An M protein coiled coil unfurls and exposes its hydrophobic core to capture LL-37

Piotr Kolesiński[1], Kuei-Chen Wang[1], Yujiro Hirose[2], Victor Nizet[1], Partho Ghosh[1,*]

[1]Department of Chemistry & Biochemistry, University of California, San Diego, La Jolla, CA 92093,

[2]Division of Host-Microbe Systems and Therapeutics, Department of Pediatrics, University of California, San Diego, La Jolla, CA 92093

*To Whom Correspondence Should be Addressed: pghosh@ucsd.edu

ABSTRACT

Surface-associated, coiled-coil M proteins of *Streptococcus pyogenes* (Strep A) disable human immunity through interaction with select proteins. However, coiled coils lack features typical of protein-protein interaction sites, and it is therefore challenging to understand how M proteins achieve specific binding, for example, with the human antimicrobial peptide LL-37, leading to its neutralization. The crystal structure of a complex of LL-37 with M87 protein, an antigenic M protein variant from a strain that is an emerging threat, revealed a novel interaction mode. The M87 coiled coil unfurled and asymmetrically exposed its hydrophobic core to capture LL-37. A single LL-37 molecule bound M87 in the crystal, but in solution recruited additional LL-37 molecules, consistent with a ‘protein trap’ neutralization mechanism. The interaction mode visualized crystallographically was verified to contribute significantly to LL-37 resistance in an M87 Strep A strain, and was identified to be conserved in a number of other M protein types that are prevalent in human populations. Our results provide specific detail for therapeutic inhibition of LL-37 neutralization by M proteins.

29 **INTRODUCTION**

30 M proteins are the major surface-localized virulence factor of the widespread and
31 potentially deadly bacterial pathogen *Streptococcus pyogenes* (Group A *Streptococcus* or Strep
32 A) (1). One of the primary functions of M proteins is to enable Strep A to evade human innate
33 and adaptive immune responses. This is brought about by interaction of M proteins with select
34 human proteins. M proteins are antigenically sequence variable, with over 220 different types
35 having been identified (2). Despite this variation, the primary sequences of M proteins generally
36 have a propensity to form dimeric, α-helical coiled coils, as verified by direct experimental
37 evidence (3-5). This propensity is easily distinguishable by the presence of heptad repeats (6),
38 in which amino acids in the *a* and *d* positions of the heptad are usually small and hydrophobic,
39 and form the hydrophobic core of the coiled coil. In contrast to the usually complex topography
40 of globular proteins, demarcated by pockets and cavities that enable specific protein-protein
41 interactions, the simple fibrillar structure of M protein coiled coils raises the question of how M
42 proteins achieve specific binding with their human targets.

43 A particular challenge lies in understanding how M proteins specifically bind and thereby
44 neutralize the human antimicrobial peptide LL-37. This is because LL-37 also has a simple
45 topography: LL-37 consists of a short amphipathic α-helix. LL-37 is a member of the cathelicidin
46 antimicrobial peptide family, and constitutes a major host immune defense against Strep A as
47 well as other pathogens (7-9). The 37-amino acid peptide is proteolytically generated from the
48 precursor protein hCAP-18, which is produced by neutrophils, macrophages, mast cells, and
49 keratinocytes along with other epithelial cell types (7, 10, 11). Like other amphipathic α-helical
50 antimicrobial peptides, LL-37 functions by inserting into and lysing bacterial plasma membranes
51 (12-14). Notably, M1 protein was shown to neutralize the antimicrobial activity of LL-37 by
52 sequestering it in a ‘protein trap’ on the Strep A surface and away from its target of action, the
53 bacterial membrane (8, 15). Not only does M1 protein attached to bacterial surface have this
54 neutralization capacity, but so too does M1 protein released in soluble form from the Strep A

55 surface, as occurs during infection (8, 16). M1 protein also binds the LL-37 precursor hCAP-18,
56 and consequently prevents the proteolytic generation of LL-37 (8).
57 To determine the mechanism of LL-37 binding and neutralization by M proteins, we
58 pursued co-crystallization. While M1 protein proved recalcitrant to co-crystallization, several new
59 M protein types that bind and neutralize LL-37 were identified. LL-37 was co-crystallized with
60 M87 protein, which is an M type from strain that is an emerging health threat (17). The structure
61 revealed a remarkable and novel mode of interaction for a coiled coil, in which a portion of the
62 M87 protein coiled coil unfurled and exposed its hydrophobic core for interaction with LL-37.
63 The LL-37-binding motif visualized in M87 protein was identified in a number of other M protein
64 types, many of which are prevalent in human populations. Experimental evidence showed that
65 some of these other M types bound and neutralized LL-37 similarly to M87 protein. Our results
66 provide specific detail for inhibiting the interaction of M proteins with LL-37.
67
68 **RESULTS**
69 **Structure of the M87/LL-37 Complex**
70 To determine the mechanism of binding of LL-37 by M proteins, we pursued co-
71 crystallization beginning with M1 protein but were unable to obtain co-crystals despite various
72 attempts. A number of other M protein constructs (M4, M5, M22, M28, M44, M58, M77, M87,
73 and M89), which were available in the laboratory for other purposes, were tried next. These
74 consisted of the N-terminal 100 amino acids (denoted below with superscripted “N100”) of these
75 M proteins. $M58^{N100}$ and $M87^{N100}$ bound LL-37, as did the $M1^{HB}$ fragment (Figure 1A), while the
76 others did not bind LL-37 above background level (Figure 1 — figure supplement 1). The level
77 of LL-37 binding by $M58^{N100}$, $M87^{N100}$, $M1^{HB}$ was approximately similar. As a point of
78 comparison, the K_D for the M1/LL-37 complex is ~1 μM (9).
79 To our knowledge, M58 and M87 proteins are the first M proteins besides M1 protein
80 identified to bind LL-37. To determine whether this binding was functionally significant, $M58^{N100}$

81 or M87^{N100} was exogenously added to an M1 Strep A strain in which *emm1* had been deleted
82 (*Δemm1*) (15). As previously shown (8), *Δemm1* was sensitive to the antimicrobial action of LL-
83 37, and exogenously added M1^{HB} increased resistance against LL-37, although in this particular
84 experiment the effect was not statistically significant (Figure 1B). Likewise, M58^{N100} or M87^{N100}
85 exogenously added to *Δemm1* increased the LL-37 MIC, and in this case, these effects were
86 statistically significant and M87^{N100} had the greatest effect (Figure 1B).

87 Various fragments of M58 and M87 were tried in co-crystallization trials with LL-37.
88 Success was had with a version of M87 protein (amino acids, aa 68-105) that had a portion of
89 the canonical coiled-coil protein GCN4 (aa 250-278) (8, 15, 18) fused in register to its N-
90 terminus. GCN4 fusion was pursued to stabilize coiled-coil formation in the M87 protein
91 fragment, a technique that has proven successful in the crystallization of a number of coiled-coil
92 proteins (19, 20). The structure of the GCN4-M87/LL-37 complex was determined through
93 molecular replacement to 2.1 Å resolution limit (Table I, Figure 2, Figure 2 — figure supplement
94 1A). The GCN4-M87 fusion protein was verified to bind and neutralize LL-37, as gauged by co-
95 precipitation and MIC assays, respectively (Figures 1B, C). GCN4-M87 bound LL-37 and led to
96 a statistically significant increase in the LL-37 MIC, to an even greater extent than observed for
97 M87^{N100}.

