Structure and Function of Gli123 Involved in *Mycoplasma mobile* Gliding

Daiki Matsuike,[a] Yuhei O Tahara,[a,b] Takahiro Nonaka,[c] Heng Ning Wu,[c] Tasuku Hamaguchi,[d] Hisashi Kudo,[e,f] Yuuki Hayashi,[e,g] Munehito Arai,[e,h] and Makoto Miyata[a,b]

[a] Graduate School of Science, Osaka Metropolitan University, Sumiyoshi-ku, Osaka, Japan

[b] The OCU Advanced Research Institute for Natural Science and Technology (OCARINA), Osaka Metropolitan University, Sumiyoshi-ku, Osaka, Japan

[c] Graduate School of Science, Osaka City University, Sumiyoshi-ku, Osaka, Japan

[d] Institute of Multidisciplinary Research for Advanced Materials (IMRAM), Tohoku University, Aoba-ku, Miyagi, Japan

[e] Department of Life Sciences, The University of Tokyo, Meguro Tokyo, Japan

[f] Graduate School of Science, Technology and Innovation, Kobe University, Nada, Kobe, Japan

[g] Environmental Science Center, The University of Tokyo, Bunkyo, Tokyo, Japan

[h] Department of Physics, The University of Tokyo, Meguro, Tokyo, Japan

Running Title: *Mycoplasma* surface gliding protein

[#] Address correspondence to Makoto Miyata, miyata@omu.ac.jp

ABSTRACT

Mycoplasma mobile is a fish pathogen that glides on solid surfaces by means of a unique mechanism. The gliding machinery of *M. mobile* is composed of internal and surface structures. In the present study, we focused on the function and structure of Gli123, a surface protein that is essential for the localization of other surface proteins. The amino acid sequence of Gli123, which is 1128 amino acids long, contains lipoprotein-specific repeats. We isolated the native Gli123 protein from *M. mobile* cells and a recombinant protein, rGli123, from *Escherichia coli*. The isolated rGli123 complemented a non-binding and non-gliding mutant of *M. mobile* that lacked Gli123. Circular dichroism and rotary-shadowing electron microscopy (EM) showed that rGli123 has a structure that is not significantly different from that of the native protein. Rotary-shadowing EM suggested that the molecules changed their shape between globular and rod-like structures, depending on the ionic strength of the solution. Negative-staining EM coupled with single-particle analysis revealed that Gli123 forms a globular structure featuring a small protrusion with dimensions of 20.0, 14.5, and 16.0 nm. Small-angle X-ray scattering analyses indicated a rod-like structure composed of several tandem globular domains with total dimensions of approximately 34 nm length and 4 nm width. Both molecular structures were suggested to be dimers based on the

predicted molecular size and structure. Gli123 may have evolved by multiplication of repeating lipoprotein units and acquired clumping role of surface proteins.

IMPORTANCE

Mycoplasmas are pathogenic bacteria that are widespread in animals. They are characterized by small cell and genome sizes but are equipped with unique abilities to escape host immunity, such as surface variation and gliding. Here, we focused on a surface-localizing protein that is essential for *Mycoplasma mobile* gliding. The findings of this study suggested that the protein undergoes drastic conformational changes between its rod-like and globular structures. These changes may be caused by a repetitive structure common in the surface proteins that is responsible for the modulation of the cell surface structure and related to the assembly process for the surface gliding machinery. An evolutionary process for this unique mycoplasma gliding mechanism has also been suggested in the present study.

KEYWORDS

electron microscopy, small-angle X-ray scattering, repeat sequence, single-particle analysis, mollicutes, lipoprotein, conformational change, motility

59 **INTRODUCTION**

60 Class Mollicutes, which includes the genus *Mycoplasma*, consists of organisms fit to
61 parasitic or commensal life cycles, as represented by their small genome size, lack of
62 peptidoglycan layer, and antigenic modulation (1-3). Interestingly, more than 20
63 *Mycoplasma* species are known to glide on solid surfaces through unique mechanisms
64 (4). *Mycoplasma* gliding mechanisms can be classified into two types, as represented
65 by either *Mycoplasma mobile* or *Mycoplasma pneumoniae* (4, 5). Four species are
66 known to perform mobile-type gliding, including *M. mobile*, *Mycoplasma pulmonis*,
67 *Mycoplasma testudineum*, and *Mycoplasma agassizii*, which serve as pathogens for
68 freshwater fish, mice, turtles, and turtles, respectively (6, 7). The gliding mechanism has
69 been studied with a focus on *M. mobile*, which glides at speeds of 2.0 to 4.5 μm/s (4, 8).
70 The gliding machinery, which forms as a protrusion, can be divided into two parts: the
71 inside and surface structures (4, 9, 10). Its surface structure is composed of four
72 proteins: Gli123, Gli349, Gli521, and Gli42. Gli123, Gli349, and Gli521 have been
73 reported as being essential for binding and gliding, based on analyses of mutants
74 missing each of these proteins (11-13). Gli349 is a flexible rod-shaped, 95-nm-long
75 protein with a C-terminal globular domain (11, 14, 15). It catches sialylated
76 oligosaccharides on the surfaces of hosts as a "leg" for gliding, and pulls the cell (16,

77 17). Gli521 is a rather rigid rod-shaped 120-nm-long protein with N-terminal globular
78 and C-terminal hook domains (18, 19) that is suggested to transmit the gliding force
79 generated in an internal motor to Gli349, like a "crank" (5, 18, 19). Gli123 localizes
80 Gli349 and Gli521 correctly to the gliding machinery, as a "mount" (13). Gli349 and
81 Gli521 are conserved in all four species of *M. mobile*, *M. pulmonis*, *M. testudineum*, and
82 *M. agassizii*, whereas Gli123 and Gli42 are only found in *M. mobile* and *M. pulmonis*
83 (20). In this study, we isolated Gli123 and found that a mutant lacking this protein could
84 be complemented by the isolated protein instead. We analyzed Gli123 structures using
85 electron microscopy (EM), small-angle X-ray scattering (SAXS), and other biochemical
86 methods, and suggested a plausible structural model and its conversion, which may be
87 related to the role of the protein in gliding motility.
88

89 **RESULTS**

90 **Gli123 orthologs in the *Mycoplasma* species**

91 Orthologs of *gli123* and *gli42* genes were not found in *M. agassizii* and *M. testudineum*,
92 unlike the orthologs of nine other genes involved in *M. mobile* gliding (20). The orthologs
93 of *gli349* and *gli521* from *M. testudineum* and *M. agassizii* have been found to be
94 positioned closer to *M. pulmonis* than to *M. mobile* in the phylogenetic tree (20).

95 Therefore, we performed a PSI-BLAST search (21) using the amino acid sequences of
96 Gli123 and Gli42 of *M. pulmonis*, and found the orthologs of *gli123* and *gli42* in the
97 genomes of *M. agassizii* and *M. testudineum* (Table S1). The phylogenetic trees
98 constructed for *gli123*, *gli42*, *gli349*, and *gli521* genes using the maximum likelihood
99 method shared their topology, as did the phylogenetic tree of 16S ribosomal DNA
100 sequences (Fig. S1) (20). Next, we examined the localization of these genes in the
101 genome. In *M. mobile* and *M. pulmonis*, the genes were clustered in one locus (13, 20),
102 while in *M. agassizii* and *M. testudineum*, they were split into two loci, *gli123-gli42* and
103 *gli349-gli521*, although the actual distances are unknown in the unassembled genome
104 sequences (Fig. S2).
105
106 **Sequence analysis of Gli123 and the lipoprotein-17 domain**
107 In this study, we found that the amino acid sequences of Gli123 protein and its orthologs
108 include four to six lipoprotein-17 domains, which are composed of approximately 100
109 amino acids (Fig. 1A) (22, 23). The lipoprotein-17 domain is found in 233 proteins of 43
110 bacterial species, mostly in lipoproteins that play a role in antigenic variation or
111 adhesion. In *Mycoplasma,* the lipoprotein-17 domain was identified in 221 proteins from
112 32 species (Table S2). This domain was found in 15 Mvsps, Gli123, and Gli349 in *M.*

113 *mobile* (23, 24). The Gli123 protein of *M. mobile* had four lipoprotein-17 domains, while

114 that of *M. pulmonis* and *M. testudineum* had five and that of *M. agassizii* had six (Fig. 1A

115 and Table S3). Mvsps have a number of lipoprotein-17 domains, up to 16 (Fig. 1A and

116 Table S3). This may suggest that the surface proteins adjust their lengths by means of

117 repetitive fusion of the domain (4). The hidden Markov model scores for the

118 lipoprotein-17 domain were distributed between 30 and 70 for the Mvsps repeats and

119 between 10 and 30 for the Gli123 and Gli349 repeats (Fig. 1B).

120

121 **MvspI mutant strains of *M. mobile***

122 The isolation of native Gli123 (nGli123) from *M. mobile* cell lysates is a long process

123 because the lipoprotein, MvspI encoded as MMOB3340, behaves in a similar way in the

124 isolation process. MvspI is characterized by a mass of 220 kDa, with 16 copies of

125 lipoprotein-17 domains (24). We isolated a spontaneous MvspI-deficient mutant to

126 simplify the isolation process of Gli123. Colonies that were non-reactive to a monoclonal

127 antibody against MvspI were screened and isolated through colony blotting and Amido

128 Black staining (Fig. S3A). Two colonies that were non-reactive to the antibody were

129 isolated from an approximate of 10000 colonies. Sodium dodecyl

130 sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) analysis of the isolate revealed

131 a lack of MvspI protein (Fig. S3B). The band intensities of the surface protein, MvspF
132 encoded as MMOB3300, was higher in the isolated mutant than in the wild-type (WT)
133 strain (Fig. S3B). Genome analysis revealed that the *mvspI* gene of the isolate had a
134 frameshift mutation, resulting in a truncated peptide of 97 amino acids (Table S4). In
135 addition to mutations in the *mvspI* gene, nine other mutations were identified in *mvsps*
136 and others (Table S4). This mutant grew constantly through serial cultivation of large
137 volumes, without recovery of MvspI. The mutant cells glided normally, but the culture
138 tended to form clumps, suggesting different surface interactions from those of the WT.
139 This mutant was named the MvspI^{-} strain and used in subsequent experiments.
140

141 **Isolation of Gli123 from *M. mobile***

142 We isolated the nGli123 protein from cultured cells of the *M. mobile* MvspI^{-} strain (Fig.
143 2A). The collected cells were suspended in a buffer containing 1% Triton™ X-100 and
144 treated on ice, for 1 h. Following that, the lysates were separated by means of
145 centrifugation into a precipitate and a soluble fraction containing nGli123, Gli349, Gli521,
146 and MMOB1650 (Fig. 2B). The soluble fraction was then subjected to anion-exchange
147 chromatography. The nGli123 protein was bound to an anion-exchange column with 20
148 mM NaCl and eluted using a gradient of NaCl. Gli521 and MMOB1650 were separated

