: electron cryo-tomography (cryoET) and

77   single-particle cryoEM, with functional data to study the T4P of *T. thermophilus,* which is a

78   well-established model organism*.* Surprisingly, we detect two forms of T4P, a wider and a

79   narrower form, assembled through the same machinery. We determine structures of the two

80   filaments at so-far unprecedented resolution (3.2 Å and 3.5 Å, respectively). This has enabled

81   us to visualise near atomic-level detail and build atomic models for each filament *ab initio*.

82   Our data unambiguously demonstrate that the wider pilus is composed of the major pilin

83   PilA4. Proteomics and knock-out mutants reveal that the narrow pilus consists of a previously

84   unknown pilin, which we name PilA5. Functional experiments confirm that PilA4 is involved in

85   natural transformation whereas PilA5 is essential for twitching motility [26]. Our results shed

86   new light on bacterial motility and gene transfer, and will help to guide the development of

87   new drugs to fight microbial pathogens.

88

89   ***Two types of T4P are assembled from the same machinery***

90   Cells of *T. thermophilus* strain HB27 assemble T4P pili on their surface [27], predominantly at

91   the cell poles [11]. Performing cryoET on cells grown at the optimal growth temperature of 68

92   °C revealed two types of pilus, with differences in their diameter. They emerge from the same

93   assembly machinery (Fig. 1a, b), suggesting that they are both T4P. We have previously shown

94   that transcription of the major pilin gene, *pilA4*, is upregulated at low temperature [14]. To

95   address the question of whether the growth temperature affects the assembly of the two

96   forms of pilus, we analysed the pili emerging from cells grown at the sub-optimal growth

97   temperature of 58 °C by cryoET. Again, two types of pilus were observed (Fig. 1c, d). Pili

98   emerge from T4P complexes only sporadically [11], thus filaments were isolated from cells in

99   order to investigate their structure in more detail. Both wide and narrow forms of the filament

100   were detected in these preparations (Fig. S1a-d).

101

102   To investigate the composition of the two pilus forms, we performed a quantitative

103   bottom-up proteomics analysis. Protein abundance was evaluated by Label-Free Quantitation

104   (LFQ) to determine relative enrichment or loss of particular pilus proteins at either 68 °C or 58

105   °C. At both temperatures, the major pilin subunit PilA4 was identified as the most abundant

106   protein component (Fig. S2a). The amount of PilA4 increased significantly at 58 °C, in line with

107   the hyperpiliation phenotype [14]. The second most abundant protein was the uncharacterised

108   protein TT_C1836; the relative abundance was also significantly increased at 58 ˚C.

109

110   To refine the identification of pilins, we performed gel based proteomic analysis (Fig. S2b,

111   Table S1). In order to increase the hyperpiliation phenotype, we further reduced the growth

112   temperature to 55 °C. At this temperature we could identify PilA4 and TT_C1836 as the most

113   abundant proteins in the lower molecular weight bands, likely representing the pilin

114   monomers. In contrast, at the optimal growth temperature of 68 °C, only PilA4 was identified

115   reliably. We questioned whether the two T4P were expressed due to differences in

116   temperature or growth phase. To quantify the abundance of different filaments, cells were

117   grown under different conditions and analysed in the electron microscope. Pili at both cell

118   poles were selected for 2D classification, which enables grouping of similar structures (Fig.

119   S3a). For cells grown on plates, both wide and narrow pili were present at a similar level,

120   whereas the ratio of the two was shifted towards the wider form for cells grown in liquid

121   medium. At 55 °C, only the total number of pili per cell increased, while the ratio between

122   wide and narrow pili was similar between the two temperatures (Fig. S3b).

123

124   To analyse the role of PilA4 and TT_C1836 in pilus assembly, we investigated the number of

125   wide and narrow pili per cell in deletion strains grown in liquid media to exponential phase

126   (Fig. S3c). PilA4 deficient cells (*pilA4::km*) were not able to assemble any pili reliably, whereas

127   TT_C1836 deficient cells (*TT_C1836::km*) were only defective in their ability to assemble

128   narrow pili. These findings suggest that PilA4 has a role in producing both pilus forms, while

129   TT_C1836 appears to be crucial for the formation of narrow pili only. However, these data do

130   not allow us to discriminate whether the proteins have structural roles in comprising the

131   filaments, or have a more functional role in their assembly mechanism.

132

133   ***High-resolution maps determine the 3D structure of both filaments***

134   In order to investigate the architecture and protein composition of T4P at high resolution,

135   both filaments were subjected to analysis by single particle cryoEM and helical reconstruction.

136   In our micrographs, the wide pili appeared very straight while the narrow pili showed a much

137   higher degree of curvature, up to 2 μm radius (arrowhead in Fig. S4A, B). Based on 2D classes

138   (Fig. S4c, d) we determined the helical parameters for the wide and the narrow pilus (Fig. S5a–

139   d, see methods section for details). The rise and twist of the wide pilus measured 9.33 Å and

140   92.5° respectively, and the narrow pilus had a rise of 11.30 Å and a twist of 84.3°. Helical

141   reconstruction [28] resulted in maps at 3.22 Å (wide) and 3.49 Å (narrow) (Fig. 2a, b, S5e, f). The

142   diameter of the wide fibre is 70 Å (Fig. 2a) and is roughly cylindrical. In contrast, the narrow

143   filament has a zigzag-like appearance in projection, owing to a 15 Å-wide groove that winds

144   through the fibre. The diameter at any position along the long axis of the fibre axis is therefore

145   only 45 Å (Fig. 2b). Both structures were in good agreement with the data obtained *in situ* by

5

146   cryoET (Fig. 1). A low-resolution structure of the T4P from *T. thermophilus* was previously

147   determined by cryoET and sub-tomogram averaging, with a diameter of ~3.5 nm [11]. It seems

148   likely that this conformation represents the narrower form of the pilus.

149

150   Both maps clearly resolved individual pilin monomers. The peptide backbone could be easily

151   traced throughout each subunit and large side chains were visible (Fig. 2c, d). The centre of

152   each filament is formed by a bundle of long N-terminal α-helices, as has been demonstrated

153   for other filaments [17,23–25,29]. In both maps, each α-helix is interrupted by an unfolded stretch

154   (brackets in Fig. 2c, d), a conserved feature observed in the N-terminal domains of all available

155   T4P structures. Interestingly, the unfolded region is significantly longer for the pilin comprising

156   the narrow filament, resulting in a longer N-terminal stalk. The outer regions of both filaments

157   are formed by globular domains consisting of β-strands, a typical hallmark of the T4P C-

158   terminal domain [30]. Whilst the C-terminal head domains of both pili are comprised of central

159   β-sheets, the domain size and the region linking to the N-terminal α-helix appear different.

160   These findings suggest that the two pili are not only distinct with regards to their helical

161   parameters, but also consist of different proteins.

162

163   As shown, we observe both filaments emerging from the T4P machinery (Fig. 1). This is

164   supported by previous studies showing that mutants defective in the PilQ secretin channel do

165   not extrude pili [27]. In addition, previous studies have shown that a mutant defective in the

166   assembly ATPase PilF is non-piliated [9]. This suggests that both pili are extruded by the same

167   assembly machinery. The aperture within the central channel of PilQ is of sufficient dimension

168   to accommodate either form (Fig. 2e, f).

169

170   ***Atomic models of T4P***

171   The resolution and quality of both maps allowed us to unambiguously build an atomic model

172   for each filament *ab initio* (Fig. 3a - f). Guided by our mass spectrometry results, the position

173   of large side chains and clear differences in the length of the polypeptide backbones, we were

174   able to identify PilA4 as the building block for the wide pilus and the previously

175   uncharacterised protein TT_C1836 as the subunit for the narrow filament. We now propose

176   that TT_C1836 be named PilA5, in keeping with *Thermus* nomenclature.

177

178   The N-terminal α-helix, including the unfolded stretch, is comprised of the first 54 (PilA4) or

179   53 amino acids (PilA5). In both proteins the helix is disrupted by an unfolded stretch around

180   the conserved Pro22 (Fig. 3c, d, S6). The stretch in PilA4 is 4 amino acids long as opposed to

181   10 amino acids long in PilA5. The region between the N-terminal α-helix and the C-terminal

6

182 β-sheet, the so-called glycosylation loop, ranges in PilA4 from amino acids 55 to 77, with a

183 two-turn α-helix comprising amino acids 61-67. The C-terminal region is an antiparallel four-

184 stranded β-sheet with the last strand facing towards the N-terminus followed by a loop that

185 ends on the β-sheet. A disulphide bond between Cys89, which is located in the second strand,

186 and the penultimate amino acid Cys124, likely stabilises the C-terminus (Fig. 3e). The

187 glycosylation loop in PilA5 spans amino acids 54 to 71, with amino acids 62 to 65 forming a

188 one-turn helix. The C-terminal β-sheet is composed of five strands, one more than observed

189 in PilA4. Due to the additional β-strand in PilA5, the last strand faces away from the N-

190 terminus of the protein. The C-terminus of PilA5 is located between the β-sheet, the

191 glycosylation loop and the long α-helix. A disulphide bond is formed between Cys60 in the

192 glycosylation loop and Cys88 in the third β-strand (Fig. 3f). Both pilins are highly hydrophobic

193 at the N-terminal part of the α-helix, and more hydrophilic on the surface of the globular

194 domain (Fig. 3g & h). The hydrophobic helices bundle and form the hydrophobic core of the

195 assembled filament. PilA4 has no net charge but the filament displays a distinctive positively

196 charged groove along the filament axis (Fig. 3i). In contrast, PilA5 has a total of 2 negative

197 charges per subunit which leads to a patch of negative charge winding around the filament

198 (Fig. 3j).