98 The structure revealed a single LL-37 molecule bound to a dimer of GCN4-M87 (Figure
99 2A). Two nearly identical 1:2 complexes occupied the asymmetric unit of the crystal (Figure 2 —
100 figure supplement 1B, RMSD 0.70 Å). Almost the entire length of LL-37, which was
101 predominantly in α-helical conformation, was visible in the crystal structure. GCN4-M87 was
102 likewise in α-helical conformation throughout, but strikingly, only the GCN4 portion formed a
103 coiled coil while the M87 α-helices were unfurled and asymmetrically disposed. This was best
104 appreciated by comparing the structures of the complexed and free form of GCN4-M87. The
105 structure of the latter was determined to 2.4 Å resolution limit (Table I, Figure 2 — figure
106 supplement 1C). Free GCN4-M87 formed a dimeric coiled coil throughout (Figure 2B), indicating

107 that the unfurling of the M87 α-helices was unique to the bound form (Figure 2C). The greatest
108 extent of unfurling occurred at Phe91, in which the distance between Cα atoms was 8.7 Å in the
109 free form but 15.6 Å in the bound form. Only the M87 portion contacted LL-37, and therefore we
110 will refer only to M87 rather than GCN4-M87 hereafter. The two M87 α-helices and the LL-37 α-
111 helix together formed a parallel, three-helix bundle. One of the M87 α-helices made more
112 contact with LL-37 (1016 $Å^2$ total buried surface area, average of the two complexes in the
113 asymmetric unit of the crystal) (Figure 3, dark blue) while the other made less but still
114 substantial contact (491 $Å^2$ average total buried surface area) (Figure 3, cyan); to differentiate
115 between the two M87 α-helices, the one making greater contact will be denoted M87α1 and the
116 other M87α2.
117
118 **Interface dominated by hydrophobic interactions**
119 Overall, hydrophobic interactions dominated the interaction site. Most notably, a number
120 of hydrophobic M87 amino acids that occupied the core *a* or *d* positions and contacted each
121 other in the free form, as is typical of coiled coils, were exposed in the bound form and instead
122 contacted the hydrophobic face of the amphipathic LL-37 α-helix. Nearly five helical turns of LL-
123 37 (Phe5-Val21) and M87 protein (Leu74-Trp92) engaged one another. Near the N-terminus of
124 LL-37, a consecutive pair of phenylalanines, Phe5 and Phe6, were surrounded by hydrophobic
125 *a* and *d* position amino acids of M87 — Leu74 (*a*, usual position in heptad), Ala77 (*d*), and
126 Tyr81 (*a*) from M87α along with Ala77 from M87α2 (Figure 3A). The two LL-37 Phe's and M87
127 Tyr81 formed a series of π-stacks. Two helical turns later in LL-37 was Ile13 which engaged in a
128 ring of isoleucines, with M87 contributing Ile84 (*d*) from each of its helices (Figure 3B). A further
129 helical turn later in LL-37 was Phe17 which packed against a pair of M87 leucines, Leu88 (*a*),
130 one from each of the M87 helices, and π-stacked against M87α2 Phe91 (*d*) (Figure 3C). Lastly,
131 one more helical turn further in LL-37 were Ile20, which was surrounded by M87α1 Phe91 (*d*)
132 and Trp92; Val21, which packed against M87α2 Phe91; and Arg23 whose aliphatic side chain

133 atoms packed against M87α1 Trp92 (Figure 3D). These hydrophobic contacts were
134 supplemented by a sparingly few polar ones, which occurred all within a small polar break in the
135 hydrophobic face of LL-37, near its N-terminus: LL-37 Ser9 and Lys12 formed a hydrogen bond
136 and salt bridge, respectively, with M87α1 Glu85 (Figure 3E), and in one of the two complexes,
137 LL-37 Lys12 also formed a salt bridge with M87α1 Glu89.
138
139 **Importance of Hydrophobic Interactions**
140 The role of the contacts observed structurally were evaluated through site-directed
141 mutagenesis of M87 protein. To ensure that our structural observations were not limited to a
142 fragment of M87 protein, these experiments were carried out with intact M87 protein. Intact His-
143 tagged M87 protein and LL-37 were incubated at 37 °C, and their interaction was determined
144 through a Ni^{2+}-NTA agarose bead co-precipitation assay. Dual alanine-substitutions of M87
145 Tyr81 (*a*) and Ile84 (*d*), which interacted with LL-37 Phe5, Phe6, and Ile13 (Figures 3A and B),
146 significantly decreased LL-37 interaction (Figures 4A and B). M87 Y81A/I84A was verified
147 through circular dichroism (CD) to have a similar structure to wild-type M87 protein at 37 °C
148 (Figure 4 — supplement 1A). Additionally, M87 Y81A/I84A had a temperature-induced unfolding
149 profile similar to that of wild-type M87 protein (Figure 4 — supplement 1B), also as monitored by
150 CD. These results indicated that alanine-substitutions of Tyr81 and Ile84 affected LL-37 binding
151 directly rather than indirectly through compromised structure, stability, or both. Alanine-
152 substitution of M87 Leu88 (*a*) and Phe91 (*d*), which interacted hydrophobically with LL-37
153 Phe17, Ile20, and Val21 (Figures 3C and D), likewise markedly decreased LL-37 interaction
154 (Figures 4A and B). The secondary structure and stability of M87 L88A/F91A resembled that of
155 wild-type M87 protein as well (Figure 4 — figure supplements 1A and B). A substantial amount
156 of the surface area of M87 Trp92 was buried by contact with LL-37, but surprisingly, substitution
157 of this amino acid with alanine increased LL-37 interaction (Figures 3C and D). M87 Trp92 was
158 adjacent to LL-37 Arg23 (Figure 3D), and thus we asked whether Arg-substitution of M87 Trp92

159 would interfere with LL-37 binding. Indeed, M87 W92R was almost entirely deficient in LL-37
160 interaction (Figures 4A and B), while showing no changes in secondary structure or stability
161 (Figure 4 — figure supplements 1A and B).
162 The only M87 amino acid seen to make polar contacts in both complexes in the
163 asymmetric unit of the crystal was Glu85 (Figure 3E). Ala-substitution of M87 Glu85 had little
164 effect on LL-37 binding (Figures 4A and B), suggesting it did not contribute to affinity. To ask
165 whether it instead contributed to specificity, we asked whether Arg-substitution of this amino
166 acid would decrease interaction with LL-37, as M87 Glu85 was adjacent to LL-37 Lys12.
167 Consistent with our structural observations and a role in specificity, M87 E85R had significantly
168 decreased interaction with LL-37 (Figures 4A and B). This decrease was a direct effect, as M87
169 E85R had greater α-helical content than wild-type M87 protein and similar stability (Figure 4 —
170 figure supplements 1A and B).
171 These results validated the structural observations regarding the mode of interaction
172 between M87 protein and LL-37, and indicated the importance of the M87 *a* and *d* heptad
173 position amino acids to the interaction. They further indicated that the polar contact conferred by
174 M87 Glu85 conferred specificity rather than binding affinity.
175
176 **Stoichiometry of Interactions**
177 For an LL-37 neutralization mechanism that involves an M protein trap (8), the 1:2
178 LL-37:M87 stoichiometry was puzzling. LL-37 forms variably sized oligomers in solution (21-23),
179 and thus we undertook solution phase studies of the complex through size-exclusion
180 chromatography (SEC) coupled to multiangle light scattering (MALS). The molecular weight of
181 intact M87 protein alone was 73.6 ± 1.2 kDa (calc. 72.4 kDa) (Figure 4C, black). A 10-fold
182 excess of LL-37 to the intact M87 dimer was added and the sample was applied to SEC-MALS
183 almost immediately after mixing. The complex had a mass of 80.4 ± 0.6 kDa, which
184 corresponded to an M87 protein dimer bound to one or two molecules of LL-37 (calc. 4.5 kDa)

185 (Figure 4C, blue). Notably, after one hour of incubation of LL-37 with M87 protein, the mass of
186 the complex was 97.1 ± 0.6 kDa (Figure 4C, red), which corresponded to an M87 protein dimer
187 bound to five or six molecules of LL-37. Incubation of up to four hours also resulted in the same
188 increased mass, indicating self-limiting growth of the complex (Figure 4 — figure supplement 2).
189 These results suggested that the single LL-37 molecule bound to an M87 dimer could recruit
190 four or five additional LL-37 molecules.
191
192 ***emm87* Confers Resistance to LL-37**
193 We asked whether the mode of interaction visualized through crystallography was
194 applicable to M87 protein in its native conformation on the Strep A surface. An isogenic
195 *Δemm87* strain was constructed, and complemented with a plasmid expressing either wild-type
196 *emm87* or *emm87* containing the E85R substitution, the latter having greatly diminished LL-37
197 binding in solution (Figure 4A). The wild-type M87 strain was resistant to LL-37 (Figure 4D), and
198 indeed had a greater LL-37 MIC than the M1 strain (Figure 1C). Deletion of *emm87* led to
199 significantly increased sensitivity to LL-37 (Figure 4D). Notably, the *Δemm87* strain
200 complemented with wild-type M87 was restored to the LL-37 resistance of the wild-type parental
201 M87 strain, while complemented with M87 E85R remained sensitive to LL-37, similar to the level
202 seen for the uncomplemented *Δemm87* strain. These LL-37 susceptibility results provide
203 physiological validation for our structural observations.
204
205 **Conservation of M87 motif in other M types**
206 The LL-37 binding motif visualized in M87 protein was identified in the sequence variable
207 N-terminal regions of 14 other M protein types (Figure 5A). The motif consisted of two
208 consecutive ideal heptads, that is, with *a* and *d* positions occupied by canonical hydrophobic
209 amino acids (24-26), including a strictly conserved Tyr at the *a* position of the first heptad.
210 These hydrophobic amino acids in M87 protein (Tyr81, Ile84, Leu88, and Phe91) were shown