149 from nGli123 by means of this elution process (Fig. 2B). Next, we subjected the
150 fractions containing nGli123 and Gli349 to size-exclusion chromatography (SEC). The
151 proteins, which were traced by means of absorbance, were divided into four peaks,
152 other than the void-volume fraction (Fig. 2C). SDS-PAGE and mass spectrometry
153 identified nGli123 in the second peak near the position of the 150 mL elution (Fig. 2B
154 and 2C). The Stokes radius of nGli123 was calculated as 5.4 nm from the elution
155 position (Fig. 2C). The final yield of the nGli123 protein was 0.3 mg, isolated from 3 L of
156 growth medium. The N-terminal amino acid sequence of the nGli123 protein band
157 collected from the electrophoresis of the whole-cell lysate was analyzed using Edman
158 degradation and determined as Ala-Ile-Ala-Ile-Gly, thereby showing that the N-terminus
159 of 18 amino acids including the partial transmembrane segment of 11 amino acids is
160 truncated in the mature protein (Fig. 2A) (13).
161
162 **Expression and purification of recombinant Gli123**
163 The yield of nGli123 was not high enough to perform various experiments. Thus, we
164 expressed recombinant Gli123 (rGli123) with a 6×His-tag at the C-terminus of mature
165 Gli123 in *E. coli* (Fig. 2A) (13) and purified it using three steps: (i) nickel-nitrilotriacetic
166 acid (Ni^{2+}-NTA) affinity chromatography, (ii) hydrophobic interaction chromatography,

167 and (iii) SEC (Fig. 2B). The purified fraction was recovered by means of SEC, as a
168 single peak that included only the rGli123 protein, which was then confirmed using
169 SDS-PAGE and visualized using Coomassie Brilliant Blue (CBB) staining and mass
170 spectrometry (Fig. 2B and 2C). The Stokes radius of rGli123 was estimated as 5.4 nm,
171 from the elution position of SEC (Fig. 2C), which agreed with that for the nGli123
172 isolated from *M. mobile* cells. The final yield of the rGli123 protein was approximately 10
173 mg per 1 L of *E. coli* culture. The circular dichroism spectrum of rGli123 showed that the
174 secondary structural contents included 14.0% α-helices and 35.1% β-sheets (Fig. 2D)
175 (25, 26), consistent with the predictions made from the amino acid sequence using
176 XtalPred, *i.e.*, 14.0% for α-helices and 43.0% for β-sheets. The purified rGli123 protein
177 was subjected to limited proteolysis with trypsin, to identify the flexible and rigid parts
178 (Fig. 2A). The SDS-PAGE profile of the cleaved fragment showed that the rGli123
179 protein finally converged into three products through several fragmentation steps (Fig.
180 S4A). Mass spectrometry results showed that the three final fragments (f, g, and h)
181 corresponded at least to the 405–1128 (+8), 565–1128 (+8), and 559–789 amino acid
182 sequences, respectively. Based on these results, we believe that some components of
183 the molecule are rigid, as presented in Fig. 2A.
184

185 **Complementation of the binding and gliding abilities of the *M. mobil*e mutant with**

186 **additional Gli123 proteins**

187 To examine the complementation activity of Gli123 for the Gli123-lacking mutant, m12

188 (13), we added purified rGli123 protein to the cell suspension and incubated it for 3 h at

189 25°C. The cells were inserted into the tunnel slides with video recordings of

190 phase-contrast optical microscopy (27). m12 cells were unable to bind to the glass

191 surface (13), whereas cells with rGli123 protein showed binding to the glass surface

192 (Fig. 3A). The average number of bound cells in a 100 μm^2 field on glass was 15.1 ± 3.4

193 cells (n = 15), which was 69.1% of that of the WT. 14.9 ± 8.5% of the bound cells

194 showed unidirectional gliding, similar to the WT strain, with a speed of 1.5 ± 0.8 μm/s (n

195 = 45), comparable with that of the WT strain, 2.7 ± 0.5 μm/s (n = 45) (Fig. 3B and 3C),

196 suggesting that the rGli123 protein binds and complements the gliding machinery (13).

197 Next, we examined the binding of the rGli123 proteins to the mutant cells. Mutant cells

198 incubated with rGli123 were recovered by means of centrifugation and analyzed using

199 SDS-PAGE (Fig. 3D). This result showed that rGli123 bound to the mutant cells at an

200 amount comparable to that of the WT strain.

201

202 **Conformational changes in Gli123**

203 We observed the structures of the nGli123 and rGli123 proteins using rotary-shadowing
204 EM. In rotary-shadowing EM, protein molecules sprayed on a rotating mica surface are
205 shadowed by platinum particles from a low angle, to observe their molecular shape (19,
206 24). We observed particles of similar sizes and shapes in the field of view. The nGli123
207 and rGli123 proteins showed similar rod-like structures, with lengths of 35.6 ± 5.0 and
208 35.9 ± 4.1 nm, respectively, under conditions including 100 mM ammonium acetate (Fig.
209 4A and 4B). Next, we observed the nGli123 and rGli123 proteins in the presence of 500
210 mM ammonium acetate. Interestingly, both proteins showed globular structures with
211 long axes of 21.0 ± 3.2 and 20.1 ± 4.1 nm (Fig. 4A and 4B). We then examined the
212 molecular ratios of the two structures in 100, 250, and 500 mM ammonium acetate. The
213 proportions of rod-like structures at each concentration were 89.7%, 47.0%, and 5.7%,
214 respectively, and the ratio changed from rod-like to globular structures depending on the
215 ammonium acetate concentration (Fig. 4C). Furthermore, we measured the light
216 scattering of rGli123. Light scattering intensity measured at different concentrations
217 increased with ammonium acetate concentration, in the range of 50–650 mM,
218 suggesting that the conformation of rGli123 molecules depends on the concentration of
219 ammonium acetate (Fig. 4D). The 50% effective concentration (EC_{50}) was 261.0 ± 11.3,
220 238.3 ± 16.1, and 292.9 ± 4.9 mM at protein concentrations of 0.23, 0.75, and 1.23

221 mg/mL, respectively (Fig. 4D). The changes in the light scattering intensities of rGli123
222 proteins due to changes in ionic strength were in good agreement with the ratios of
223 globular and rod-like structures measured using rotary-shadowing EM (Fig. 4C and 4D).
224

225 **Globular structure of Gli123**

226 Negative-staining EM was performed to visualize the three-dimensional structure of
227 globular nGli123 with better resolution. The field images of EM showed uniform globular
228 particles of 15–20 nm (Fig. 5A), and their density changed depending on the protein
229 concentration used. In addition, many rod-like structures were observed in the
230 background. We selected globular particles (Fig. 5A) and obtained 2D-averaged images
231 using the structural analysis software EMAN2.91 (Fig. 5B). The 2D-averaged images
232 showed that the Gli123 molecules had dimensions of approximately 17 and 14 nm for
233 the major and minor axes, respectively. Based on the averaged image of the Gli123
234 protein, a three-dimensional image was reconstructed, without C1 and with C2, the
235 two-rotational symmetry enforcement. The three-dimensional (3D)-reconstructed
236 globular structure of Gli123 had a protrusion that was reminiscent of a ‘spinning top’ (Fig.
237 5C). Considering the globular molecular shape, the molecular weight of Gli123 was
238 estimated to be 244 kDa from the elution position of SEC, indicating that the protein

239 behaves as a dimer (Fig. 2C).

240

241 **Rod-like structure of Gli123**

242 SAXS was used to obtain the 3D rod-like structure of Gli123. The rGli123 protein under

243 low ionic strength condition was subjected to SEC-SAXS. The protein showed an

244 elution curve with a single peak (Fig. 6A). We analyzed the scattering curve at the peak

245 top of the elution profile with a Guinier plot [ln $I(Q)$ *vs.* Q^2, where $I(Q)$ is the scattering

246 intensity at scattering vector Q], to estimate the size of the rGli123 protein (Fig. 6B). The

247 Guinier plot was based on the scattering curve at small angles, and the square of the

248 radius of gyration (R_g) was obtained from the slope of the approximate line in the

249 Guinier region ($Q \times R_g < 1.3$) as $\ln I(Q) = \ln I(0) - R_g^2 Q^2/3$, where $I(0)$ is the zero-angle

250 scattering intensity (28). R_g was calculated as 8.24 nm from the slope of the

251 approximate line (Fig. 6B). Next, to discuss the shape of the rGli123 protein, the

252 scattering curve was analyzed using a Kratky plot [$I(Q)Q^2$ *vs.* Q] (Fig. 6C), which shows

253 the characteristics of the protein shape (29). The plot shows a mountainous curve with a

254 peak around $Q = 0.1$ Å^{-1}, suggesting that the rGli123 protein is not in a random coil but

255 has a folded structure. For an ideal sphere, the Kratky plot shows a hyperbolic curve

256 with a single peak in the small-angle regions. However, in the present data, the plot

257 shows a shoulder around Q = 0.02 Å^{-1}, marked by an inverted black triangle in the figure

258 (Fig. 6C), suggesting the existence of another hidden peak. The bimodal character of

259 the Kratky plot suggests a dumbbell shape for the protein; the innermost shoulder

260 reflects the globularity of the whole molecule, while the peak reflects that of each

261 domain (30). According to the pair-distance distribution function, $P(r)$, the maximum

262 length (D_{max}) of the rGli123 protein was approximately 31.6 nm (Fig. 6D). The 3D

263 structure of the rGli123 protein was modeled from SAXS data using PRIMUSQT

264 software (Fig. 6E) (31). The structural model was rod-shaped and featured several

265 narrow parts, similar to an articulated robot. This structure was consistent with the

266 images obtained using rotary-shadowing EM with low ionic intensity (Fig. 4B). In the

267 cross-section plot [ln $I(Q)Q$ *vs.* Q^2], which tests the validity of the rod-like structure in the

268 SAXS model (32), the slope of the plot showed high linearity (Fig. S5). The radius of

269 gyration of cross-section R_c was 1.14 nm, which was obtained by fitting the

270 cross-section plot to the equation ln [$I(Q)Q$] = ln [$I(0)Q$] – $R_c^2Q^2/2$ (33). The radius of a

271 cylinder R_p (a short-axis radius) was 1.61 nm, as calculated using the equation R_p =

272 $\sqrt{2}R_c$ (33). The height of the cylinder L (twice the major-axis radius), determined from

273 the equation $R_g^2 = R_p^2/2 + L^2/12$ (33), was 28.2 nm. The L value obtained from the

274 cross-section plot was consistent with the length (D_{max} = 31.6 nm) estimated from the

275 *P*(*r*) function, indicating the high validity of the rod-like shape of the SAXS model. We
276 also believe that the rod-like structure is likely to be a dimer, because the SAXS model
277 has a molecular volume that corresponds to two molecules of a model predicted from
278 the amino acid sequence using Robetta (Fig. 6F) (34). SAXS measurements were also
279 performed using static cells at three different concentrations of rGli123 (Fig. S6). These
280 results were consistent with those obtained from the SEC-SAXS measurements, and
281 supported the absence of aggregation in the SAXS measurements.
282

283 **DISCUSSION**

284 **Surface gliding proteins**

285 In this study, *gli123* and *gli42* genes were found in the genomes of *M. testudineum* and
286 *M. agassizii* (Table S3), indicating that the two surface proteins of *M. mobile*, Gli123 and
287 Gli42, are also essential for gliding, similar to Gli349 and Gli521 (20). The topologies of
288 phylogenetic trees were common for all these proteins and 16S rDNA (20), suggesting
289 that these mycoplasmas have acquired gliding motility in evolution, not through
290 horizontal transfer (Fig. S1). In addition, the orthologs of *gli349* and *gli521* were
291 positioned adjacent to the genomes of all gliding *mycoplasmas* (Fig. S2) (20), and the
292 localization of Gli349 is dependent on that of Gli521 (13). These gliding proteins may

293 form a complex immediately after synthesis and establish a gliding machinery in which

294 binding and pulling are smoothly interlocked.