199

200 A network of cooperative interactions between pilin subunits holds the fibres together. Each

201 subunit has 6 (wide form) or 7 (narrow form) physical interaction partners in each direction

202 of the fibre (thus 12 or 14 total interaction partners) spread in side-by-side or top-to-bottom

203 directions (Fig. S7a, b). Most of the interactions involve a large portion of the N-terminal α-

204 helices within the hydrophobic core as well as the head domains. In the wide (PilA4) filament,

205 each subunit (subunit A) interacts with the N-termini that project down from the next two

206 subunits (B and C) above and from the subunits which are six and seven subunits above (G

207 and H). A second interaction takes place between the upper part of the α-helix in subunit A

208 and the glycosylation loop in subunit B (Fig. S7a). In the narrow (PilA5) filaments, subunit A

209 interacts via its α-helix with the N-termini of subunits B, C, F, G and H, while there is an

210 additional interaction between the upper part of the α-helix in subunit A with the

211 glycosylation loop in subunit B (Fig. S7b).

212

213 For both types of pili, the largest interaction interface is between subunits which are 3 or 4

214 subunits apart (subunit A with subunit D and E). Thus each pilin subunit has a large interaction

215 interface with 6 other subunits (B, D and E in Fig. S7a, b) and 6 or 8 smaller interaction sites

216 (C, G, H in Fig. S7a and C, G, H, F in Fig. S7b). Most interactions involve the hydrophobic

217 sidechains in the centre of the filament and appear to be nonspecific, likely allowing sliding

218   movements between the subunits when the filaments are stretched. This is in accordance

219   with the observation that pili can stretch up to threefold upon force [31]. In addition, PilA4

220   contains salt bridges between Asp53 in subunit A and Arg30 in subunit D, and between Glu48

221   in subunit A and Arg28 in subunit E. In contrast, PilA5 contains a single salt bridge between

222   Glu68 in subunit A and Arg23 in subunit D (Fig. S7c). The conserved Glu5 is likely required to

223   neutralise the positive charge of the N-terminus within the hydrophobic core of the filament

224   [4,32,33]. A salt bridge is also found between Glu5 and the N-terminus of the neighbouring subunit

225   in other T4P [23,24]. For both *T. thermophilus* fibres the distance between Glu5 and adjacent N-

226   termini is too far to form a salt bridge. Instead, Glu5 forms an intramolecular salt bridge to

227   the N-terminus in the same subunit (Fig. S7c). This was also modelled for the related *Klebsiella*

228   *oxytoca* pseudopilus [34].

229

230   ***Posttranslational modification***

231   Densities were observed in both EM maps that protrude into the solvent and cannot be

232   attributed to the polypeptide backbone and were too large to account for an amino acid side

233   chain (Fig. 4). Interestingly, these densities co-localised with serine residues and were similar

234   in appearance to previously published densities attributed to glycosylation sites [29,35,36]. We

235   suggest that these densities correspond to O-linked glycosylation, which is consistent with the

236   previous finding that the major pilin PilA4 of *T. thermophilus* is glycosylated [37]. Glycosylation

237   has also been observed in similar locations in the X-ray structure of *N. gonorrhoeae* type IV

238   pilin head domain (PDB: 2HI2) [16,17,38]. In PilA4, we found extra densities at 3 serine residues in

239   the glycosylation loop (Ser59, Ser66 and Ser71), while only one serine appeared to be

240   modified in PilA5 (Ser73). Interestingly the density was much more pronounced in PilA5 than

241   in PilA4. This may be due to a less flexible glycan moiety in PilA5, allowing for improved

242   resolution, or by a different composition of the sugar residues entirely. Glycosylation may

243   enhance temperature stability via additional hydrogen bonds, increase adhesive properties

244   (either to surfaces or to small molecules such as DNA) or act as recognition tags for cell-cell

245   communication [39].

246

247   ***The functional importance of two types of pili***

248   A key outstanding question pertains to the functional relevance of the two types of T4P.

249   To investigate this question, we performed various functional analyses on PilA4 and PilA5

250   deletion strains. We assessed cell lines without pili (*pilA4::km*), with wide pili only (*pilA5::km*),

251   or a mixed population of wide and narrow pili (wild type).

252

253   We analysed cellular motility by twitching assays at 68 °C and 55 °C. Wild-type cells formed

254   characteristic twitching zones of ~2 cm and ~1.2 cm in diameter, respectively. The mutants

255   *pilA4::km* and *pilA5::km* did not exhibit any twitching motility (Fig. 5a). Since the immotile

256   *pilA5*::km cells could still produce wide pili comprised of PilA4, we deduce that PilA5 is

257   required to promote cell movement. Cells lacking both types of pili in the *pilA4::km* mutant

258   were completely defective in natural transformation, in agreement with our former finding [40].

259   Transformation efficiency was only partially reduced in the *pilA5::km* mutant (~30%), which

260   expresses wide PilA4 pili (Fig. 5b). This corresponds to our previous finding that a *pilA5::km*

261   mutant is still transformable [40] and demonstrate that the narrow pili are dispensable for DNA

262   uptake.

263

264   In summary, we conclude that PilA4 has three known roles: it promotes pilus formation for

265   both wide and narrow filaments, it comprises the main structural element of the wide T4P,

266   and it plays a role in natural transformation. We determine two functions for PilA5: it forms

267   the basis of the narrow T4P and is a requirement for cell motility.

268

269

270   **Discussion**

271   We have determined the first cryoEM structures of T4P that have allowed atomic models to

272   be built *ab initio*. Moreover, we have discovered two distinct T4P filaments, which are

273   composed of different proteins. Our data provide compelling evidence that PilA5 is essential

274   for twitching motility and confirm the previous finding that PilA4 is involved in natural

275   transformation [40]. In addition, we find that PilA4 is essential for the assembly of both wide and

276   narrow pili. PilA4 may therefore play a crucial regulatory role, could initiate pilus formation,

277   or even form a capping structure. In many bacteria, minor pilins are thought to prime pilus

278   assembly by reducing the energy barrier to the extraction of pilins from the membrane [41]. In

279   *Thermus*, PilA4 may perform this role.

280

281   The unique functionality of PilA4 and PilA5 is hardcoded in their distinct structural features.

282   Both filaments follow the conserved T4P blueprint, encompassing a central bundle of

283   hydrophobic N-terminal α-helices and a hydrophilic C-terminal β-strand globular domain [39].

284   Structural variations in PilA4 and PilA5 determine distinct inter-subunit interactions, helical

285   parameters, mechanical properties, adhesiveness and binding affinity. Narrow pili comprised

286   of PilA5 are more flexible than those comprised of PilA4. In line with their predicted role in

287   twitching, narrow pili that can bend and flex would enable the filaments to curve from the

288   surface of cells to interact with surfaces, negotiate obstacles and increase the exploratory

289 range of the cell. Their overall net negative surface charge would enhance the adhesive
290 properties of the fibre and facilitate surface adhesion.

291

292 In accordance with the role of PilA4 in natural transformation, the surface of wide PilA4
293 filaments show a striking line of positively-charged residues along the long axis. We speculate
294 that these may be involved in binding the negatively charged DNA backbone. A double
295 stranded DNA molecule fits into a right-handed helical groove flanked by glycan residues. This
296 would allow interaction of the DNA backbone phosphates with positive charges in the PilA4
297 filament at regular intervals. Flanking glycans and negative charges may coordinate proper
298 DNA binding (Fig. 6). This suggestion is in good agreement with the well characterised roles of
299 T4P in DNA binding and uptake in *Neisseria*, and of DNA binding in *P. aeruginosa* [42,43]. Our
300 proposal is also consistent with that of Craig *et al,* who suggest that the positively charged
301 groove of gonoccocal T4P is wide enough to bind the negatively charged backbone of dsDNA
302 [17]. Interestingly, previous experiments have shown that pili are not solely important for
303 natural transformation. For example, a mutant defective in the PilF assembly ATPase was
304 impaired in piliation and was hypertransformable, and a mutant carrying a deletion in a
305 domain of the secretin PilQ was impaired in piliation but exhibited wildtype transformation
306 frequencies [10,44]. Taken together, we suggest that wide pili comprised of PilA4 may capture
307 DNA rather like a fishing net, thus improving the efficiency of DNA uptake by increasing the
308 local concentration of DNA near the outer membrane.