211 above to be crucial to interaction with LL-37 (Figures 4A and B). Along with these, the motif had
212 a hydrogen bond acceptor at the *e* position of the first heptad, which in M87 protein (Glu85)
213 provided specificity by contacting LL-37 Lys12 (Figure 3). In addition, a positively charged
214 amino acid was excluded from the *e* position of the second heptad, which in M87 protein was
215 Trp92 and proximal to LL-37 Arg23 (Figure 3). Preceding these two ideal heptads was a nearly
216 ideal heptad, with the *a* position occupied by canonical amino acids (i.e., Leu or Tyr) while the *d*
217 position was tolerant of less than ideal amino acids (i.e., Ala77 as in M87, and also Gln) (25,
218 26).

219 Among the M proteins identified to have this motif were M25, M58, and M68 proteins.
220 We showed above that $M58^{N100}$ bound LL-37 (Figure 1A). Similar fragments consisting of the N-
221 terminal 100 amino acids were constructed for M25 and M68. $M68^{N100}$ bound LL-37 as did
222 $M25^{N100}$ (Figure 5B). To test whether the LL-37 binding mode observed for M87 protein was
223 conserved in these M proteins, Arg-substitutions of the equivalent of M87 E85 were constructed.
224 As the E85R substitution had been evaluated only in intact M87 protein, it was introduced into
225 the $M87^{N100}$ fragment, which resulted in significantly decreased LL-37 binding as well (Figures
226 5B and D). Arg-substitution of the equivalent Glu in $M25^{N100}$ (E87) and $M68^{N100}$ (E79) also led to
227 significantly decreased LL-37 binding (Figures 5B and D). While Arg-substitution of $M58^{N100}$
228 (E97) decreased LL-37 binding, this effect was not statistically significant. Modeling suggested
229 that an Arg at this position in M58 was capable of salt bridging with M58 D94, thereby
230 attenuating its LL-37 repulsion. In the case of M25 and M58 proteins, we noticed that the amino
231 acid preceding the strictly conserved Tyr was Gly, which is a helix-breaking amino acid.
232 Therefore, we substituted the Gly in M25 and M58 proteins with Asp, which is the equivalent
233 amino acid in M87 (D80), and found that $M25^{N100}$ G82D and $M58^{N100}$ G92D had significantly
234 higher levels of LL-37 binding as compared to their wild-type counterparts (Figures 5C and D).

235 These results are consistent with M25, M58, and M68 binding LL-37 in the same mode
236 as observed for M87 protein, and support the hypothesis that the M87 protein LL-37 binding

237 motif is conserved in at least fourteen other M proteins, which belong to the E2 or E3 cluster of
238 M proteins (27).
239
240 **DISCUSSION**
241 We sought to understand how M proteins achieve specific binding of LL-37, which is
242 essential to neutralization of this human antimicrobial peptide by the M1 strain of Strep A and
243 likely by other strains as well (7, 8, 15). This question was challenging as the fibrillar structure of
244 M proteins excludes features commonly present in globular proteins that enable specific
245 binding, such as pockets or cavities. Adding to the challenge, the structure of LL-37 lacks
246 complexity, consisting simply of an amphipathic α-helix. We discovered a remarkable mode of
247 interaction of M87 protein with LL-37 through structure determination. The two α-helices of M87
248 protein in the free state formed a coiled coil, with amino acids in the core *a* and *d* positions
249 engaged in 'knobs-into-holes' packing. In the bound state, the two helices did not form a coiled
250 coil, which was instead unfurled with the helices asymmetrically disposed. Intriguingly, analysis
251 of the ideality of 'knobs-into-holes' packing of M87 in the free state showed that A77 and Y81
252 were not ideally packed (28). This suggested that the segment containing A77 and Y81 may
253 initiate unfurling to bind LL-37. Most significantly, in the bound state, hydrophobic amino acids at
254 the M87 *a* and *d* positions were exposed and formed a continuous patch that contacted the
255 hydrophobic face of the LL-37 α-helix. These hydrophobic contacts dominated the interaction
256 interface and were shown to be essential by Ala-substitution mutagenesis. Recently, a coiled-
257 coil dimer that also binds an α-helical partner was reported (29), but in that complex coiled-coil
258 character is maintained, unlike the loss of coiled-coil character observed for the M87/LL-37
259 complex. In addition, while coiled-coil asymmetry coupled to cognate partner binding has been
260 noted in some proteins (30, 31), this has taken the form of a helical stagger with coiled-coil
261 character maintained. To our knowledge, the unfurling and asymmetric exposure of consecutive

262 *a* and *d* position amino acids to form a continuous interaction site is a novel binding mechanism
263 for a coiled-coil protein.

264 The LL-37 binding site in M87 protein was formed by two ideal heptads preceded by a
265 nearly ideal heptad (Ala and Gln being tolerated at the *a* position). M proteins identified or
266 predicted to bind LL-37 in the same mode also had at least three such consecutive repeats, with
267 the first heptad being ideal in most cases (Figures 6A and Figure 6 — figure supplement 1A). In
268 contrast, M proteins identified not to bind LL-37 lacked consecutive ideal or near ideal heptads.
269 Instead ideal heptads occurred as isolated singletons in the midst of non-ideal heptads. Notably,
270 M1 protein has only an isolated ideal repeat (Figure 6A), indicating that there are additional
271 mechanisms for binding LL-37, different from those observed for M87 protein. In general, the
272 occurrence of two or more consecutive ideal heptads is rare in M protein variable regions
273 (Figure 6 — figure supplement 1), whose sequences frequently have coiled-coil destabilizing
274 amino acids at the *a* and *d* positions (e.g., Glu or Lys) (3, 32). In the case of the M1 B-repeats
275 region, non-ideal heptads have been shown to be functionally essential for creating protein
276 dynamics, which are required for binding fibrinogen in a “capture-and-collapse” mechanism (33).
277 Dynamics in the consecutive ideal repeats of M87 and related M proteins are likely to be much
278 lower in magnitude than in the non-ideal M1 B-repeats, but nevertheless, it appears that enough
279 breathing motion exists even in these ideal coiled coils for the infiltration of LL-37 and disruption
280 of coiled-coil structure. The loose ‘knobs-into-holes’ packing of Y81 in the free form of M87
281 along with the conservation of the Tyr in M proteins shown or implicated to bind LL-37 suggests
282 that this may be a general site for initiating unfurling.

283 The single molecule of LL-37 bound to M87 protein acted as a nucleator for the
284 recruitment of additional LL-37 molecules. One or two molecules of LL-37 molecule bound to
285 M87 protein in solution at an initial time point, but grew to five or six molecules over time, as
286 evidenced by SEC-MALS analysis. This is consistent with observations that LL-37 oligomerizes
287 in solution (21-23). Plausible mechanisms for the recruitment of additional LL-37 molecules are

288 suggested by crystal structures of LL-37 alone. In one crystal structure, the polar faces of two
289 LL-37 molecules contact one another in anti-parallel orientation (13), and in another, the
290 hydrophobic faces do likewise (12). Thus, it is possible that the single LL-37 molecule bound
291 through its hydrophobic face to M87 protein is able to recruit a second molecule of LL-37
292 through polar face-polar face interactions. This second LL-37 molecule would then be able to
293 recruit a third molecule of LL-37 through hydrophobic face-hydrophobic face interactions (Figure
294 6B). While this sort of growth is not self-limiting, we observed that the LL-37/M87 complex was
295 self-limiting at five or six molecules of LL-37 per M87 dimer, and thus the specific details of LL-
296 37/LL-37 interactions are likely to differ. The excess of LL-37 bound by M87 protein is consistent
297 with a 'protein trap' neutralization mechanism (8).