295

296 **Structure and conformational changes of Gli123**

297 We isolated the Gli123 protein and found that its molecular structure changed from

298 rod-like to globular depending on the ionic strength. This conformational change may

299 occur in the flexible regions identified through limited proteolysis (Fig. S4). A pair of

300 Gli123 molecules predicted from the amino acid sequence fit into both the rod-like

301 structure obtained using SAXS and the globular structure obtained by means of

302 single-particle analysis, suggesting that these structures of Gli123 are dimers (Figs. 5D

303 and 6F).

304 The structure of the isolated molecule changes drastically with ionic strength;

305 however, *M. mobile* can glide under different ionic strength environments, suggesting

306 that this conformational change is unlikely to occur in the completed machinery (13). We

307 observed that the purified rGli123 protein complemented the gliding of mutant cells

308 lacking Gli123. Conformational changes in the Gli123 molecule may be related to this

309 observation. These three gliding proteins have a localization hierarchy of

310 Gli123-Gli521-Gli349, suggesting a physical interaction between these proteins at the

311 cell surface (13). Gli123 might reach the target position on the cell surface as a flexible
312 rod shape and bundle Gli521 and Gli349 molecules on the cell surface through a drastic
313 conformational change, which is induced by the electrostatic fields made by the
314 neighboring proteins, Gli349 and Gli521.
315
316 **Roles of the lipoprotein-17 domain**
317 We found that a common repeat sequence of the lipoprotein-17 domain was shared by
318 the Mvsps and gliding proteins. The relationship between the gliding machinery and
319 Mvsps has been proposed based on the existence of a repeat comprising of
320 approximately 100 amino acids (22, 24). In this study, we succeeded in determining the
321 similarity between the repeat sequences. The existence of the lipoprotein-17 domain in
322 Gli123, Gli349, and their orthologs may suggest that they originated from a common
323 surface protein (Figs. 1A and 7A). The structure of the lipoprotein-17 domain in Gli123
324 and Mvsp predicted using AlphaFold2 featured three β-strands, which is consistent with
325 previously reported structural features determined using nuclear magnetic resonance
326 (NMR) and crystallography (Fig. S7A and S7B) (24). The predicted Mvsp structures
327 suggest that an α-helix with a short loop connects the lipoprotein-17 domains to form a
328 flexible chain (Fig. S7C). Although the detailed structures of Gli123 and Gli349 are still

329 unknown, their flexibility may be derived from the structures common with the
330 lipoprotein-17 domain and adjacent regions. Lipoprotein-17 domain may be useful for
331 adjusting molecular length and flexibility (Fig. 7A). It may provide length variation
332 adjustment to Mvsps and Gli349, which are involved in antigenic modulation and gliding
333 leg, respectively. Flexible and long leg proteins may be advantageous for catching the
334 host cell surface (17). In case of Gli123, it may provide length and flexibility for its role in
335 clamping other surface gliding proteins on the machinery surface, where as many as
336 450 gliding units are packed (Fig. 7B). The gliding units fixed by Gli123 can be aligned
337 in an ordered way for smooth and efficient gliding (13, 27).
338

339 **MATERIALS AND METHODS**

340 **Sequence analyses**

341 The genome sequences of *M. mobile* (163K), *M. pulmonis* (UAB CTIP), *M. agassizii*
342 (PS6), and *M. testudineum* (ATCC700618) were obtained from the National Center for
343 Biotechnology Information (NCBI) GenBank database
344 (https://www.ncbi.nlm.nih.gov/genome/). A homology search against the NCBI
345 non-redundant protein sequence databank was then performed using PSI-BLAST
346 (https://blast.ncbi.nlm.nih.gov/Blast.cgi), with an expected threshold of 100.

347 Phylogenetic trees of the proteins were constructed using neighbor-joining and
348 maximum likelihood methods. The reliability of the topology of each phylogenetic tree
349 was examined with 1000 bootstrap replicates, using the MEGAX package
350 (https://www.megasoftware.net/). The secondary structure of Gli123 was predicted
351 using XtalPred (https://xtalpred.godziklab.org/XtalPred-cgi/xtal.pl). Hidden Markov
352 model scores were obtained from Pfam (http://pfam.xfam.org/family/PF04200) and
353 UniProtKB (https://ftp.uniprot.org/pub/databases/uniprot/previous_releases/2021_03),
354 based on the HMMER3 package (http://hmmer.org/).
355
356 **Strains and culture conditions**
357 *M. mobile* 163K (ATCC 43663) was grown at 25°C in Aluotto medium, to an optical
358 density at the wavelength of 600 nm (OD_{600}) of around 0.08, as previously described
359 (35). *The M. mobile* MvspI^{-} strain was isolated as follows. Mutants that did not react with
360 monoclonal antibodies against MvspI were selected using colony blotting and Amido
361 Black staining (36). The colonies were isolated and cultured in Aluotto liquid medium,
362 following which SDS-PAGE analysis of the isolates was carried out, which confirmed the
363 absence of MvspI protein. *E. coli* DH5α and BL21(DE3) strains were used for DNA
364 recombination and protein expression, respectively.

365

366 **Isolation of Gli123 from *M. mobile* cells**

367 The nGli123 protein was purified from the *M. mobile* MvspI$^-$ strain cells. Cells stored at –
368 80°C were precultured and grown in 3 L of Aluotto liquid medium, dispensed into nine
369 300 cm^2 tissue culture flasks, and statically cultured at 25°C until the OD_{600} reached
370 0.08. The cells were collected by means of centrifugation at 18,800 × *g* and 4°C, for 30
371 min, and washed twice with phosphate-buffered saline (PBS; 58 mM Na_2HPO_4, 17 mM
372 NaH_2PO_4, and 68 mM NaCl). The precipitate was suspended in 30 mL of buffer A (20
373 mM Tris-HCl pH 8.0 and 0.1 mM phenylmethylsulfonyl fluoride) and sonicated for 2 min,
374 at intervals of 5 s, using a sonicator (US-600T, Nippon Seiki Co., Niigata, Japan), on ice.
375 The cell lysate was combined with 30 mL of buffer A containing 2% Triton™ X-100 (v/v)
376 and shaken on ice for 1 h, to solubilize the proteins. The cell lysate was separated by
377 means of centrifugation at 32,300 × *g* and 4°C, for 30 min, following which the obtained
378 supernatant was applied to an anion-exchange column (HiTrap™ Q HP 5 mL; Cytiva,
379 Tokyo, Japan) equilibrated with buffer A. Proteins bound to the column were eluted with
380 a NaCl concentration gradient of 0–500 mM, at 4°C. The eluted fraction containing
381 Gli123 was concentrated to a volume of 10 mL and subjected to a SEC column
382 (HiLoad® 26/600 Superdex® 200 prep grade, Cytiva) equilibrated with SEC buffer (20

383 mM Tris-HCl pH 8.0 and 150 mM NaCl), at 4°C. Thyroglobulin (669 kDa), ferritin (440

384 kDa), aldolase (158 kDa), and ovalbumin (43 kDa), with Stokes radii of 8.5, 6.1, 4.8, and

385 3.1 nm, respectively, were used as standards. Protein fractions were analyzed using

386 SDS-PAGE with a 10% polyacrylamide gel, followed by CBB staining.

387

388 **Expression and purification of recombinant Gli123**

389 The N-terminal sequence of Gli123 was determined using Edman degradation, as

390 previously described (11, 12). According to the sequencing results, the DNA sequence

391 encoding the amino acid region from the 19th to 1128th position was codon-optimized for

392 expression in *E. coli* and inserted between the *Bam*HI and *Xho*I sites in the multicloning

393 site of the pET21a plasmid (Novagen, Madison, WI). The recombinant protein, rGli123,

394 was expressed in *E. coli* BL21(DE3) and induced using 1 mM

395 isopropylthio-β-galactoside at 15°C, for 18 h. Following that, the cells were collected,

396 suspended in His-buffer A (20 mM Tris-HCl pH 8.0 and 20 mM NaCl), and washed twice.

397 The cells were suspended in 60 mL His-buffer A with 60 μL of 100 mM

398 phenylmethylsulfonyl fluoride and sonicated for 4 min, at intervals of 5 s, on ice, as

399 described above. The lysate was centrifuged at 32,300 × *g* and 4°C, for 30 min,

400 following which the supernatant was subjected to Ni^{2+}-NTA affinity chromatography

401 (HisTrap™ HP 5 mL, Cytiva). Proteins bound to the column were eluted with an
402 imidazole concentration gradient of 0–250 mM, at 4°C. The fraction containing rGli123
403 was adjusted to 20% saturation of ammonium sulfate, kept on ice for 1 h, following
404 which the obtained supernatant was collected after centrifugation at 32,300 × *g* and 4°C,
405 for 30 min. The supernatant was subjected to hydrophobic interaction chromatography
406 column (Phenyl Sepharose HP 1 mL, Cytiva) and eluted in a gradient of 20%–0%
407 ammonium sulfate. The rGli123 protein was concentrated to 5 mL by means of
408 centrifugation at 5,000 × *g* and 4°C, using a 50k Amicon tube (Amicon Ultra, Merck,
409 Darmstadt, Germany), and subjected to SEC. The protein concentration was adjusted
410 using a 50k Amicon tube, if necessary, for analyses.
411
412 **Optical analyses**
413 The circular dichroism spectra of the proteins were measured in the range of 200–250
414 nm using a spectropolarimeter (J-805, Jasco International Co., Tokyo, Japan), at 25°C.
415 A quartz cuvette with a 1-mm path length was filled with 1 mg/mL of rGli123 in SEC
416 buffer. The conditions for the measurements have been previously described (37). The
417 measured spectrum was analyzed using BeStSel (26), followed by estimation of the
418 α-helix, β-strand, and turn contents (25). The secondary structure was predicted from

419 the amino acid sequence using XtalPred (38). For light scattering, the rGli123 protein
420 dialyzed against 50 mM ammonium acetate was placed into a 10-mm quartz cell, and
421 the scattered light at the wavelength of 400 nm was measured at 25°C using a
422 fluorescence meter (FP-6200, Jasco International Co.). The ammonium acetate
423 concentration was adjusted by means of addition of 3 M ammonium acetate solution.
424