309

310 In evolutionary terms, bacteria appear to have reused the T4P blueprint to develop a system
311 that can assemble two different filaments with unique properties. This could enable tasks to
312 be performed more effectively and at reduced energy cost to the cell. It will now be interesting
313 to discover if this principle occurs in other bacterial species, and excitingly, will open avenues
314 to the development of vaccines or therapeutics targeting a particular T4P mechanism.

315

316

317 **Experimental procedures**

318 **Cultivation of organisms**

319 *T. thermophilus* HB27 was grown in TM$^+$ medium (8 g/l tryptone, 4 g/l yeast extract, 3 g/l NaCl,
320 0.6 mM MgCl$_2$, 0.17 mM CaCl$_2$) [45] at 55 °C, 58 °C or 68 °C. Antibiotics were added when
321 appropriate (kanamycin, 80 mg/ml; streptomycin, 100 mg/ml in solid medium (containing 2
322 % agar [w/vol]) or kanamycin, 60 mg/ml; streptomycin 100 mg/ml in liquid medium).
323 Disruption of *TT_C1836* and *pilA4* was performed by insertion of a kanamycin marker [27].

324

325   **Purification of pili**

326   *T. thermophilus* HB27 cells were grown for two days on TM$^+$ medium at 68 °C or three days at

327   55 °C.. Cells were scraped off and resuspended by pipetting and shaking in ethanolamine

328   buffer (0.15 M ethanolamine, pH 10.5). Cells were sedimented by centrifugation (16.200 *x g*,

329   1 h, 4 °C). The supernatant was gently mixed with saturated ammonium sulfate solution [10/1

330   (v/v)] and incubated on ice for 12 h. Pili were precipitated by centrifugation for 10 min (16,200

331   *x g*, 4 °C). The resulting pellet was washed twice with TBS buffer (50 mM Tris/HCl, 150 mM

332   NaCl, pH 7.5). The pellet was resuspended by incubation in distilled water for 4 h. 10x buffer

333   (500 mM Tris/HCl, 500 mM NaCl, 10 mM CaCl$_2$, 10 mM MgCl$_2$, pH 7.5) was added prior to

334   structural analyses.

335

336   **CryoET sample preparation and imaging**

337   Cubes of agar with growing *T. thermophilus* HB27 cells were cut out, placed into EDTA buffer

338   (20 mM Tris/HCl, 100 mM EDTA, pH 7.4) and gently agitated for 1 hour at room temperature.

339   Samples were mixed 1:1 with 10 nm protein A-gold (Aurion, Wageningen, The Netherlands)

340   as fiducial markers and glow-discharged R2/2 Cu 300 mesh holey carbon-coated support grids

341   (Quantifoil, Jena, Germany) were dipped into the solution. For analysis of isolated pili by

342   cryoET, preparations were mixed 1:1 with 10 nm protein A-gold fiducial markers and solutions

343   were gently pipetted onto the grids. All grids were blotted using Whatman 41 filter paper for

344   ~4 s in a humidified atmosphere and plunge-frozen in liquid ethane in a home-built device or

345   using a Vitrobot Mark IV (Thermo Fisher, Waltham, USA). Grids were maintained under liquid

346   nitrogen and transferred into the electron microscope at liquid nitrogen temperature.

347

348   Tomograms were typically collected from +60° to -60° at tilt steps of 2° and 5 - 7 μm

349   underfocus (whole cells), or at 3 μm underfocus (isolated pili), using either a Tecnai Polara,

350   Titan Krios (Thermo Fisher) or JEM-3200FSC (JEOL, Tokyo, Japan) microscope, all equipped

351   with field emission guns operating at 300 keV. All instruments were fitted with energy filters

352   and K2 Summit direct electron detector cameras (Gatan, Pleasanton, USA). Dose-fractionated

353   data (3-5 frames per projection image) were collected using Digital Micrograph (Gatan).

354   Magnifications varied depending on microscope; pixel sizes were within the range 3.8 – 4.2 Å.

355   The total dose per tomogram was < ~140e$^-$/Å^2. Tomograms were aligned using gold fiducial

356   markers and volumes reconstructed by weighted back-projection using the IMOD software

357   (Boulder Laboratory, Boulder, USA) [46]. Contrast was enhanced by non-linear anisotropic

358   diffusion (NAD) filtering in IMOD [47].

359

360   **Calculations of pilus diameter**

11

361   Slices through tomograms were analysed by drawing a plot profile of grey values in ImageJ [48],

362   which could be exported as a function of distance. Statistical analysis of pilus diameters was

363   conducted with a sample size of 8 tomograms containing ~60 isolated pili.

364

365   **Negative-stain electron microscopy**

366   Two microliters of purified pili were pipetted onto glow-discharged carbon-coated Cu 400

367   mesh support grids (Sigma-Aldrich) for 2 minutes. Grids were blotted with Whatman No 41

368   filter paper and stained with 5 % ammonium molybdate for 60 s. Images were recorded with

369   a Tecnai Spirit microscope (Thermo Fisher) operated at 120 keV and a OneView CMOS camera

370   (Gatan). Images were analysed for fibre quality, size, sample density and homogeneity using

371   EMAN2 [49]. For whole cell samples of *T. thermophilus*, either liquid culture was used directly or

372   some cells were carefully scraped off the plates and resuspended in TBS. If required, cells were

373   diluted in TBS. Negative staining was performed as described above. For the quantitative

374   analysis of number and type of pili per cell, filaments from each cell pole were counted and

375   helices of equal length were selected using e2helixboxer (EMAN2) and subsequently classified

376   in 2D using RELION [50]. The percentage of wide, narrow and unassigned pili was calculated

377   based on the number of particles in each class.

378

379   **CryoEM sample preparation and imaging of fibres**

380   Three microliters of isolated pilus suspension were pipetted onto a glow-discharged R2/2 Cu

381   300 mesh holey carbon-coated grids (Quantifoil). Grids were plunge frozen in liquid ethane

382   after blotting using a Vitrobot Mark IV (Thermo Fisher) and stored in liquid nitrogen. Cryo

383   images were collected with a Titan Krios microscope (Thermo Fisher) at the UK national

384   electron bio-imaging centre (eBIC), equipped with a field emission gun operating at 300 keV .

385   The microscope was fitted with K2 Summit direct electron detector and Quantum energy filter

386   (both Gatan, Pleasanton, USA). Dose-fractionated data were collected at 1.5 – 4 µm defocus

387   using EPU (Thermo Fisher). 3138 micrographs containing both forms of pili were collected as

388   40-frame movies, corresponding to 8 seconds at a frame rate of 1 frame for every 0.2 seconds.

389   The total dose was 48 electrons/$Å^2$ at a magnification of 130,000 x, corresponding to a pixel

390   size of 1.048 Å.

391

392   **Image processing, symmetry determination and helical reconstruction**

393   Drift correction was performed using UNBLUR [51]. Straight sections of thin and wide fibres were

394   boxed separately from the drift corrected images using the helixboxer function of EMAN2,

395   such that the filaments were centred in each rectangular box. Helical reconstruction was

396   performed using the boxed filaments and SPRING as follows [52]. Contrast transfer function

12

397   (CTF) correction was performed using CTFFIND [53]. In order to determine the helical

398   parameters of the wide filaments a subset of the boxed filaments were cut into small

399   segments of 373 Å (a multiple of the helical rise) with an 80 % overlap, yielding a total of 9576

400   segments, which were classed in 2D (Fig. S4a). Close examination of the segments indicated a

401   filament diameter of 75 Å. The calculated power spectrum from the total segments indicated

402   clear layer lines that could be indexed. A meridional reflection at approximately 9 Å and a

403   layer line of order 1 at approximately 36 Å indicated that there are approximately 4 subunits

404   per turn. The ninth layer was found to be of order 1, suggesting that the helix repeats exactly

405   after nine turns, with a non-integer number of subunits per turn. The

406   SEGMENTCLASSRECONSTRUCT module in SPRING on class averages was used to determine

407   the accurate helical symmetry (Fig. S5a). The suggested output was determined to be either

408   4.10 or 3.89 subunits in a helical pitch of 36.3 Å. In order to determine the helical parameters

409   for the narrow filaments a subset of the boxed filaments were cut into segments of 800 Å with

410   a step size of 330 Å and classed in 2D (Fig. S4b). The helical pitch could be determined directly

411   as 48.1 Å. A meridional reflection and thus a helical rise at 11.3 Å could be identified. These

412   parameters allow calculation of a helical rotation of 84.6° and 4.26 subunits per turn. The

413   SEGMENTCLASSRECONSTRUCT module in SPRING was again used to determine the accurate

414   helical symmetry (Fig. S5b). The suggested output was determined to be either 4.11, 4.14,

415   4.27 or 4.30 subunits with a helical pitch of 48.1 Å. 3D reconstruction was performed using

416   the above parameters by iterative projection matching and back projection as implemented

417   in the SEGMENTREFINE3D of SPRING, starting from a solid cylinder of 75 Å as a reference.