298 The physiological relevance of the mode of LL-37 binding by M87 protein was
299 established through deletion and complementation experiments. A deletion of *emm87* resulted
300 in significant sensitivity to LL-37, as seen in the decrease in the LL-37 MIC. Resistance against
301 LL-37 was restored to the *Δemm87* strain by a plasmid encoding wild-type *emm87* but not
302 *emm*87 (E85R). *In vitro* studies using purified proteins showed that M87 E85R does not bind LL-
303 37, and structural studies provided an explanation for this — M87 E85 is positioned next to LL-
304 37 Lys12.

305 M87 Strep A strains are prevalent in human populations and are an emerging cause of
306 human clinical disease (17, 34, 35). The genomes of almost all M87 strain isolates contain a
307 recombination event that increases the expression of Strep A toxin genes (NADase and
308 streptolysin O) (17). Importantly, the same recombination event is present in the M1 strain that
309 caused a global pandemic starting in the 1980's and an M89 strain that is responsible for an
310 ongoing epidemic (36). A few isolates of the M82 strain also carry this recombination event (17).
311 The M82 strain, along with a number of the other M types identified to have the M87 motif, are
312 prevalent in human populations, including M25, M68, M90, and M103 (17, 34, 35). These

313 observations provide motivation for pursuing therapeutic inhibition of LL-37 binding by M
314 proteins, and our results provide specific detail to achieve this end.
315
316 **Materials and methods**
317 **Bacterial strains and culture conditions.** *Streptococcus pyogenes* (Strep A) strains M87
318 20161436 (NCBI SRA accession: SAMN07154152) and its isogenic *Δemm87* strain (37), and
319 M1T1 5448 and its isogenic Δ*emm1* strain (15) were used. *S. pyogenes* was grown as standing
320 cultures in Todd-Hewitt broth in ambient atmosphere at 37 ˚C. *Escherichia coli* was cultured in
321 Lysogeny Broth at 37 ˚C with agitation. For selection and maintenance of strains, antibiotics
322 were added to the medium at the following concentrations: erythromycin 500 µg/mL for *E. coli*
323 and 2 µg/mL for *S. pyogenes*; chloramphenicol, 2 µg/mL for *S. pyogenes*.
324
325 **DNA manipulation.** Coding sequences for M proteins were cloned into a modified pET28b
326 vector (Novagen) that contained sequences encoding an N-terminal His_6-tag followed by a
327 PreScission protease cleavage site (5). Amino acid substitutions and deletions were introduced
328 into pET28b vectors with the QuickChange II Site-Directed mutagenesis kit (Stratagene),
329 according to the manufacturer's directions, and into pM87 E85R (37), which was used for
330 expression in *S. pyogenes*, with the Phusion Site Directed Mutagenesis Kit (ThermoScientific,
331 Waltham, MA, USA). The coding sequence for GCN4 250-278 was subcloned from
332 *Saccharomyces cerevisiae*, and was fused to M87 68-105 through strand overlap extension
333 PCR. An M87 protein-expressing vector (pM87) was constructed by insertion of *emm87* into
334 pDCerm (37).
335
336 **Peptides and Proteins.** LL-37, which was chemically synthesized and lyophilized as a fluoride
337 salt (Genscript, 95% purity), was solubilized at 5 mg/ml in sterile deionized water for MIC assays
338 or 100 mM NaCl, 20 mM HEPES-NaOH, pH 7.5 (HS) for other experiments.

339 Expression and purification of M proteins constructs was carried out as previously
340 described (3, 5), except for the following minor modifications. After PreScission protease
341 digestion, GCN4-M87 was subjected to gel filtration chromatography using a Superdex 200 (GE
342 Healthcare) column equilibrated with HS buffer. For formation of the GCN4-M87/LL-37 complex,
343 GCN4-M87 (2 mg/ml) was mixed with a three-fold molar excess of LL-37 (3 mg/ml), both in HS
344 buffer, and the complex was purified by gel filtration chromatography using a Superdex 200
345 column that had been equilibrated with 100 mM NaCl, 20 mM MES-NaOH, pH 6.5. Intact wild-
346 type and mutant M87 proteins, following Ni^{2+}-NTA agarose bead purification, were subjected to
347 gel filtration chromatography on a Superdex 200 column that had been equilibrated with HS
348 buffer. For CD measurements, the His-tag on intact wild-type and mutant M87 proteins was
349 removed by PreScission protease digestion, and the cleaved product was further purified by
350 reverse Ni^{2+}-NTA chromatography. Protein concentrations of M proteins were determined by
351 measuring A_{280} with the sample in 6 M guanidine hydrochloride, 20 mM Tris, pH 8.5, and using
352 a calculated molar 280 nm extinction coefficient. The concentration of LL-37 and GCN4-M87/LL-
353 37 complex was measured using the Bradford assay (Bio-Rad) with BSA as a standard.
354
355 **Co-precipitation assay.** One and half nmol of His_6-tagged M protein constructs (2-10 µl) were
356 added to 50 µl of Ni^{2+}-NTA agarose beads that had been pre-equilibrated with HS buffer and
357 incubated with gentle agitation for 10 min at RT. Beads were centrifuged (3,000 x *g*, 30 s, RT)
358 and the supernatant was removed. Six nmol of LL-37 in 150 µl of HS buffer was added to the
359 beads and incubated with gentle agitation for 30 min at 37 °C. The beads were washed three
360 times each with 1 ml HS buffer containing 5 mM imidazole, pH 8.0. For the washes, the resin
361 was mixed with the wash solution by gentle agitation, incubated for 1 min at RT and then
362 centrifuged (3,000 x *g*, 30 s, RT). Bound proteins were eluted with 50 µl HS containing 400 mM
363 imidazole, pH 8.0. Protein samples were resolved by SDS-PAGE and stained with InstaBlue
364 (APExBIO). The intensity of gel bands was quantified with ImageJ (38) and confirmed to be

365 within the linear range of detection. As the intensity on each gel depended on the length of time
366 of staining and destaining, a normalization factor was used. For this, the ratio of the LL-37 band
367 intensity to wild-type M protein band intensity from a particular gel lane was determined from
368 three or more independent gels, and the mean of these ratios (mean of LL-37/M_{WT}) was used as
369 the normalization factor. Accordingly, for each data point (i.e., gel lane), the ratio of the LL-37
370 band intensity to M protein (wild-type or mutant) band intensity was quantified, and divided by
371 the normalization factor (mean of LL-37/M_{WT}).
372
373 **Minimal Inhibitory Concentration.** *S. pyogenes* that had been grown overnight were
374 inoculated into Todd-Hewitt broth at 1:100 dilution and grown at 37 °C to an OD_{600} of 0.4. The
375 culture was diluted to an OD_{600} of 0.1, and 5 µl (~10^5 CFU) was mixed into 100 µl of RPMI 1640
376 medium with glutamine, which contained 0, 2, 4, 8, 12, 18, or 32 µM LL-37. In some
377 experiments, the medium also contained 10 µM $M1^{HB}$, $M58^{N100}$, or $M87^{N100}$ protein. *S. pyogenes*
378 were grown in individual wells of a 96-well plate for 24 h at 37 °C. *S. pyogenes* viability was
379 assessed at this time point by the color of the RPMI medium, where yellow indicated bacterial
380 growth and red no bacterial growth. The MIC was defined as the LL-37 concentration at which
381 no growth was detectable at 24 h.
382
383 **Molecular mass determination.** Intact His-tagged M87 protein (2.5 mg/ml) alone or mixed with
384 LL-37 (1.6 mg/ml; 10-fold molar excess over M87 dimer) in HS (100 mM NaCl, 20 mM HEPES-
385 NaOH, pH 7.5) was centrifuged (10 min, 20,000 x *g*, 20 °C) to remove aggregates. Samples
386 (100 µl) were then either immediately applied to a Superdex 200 10/300 column that had been
387 pre-equilibrated in HS, or incubated 1-4 h at RT before application to the column. Samples
388 eluting from the column were monitored with a light scattering detector (DAWN HELEOS II,
389 Wyatt Technology, Santa Barbara, CA, USA) and a differential refractometer (Optilab T-rEX;