425 **Optical and electron microscopy**

426 Rotary-shadowing and negative-staining EM were performed as previously described
427 (10, 19, 24, 39). Protein images were analyzed using the image analysis software
428 EMAN2.91 (https://blake.bcm.edu/emanwiki/EMAN2). The approximately 10,000
429 particle images were selected and classified into 50 groups, averaged, and then
430 reconstructed into 3D structures from 3,202 particle images. For optical microscopy, *M.*
431 *mobile* cells were inserted into a tunnel slide (20 μm) precoated with Aluotto medium, for
432 60 min. After 1 min, the floating cells were removed. The cells on the glass surface were
433 observed using phase-contrast microscopy and analyzed using ImageJ v1.51w
434 (http://rsb.info.nih.gov/ij/), as previously described (17, 20, 40).
435

436 **Assessment of binding of rGli123 to mycoplasma cells**

437 Cultured *M. mobile* WT and m12 cells were washed twice and suspended in PBS. The
438 suspended cells were kept in PBS containing 0.5 mg/mL of rGli123 protein for 3 h,
439 following which they were collected by means of centrifugation at 12,000 × *g* and 25°C,
440 for 5 min, suspended in PBS, sonicated, and analyzed using SDS-PAGE with CBB
441 staining. Protein band intensities were analyzed using ImageJ v1.51w.
442
443 **SAXS measurements**
444 SEC-SAXS measurements were performed at beamline 10C of the Photon Factory of
445 the High Energy Accelerator Research Organization (KEK), Tsukuba, Japan. The
446 rGli123 proteins were adjusted to a concentration of 5.33 mg/mL and subjected to a
447 SEC column (Superdex® 200 Increase 3.2/300, Cytiva) equilibrated with 100 mM
448 ammonium phosphate, at 4°C. The camera configuration has been previously described
449 (37). SAXS measurements using a static cell were performed at three different protein
450 concentrations. The measurements of the two-dimensional scattering images were
451 circularly averaged and converted to a one-dimensional scattering curve using the
452 SAngler software (41). The scattering curves immediately before the elution of the
453 protein were used as those of the blank samples. The average scattering curve of the
454 blanks was subtracted from the scattering curve of the protein solution to obtain the

455 scattering profile I(Q) of the protein. R_g and I(0) were obtained based on Guinier

456 approximation within the Guinier region (R_gQ < 1.3). The pairwise distance distribution

457 function $P(r)$ was calculated using GNOM software (31). The *ab initio* structural model

458 was simulated using the DAMMIF and DAMMIN softwares (31). Briefly, the outline of the

459 3D structure was calculated 10 times independently with DAMMIF, using dummy atoms.

460 The structural models were averaged using DAMAVER, and the averaged structural

461 models were calculated again using DAMMIF, to obtain the final structural model. The

462 constructed model was displayed in UCSF Chimera 1.10 (42). In the cross-section plot

463 (Fig. S5), the radius of gyration of the cross-section (R_c) was obtained by applying a

464 linear approximation to a plot of the linear region in the range after $Q = 1/R_g$ (33). The

465 radius of the short axis (Rp) was calculated using the equation $R_p = \sqrt{2}R_c$ (33). The

466 height of the cylinder (L) was calculated by substituting the values of R_g and R_p in $R_g^2 =$

467 $R_p^2/2 + L^2/12$ (33).

468

469 **Limited proteolysis**

470 The rGli123 protein in SEC-buffer, at the concentration of 1 mg/mL, was treated with

471 0.001 mg/mL trypsin, at 25°C, for various reaction times, following which the reaction

472 was stopped by means of heat treatment at 95°C for 5 min, in SDS-sample solution

473 containing 5% glycerol, 0.025% Bromophenol Blue, 62.5 mM Tris-HCl pH 6.8, 2.5%
474 SDS, and 5% β-mercaptoethanol. The protein digests were analyzed using SDS-PAGE
475 with a 10% acrylamide gel and peptide mass fingerprinting, as described previously (19,
476 39).
477
478 **Figure Legends**
479 **FIG. 1** Lipoprotein-17 domains in Mvsps, Gli123, and Gli349. (A) Distributions of
480 lipoprotein-17 domains in Mvsps, Gli349, and Gli123 sequences have been shown
481 using white boxes. MYPU_2160, CJF60_02770, and EI74_0698 are Gli123 orthologs of
482 *M. pulmonis*, *M. agassizii*, and *M. testudineum*, respectively. (B) Column distribution of
483 the scores calculated using the hidden Markov model (HMM) for each lipoprotein-17
484 domain. The scores for each lipoprotein-17 domain of Gli123 and Gli349 and orthologs
485 in *M. mobile, M. pulmonis*, *M. agassizii*, and *M. testudineum* and Mvsps in *M. mobile* are
486 shown.
487
488 **FIG. 2** Amino acid sequence and isolation process of native and recombinant Gli123
489 proteins (nGli123 and rGli123, respectively). (A) Schematic illustration of the amino acid
490 sequences of nGli123 and rGli123. Transmembrane segment and tagging sequences

491 are marked in red and green, respectively. The blue line and triangle indicate the
492 truncated part of the mature protein. The hatched boxes indicate the lipoprotein-17
493 domains. The bottom diagram for limited proteolysis shows overlap of degradation
494 products in the digest, suggesting rigid areas, as indicated by the four shades. (B) Left:
495 protein profiles of the fractions obtained in the process of isolation of nGli123 from *M.*
496 *mobile*. Lane 1, whole-cell lysate; Lane 2, Triton™-insoluble fraction; Lane 3,
497 Triton™-soluble fraction; Lane 4, nGli123 fraction eluted by means of anion-exchange
498 chromatography; Lane 5, nGli123 fraction eluted by means of size-exclusion
499 chromatography (SEC). The protein bands of Gli521, Gli349, MMOB1650, and nGli123
500 are marked using green, red, purple, and magenta circles, respectively. Right: protein
501 profiles of fractions in the rGli123 purification process. Lane 1, whole-cell lysate; Lanes
502 2 and 3, insoluble and soluble fractions of the lysate after sonication, respectively; Lane
503 4, rGli123 fraction after Ni^{2+}-NTA affinity chromatography; Lanes 5, rGli123 fraction after
504 hydrophobic interaction chromatography; Lanes 6, rGli123 fraction after SEC. The
505 protein fractions were applied to SDS-PAGE with a 10% acrylamide gel and stained with
506 Coomassie Brilliant Blue. The molecular masses are indicated on the left. (C) SEC of
507 nGli123 and rGli123. Thyroglobulin (669 kDa, 8.5 nm), ferritin (440 kDa, 6.1 nm),
508 aldolase (158 kDa, 4.8 nm), and ovalbumin (43 kDa, 3.1 nm) were used as the markers

509 for molecular mass and Stokes radius. The triangles mark the peaks of nGli123 and

510 rGli123. (D) CD spectrum of rGli123.

511

512 **FIG. 3** Recovery of gliding and binding by Gli123 proteins in m12 mutant cells. (A) Cell

513 movements analyzed using optical microscopy. The cells were observed using

514 phase-contrast microscopy (left) and a video was recorded. The video frame (at every

515 0.03 s) was overlaid as a stack for 10 s, changing its color from red to blue (right). (B)

516 Number of cells binding and gliding on glass surfaces. (C) Gliding speeds. (D) Left:

517 Protein profiles of m12 mutant and m12 wild-type cells, after addition of rGli123 protein.

518 The protein bands for Gli521, Gli349, and Gli123 are marked using green, red, and

519 magenta circles, respectively. Right: Molecular ratios of Gli123, Gli349, and Gli521 were

520 calculated from the band intensities normalized to the value of MvspI.

521

522 **FIG. 4** Ionic strength-induced conformational changes in the nGli123 and rGli123

523 proteins. Field images of the nGli123 (A) and rGli123 (B) proteins under the conditions

524 of 100 mM (left) and 500 mM (right) ammonium acetate. Scale bar: 50 nm. The

525 representative particle images are magnified in the bottom. Scale bar: 10 nm. The

526 distributions of diameter and length are shown for the globular and rod-like structures on

527 the right. (C) Ratio of globular and rod-like structures of the rGli123 protein under the

528 conditions of 100, 250, and 500 mM ammonium acetate. (D) Light scattering of rGli123

529 proteins, as a function of ammonium acetate concentration.

530

531 **FIG. 5** Globular structure of Gli123 protein, as observed by means of negative-staining

532 EM. (A) Field image of Gli123 molecules that were negatively stained with 2% uranyl

533 acetate. Representative particle images from the fields are shown in the bottom.

534 Single-particle analysis was performed using EMAN2.91 software. (B) Representative

535 averaged images of Gli123 were selected from 50 classes which were obtained by

536 averaging approximately 10,000 particles using the command "2D Analysis". (C)

537 Three-dimensional reconstruction of Gli123 molecule by means of *ab initio*

538 reconstruction with the parameters; C1, non-symmetrical; C2, two-rotational symmetry.

539 (D) Superposition of the C2 model structures obtained from the single-particle analysis

540 and those predicted from the amino acid sequence (1–1000 aa) using Robetta.

541

542 **FIG. 6** Rod-like structure of rGli123 protein, as predicted using SEC-SAXS. (A) SEC of

543 rGli123 was carried out by measuring the UV absorbance at the wavelength of 280 nm,

544 in the presence of 100 mM ammonium acetate. (B) Scattering intensities transformed as

545 a Guinier plot [ln I(Q) *versus* Q^2]. Data-points in a Guinier region (R_gQ < 1.3), which are

546 shown in the figure using filled circles, were then used to determine the R_g value and

547 zero-angle scattering intensity I(0) with Guinier approximation. (C) Scattering intensities

548 transformed as a Kratky-plot [$I(Q)Q^2$ *versus* Q]. White and black triangles indicate the

549 peak and hidden peak (shoulder), respectively. (D) Pair-distance distribution function

550 $P(r)$ calculated from the scattering curve. (E) SAXS models of rGli123. (F) Overlay of

551 SAXS models and the structures predicted from the amino acid sequence (1–1000 aa)

552 using Robetta.

553

554 **FIG. 7** Role of lipoprotein-17 domain repeats and Gli123. (A) The repeat sequence is

555 possibly used to adjust the length and flexibility of surface molecules that are involved in

556 antigenic modulation (Mvsp) and gliding (Gli349 and Gli123). (B) The flexibility and

557 molecular length of Gli123 allow for conformational changes in it, which facilitate its

558 clamping to other gliding proteins at appropriate positions in the machinery.