418   Examination of the Fourier transforms simulated from the reconstructed volume to that

419   experimentally calculated from fibres indicated that 3.89 subunits in a pitch of 36.3 Å

420   (accounting for a helical rise of 9.33 Å and a helical rotation of 92.5 degrees) is correct for the

421   wide filaments, and 4.27 subunits in a pitch of 48.1 Å (accounting for a helical rise of 11.26 Å

422   and a helical rotation of 84.3 degrees) is correct for the narrow filaments (Fig. S5c, d). For the

423   final maps 300 Å segments with a step size of three times the helical rise from 400 images

424   were extracted. Doubling the number of used images did not further increase the final

425   resolution. The calculated final maps were determined at 3.22 Å resolution from 98,415

426   asymmetric units for wide filaments, and 3.49 Å from 76,866 asymmetric units for narrow

427   filaments (Fig. S5e, f) using Fourier shell correlation (0.143 cut-off). For the final maps a B-

428   factor of -60 Å^2 was applied. Figures were drawn in Chimera, Coot and CCP4mg [54–56].

429

430   **Model building**

431   Atomic models for both forms of pili were built manually *de novo* in Coot. We assumed that

432   one of the two pili consists of the major pilin PilA4, which is 125 amino acids in length. The

433   backbone as well as all large side chain densities of PilA4 match the density map of the wider

434   form of the pilus. While tracing the backbone into the density maps it became apparent that

435   the subunits forming the narrower pili are ~10% smaller than the subunits of the wider pili.

436   Following the results of mass spectrometry and deletion experiments we modelled the second

437   most abundant protein, TT_C1836 (111 amino acids) into the density map of thin pili. The

438   backbone and visible side chains fit perfectly into the density map. The structure was

439   iteratively refined by Refmac5 [57] followed by manual rebuilding in Coot and ISOLDE [58]. The

440   final models contain all amino acid residues of the mature protein. The double stranded DNA

441   in Fig. 6 (based on PDB: 1bna) was modelled around the wide pilus using Chimera and Coot.

442

Data deposition

444   The cryo-EM maps were deposited in the Electron Microscopy Data Bank with accession codes

445   EMD-XXXX (wide pilus) and EMD-YYYY (narrow pilus). The structure coordinates of the atomic

446   models of the wide and the narrow pilus were deposited in the Protein Data Bank with

447   accession numbers WWWW and ZZZZ, respectively.

448

Sequence Alignment

450   Sequence alignment was performed using the PRALINE server [59] with the default settings

451   (weight matrix: BLOSSUM62, gap opening penalty: 12, gap extension penalty: 1).

452

Mass spectrometry

454   Purified pilus preparations were processed using a modified FASP workflow [60] as described

455   previously [61]. In brief, reduced and alkylated protein extracts were digested sequentially with

456   Lys-C and trypsin on Microcon-10 filters (Merck Millipore, # MRCPRT010 Ultracel YM-10).

457   Digested samples were desalted using ZipTips according to the manufacturer's instructions,

458   dried in a Speed-Vac and stored at -20 °C until LC/MS-MS analysis. Dried peptides were

459   dissolved in 5 % acetonitrile with 0.1 % formic acid, and subsequently loaded using a nano-

460   HPLC (Dionex U3000 RSLCnano) on reverse-phase columns (trapping column: particle size

461   3μm, C18, L=20mm; analytical column: particle size <2μm, C18, L=50cm; PepMap,

462   Dionex/Thermo Fisher). Peptides were eluted in gradients of water (buffer A: water with 5 %

463   v/v acetonitrile and 0.1 % formic acid) and acetonitrile (buffer B: 20 % v/v water and 80 % v/v

464   acetonitrile and 0.1 % formic acid). All LC-MS-grade solvents were purchased from Fluka.

465   Gradients were ramped from 4 % to 48 % B in 120 minutes at flow rates of 300 nl/min.

466   Peptides eluting from the column were ionised online using a Thermo nanoFlex ESI-source

467   and analysed in a Thermo "Q Exactive Plus" mass spectrometer. Mass spectra were acquired

468   over the mass range 350-1400m/z (Q Exactive Plus) and sequence information was acquired

14

469   by computer-controlled, data-dependent automated switching to MS/MS mode using

470   collision energies based on mass and charge state of the candidate ions.

471

472   Raw MS data were processed and analysed with MaxQuant [62] . In brief, spectra were matched

473   to the full 15 nalyse.org database (reviewed and non-reviewed, downloaded on the

474   13/05/2016) and a contaminant and decoy database. Precursor mass tolerance was set to 4.5

475   ppm, fragment ion tolerance to 20 ppm, with fixed modification of Cys residues

476   (carboxyamidomethylation +57.021) and variable modifications of Met residues (Ox +15.995),

477   Lys residues (Acetyl +42.011), Asn and Gln residues (Deamidation +0.984) and of N-termini

478   (carbamylation +43.006). Peptide identifications were calculated with FDR = 0.01, and

479   proteins with one peptide per protein included for subsequent analyses. Proteomics data

480   associated with this manuscript have been uploaded to PRIDE [63]. Anonymous reviewer access

481   is available upon request. Peptide intensities (label free quantitation) were analysed using

482   MaxQuant and Perseus [62]. Differential abundance of proteins (detected in at least 3 of 4

483   replicates in each condition) was analysed using a two-sided t test with a FDR of 0.01 and s0

484   = 0.05.

485

486   For gel-based MS, purified pili were separated by SDS-PAGE (Mini Protean TGX 4-15 %, Biorad,

487   Hercules, USA) and proteins were stained using Bio-Safe Coomassie Stain (Biorad, Hercules,

488   USA). Bands were cut out and analysed by MS at the University of Bristol Proteomics Facility.

489

490   **Twitching motility**

491   *T. thermophilus* HB27 strains were grown at 68 °C for 3 days and at 58 °C for 7 days under

492   humid conditions on minimal medium agar plates [10] containing 0.1 % bovine serum albumin.

493   Plates were then stained with Coomassie blue and cells washed off to reveal twitching zones.

494   The contrast has been inverted in the images.

495

496   **Natural transformation**

497   *T. thermophilus* wild-type, *pilA4::km* and *TT_C1836::km* mutants were cultured in TM$^+$ media

498   containing appropriate antibiotics for 24h at 68 °C, 150 rpm (New Brunswick Innova 42,

499   Eppendorf, Hamburg, Germany). These cultures were used to inoculate 10 ml TM$^+$ media (with

500   appropriate antibiotics) to a starting OD600 = 0.2 and incubated until OD600 = 0.5 was

501   reached. 30 µl of the cultures were transferred into 370 µl prewarmed TM$^+$ medium, and 10

502   µg genomic DNA from a spontaneous streptomycin-resistant HB27 mutant was added (HB27

503   Strep). The cultures were incubated for 30 min at 68 °C, 150 rpm. These were subsequently

15

504   diluted (400 µl in 3 ml TM⁺) and incubated for an additional 3 h at 68 °C, 150rpm.

505   Transformations were plated in suitable dilutions onto TM⁺ agar plates containing 100 µg/ml

506   streptomycin to determine the number of transformants. The number of viable cells was

507   determined by plating on TM⁺ agar plates lacking streptomycin. Following incubation for 2

508   days at 68°C, colonies were counted. The transformation efficiency was calculated as the

509   number of transformations per number of living cells.

510

523

Author contributions

525   Major contributions to (i) the conception or design of the study (AN, MS, RS, BD, BA, VAMG)

526   (ii) the acquisition, analysis, or interpretation of the data (AN, MS, RS, KK, KS, JDL, BD, BA,

527   VAMG); and (iii) writing of the manuscript (AN, MS, BD, VAMG). All authors commented on

528   the manuscript.

529

Conflict of interest

531   The authors declare no conflict of interest.

References

1. Hobbs, M. & Mattick, J. S. Common components in the assembly of type 4 fimbriae, DNA transfer systems, filamentous phage and protein-secretion apparatus: a general system for the formation of surface-associated protein complexes. *Mol. Microbiol.* (1993). doi:10.1111/j.1365-2958.1993.tb01949.x

2. Kaiser, D. Bacterial motility: How do pili pull? *Current Biology* (2000). doi:10.1016/S0960-9822(00)00764-8

3. Averhoff, B. & B., A. Shuffling genes around in hot environments: The unique DNA transporter of Thermus thermophilus. in *FEMS microbiology reviews* **33**, 611–626 (2009).

4. Strom, M. S. & Lory, S. Amino acid substitutions in pilin of Pseudomonas aeruginosa. Effect on leader peptide cleavage, amino-terminal methylation, and pilus assembly. *J. Biol. Chem.* (1991).

5. Paetzel, M., Dalbey, R. E. & Strynadka, N. C. J. Crystal structure of a bacterial signal peptidase apoenzyme. Implications for signal peptide binding and the Ser-Lys dyad mechanism. *J. Biol. Chem.* (2002). doi:10.1074/jbc.M110983200

6. Karuppiah, V., Collins, R. F., Thistlethwaite, A., Gao, Y. & Derrick, J. P. Structure and assembly of an inner membrane platform for initiation of type IV pilus biogenesis. *Proc. Natl. Acad. Sci. U. S. A.* **110**, E4638-47 (2013).