390 Wyatt Technology). Data processing and molecular mass calculation were performed with
391 ASTRA software (Wyatt Technology).
392
393 **Crystallization and data collection**. Crystallization trials were carried out at 293 K using the
394 hanging drop vapor diffusion method. The GCN4-M87/LL-37 complex was concentrated by
395 ultrafiltration using a 3,500 MWCO membrane (Millipore; 4,500 x *g*, 30 min, 15 °C) to 8 mg/ml.
396 GCN4-M87 alone was concentrated to 10 mg/ml by ultrafiltration through 3,500 MWCO
397 membrane (Millipore; 4,500 x *g*, 30 min, 4 °C) The complex was brought to RT before
398 introduction into crystallization drops to overcome its low solubility at 4 °C.
399 The GCN4-M87/LL-37 complex (0.8 µl) was mixed in a 1:1 ratio with 5% (v/v)
400 acetonitrile, 0.1 M MES-NaOH, pH 6.5. Microclusters of plates (ca. 20 x 20 x 5 μm^3) that had
401 grown after 2-4 days were crushed, diluted 125-fold with the precipitant solution, and
402 centrifuged (2,000 x *g*, 30 s, RT). The supernatant was collected and used as a seed stock.
403 GCN4-M87/LL-37 (0.9 µl) was mixed with 0.9 µl of 10% (v/v) acetonitrile, 0.1 M MES-NaOH, pH
404 6.5 and 0.2 µl of the seed stock. Clusters of diffraction-sized crystals (50–150 µm in each
405 dimension) which were obtained after 3-7 days were crushed once again and used for a
406 subsequent round of seeding, carried out in the same manner as the first round. These two
407 rounds of seeding yielded single crystals (200 x 200 x 10 μm^3) that were cryo-preserved by
408 three serial transfers to the precipitant solution supplemented with 10, 20, and 30% ethylene
409 glycol, respectively, and flash-cooled in liquid N_2.
410 GCN4-M87 (0.8 µl) was mixed in a 1:1 ratio with 0.6 M monobasic ammonium
411 phosphate, 0.1 M Tris-HCl, pH 8.0,. Single crystals (300 x 150 x 150 μm^3), which were obtained
412 after 4-7 days, were cryo-protected in the precipitant solution supplemented with 30% ethylene
413 glycol and flash cooled in liquid N_2 prior to data collection.
414 Diffraction data were collected at SSRL (beamline 9-2) at 0.979 Å, integrated with
415 Mosflm (GCN4-M87/LL-37) (39) or DIALS (GCN4-M87) (40) and scaled with Aimless (41)

416 (Table I). Because of the highly anisotropic nature of both datasets, the resolution cutoffs for
417 both were determined using anisotropic $CC_{1/2}$.
418
419 **Structure determination and refinement**. Phases for the GCN4-M87/LL-37 complex and free
420 GCN4-M87 were determined by molecular replacement using Phaser (42). The search model
421 was generated from the coiled-coil dimer structure of GCN4 fused to the coiled-coil dimer
422 structure of striated muscle α-tropomyosin (PDB 1KQL) using Sculptor (43) with default settings.
423 Extensive model modification and building were performed with Coot and guided by the
424 inspection of σA-weighted 2mFo-DFc and mFo-DFc omit maps (Figure 2 — figure supplement
425 1).
426 LL-37 in the GCN4-M87/LL-37 complex was manually modelled into well-defined
427 difference electron density that was visible after a few rounds of refinement of the search model.
428 The asymmetric unit of the GCN4-M87/LL-37 crystal contained two heteromeric assemblies,
429 each composed of two chains of GCN4-M87 and one of LL-37. Refinement was performed
430 using Refine from the Phenix suite (44) with default settings. At the final stages of refinement,
431 TLS parameters were applied. TLS groups were applied as follows: chain A; chain B; chain C
432 aa 41-55 and 56-93' chain D aa 38-55 56-104; chain E; and chain F. Side chains with no
433 corresponding electron density were truncated to Cβ. Interaction interfaces between M87 and
434 LL-37 were analyzed with PISA (45).
435 The GCN4-M87 dimer was modeled into continuous electron density. The asymmetric
436 unit of the GCN4-M87 crystal contained a single coiled-coil dimer. Refinement of GCN4-M87
437 was performed using Refine with default settings in addition to the application of two-fold NCS
438 restraints and TLS parameters. Each chain constituted a single TLS group. Three and six amino
439 acids at N- and C-termini, respectively, of chain A and a single N-terminal amino acid of chain B
440 lacked electron density and were not modeled.
441 Molecular figures were generated with PyMol (http://pymol.sourceforge.net).

442 Structures have been deposited in the PDB: 7SAY for GCN4-M87/LL-37 and 7SAF for
443 GCN4-M87.
444
445 **CD spectroscopy.** CD spectra were measured on an Aviv 215 Circular Dichroism
446 Spectrometer using a quartz cell with 1 mm path length. Protein samples were ~0.125-0.250
447 mg/ml in 5 mM sodium phosphate, pH 7.9. Wavelength spectra were recorded in a range of
448 190-260 nm at 37 °C at 0.5 nm intervals with a 0.5 s averaging time per data point. Melting
449 curves were determined at 222 nm between 20-75 °C with 1 °C increments and 30 s
450 equilibration time for each temperature point. Two independent experiments were carried out for
451 each sample, and the data were averaged and presented as a mean molar residue ellipticity.
452
453 **Identification of M87 LL-37-binding motif in other M proteins.** The sequence of the N-
454 terminal 250 amino acids of mature M87 was aligned against the sequence of the N-terminal
455 250 amino acids of the mature form of 179 M proteins using Clustal Omega (46). Alignments
456 were manually curated for the presence of a Tyr (or Phe, although none were found) occupying
457 the position equivalent to M87 Tyr81, the occurrence of hydrophobic amino acids at the *d*, *a,*
458 and *d* positions following the Tyr, and a hydrogen bond acceptor at the *e* position following the
459 Tyr.
460
461 **Acknowledgements**
462 We thank S. Rees for his help with crystallographic data collection. This work was supported by
463 NIH 1R21AI144901 (P.G.).
464
465 **Competing Interest Statement**
466 The authors declare that they have no competing interests.
467

468 **References**

469 1. Ghosh P. Variation, Indispensability, and Masking in the M protein. Trends Microbiol.
470 2018;26(2):132-44. PMC5794653. PubMed 28867148

471 2. McMillan DJ, Dreze PA, Vu T, Bessen DE, Guglielmini J, Steer AC, Carapetis JR, Van
472 Melderen L, Sriprakash KS, Smeesters PR. Updated model of group A *Streptococcus* M
473 proteins based on a comprehensive worldwide study. Clin Microbiol Infect.
474 2013;19(5):E222-9. PubMed 23464795

475 3. McNamara C, Zinkernagel AS, Macheboeuf P, Cunningham MW, Nizet V, Ghosh P.
476 Coiled-coil irregularities and instabilities in group A *Streptococcus* M1 are required for
477 virulence. Science. 2008;319(5868):1405-8. PMC2288698. PubMed 18323455

478 4. Macheboeuf P, Buffalo C, Fu CY, Zinkernagel AS, Cole JN, Johnson JE, Nizet V, Ghosh
479 P. Streptococcal M1 protein constructs a pathological host fibrinogen network. Nature.
480 2011;472(7341):64-8. PMC3268815. PubMed 21475196

481 5. Buffalo CZ, Bahn-Suh AJ, Hirakis SP, Biswas T, Amaro RE, Nizet V, Ghosh P.
482 Conserved patterns hidden within group A Streptococcus M protein hypervariability
483 recognize human C4b-binding protein. Nat Microbiol. 2016;1:16155. PMC5014329.
484 PubMed 27595425

485 6. Manjula BN, Fischetti VA. Tropomyosin-like seven residue periodicity in three
486 immunologically distinct streptococal M proteins and its implications for the
487 antiphagocytic property of the molecule. J Exp Med. 1980;151(3):695-708.
488 PMC2185799. PubMed 6987328

489 7. Nizet V, Ohtake T, Lauth X, Trowbridge J, Rudisill J, Dorschner RA, Pestonjamasp V,
490 Piraino J, Huttner K, Gallo RL. Innate antimicrobial peptide protects the skin from
491 invasive bacterial infection. Nature. 2001;414(6862):454-7. PubMed 11719807