559

560 **ACKNOWLEDGEMENTS**

561 We thank Dr. Nobutaka Shimizu (KEK), Dr. Shinya Saijo (KEK), and the students of Arai

562 Lab (Department of Life Sciences, The University of Tokyo) for their assistance with the

563 SAXS measurements. We also thank Toshiaki Arata (Graduate School of Science,
564 Osaka City University) and Ikuko Fujiwara (Department of Bioengineering, Nagaoka
565 University of Technology) for their experimental advice in the light scattering analyses,
566 and Ms. Aya Takamori (Graduate School of Science, Osaka City University) for her
567 technical assistance with the MALDI-TOF MASS spectrometry. The SAXS
568 measurements were performed with the approval of the Photon Factory Program
569 Advisory Committee. This study was supported by Grants-in-Aid for Scientific Research
570 on the Innovative Area Harmonized Supramolecular Motility Machinery and Its Diversity
571 (MEXT KAKENHI grant numbers JP24117002, JP25117504, and JP15H01311),
572 grants-in-aid for scientific research (B) and (A) (MEXT KAKENHI grant numbers
573 JP24390107, JP17H01544, and JP19H02521), and JST CREST grant number
574 JPMJCR19S5 to MM.

575

576 **REFERENCES**

577 1. Kirchhoff H, Rosengarten R. 1984. Isolation of a motile mycoplasma from fish. J
578 Gen Microbiol 130:2439-45.
579 2. Razin S, Yogev D, Naot Y. 1998. Molecular biology and pathogenicity of
580 mycoplasmas. Microbiol Mol Biol Rev 62:1094-156.

581 3. Grosjean H, Breton M, Sirand-Pugnet P, Tardy F, Thiaucourt F, Citti C, Barre A,
582 Yoshizawa S, Fourmy D, de Crecy-Lagard V, Blanchard A. 2014. Predicting the
583 minimal translation apparatus: lessons from the reductive evolution of *mollicutes*.
584 PLoS Genet 10:e1004363.
585 4. Miyata M, Hamaguchi T. 2016. Prospects for the gliding mechanism of
586 *Mycoplasma mobile*. Curr Opin Microbiol 29:15-21.
587 5. Miyata M, Hamaguchi T. 2016. Integrated information and prospects for gliding
588 mechanism of the pathogenic bacterium *Mycoplasma pneumoniae*. Front
589 Microbiol 7:960.
590 6. Hatchel JM, Jurkovic DA, Balish RS, Balish MF. 2006. Differences and
591 correlations in gliding motility and cytadherence between species and strains of
592 the *Mycoplasma pneumoniae* group. Abstract book for 16th International
593 Congress of International Organization for Mycoplasmology:p137.
594 7. Jurkovic DA, Friedberg AJ, Hatchel JM, Relich RF, Balish MF. 2006. Adherence
595 and gliding motility at low temperatures by tortoise-associated mycoplasmas.
596 Abstract book for 16th International Congress of International Organization for
597 Mycoplasmology:p24.
598 8. Miyata M. 2008. Centipede and inchworm models to explain *Mycoplasma* gliding.

599 Trends Microbiol 16:6-12.

600 9. Nishikawa MS, Nakane D, Toyonaga T, Kawamoto A, Kato T, Namba K, Miyata
601 M. 2019. Refined mechanism of *Mycoplasma mobile* gliding based on structure,
602 ATPase activity, and sialic acid binding of machinery. *mBio* 10.

603 10. Toyonaga T, Kato T, Kawamoto A, Kodera N, Hamaguchi T, Tahara YO, Ando T,
604 Namba K, Miyata M. 2021. Chained structure of dimeric F1-like ATPase in
605 *Mycoplasma mobile* gliding machinery. *mBio* 12:e0141421.

606 11. Uenoyama A, Kusumoto A, Miyata M. 2004. Identification of a 349-kilodalton
607 protein (Gli349) responsible for cytadherence and glass binding during gliding of
608 *Mycoplasma mobile*. J Bacteriol 186:1537-45.

609 12. Seto S, Uenoyama A, Miyata M. 2005. Identification of a 521-kilodalton protein
610 (Gli521) involved in force generation or force transmission for *Mycoplasma*
611 *mobile* gliding. J Bacteriol 187:3502-10.

612 13. Uenoyama A, Miyata M. 2005. Identification of a 123-kilodalton protein (Gli123)
613 involved in machinery for gliding motility of *Mycoplasma mobile*. J Bacteriol
614 187:5578-84.

615 14. Adan-Kubo J, Uenoyama A, Arata T, Miyata M. 2006. Morphology of isolated
616 Gli349, a leg protein responsible for *Mycoplasma mobile* gliding via glass

617 binding, revealed by rotary shadowing electron microscopy. J Bacteriol
618 188:2821-8.
619 15. Hamaguchi T, Kawakami M, Furukawa H, Miyata M. 2019. Identification of novel
620 protein domain for sialyloligosaccharide binding essential to *Mycoplasma mobile*
621 gliding. FEMS Microbiol Lett 366.
622 16. Mizutani M, Miyata M. 2017. Force measurement on *Mycoplasma mobile* gliding
623 using optical tweezers. Bio Protoc 7:e2127.
624 17. Mizutani M, Tulum I, Kinosita Y, Nishizaka T, Miyata M. 2018. Detailed analyses
625 of stall force generation in *Mycoplasma mobile* gliding. Biophys J
626 114:1411-1419.
627 18. Uenoyama A, Seto S, Nakane D, Miyata M. 2009. Regions on Gli349 and Gli521
628 protein molecules directly involved in movements of *Mycoplasma mobile* gliding
629 machinery, suggested by use of inhibitory antibodies and mutants. J Bacteriol
630 191:1982-5.
631 19. Nonaka T, Adan-Kubo J, Miyata M. 2010. Triskelion structure of the Gli521
632 protein, involved in the gliding mechanism of *Mycoplasma mobile*. J Bacteriol
633 192:636-42.
634 20. Tulum I, Kimura K, Miyata M. 2020. Identification and sequence analyses of the

635 gliding machinery proteins from *Mycoplasma mobile*. Sci Rep 10:3792.

636 21. Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ.

637 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database

638 search programs. Nucleic Acids Res 25:3389-402.

639 22. Mani R, Vorobiev S, Swapna GV, Neely H, Janjua H, Ciccosanti C, Xiao R,

640 Acton TB, Everett JK, Hunt J, Montelione GT. 2011. Solution NMR and X-ray

641 crystal structures of membrane-associated lipoprotein-17 domain reveal a novel

642 fold. J Struct Funct Genomics 12:27-32.

643 23. Wu HN, Kawaguchi C, Nakane D, Miyata M. 2012. "Mycoplasmal antigen

644 modulation," a novel surface variation suggested for a lipoprotein specifically

645 localized on *Mycoplasma mobile*. Curr Microbiol 64:433-40.

646 24. Adan-Kubo J, Yoshii SH, Kono H, Miyata M. 2012. Molecular structure of

647 isolated MvspI, a variable surface protein of the fish pathogen *Mycoplasma*

648 *mobile*. J Bacteriol 194:3050-7.

649 25. Micsonai A, Wien F, Kernya L, Lee YH, Goto Y, Refregiers M, Kardos J. 2015.

650 Accurate secondary structure prediction and fold recognition for circular

651 dichroism spectroscopy. Proc Natl Acad Sci U S A 112:E3095-103.

652 26. Micsonai A, Wien F, Bulyaki E, Kun J, Moussong E, Lee YH, Goto Y, Refregiers

653 M, Kardos J. 2018. BeStSel: a web server for accurate protein secondary
654 structure prediction and fold recognition from the circular dichroism spectra.
655 Nucleic Acids Res 46:W315-W322.
656 27. Morio H, Kasai T, Miyata M. 2016. Gliding direction of *Mycoplasma mobile*. J
657 Bacteriol 198:283-90.
658 28. Trewhella J, Duff AP, Durand D, Gabel F, Guss JM, Hendrickson WA, Hura GL,
659 Jacques DA, Kirby NM, Kwan AH, Perez J, Pollack L, Ryan TM, Sali A,
660 Schneidman-Duhovny D, Schwede T, Svergun DI, Sugiyama M, Tainer JA,
661 Vachette P, Westbrook J, Whitten AE. 2017. 2017 publication guidelines for
662 structural modelling of small-angle scattering data from biomolecules in solution:
663 an update. Acta Crystallogr D Struct Biol 73:710-728.
664 29. Arai M, Kondrashkina E, Kayatekin C, Matthews CR, Iwakura M, Bilsel O. 2007.
665 Microsecond hydrophobic collapse in the folding of Escherichia coli dihydrofolate
666 reductase, an alpha/beta-type protein. J Mol Biol 368:219-29.
667 30. Kataoka M, Nishii I, Fujisawa T, Ueki T, Tokunaga F, Goto Y. 1995. Structural
668 characterization of the molten globule and native states of apomyoglobin by
669 solution X-ray scattering. J Mol Biol 249:215-28.
670 31. Manalastas-Cantos K, Konarev PV, Hajizadeh NR, Kikhney AG, Petoukhov MV,

671 Molodenskiy DS, Panjkovich A, Mertens HDT, Gruzinov A, Borges C, Jeffries
672 CM, Svergun DI, Franke D. 2021. ATSAS 3.0: expanded functionality and new
673 tools for small-angle scattering data analysis. J Appl Crystallogr 54:343-355.
674 32. Arai K, Horikawa Y, Shikata T, Iwase H. 2020. Reconsideration of the
675 conformation of methyl cellulose and hydroxypropyl methyl cellulose ethers in
676 aqueous solution. RSC Adv 10:19059-19066.
677 33. Fukuyama K, Kasahara Y, Kasahara N, Oya A, Nishikawa K. 2001. Small-angle
678 X-ray scattering study of the pore structure of carbon fibers prepared from a
679 polymer blend of phenolic resin and polystyrene. Carbon 39:287-290.
680 34. Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O,
681 Tunyasuvunakool K, Bates R, Zidek A, Potapenko A, Bridgland A, Meyer C, Kohl
682 SAA, Ballard AJ, Cowie A, Romera-Paredes B, Nikolov S, Jain R, Adler J, Back
683 T, Petersen S, Reiman D, Clancy E, Zielinski M, Steinegger M, Pacholska M,
684 Berghammer T, Bodenstein S, Silver D, Vinyals O, Senior AW, Kavukcuoglu K,
685 Kohli P, Hassabis D. 2021. Highly accurate protein structure prediction with
686 AlphaFold. Nature 596:583-589.
687 35. Aluotto BB, Wittler RG, Williams CO, Faber JE. 1970. Standardized
688 bacteriologic techniques for the characterization of *Mycoplasma* species. Int J

689 Syst Bacteriol 20:35-58.

690 36. Goldman A, Harper S, Speicher DW. 2016. Detection of Proteins on Blot

691 Membranes. Curr Protoc Protein Sci 86:10 8 1-10 8 11.

692 37. Kenri T, Kawakita Y, Kudo H, Matsumoto U, Mori S, Furukawa Y, Tahara YO,

693 Shibayama K, Hayashi Y, Arai M, Miyata M. 2019. Production and

694 characterization of recombinant P1 adhesin essential for adhesion, gliding, and

695 antigenic variation in the human pathogenic bacterium, *Mycoplasma*

696 *pneumoniae*. Biochem Biophys Res Commun 508:1050-1055.

697 38. Slabinski L, Jaroszewski L, Rychlewski L, Wilson IA, Lesley SA, Godzik A. 2007.

698 XtalPred: a web server for prediction of protein crystallizability. Bioinformatics

699 23:3403-5.