7. Kruse, K., Salzer, R. & Averhoff, B. The traffic ATPase PilF interacts with the inner membrane platform of the DNA translocator and type IV pili from Thermus thermophilus. *FEBS Open Bio* (2019). doi:10.1002/2211-5463.12548

8. Rose, I. *et al.* Identification and characterization of a unique, zinc-containing transport ATPase essential for natural transformation in Thermus thermophilus HB27. *Extremophiles* (2011). doi:10.1007/s00792-010-0343-2

9. Salzer, R., Joos, F. & Averhoff, B. Type IV pilus biogenesis, twitching motility, and DNA uptake in Thermus thermophilus: Discrete roles of antagonistic ATPases PilF, PilT1, and PilT2. *Appl. Environ. Microbiol.* **80**, 644–652 (2014).

10. Burkhardt, J., Vonck, J., Langer, J. D., Salzer, R. & Averhoff, B. Unusual N-terminal ααβαββα fold of PilQ from Thermus thermophilus mediates ring formation and is essential for piliation. *J. Biol. Chem.* **287**, 8484–8494 (2012).

11. Gold, V. A. M., Salzer, R., Averhoff, B. & Kühlbrandt, W. Structure of a type IV pilus machinery in the open and closed state. *Elife* **4**, (2015).

12. Salzer, R. *et al.* Topology and structure/function correlation of ring-and gate-forming domains in the dynamic secretin complex of thermus thermophilus. *J. Biol. Chem.* **291**, (2016).

17

568   13.   D'Imprima, E. *et al.* Cryo-EM structure of the bifunctional secretin complex of

569   Thermus thermophilus. *Elife* (2017). doi:10.7554/elife.30483

570   14.   Salzer, R., Kern, T., Joos, F. & Averhoff, B. Environmental factors affecting the

571   expression of type IV pilus genes as well as piliation of Thermus thermophilus. *FEMS*

572   *Microbiol. Lett.* (2014). doi:10.1111/1574-6968.12506

573   15.   Chang, Y.-W. W. *et al.* Architecture of the type IVa pilus machine. *Science (80-.).* **351**,

574   aad2001 (2016).

575   16.   Parge, H. E. *et al.* Structure of the fibre-forming protein pilin at 2.6 Å resolution.

576   *Nature* (1995). doi:10.1038/378032a0

577   17.   Craig, L. *et al.* Type IV Pilus Structure by Cryo-Electron Microscopy and

578   Crystallography: Implications for Pilus Assembly and Functions. *Mol. Cell* **23**, 651–662

579   (2006).

580   18.   Fukakusa, S. *et al.* Structure of the CFA/III major pilin subunit CofA from human

581   enterotoxigenic Escherichia coli determined at 0.90 Å resolution by sulfur-SAD

582   phasing. *Acta Crystallogr. Sect. D Biol. Crystallogr.* (2012).

583   doi:10.1107/S0907444912034464

584   19.   Kolappan, S., Roos, J., Yuen, A. S. W., Pierce, O. M. & Craig, L. Structural

585   characterization of CFA/III and longus type IVb Pili from enterotoxigenic Escherichia

586   coli. *J. Bacteriol.* (2012). doi:10.1128/JB.00282-12

587   20.   Gorgel, M. *et al.* High-resolution structure of a type IV pilin from the metal-reducing

588   bacterium Shewanella oneidensis. *BMC Struct. Biol.* (2015). doi:10.1186/s12900-015-

589   0031-7

590   21.   Piepenbrink, K. H. *et al.* Structural diversity in the type IV Pili of multidrug-resistant

591   acinetobacter. *J. Biol. Chem.* (2016). doi:10.1074/jbc.M116.751099

592   22.   Karuppiah, V., Thistlethwaite, A. & Derrick, J. P. Structures of type IV pilins from

593   Thermus thermophilus demonstrate similarities with type II secretion system

594   pseudopilins. *J. Struct. Biol.* (2016). doi:10.1016/j.jsb.2016.08.006

595   23.   Kolappan, S. *et al.* Structure of the neisseria meningitidis type IV pilus. *Nat. Commun.*

596   1–12 (2016). doi:10.1038/ncomms13015

597   24.   Wang, F. *et al.* Cryoelectron Microscopy Reconstructions of the Pseudomonas

598   aeruginosa and Neisseria gonorrhoeae Type IV Pili at Sub-nanometer Resolution.

599   *Structure* **25**, 1423-1435.e4 (2017).

600   25.   Bardiaux, B. *et al.* Structure and Assembly of the Enterohemorrhagic Escherichia coli

601   Type 4 Pilus. *Structure* (2019). doi:10.1016/j.str.2019.03.021

602   26.   Schwarzenlander, C., Haase, W. & Averhoff, B. The role of single subunits of the DNA

603   transport machinery of Thermus thermophilus HB27 in DNA binding and transport.

18

604   *Environ. Microbiol.* **11**, 801–808 (2009).

605   27.   Friedrich, A., Prust, C., Hartsch, T., Henne, A. & Averhoff, B. Molecular analyses of the

606   natural transformation machinery and identification of pilus structures in the

607   extremely thermophilic bacterium Thermus thermophilus strain HB27. *Appl. Environ.*

608   *Microbiol.* **68**, 745–755 (2002).

609   28.   Fromm, S. A. & Sachse, C. Cryo-EM Structure Determination Using Segmented Helical

610   Image Reconstruction. in *Methods in Enzymology* (2016).

611   doi:10.1016/bs.mie.2016.05.034

612   29.   Daum, B. *et al.* Structure and in situ organisation of the pyrococcus furiosus

613   archaellum machinery. *Elife* **6**, (2017).

614   30.   Craig, L. & Li, J. Type IV pili: paradoxes in form and function. *Current Opinion in*

615   *Structural Biology* **18**, 267–277 (2008).

616   31.   Biais, N., Higashi, D. L., Brujic, J., So, M. & Sheetz, M. P. Force-dependent

617   polymorphism in type IV pili reveals hidden epitopes. *Proc. Natl. Acad. Sci. U. S. A.*

618   **107**, 11358–63 (2010).

619   32.   Aas, F. E. *et al.* Substitutions in the N-terminal alpha helical spine of Neisseria

620   gonorrhoeae pilin affect Type IV pilus assembly, dynamics and associated functions.

621   *Mol. Microbiol.* (2007). doi:10.1111/j.1365-2958.2006.05482.x

622   33.   Li, J., Egelman, E. H. & Craig, L. Structure of the Vibrio cholerae Type IVb pilus and

623   stability comparison with the Neisseria gonorrhoeae Type IVa pilus. *J. Mol. Biol.* **418**,

624   47–64 (2012).

625   34.   Campos, M., Nilges, M., Cisneros, D. A. & Francetic, O. Detailed structural and

626   assembly model of the type II secretion pilus from sparse data. *Proc. Natl. Acad. Sci.*

627   (2010). doi:10.1073/pnas.1001703107

628   35.   Poweleit, N. *et al.* CryoEM structure of the Methanospirillum hungatei archaellum

629   reveals structural features distinct from the bacterial flagellum and type IV pili. *Nat.*

630   *Microbiol.* **2**, 16222 (2016).

631   36.   Meshcheryakov, V. A. *et al.* High-resolution archaellum structure reveals a conserved

632   metal-binding site. *EMBO Rep.* (2019). doi:10.15252/embr.201846340

633   37.   Rumszauer, J., Schwarzenlander, C. & Averhoff, B. Identification, subcellular

634   localization and functional interactions of PilMNOWQ and PilA4 involved in

635   transformation competency and pilus biogenesis in the thermophilic bacterium

636   Thermus thermophilus HB27. *FEBS J.* **273**, 3261–3272 (2006).

637   38.   Marceau, M., Forest, K., Béretti, J. L., Tainer, J. & Nassif, X. Consequences of the loss

638   of O-linked glycosylation of meningococcal type IV pilin on piliation and pilus-

639   mediated adhesion. *Mol. Microbiol.* (1998). doi:10.1046/j.1365-2958.1998.00706.x

640   39.   Daum, B. & Gold, V. Twitch or swim: towards the understanding of prokaryotic

641   motion based on the type IV pilus blueprint. *Biol. Chem.* (2018). doi:10.1515/hsz-

642   2018-0157

643   40.   Friedrich, A., Rumszauer, J., Henne, A. & Averhoff, B. Pilin-like proteins in the

644   extremely thermophilic bacterium Thermus thermophilus HB27: Implication in

645   competence for natural transformation and links to type IV pilus biogenesis. *Appl.*

646   *Environ. Microbiol.* **69**, 3695–3700 (2003).