492 8. LaRock CN, Dohrmann S, Todd J, Corriden R, Olson J, Johannssen T, Lepenies B,
493 Gallo RL, Ghosh P, Nizet V. Group A Streptococcal M1 Protein Sequesters Cathelicidin
494 to Evade Innate Immune Killing. Cell Host Microbe. 2015;18(4):471-7. PMC4636435.
495 PubMed 26468750

496 9. LaRock CN, Nizet V. Cationic antimicrobial peptide resistance mechanisms of
497 streptococcal pathogens. Biochim Biophys Acta. 2015;1848(11 Pt B):3047-54.
498 PMC4539303. PubMed 25701232

499 10. Zaiou M, Nizet V, Gallo RL. Antimicrobial and protease inhibitory functions of the human
500 cathelicidin (hCAP18/LL-37) prosequence. J Invest Dermatol. 2003;120(5):810-6.
501 PubMed 12713586

502 11. Wong JH, Ye XJ, Ng TB. Cathelicidins: peptides with antimicrobial, immunomodulatory,
503 anti-inflammatory, angiogenic, anticancer and procancer activities. Curr Protein Pept Sci.
504 2013;14(6):504-14. PubMed 23968350

505 12. Sancho-Vaello E, Francois P, Bonetti EJ, Lilie H, Finger S, Gil-Ortiz F, Gil-Carton D,
506 Zeth K. Structural remodeling and oligomerization of human cathelicidin on membranes
507 suggest fibril-like structures as active species. Sci Rep. 2017;7(1):15371. PMC5684418.
508 PubMed 29133814

509 13. Sancho-Vaello E, Gil-Carton D, Francois P, Bonetti EJ, Kreir M, Pothula KR,
510 Kleinekathofer U, Zeth K. The structure of the antimicrobial human cathelicidin LL-37
511 shows oligomerization and channel formation in the presence of membrane mimics. Sci
512 Rep. 2020;10(1):17356. PMC7562864. PubMed 33060695

513 14. Schneider VA, Coorens M, Ordonez SR, Tjeerdsma-van Bokhoven JL, Posthuma G, van
514 Dijk A, Haagsman HP, Veldhuizen EJ. Imaging the antimicrobial mechanism(s) of
515 cathelicidin-2. Sci Rep. 2016;6:32948. PMC5021996. PubMed 27624595

516 15. Lauth X, von Kockritz-Blickwede M, McNamara CW, Myskowski S, Zinkernagel AS,
517 Beall B, Ghosh P, Gallo RL, Nizet V. M1 protein allows Group A streptococcal survival in
518 phagocyte extracellular traps through cathelicidin inhibition. J Innate Immun.
519 2009;1(3):202-14. PubMed 20375578

520 16. Herwald H, Cramer H, Morgelin M, Russell W, Sollenberg U, Norrby-Teglund A,
521 Flodgaard H, Lindbom L, Bjorck L. M protein, a classical bacterial virulence determinant,
522 forms complexes with fibrinogen that induce vascular leakage. Cell. 2004;116(3):367-79.
523 PubMed 15016372

524 17. Turner CE, Holden MTG, Blane B, Horner C, Peacock SJ, Sriskandan S. The
525 Emergence of Successful Streptococcus pyogenes Lineages through Convergent
526 Pathways of Capsule Loss and Recombination Directing High Toxin Expression. mBio.
527 2019;10(6):e02521-19. PMC6904876. PubMed 31822586

528 18. O'Shea EK, Klemm JD, Kim PS, Alber T. X-ray structure of the GCN4 leucine zipper, a
529 two-stranded, parallel coiled coil. Science. 1991;254(5031):539-44. PubMed 1948029

530 19. De N, Navarro MV, Raghavan RV, Sondermann H. Determinants for the activation and
531 autoinhibition of the diguanylate cyclase response regulator WspR. J Mol Biol.
532 2009;393(3):619-33. PMC2760619. PubMed 19695263

533 20. Li Y, Mui S, Brown JH, Strand J, Reshetnikova L, Tobacman LS, Cohen C. The crystal
534 structure of the C-terminal fragment of striated-muscle alpha-tropomyosin reveals a key
535 troponin T recognition site. Proc Natl Acad Sci U S A. 2002;99(11):7378-83.
536 PMC124239. PubMed 12032291

537 21. Johansson J, Gudmundsson GH, Rottenberg ME, Berndt KD, Agerberth B.
538 Conformation-dependent antibacterial activity of the naturally occurring human peptide
539 LL-37. J Biol Chem. 1998;273(6):3718-24. PubMed 9452503

540 22. Oren Z, Lerman JC, Gudmundsson GH, Agerberth B, Shai Y. Structure and organization
541 of the human antimicrobial peptide LL-37 in phospholipid membranes: relevance to the
542 molecular basis for its non-cell-selective activity. Biochem J. 1999;341 (Pt 3):501-13.
543 PMC1220385. PubMed 10417311

544 23. Xhindoli D, Pacor S, Benincasa M, Scocchi M, Gennaro R, Tossi A. The human
545 cathelicidin LL-37--A pore-forming antibacterial peptide and host-cell modulator. Biochim
546 Biophys Acta. 2016;1858(3):546-66. PubMed 26556394

547 24. Wagschal K, Tripet B, Lavigne P, Mant C, Hodges RS. The role of position a in
548 determining the stability and oligomerization state of alpha-helical coiled coils: 20 amino
549 acid stability coefficients in the hydrophobic core of proteins. Protein Sci.
550 1999;8(11):2312-29. PMC2144206. PubMed 10595534

551 25. Tripet B, Wagschal K, Lavigne P, Mant CT, Hodges RS. Effects of side-chain
552 characteristics on stability and oligomerization state of a de novo-designed model coiled-
553 coil: 20 amino acid substitutions in position "d". J Mol Biol. 2000;300(2):377-402.
554 PubMed 10873472

555 26. Kwok SC, Hodges RS. Stabilizing and destabilizing clusters in the hydrophobic core of
556 long two-stranded alpha-helical coiled-coils. J Biol Chem. 2004;279(20):21576-88.
557 PubMed 15020585

558 27. Sanderson-Smith M, De Oliveira DM, Guglielmini J, McMillan DJ, Vu T, Holien JK,
559 Henningham A, Steer AC, Bessen DE, Dale JB, Curtis N, Beall BW, Walker MJ, Parker
560 MW, Carapetis JR, Van Melderen L, Sriprakash KS, Smeesters PR, Group MPS. A
561 systematic and functional classification of *Streptococcus pyogenes* that serves as a new
562 tool for molecular typing and vaccine development. J Infect Dis. 2014;210(8):1325-38.
563 PubMed 24799598

564 28. Kumar P, Woolfson DN. Socket2: A Program for Locating, Visualising, and Analysing
565 Coiled-coil Interfaces in Protein Structures. Bioinformatics. 2021. PMC8652024. PubMed
566 34498035

567 29. Dimitrova-Paternoga L, Jagtap PKA, Cyrklaff A, Vaishali, Lapouge K, Sehr P, Perez K,
568 Heber S, Low C, Hennig J, Ephrussi A. Molecular basis of mRNA transport by a kinesin-
569 1-atypical tropomyosin complex. Genes Dev. 2021;35(13-14):976-91. PMC8247599.
570 PubMed 34140355

571 30. Sato Y, Fukai S, Ishitani R, Nureki O. Crystal structure of the Sec4p·Sec2p complex in
572 the nucleotide exchanging intermediate state. Proceedings of the National Academy of
573 Sciences. 2007;104(20):8305-10.