700 39. Kawakita Y, Kinoshita M, Furukawa Y, Tulum I, Tahara YO, Katayama E, Namba

701 K, Miyata M. 2016. Structural study of MPN387, an essential protein for gliding

702 motility of a human-pathogenic bacterium, *Mycoplasma pneumoniae*. J Bacteriol

703 198:2352-9.

704 40. Mizutani M, Miyata M. 2019. Behaviors and energy source of *Mycoplasma*

705 *gallisepticum* gliding. J Bacteriol 201.

706 41. Shimizu N, Yatabe K, Nagatani Y, Saijyo S, Kosuge T, Igarashi N. 2016.

Software development for analysis of small-angle x-ray scattering data. AIP Conference Proceedings 1741.

42. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE. 2004. UCSF Chimera--a visualization system for exploratory research and analysis. J Comput Chem 25:1605-12.

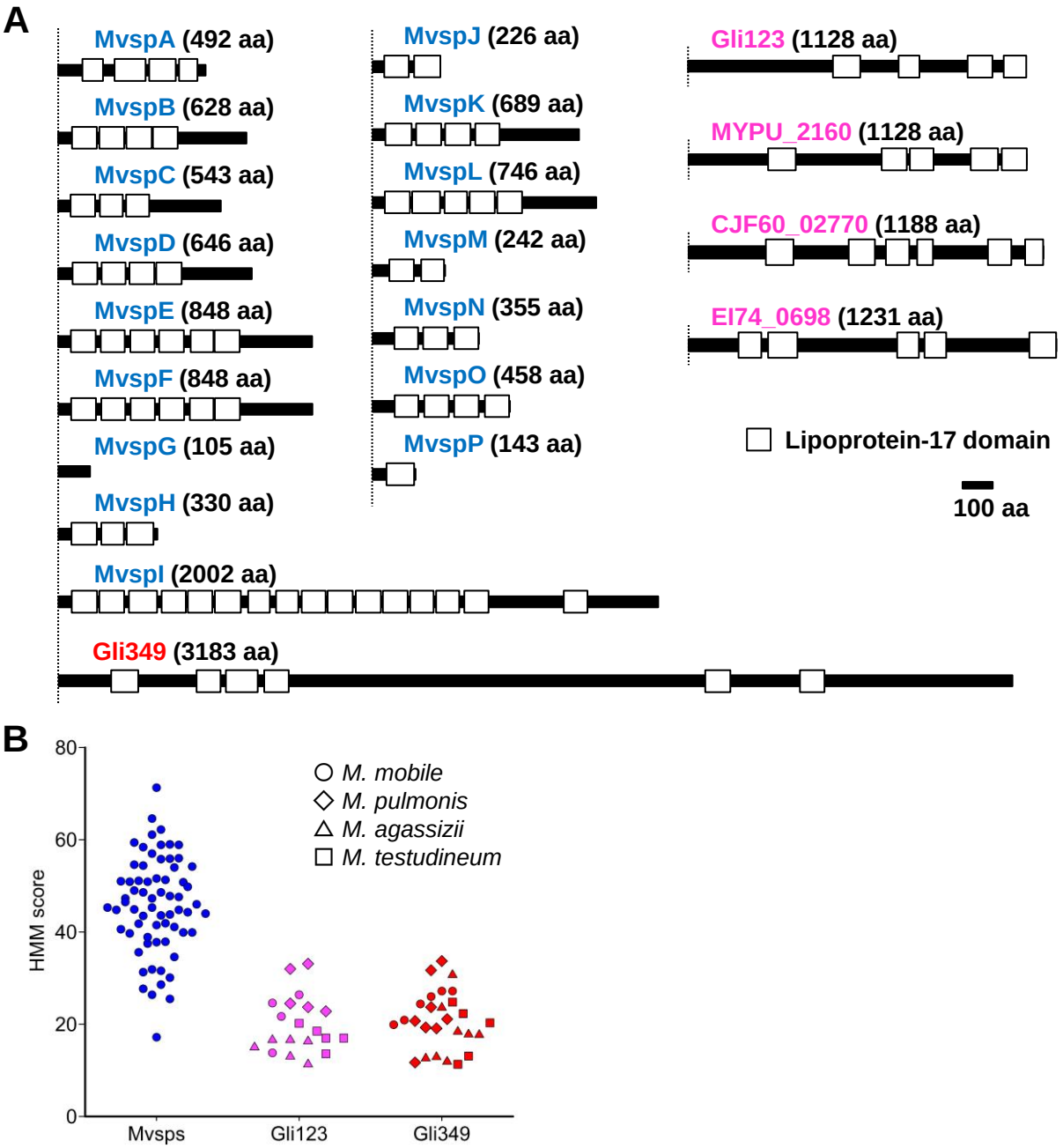

FIG. 1 Lipoprotein-17 domains in Mvsps, Gli123, and Gli349. (A) Distributions of lipoprotein-17 domains in Mvsps, Gli349, and Gli123 sequences have been shown using white boxes. MYPU_2160, CJF60_02770, and EI74_0698 are Gli123 orthologs of *M. pulmonis*, *M. agassizii*, and *M. testudineum*, respectively. (B) Column distribution of the scores calculated using the hidden Markov model (HMM) for each lipoprotein-17 domain. The scores for each lipoprotein-17 domain of Gli123 and Gli349 and orthologs in *M. mobile, M. pulmonis*, *M. agassizii*, and *M. testudineum* and Mvsps in *M. mobile* are shown.

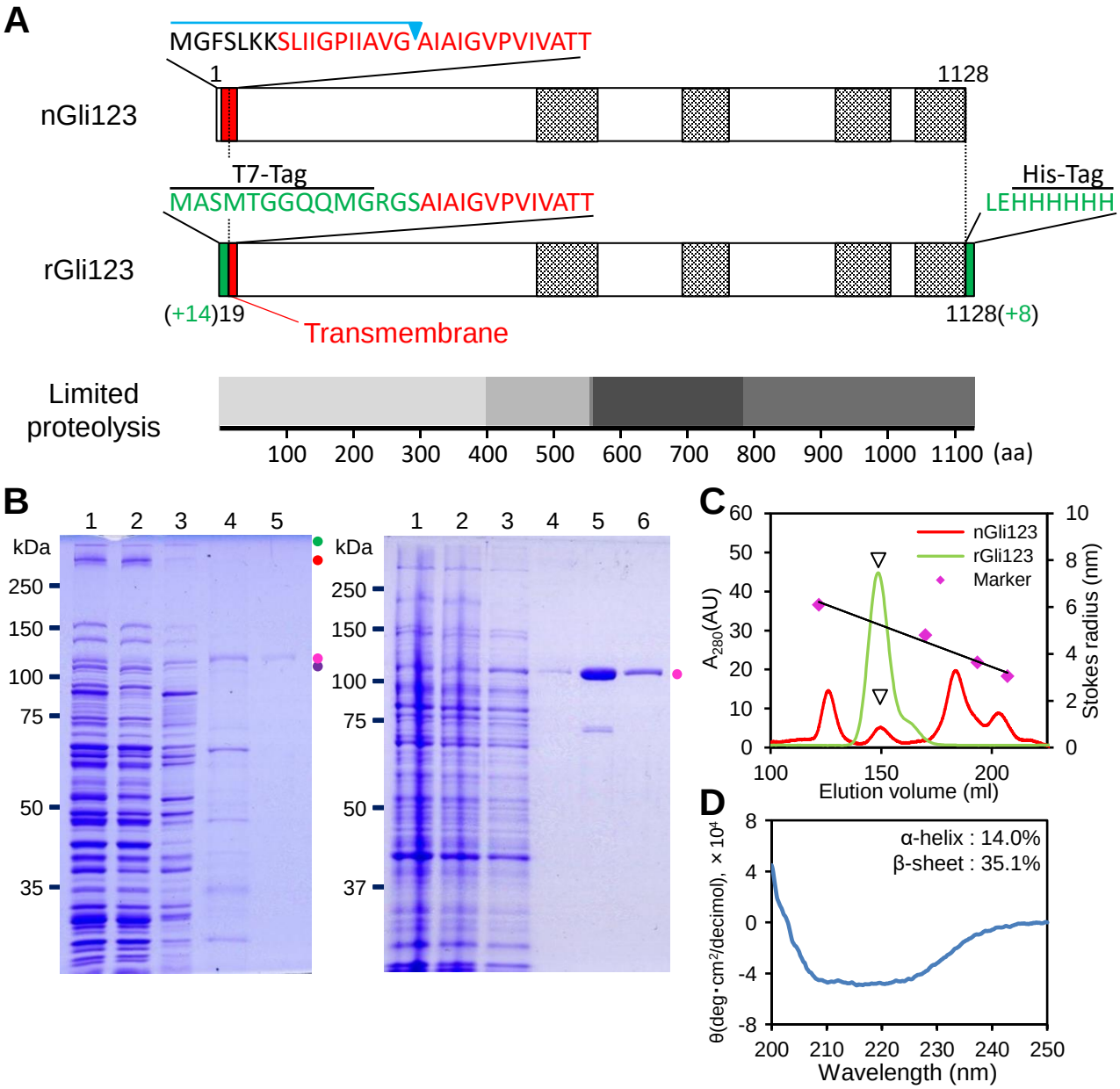

FIG. 2 Amino acid sequence and isolation process of native and recombinant Gli123 proteins (nGli123 and rGli123, respectively). (A) Schematic illustration of the amino acid sequences of nGli123 and rGli123. Transmembrane segment and tagging sequences are marked in red and green, respectively. The blue line and triangle indicate the truncated part of the mature protein. The hatched boxes indicate the lipoprotein-17 domains. The bottom diagram for limited proteolysis shows overlap of degradation products in the digest, suggesting rigid areas, as indicated by the four shades. (B) Left: protein profiles of the fractions obtained in the process of isolation of nGli123 from *M. mobile*. Lane 1, whole-cell lysate; Lane 2, Triton™-insoluble fraction; Lane 3, Triton™-soluble fraction; Lane 4, nGli123 fraction eluted by means of anion-exchange chromatography; Lane 5, nGli123 fraction eluted by means of size-exclusion chromatography (SEC). The protein bands of Gli521, Gli349, MMOB1650, and nGli123 are marked using green, red, purple, and magenta circles, respectively. Right: protein profiles of fractions in the rGli123 purification process. Lane 1, whole-cell lysate; Lanes 2 and 3, insoluble and soluble fractions of the lysate after sonication, respectively; Lane 4, rGli123 fraction after Ni^{2+}-NTA affinity chromatography; Lanes 5, rGli123 fraction after hydrophobic interaction chromatography; Lanes 6, rGli123 fraction after SEC. The protein fractions were applied to SDS-PAGE with a 10% acrylamide gel and stained with Coomassie Brilliant Blue. The molecular masses are indicated on the left. (C) SEC of nGli123 and rGli123. Thyroglobulin (669 kDa, 8.5 nm), ferritin (440 kDa, 6.1 nm), aldolase (158 kDa, 4.8 nm), and ovalbumin (43 kDa, 3.1 nm) were used as the markers for molecular mass and Stokes radius. The triangles mark the peaks of nGli123 and rGli123. (D) CD spectrum of rGli123.