647   41.   Cisneros, D. A., Bond, P. J., Pugsley, A. P., Campos, M. & Francetic, O. Minor

648   pseudopilin self-assembly primes type II secretion pseudopilus elongation. *EMBO J.*

649   (2012). doi:10.1038/emboj.2011.454

650   42.   Van Schaik, E. J. *et al.* DNA binding: A novel function of Pseudomonas aeruginosa type

651   IV pili. *J. Bacteriol.* (2005). doi:10.1128/JB.187.4.1455-1464.2005

652   43.   Hamilton, H. L. & Dillard, J. P. Natural transformation of Neisseria gonorrhoeae: From

653   DNA donation to homologous recombination. *Molecular Microbiology* (2006).

654   doi:10.1111/j.1365-2958.2005.04964.x

655   44.   Kruse, K., Salzer, R., Joos, F. & Averhoff, B. Functional dissection of the three N-

656   terminal general secretory pathway domains and the Walker motifs of the traffic

657   ATPase PilF from Thermus thermophilus. *Extremophiles* (2018). doi:10.1007/s00792-

658   018-1008-9

659   45.   Oshima, T. & Imahori, K. Description of Thermus thermophilus (Yoshida and Oshima)

660   comb. nov., a Nonsporulating Thermophilic Bacterium from a Japanese Thermal Spa.

661   *Int. J. Syst. Bacteriol.* (1974). doi:10.1099/00207713-24-1-102

662   46.   Kremer, J. R., Mastronarde, D. N. & McIntosh, J. R. Computer visualization of three-

663   dimensional image data using IMOD. *J. Struct. Biol.* (1996).

664   doi:10.1006/jsbi.1996.0013

665   47.   Frangakis, A. S. & Hegerl, R. Noise reduction in electron tomographic reconstructions

666   using nonlinear anisotropic diffusion. *J. Struct. Biol.* (2001).

667   doi:10.1006/jsbi.2001.4406

668   48.   Schneider, C. A., Rasband, W. S. & Eliceiri, K. W. NIH Image to ImageJ: 25 years of

669   image analysis. *Nat. Methods* (2012).

670   49.   Tang, G. *et al.* EMAN2: an extensible image processing suite for electron microscopy.

671   *J. Struct. Biol.* (2007). doi:10.1016/j.jsb.2006.05.009

672   50.   Scheres, S. H. W. RELION: Implementation of a Bayesian approach to cryo-EM

673   structure determination. *J. Struct. Biol.* (2012). doi:10.1016/j.jsb.2012.09.006

674   51.   Grant, T. & Grigorieff, N. Measuring the optimal exposure for single particle cryo-EM

675   using a 2.6 Å reconstruction of rotavirus VP6. *Elife* (2015). doi:10.7554/elife.06980

676   52.   Desfosses, A., Ciuffa, R., Gutsche, I. & Sachse, C. SPRING - An image processing

677   package for single-particle based helical reconstruction from electron

678   cryomicrographs. *J. Struct. Biol.* (2014). doi:10.1016/j.jsb.2013.11.003

679   53.   Mindell, J. A. & Grigorieff, N. Accurate determination of local defocus and specimen

680   tilt in electron microscopy. *J. Struct. Biol.* (2003). doi:10.1016/S1047-8477(03)00069-

681   8

682   54.   Pettersen, E. F. *et al.* UCSF Chimera--a visualization system for exploratory research

683   and analysis. *J. Comput. Chem.* (2004). doi:10.1002/jcc.20084

684   55.   Emsley, P. & Cowtan, K. Coot: Model-building tools for molecular graphics. *Acta*

685   *Crystallogr. Sect. D Biol. Crystallogr.* (2004). doi:10.1107/S0907444904019158

686   56.   McNicholas, S., Potterton, E., Wilson, K. S. & Noble, M. E. M. Presenting your

687   structures: The CCP4mg molecular-graphics software. *Acta Crystallogr. Sect. D Biol.*

688   *Crystallogr.* (2011). doi:10.1107/S0907444911007281

689   57.   Brown, A. *et al.* Tools for macromolecular model building and refinement into

690   electron cryo-microscopy reconstructions. *Acta Crystallogr. Sect. D Biol. Crystallogr.*

691   (2015). doi:10.1107/S1399004714021683

692   58.   Croll, T. I. ISOLDE : a physically realistic environment for model building into low-

693   resolution electron-density maps . *Acta Crystallogr. Sect. D Struct. Biol.* (2018).

694   doi:10.1107/s2059798318002425

695   59.   Heringa, J. Two strategies for sequence comparison: Profile-preprocessed and

696   secondary structure-induced multiple alignment. *Comput. Chem.* (1999).

697   doi:10.1016/S0097-8485(99)00012-1

698   60.   Wiśniewski, J. R., Zougman, A., Nagaraj, N. & Mann, M. Universal sample preparation

699   method for proteome analysis. *Nat. Methods* (2009). doi:10.1038/nmeth.1322

700   61.   Schanzenbächer, C. T., Sambandan, S., Langer, J. D. & Schuman, E. M. Nascent

701   Proteome Remodeling following Homeostatic Scaling at Hippocampal Synapses.

702   *Neuron* (2016). doi:10.1016/j.neuron.2016.09.058

703   62.   Tyanova, S., Temu, T. & Cox, J. The MaxQuant computational platform for mass

704   spectrometry-based shotgun proteomics. *Nat. Protoc.* (2016).

705   doi:10.1038/nprot.2016.136

706   63.   Vizcaíno, J. A. *et al.* 2016 update of the PRIDE database and its related tools. *Nucleic*

707   *Acids Res.* (2016). doi:10.1093/nar/gkv1145

708

709 **Figures**
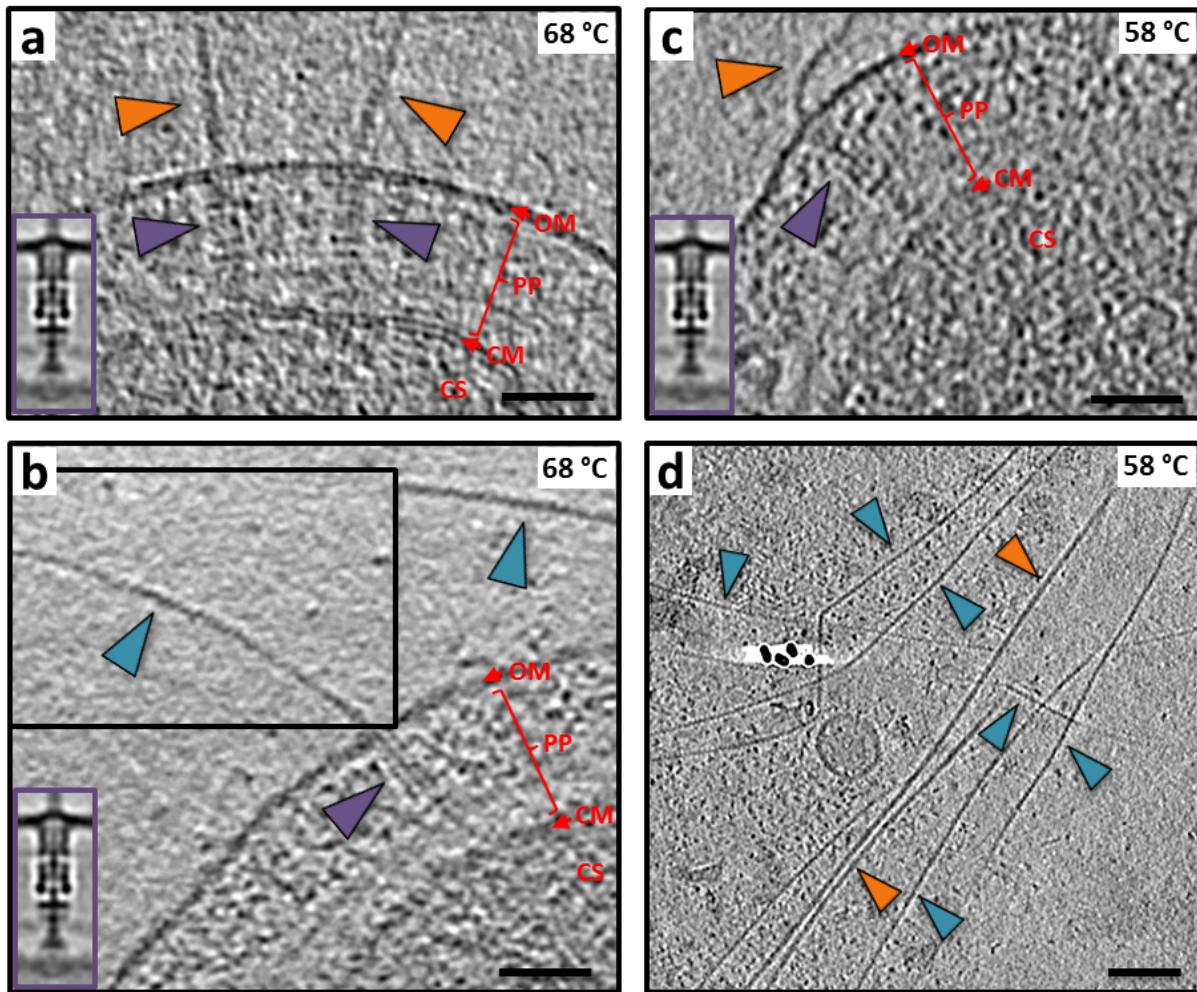
710

711 **Fig. 1: The T4P machinery assembles two types of pilus**

712 **a, b,** Tomographic slices through *T. thermophilus* cells grown at 68 °C show both wide (orange

713 arrowheads) and narrow (teal arrowheads) pili emerging from the T4P machinery (purple

714 arrowheads). The pilus emerges from the cell at an acute angle in (**b**), thus the tomographic

715 volume has been rotated to align with the T4P machinery for visualisation (upper inset box).