574 31. Noell CR, Loh JY, Debler EW, Loftus KM, Cui H, Russ BB, Zhang K, Goyal P, Solmaz
575 SR. Role of Coiled-Coil Registry Shifts in the Activation of Human Bicaudal D2 for
576 Dynein Recruitment upon Cargo Binding. J Phys Chem Lett. 2019;10(15):4362-7.
577 PMC7243283. PubMed 31306018

578 32. Nilson BH, Frick IM, Akesson P, Forsen S, Bjorck L, Akerstrom B, Wikstrom M. Structure
579 and stability of protein H and the M1 protein from Streptococcus pyogenes. Implications
580 for other surface proteins of gram-positive bacteria. Biochemistry. 1995;34(41):13688-
581 98. PubMed 7577960

582 33. Stewart CM, Buffalo CZ, Valderrama JA, Henningham A, Cole JN, Nizet V, Ghosh P.
583 Coiled-coil destabilizing residues in the group A Streptococcus M1 protein are required

584 for functional interaction. Proc Natl Acad Sci U S A. 2016;113(34):9515-20.
585 PMC5003295. PubMed 27512043

586 34. Steer AC, Law I, Matatolu L, Beall BW, Carapetis JR. Global *emm* type distribution of
587 group A *streptococci*: systematic review and implications for vaccine development.
588 Lancet Infect Dis. 2009;9(10):611-6. PubMed 19778763

589 35. Li Y, Rivers J, Mathis S, Li Z, Velusamy S, Nanduri SA, Van Beneden CA, Snippes-
590 Vagnone P, Lynfield R, McGee L, Chochua S, Metcalf BJ, Beall B. Genomic
591 Surveillance of Streptococcus pyogenes Strains Causing Invasive Disease, United
592 States, 2016-2017. Front Microbiol. 2020;11:1547. PMC7396493. PubMed 32849323

593 36. Zhu L, Olsen RJ, Nasser W, Beres SB, Vuopio J, Kristinsson KG, Gottfredsson M, Porter
594 AR, DeLeo FR, Musser JM. A molecular trigger for intercontinental epidemics of group A
595 Streptococcus. J Clin Invest. 2015;125(9):3545-59. PMC4588293. PubMed 26258415

596 37. Hirose Y, Kolesinski P, Hiraoka M, Uchiyama S, Zurich RH, Kumaraswamy M, Bjanes E,
597 Ghosh P, Kawabata S, Nizet V. Contributions of *Stretpococcus pyogenes* M87 protein to
598 innate immune resistance and virulence. Microb Pathog. 2022;In Review.

599 38. Schneider CA, Rasband WS, Eliceiri KW. NIH Image to ImageJ: 25 years of image
600 analysis. Nat Methods. 2012;9(7):671-5. PubMed 22930834

601 39. Battye TGG, Kontogiannis L, Johnson O, Powell HR, Leslie AGW. iMOSFLM: a new
602 graphical interface for diffraction-image processing with MOSFLM. Acta
603 Crystallographica Section D. 2011;67(4):271-81.

604 40. Winter G, Waterman DG, Parkhurst JM, Brewster AS, Gildea RJ, Gerstel M, Fuentes-
605 Montero L, Vollmar M, Michels-Clark T, Young ID, Sauter NK, Evans G. DIALS:
606 implementation and evaluation of a new integration package. Acta Crystallographica
607 Section D. 2018;74(2):85-97.

608 41. Evans PR, Murshudov GN. How good are my data and what is the resolution? Acta
609 Crystallographica Section D. 2013;69(7):1204-14.

610 42. McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ. Phaser
611 crystallographic software. Journal of Applied Crystallography. 2007;40:658-74.

612 43. Bunkoczi G, Read RJ. Improvement of molecular-replacement models with Sculptor.
613 Acta Crystallogr D Biol Crystallogr. 2011;67(Pt 4):303-12. PMC3069745. PubMed
614 21460448

615 44. Afonine PV, Grosse-Kunstleve RW, Echols N, Headd JJ, Moriarty NW, Mustyakimov M,
616 Terwilliger TC, Urzhumtsev A, Zwart PH, Adams PD. Towards automated
617 crystallographic structure refinement with phenix.refine. Acta Crystallographica Section
618 D. 2012;68(4):352-67.

619 45. Krissinel E, Henrick K. Inference of macromolecular assemblies from crystalline state. J
620 Mol Biol. 2007;372(3):774-97. PubMed 17681537

621 46. Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, Lopez R, McWilliam H,
622 Remmert M, Söding J, Thompson JD, Higgins DG. Fast, scalable generation of high-
623 quality protein multiple sequence alignments using Clustal Omega. Mol Syst Biol.
624 2011;7:539. PMC3261699. PubMed 21988835
625

626

Table I. Crystallographic Data Collection and Model Refinement

	GCN4-M87/LL-37	GCN4-M87
Data Collection		
Wavelength (Å)	0.979	0.979
Resolution range (Å)	45.02 - 2.10 (2.18 - 2.10)[a]	43.48 - 2.45 (2.54 - 2.45)
Space group	$P2_1$	$P4_32_12$
Cell dimensions		
a, b, c (Å)	50.2, 57.1, 71.5	97.2, 97.2, 40.7
α, β, γ (°)	90.0, 102.7, 90.0	90.0, 90.0, 90.0
Total reflections	134,344 (9,365)	91,495 (8,197)
Unique reflections	22,092 (1,694)	7,580 (738)
Multiplicity	6.1 (5.5)	12.1 (11.1)
Completeness (%)	94.51 (89.58)	99.75 (100.00)
I/σ(I)	8.50 (2.10)	31.40 (0.78)
Wilson B-factor (Å^2)	26.42	55.23
R_{meas}[b]	0.128 (0.735)	0.404 (4.911)
$CC_{1/2}$[c]	0.996 (0.878)	0.987 (0.556)
Refinement		
Resolution range (Å)	45.02 - 2.10 (2.20 - 2.10)	43.48 - 2.45 (2.80 - 2.45)
No. of reflections (work/test set)	21,958/2,426	7,563/383
R_{work}/R_{free}[d]	0.22/0.27 (0.27/0.34)	0.26/0.27 (0.29/0.38)
No. of non-hydrogen atoms	2673	1138
macromolecules	2556	1126
ligands	16 (ethylene glycol)	2 (PO_4^{3-})
solvent	101	2
R.m.s. deviations		
bonds (Å)	0.003	0.007
angles (°)	0.46	0.95
Ramachandran plot		
favoured (%)	100.00	100.00
outliers (%)	0.00	0.00
Rotamer outliers (%)	0.00	0.00
Clashscore	3.37	3.98
Average B-factor (Å^2)	46.59	76.91
macromolecules	46.80	75.60
ligands	37.52	110.74
solvent	42.81	64.27
Number of TLS groups	8	2
PDB code	7SAY	7SAF

627

[a]Values in parentheses are for the highest resolution shell.

[b]$R_{meas} = \sum_{hkl} \sqrt{\frac{n}{n-1}} \sum_{j=1}^{n} |I_{hkl\,j} - \langle I_{hkl} \rangle| / \sum_{hkl} \sum_{j} I_{hkl\,j}$

[c]$CC_{1/2}$ is the Pearson correlation coefficient between two random half datasets.

[d]$R_{work} = \sum_{hkl} |F_{obs} - F_{calc}| / \sum_{hkl} F_{obs}$

Formulas for R_{work} and R_{free} are identical except 95% of the total number of reflections was used to calculate R_{work} whereas the remaining 5% of reflection was used to calculate R_{free}.

635 **FIGURE LEGENDS**

636 **Figure 1. LL-37 binding and detoxification.**

637 **A.** LL-37 binding to His-tagged $M1^{HB}$, $M58^{N100}$, or $M87^{N100}$ as determined by Ni^{2+}-NTA agarose
638 bead co-precipitation assay at 37 °C. The last lane contains no M protein. Bound fractions were
639 resolved by SDS-PAGE and visualized by Coomassie staining. The gel is representative of at
640 least three experiments. Figure 1 — figure supplement 1 shows LL-37 binding to other M types.

641 **B.** LL-37 MIC for wild-type Strep A *emm1* and isogenic *Δemm1* alone or supplemented with 10
642 μM $M1^{HB}$, $M58^{N100}$, $M87^{N100}$, or GCN4-M87. Data from four independent experiments are
643 presented with means and standard deviations. Statistical significance was calculated using
644 one-way ANOVA with Dunnett post-hoc test. P values are as follows: ns > 0.05, * ≤ 0.05, ** ≤
645 0.01, *** ≤ 0.001 and **** ≤ 0.0001."

646 **C.** LL-37 binding to His-tagged $M87^{N100}$ or GCN-M87, determined as in panel A. The last lane
647 contains no M protein. The gel is representative of at least three experiments.

648

649 **Figure 2. Structures of GCN4-M87/LL-37 complex and free GCN4-M87.**

650 **A.** Cartoon representation of the GCN4-M87/LL-37 complex. LL-37 is in magenta. The GCN4
651 portion of GCN4-M87 is in gray, and the M87 portion in blue for the chain that makes more
652 contacts to LL-37 (M87α1), and cyan for the chain that makes fewer contacts to LL-37 (M87α2).
653 The N- and C-terminal amino acids for LL-37 and M87 are indicated. Figure 2 — figure
654 supplement 1 shows electron density and superposition of the two GCN4-M87/LL-37 complexes
655 in the asymmetric unit.