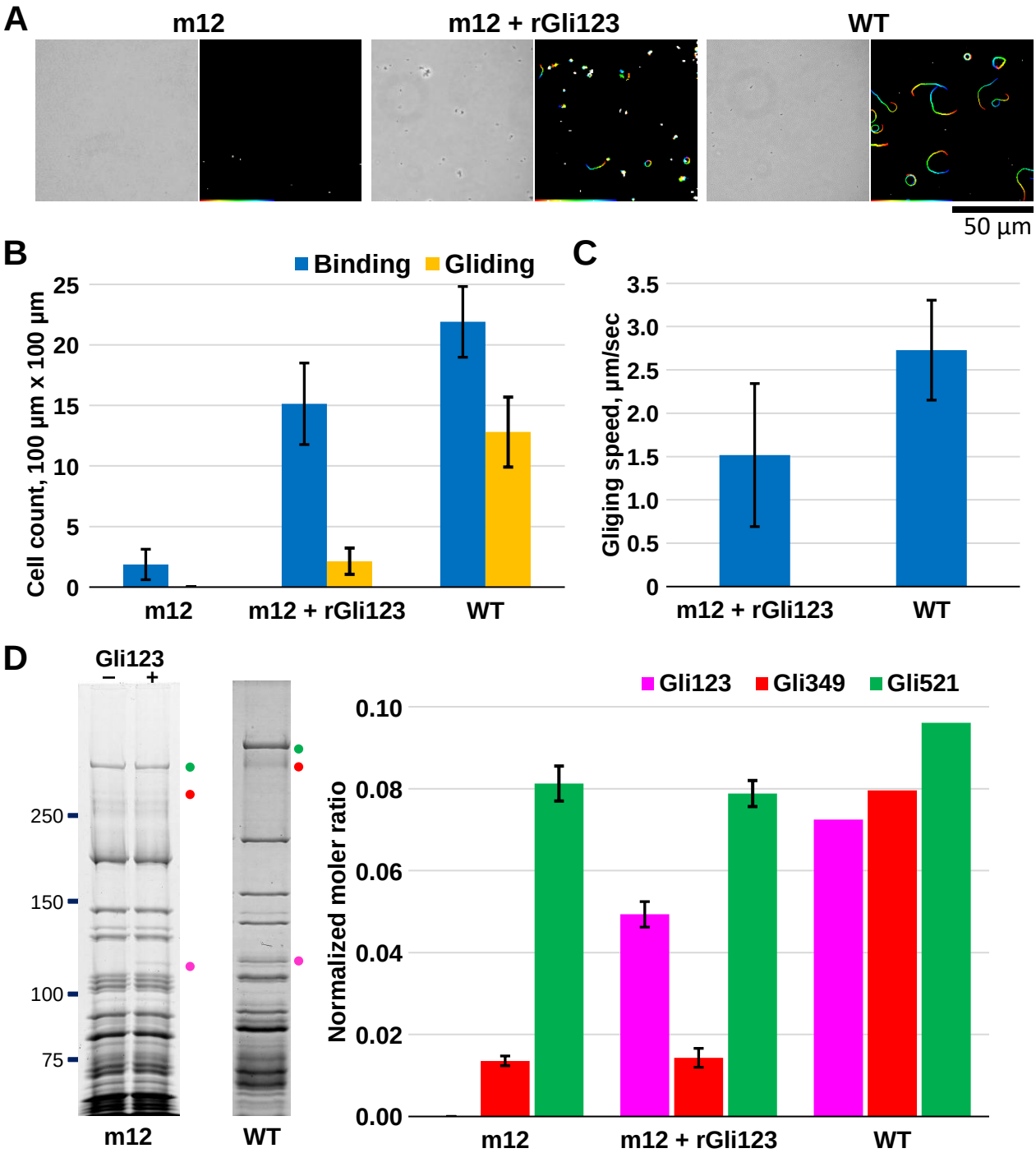

FIG. 3 Recovery of gliding and binding by Gli123 proteins in m12 mutant cells. (A) Cell movements analyzed using optical microscopy. The cells were observed using phase-contrast microscopy (left) and a video was recorded. The video frame (at every 0.03 s) was overlaid as a stack for 10 s, changing its color from red to blue (right). (B) Number of cells binding and gliding on glass surfaces. (C) Gliding speeds. (D) Left: Protein profiles of m12 mutant and m12 wild-type cells, after addition of rGli123 protein. The protein bands for Gli521, Gli349, and Gli123 are marked using green, red, and magenta circles, respectively. Right: Molecular ratios of Gli123, Gli349, and Gli521 were calculated from the band intensities normalized to the value of MvspI.

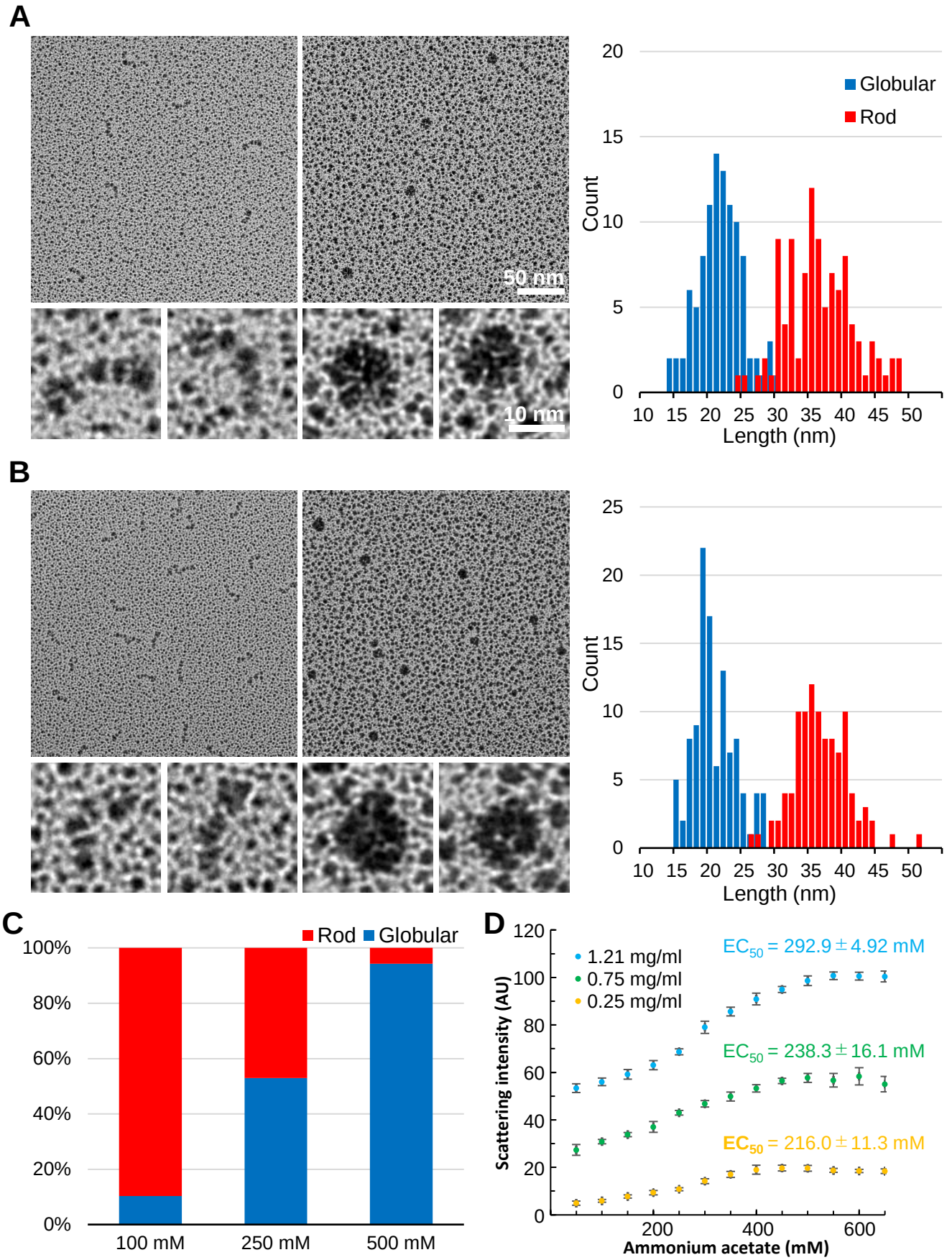

FIG. 4 Ionic strength-induced conformational changes in the nGli123 and rGli123 proteins. Field images of the nGli123 (A) and rGli123 (B) proteins under the conditions of 100 mM (left) and 500 mM (right) ammonium acetate. Scale bar: 50 nm. The representative particle images are magnified in the bottom. Scale bar: 10 nm. The distributions of diameter and length are shown for the globular and rod-like structures on the right. (C) Ratio of globular and rod-like structures of the rGli123 protein under the conditions of 100, 250, and 500 mM ammonium acetate. (D) Light scattering of rGli123 proteins, as a function of ammonium acetate concentration.