716 **c,** Tomographic slice through a *T. thermophilus* cell grown at 58 °C shows a wide pilus (orange

717 arrowhead) emerging from the T4P machinery (purple arrowhead).

718 **d,** Tomographic slice of an area containing many pili from a cell preparation grown at 58 °C.

719 Both wide (orange arrowheads) and narrow pili (teal arrowheads) are visible. Purple insets in

720 **a-c** show the subtomogram average of the T4P machinery (EMD-3023) [11]. OM, outer

721 membrane; PP, periplasm; CM, cytoplasmic membrane; CS, Cytosol. Scale bars, 50 nm.

22
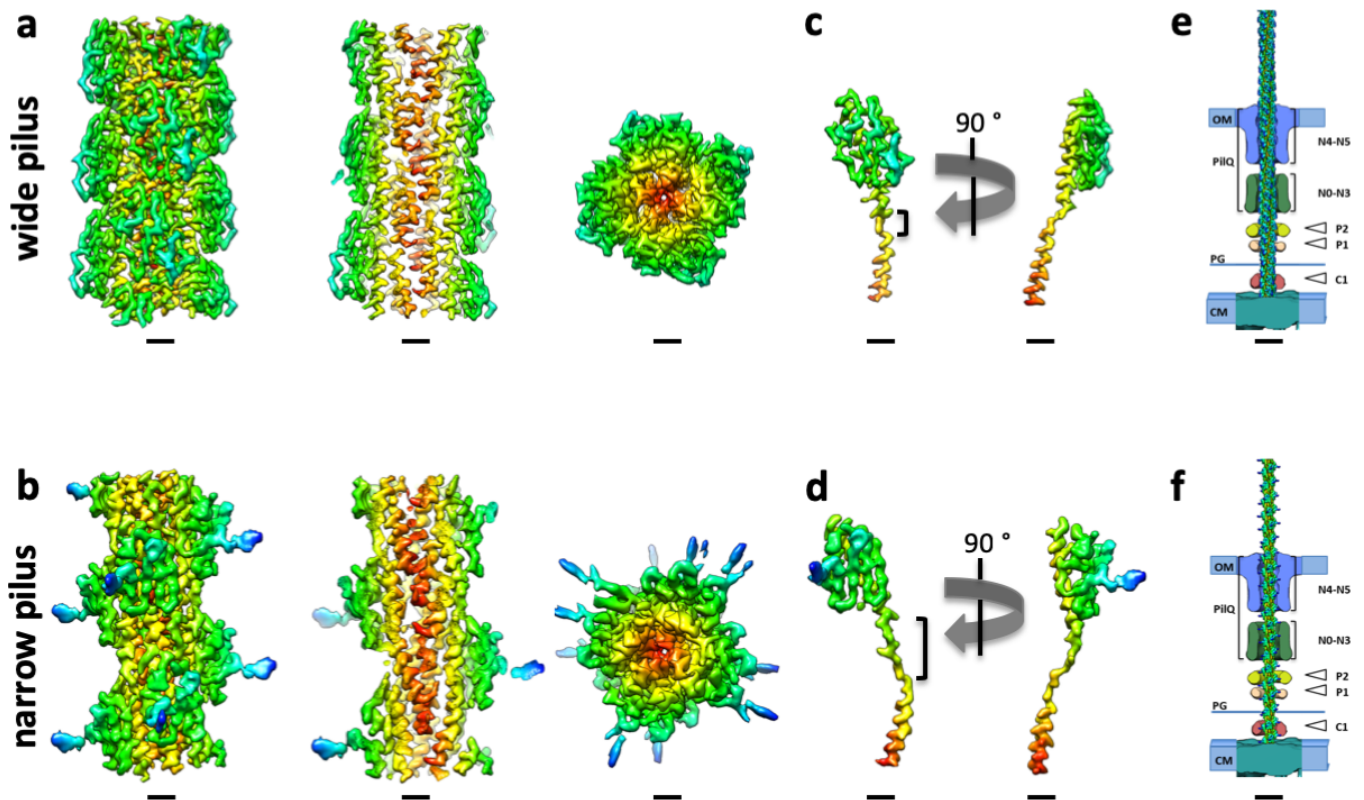
723 **Fig. 2: Helical reconstruction of *Tth*HB27 pili**

724 **a, b,** Side views, cross sections and top views of the density maps of the wide (**a**) and narrow

725 (**b**) pilus fibres. Coloured by cylinder radius from red (centre) to blue (periphery). Scale bars,

726 10 Å.

727 **c, d,** A single subunit of the wide (**c**) and narrow (**d**) pilus segmented from the maps in **a** and

728 **b**. The brackets show the melted stretch in the N-terminal α-helix. Scale bars, 10 Å.

729 **e, f,** Pilus filaments docked into the subtomogram average of the open state of the T4P

730 machinery (EMD-3023) [11]. Domains of the secretin PilQ (blue and green; N0-N5), with

731 unassigned protein densities P2 (yellow), P1 (orange) and C1 (red) are shown. OM, outer

732 membrane; PG, peptidoglycan; CM, cytoplasmic membrane. Scale bars, 10 nm.
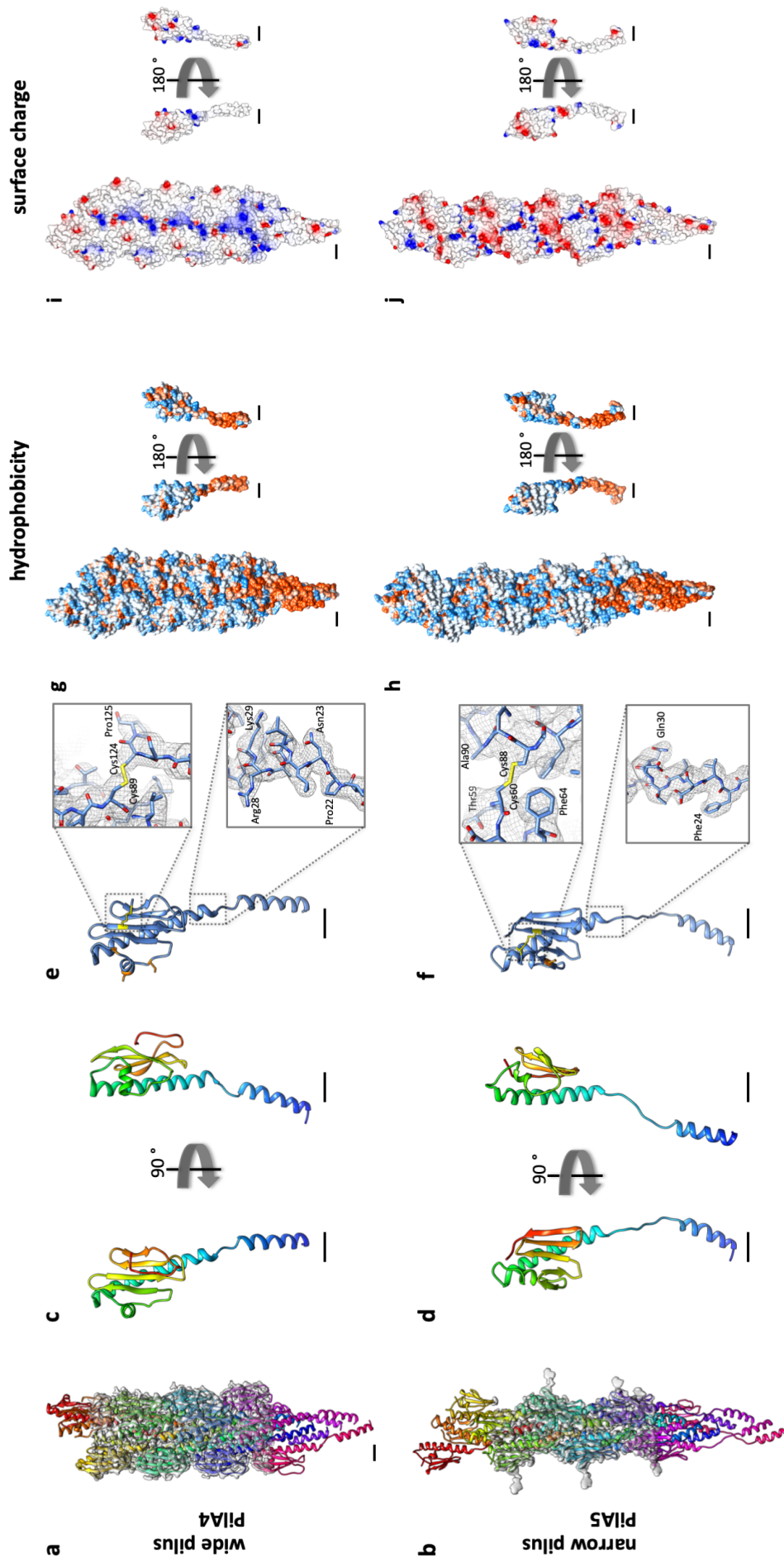
733 **Fig. 3: Molecular models of wide and narrow pili**

734   **a, b,** Molecular models of short sections of filaments (15 subunits each) with the

735   corresponding EM density maps. **a,** wide pilus comprised of PilA4; **b,** narrow pilus comprised

736   of PilA5. Scale bars, 10 Å.