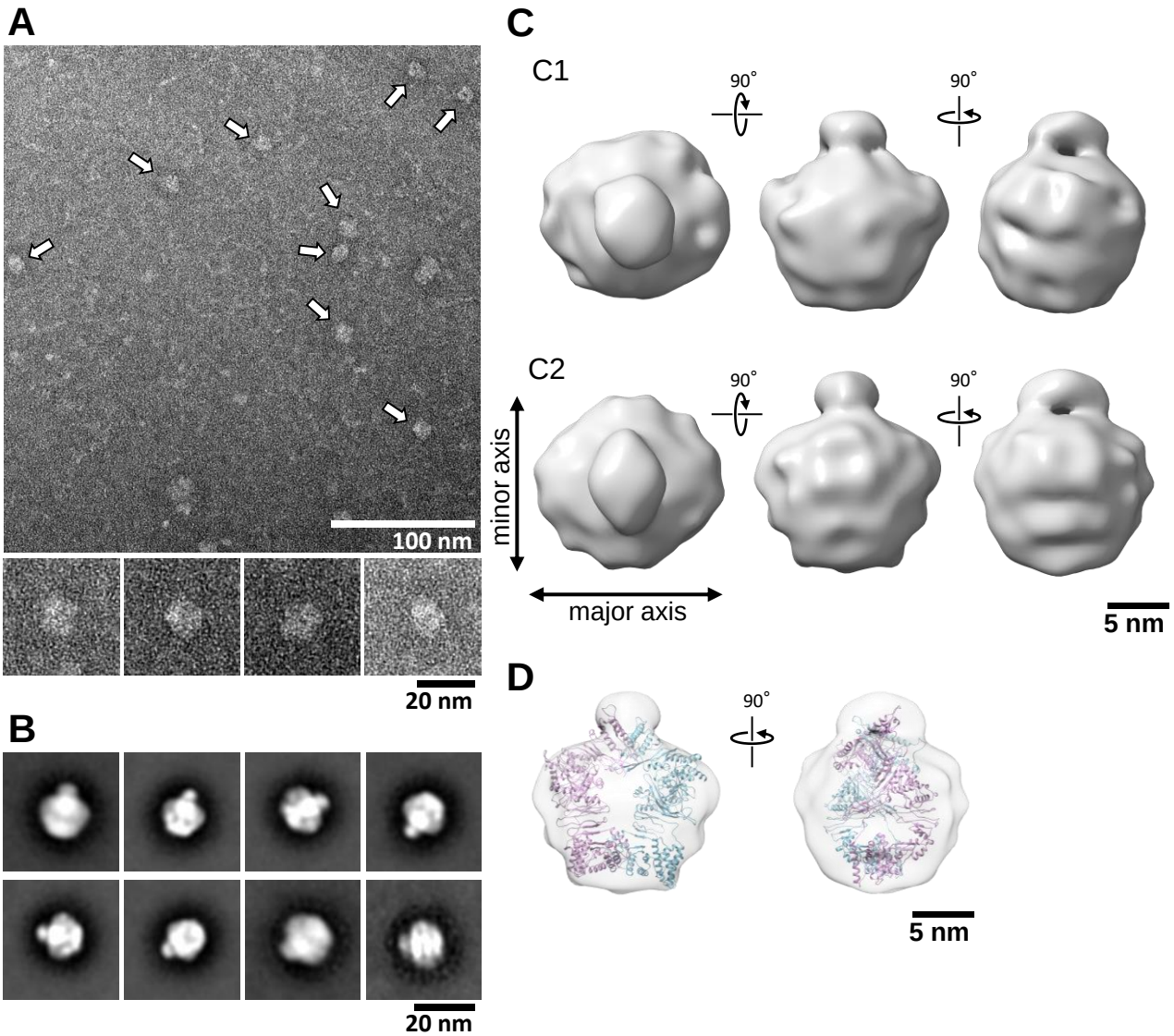

FIG. 5 Globular structure of Gli123 protein, as observed by means of negative-staining EM. (A) Field image of Gli123 molecules that were negatively stained with 2% uranyl acetate. Representative particle images from the fields are shown in the bottom. Single-particle analysis was performed using EMAN2.91 software. (B) Representative averaged images of Gli123 were selected from 50 classes which were obtained by averaging approximately 10,000 particles using the command "2D Analysis". (C) Three-dimensional reconstruction of Gli123 molecule by means of *ab initio* reconstruction with the parameters; C1, non-symmetrical; C2, two-rotational symmetry. (D) Superposition of the C2 model structures obtained from the single-particle analysis and those predicted from the amino acid sequence (1–1000 aa) using Robetta.

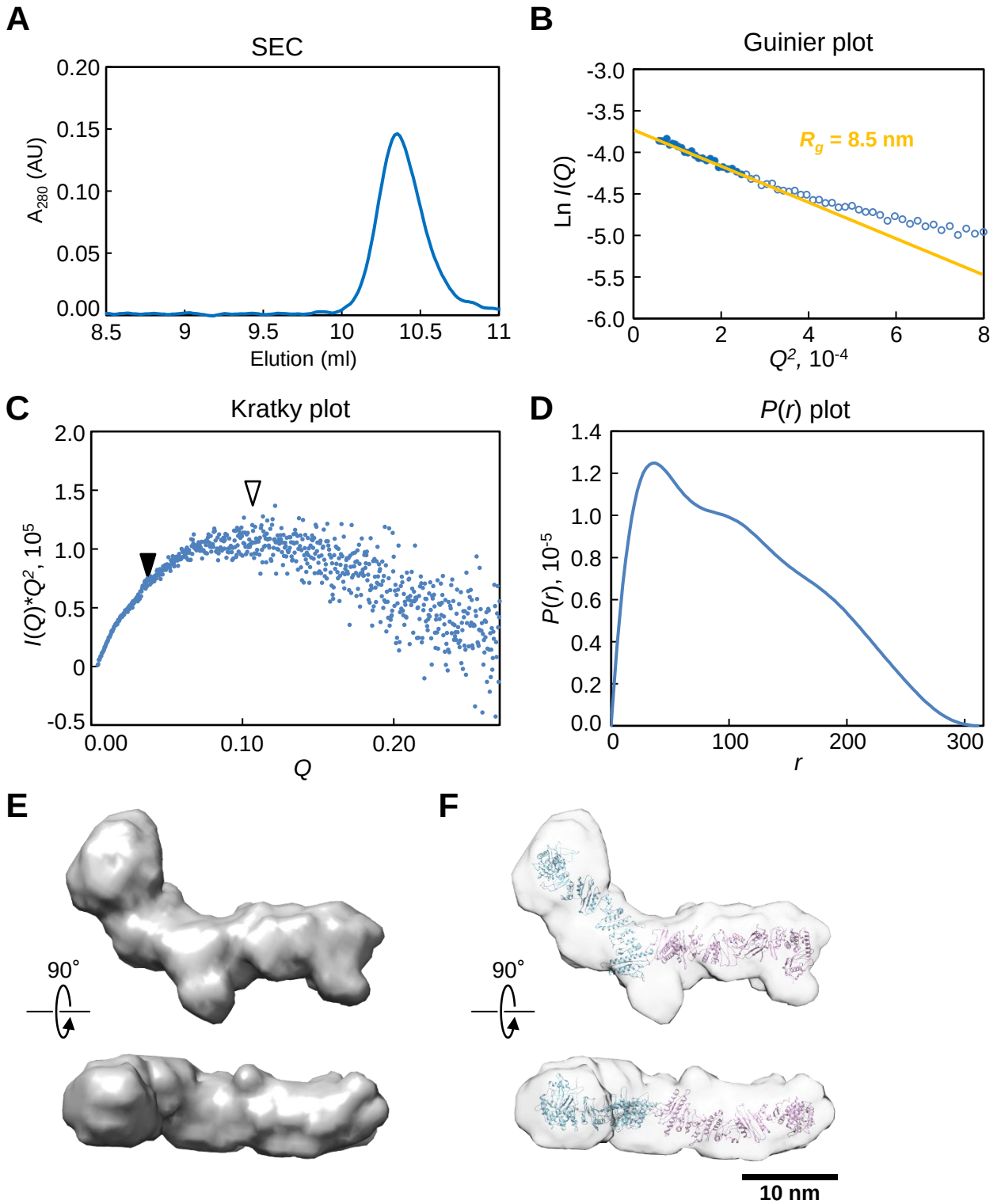

FIG. 6 Rod-like structure of rGli123 protein, as predicted using SEC-SAXS. (A) SEC of rGli123 was carried out by measuring the UV absorbance at the wavelength of 280 nm, in the presence of 100 mM ammonium acetate. (B) Scattering intensities transformed as a Guinier plot [ln I(Q) *versus* Q^2]. Data-points in a Guinier region ($R_gQ < 1.3$), which are shown in the figure using filled circles, were then used to determine the R_g value and zero-angle scattering intensity $I(0)$ with Guinier approximation. (C) Scattering intensities transformed as a Kratky-plot [$I(Q)Q^2$ *versus* Q]. White and black triangles indicate the peak and hidden peak (shoulder), respectively. (D) Pair-distance distribution function $P(r)$ calculated from the scattering curve. (E) SAXS models of rGli123. (F) Overlay of SAXS models and the structures predicted from the amino acid sequence (1–1000 aa) using Robetta.

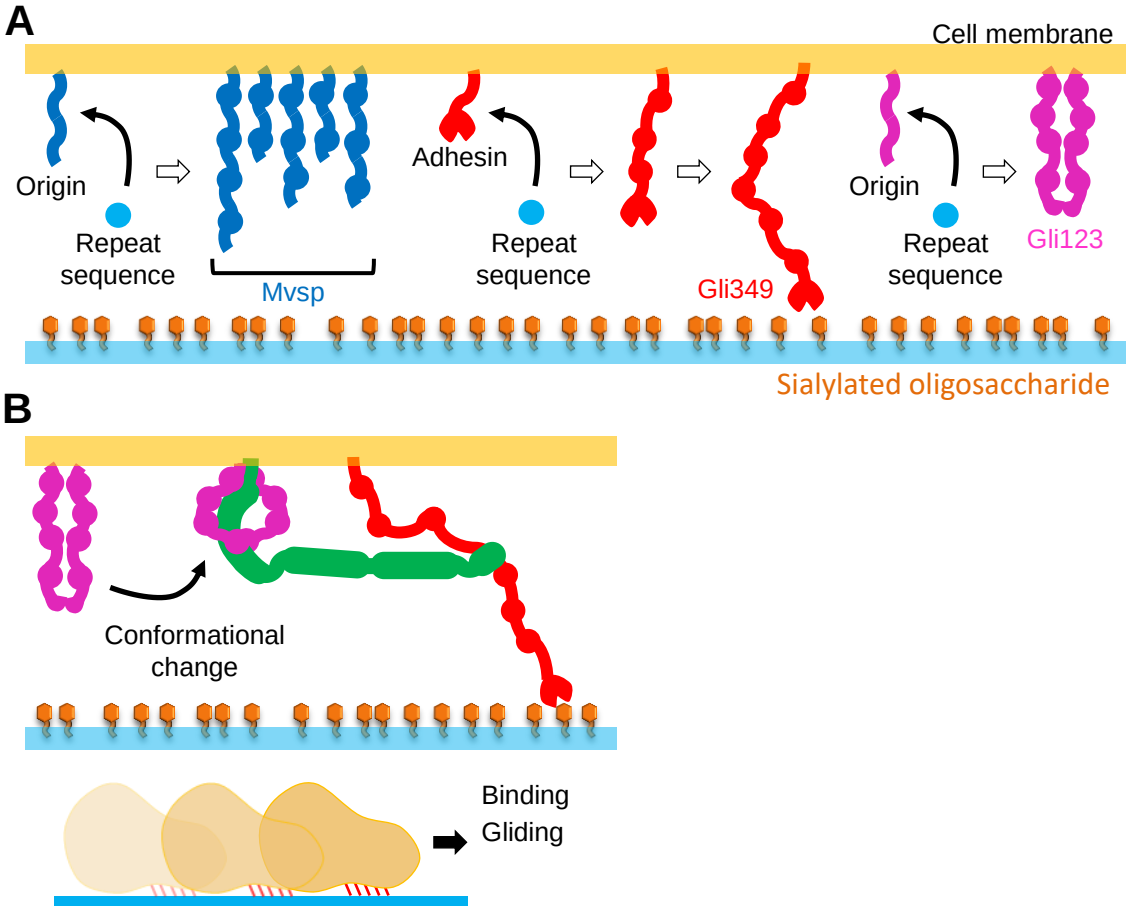

FIG. 7 Role of lipoprotein-17 domain repeats and Gli123. (A) The repeat sequence is possibly used to adjust the length and flexibility of surface molecules that are involved in antigenic modulation (Mvsp) and gliding (Gli349 and Gli123). (B) The flexibility and molecular length of Gli123 allow for conformational changes in it, which facilitate its clamping to other gliding proteins at appropriate positions in the machinery.