737   **c, d,** Ribbon representation of a single PilA4 (**c**) and PilA5 (**d**) subunit. N-terminus blue, C-

738   terminus red. Scale bars, 10 Å.

739   **e, f,** Ribbon representation of a single PilA4 (**e**) and PilA5 (**f**) subunit with selected sidechains

740   shown. Yellow, disulphide bond; orange, serines with posttranslational modification. Insets

741   show details of the molecular model within the EM density map. Top insets, disulphide bond;

742   bottom insets, side chain densities. Scale bars, 10 Å.

743   **g, h,** Hydrophobicity of PilA4 (**g**) and PilA5 (**h**) filaments (left) and single subunits (right).

744   Hydrophobic (red) and hydrophilic (blue) residues are shown. Scale bars, 10 Å.

745   **i, j,** Electrostatic surface charge of PilA4 (**i**) and PilA5 (**j**) filaments (left) and single subunits

746   (right). Negative charges (red) and positive charges (blue) are shown. Scale bars, 10 Å.

747

748

749
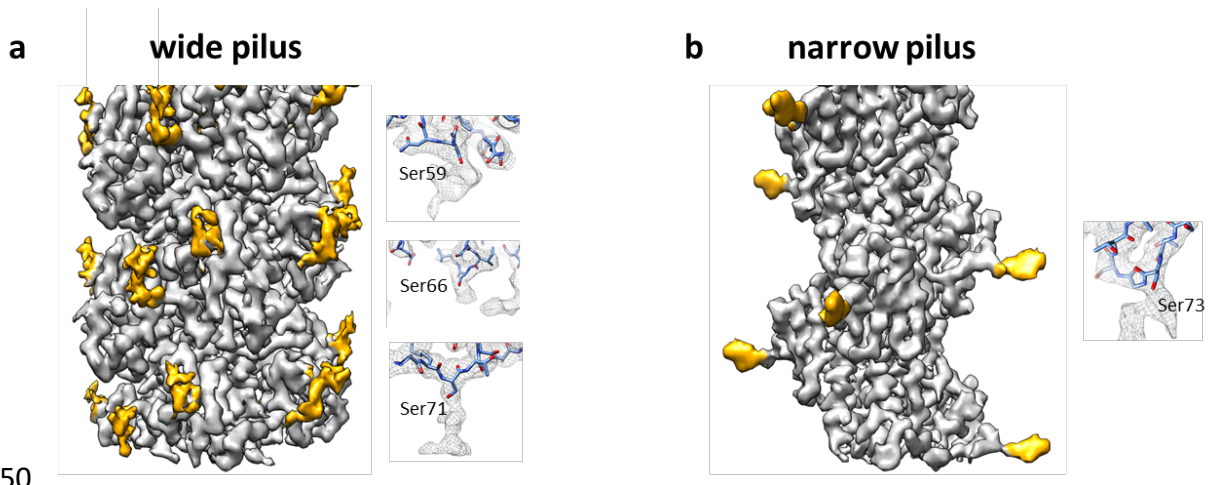
751   **Fig. 4: Post-translational modification**

752   **a, b,** Surface representation of the EM maps of the wide (**a**) and narrow (**b**) pilus showing the

753   protein model (grey) and densities that protrude into the solvent and could not be attributed

754   to the polypeptide backbone or an amino acid side chain. These are most likely glycans

755   (yellow). Insets show close-ups of large unassigned densities near serine residues. Scale bars,

756   10 Å.
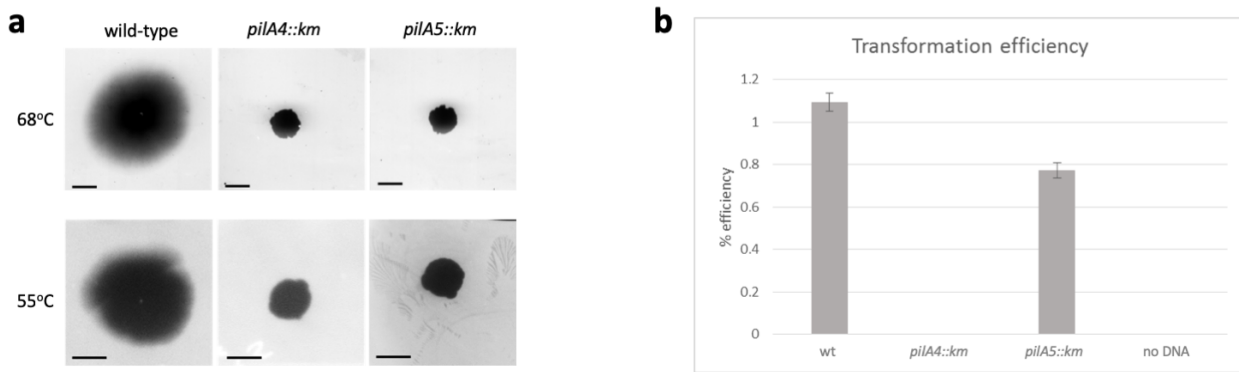
Fig. 5: Functional characterisation of *T. thermophilus* mutants

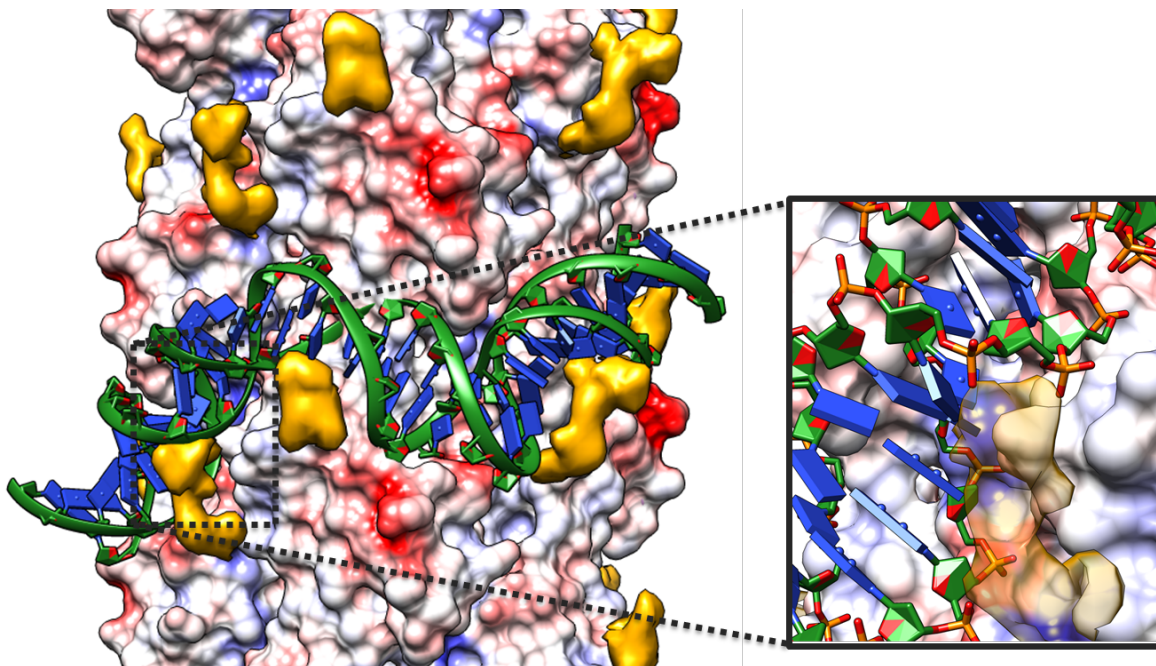**a,** Twitching motility of *T. thermophilus* HB27 strains. Only wild-type cells show twitching. Scale bars = 0.5 cm.

b, Natural transformation efficiency of *T. thermophilus* HB27 strains.

Fig. 6: Proposed DNA binding to wide pili

A double stranded DNA molecule (green) is modelled around a wide pilus shown in surface charge representation (negative charges, red; positive charges, blue). The DNA backbone fits neatly into the positively charged groove of the PilA4 filament (inset). Post-translational modifications are shown in yellow (transparent yellow in inset). Scale bar, 10 Å.