656 **B.** Cartoon representation of free GCN4-M87. The GCN4 portion is colored gray and the M87
657 portion green.

658 **C.** Superposition (based on GCN4, not depicted) of the M87 portions in bound (blue and cyan)
659 and free (green) states.

660

661 **Figure 3. M87/LL-37 Interface.**

662 LL-37 is in magenta, and the M87 α-helix that forms a greater number of contacts with LL-37 is
663 in blue (M87α1) and the one that forms fewer contacts in cyan (M87α2). Shown are contacts
664 formed with LL-37 (**A**) Phe5 and Phe6; (**B**) Ile 13; (**C**) Phe17; (**D**) Ile20, Val21, and Arg23; and
665 (**E**) Ser9 and Lys12.

666

667 **Figure 4. Evaluation of M87/LL-37 Interactions.**

668 **A.** LL-37 binding by His-tagged, intact wild-type M87 or M87 Y81A/I84A, L88A/F91A, E85A,
669 E85R, W92A, or W92R at 37 °C as determined by Ni^{2+}-NTA agarose bead co-precipitation. The
670 last lane contains no M protein. Bound fractions were resolved by SDS-PAGE and visualized by
671 Coomassie staining. The gel is representative of four experiments. Figure 4 — figure
672 supplement 1 shows CD spectra and melting curves of wild-type and mutant M87 proteins.

673 **B.** Quantification of LL-37 binding to intact wild-type and mutant M87 proteins. The ratio
674 between LL-37 and wild-type M87 protein band intensities was quantified in four independent
675 experiments, and used to determine a mean LL-37/wt-M87 ratio. Points shown are ratios
676 between LL-37 and (wild-type or mutant) M87 protein band intensities, normalized by the mean
677 LL-37/wt-M87 ratio. For each sample, means and SD are shown. Statistical significance was
678 calculated using one-way ANOVA with Dunnett post-hoc test. P values are as follows: ns >
679 0.05, * ≤ 0.05, ** ≤ 0.01, *** ≤ 0.001 and **** ≤ 0.0001.

680 **C.** Molecular weight determination of intact M87 alone (black), M87 and LL-37 added together
681 and immediately analyzed (blue), or M87 and LL-37 incubated for 1 h before being analyzed
682 (red) by size-exclusion chromatography (SEC) coupled to multiangle light scattering (MALS).
683 Dotted lines indicate calculated molecular weights across the profile. Data are representative of
684 three experiments. Figure 4 — figure supplement 2 shows SEC-MALS analysis of M87 protein
685 and LL-37 incubated for 1 or 4 h.

686 **D.** LL-37 MIC for *emm87* and *Δemm87* alone or expressing wild-type M87 (pM87) or M87 E85R
687 pM87(E85R) from a plasmid. Data from eight independent experiments are presented as mean
688 ± SD. The statistical test and P values are the same as described for panel B.
689
690 **Figure 5. Conservation of the M87LL-37-binding motif.**
691 **A.** M87 amino acids that make hydrophobic contact to LL-37 are boxed in red and those that
692 disrupt interaction with LL-37 through Arg-substitution are in blue boxes, as are amino acids that
693 are similarly positioned in other M proteins. The arrow indicates the position of M87 D80.
694 **B.** LL-37 interaction with His-tagged wild-type or mutant $M87^{N100}$, $M25^{N100}$, $M58^{N100}$, $M68^{N100}$
695 proteins determined by Ni^{2+}-NTA agarose bead co-precipitation at 37 °C. Bound fractions were
696 resolved by SDS-PAGE and visualized by Coomassie staining. Data are representative of six
697 experiments.
698 **C.** LL-37 interaction with His-tagged $M25^{N100}$, $M25^{N10}$ G82D, $M58^{N100}$, and $M58^{N100}$ G02D,
699 determined as in panel B. Data are representative of six experiments.
700 **D.** Quantification of the effect of mutagenesis on LL-37 binding. The ratio between LL-37 and
701 wild-type M protein (i.e., M25, M58, M68, or M87) band intensities was quantified in six
702 independent experiments, and used to determine a mean LL-37/wt-M ratio. Points shown are
703 ratios between LL-37 and mutant M protein band intensities, normalized by the corresponding
704 mean LL-37/wt-M ratio. For each sample, means and SD are shown. Statistical significance was
705 calculated using Student's t-test to compare mutant and the corresponding wild-type protein. P
706 values are as follows: ns > 0.05, * ≤ 0.05, ** ≤ 0.01, *** ≤ 0.001 and **** ≤ 0.0001.
707
708 **Figure 6. LL-37 binding motif and higher-order complex formation.**
709 **A.** Occurrence of ideal heptads (dark grey boxes; Ile, Leu, Met, Phe, Tyr, or Val at both *a* and *d*
710 positions) and near ideal heptads (light grey boxes; ideal amino acid at *a* position and Ala at *d*
711 position, with M25 being an exception with Gln at this position) within the N-terminal regions of

712 various M protein types. Within these boxes, yellow circles indicate Tyr at an *a* position; green
713 circles Asp, Asn, Glu, or Gln at an *e* position; and red circles Arg or Lys at an *e* position. Yellow
714 and red circles favor LL-37 binding, while the red circles disfavor LL-37 binding. Figure 6 —
715 figure supplement 1 shows the occurrence of ideal and near ideal heptads within the N-terminal
716 regions of additional M types.
717 **B.** Model of higher-order M87/LL-37 complex. M87α1 and M87α2 are in blue and cyan,
718 respectively, and LL-37 is in magenta. Left, LL-37 bound to M87 protein has its polar face free
719 to form polar contacts with a second LL-37 molecule (light gray, as seen in a structure of LL-37
720 alone, PDB 7PDC). Right, the light gray LL-37 has its hydrophobic face free to form hydrophobic
721 contacts with yet another LL-37 molecule (dark gray, as seen in another structure of LL-37
722 alone, PDB 5NNM). The light and dark gray LL-37 molecules do not contact M87 protein.
723

724 **FIGURE SUPPLEMENT LEGENDS**

725 **Figure 1 — figure supplement 1. LL-37 Binding.**

726 LL-37 binding to His-tagged $M1^{HB}$, $M4^{N100}$, $M5^{N100}$, $M22^{N100}$, $M28^{N100}$, $M44^{N100}$, $M77^{N100}$, or

727 $M89^{N100}$ as determined by a Ni^{2+}-NTA agarose bead co-precipitation assay at 37 °C. The last

728 lane contains no M protein. Bound fractions were resolved by SDS-PAGE and visualized by

729 Coomassie staining. The gel is representative of three experiments.

730

731 **Figure 2 — figure supplement 1. Structures of the GCN4-M87/LL-37 complex and free**

732 **GCN4-M87.**

733 **A.** Electron density for GCN4-M87 from a simulated annealing σA-weighted 2mFo-DFc

734 composite omit map contoured at 1.4σ. Part of the interface between M87 (left) and LL-37

735 (right) is shown.

736 **B.** Superposition of the two GCN4-M87/LL-37 complexes present in the asymmetric unit of the

737 crystal. LL-37 is in different shades of magenta. The GCN4 portions are in shades of gray, and

738 the M87 portions are in blue and cyan.

739 **C.** Electron density for free GCN-M87 from a simulated annealing σA-weighted 2mFo-DFc

740 composite omit map contoured at 1.4σ.

741

742 **Figure 4 — figure supplement 1. Structure and Stability of M87 mutant proteins.**

743 **A.** CD spectra of intact wild-type and mutant M87 proteins at 37 °C. The data are an average of

744 two independent measurements.

745 **B.** Melting curves of intact wild-type and mutant M87 proteins, as monitored by the CD mean

746 molar residue ellipticity at 222 nm. The data are an average of two independent measurements.

747

748 **Figure 4 — figure supplement 2. Self-limiting growth of the M87/LL-37 complex.**

749 Molecular weights of intact M87 protein and LL-37 added together and incubated for 1 h (black)
750 or 4 h (blue) were determined by size-exclusion chromatography coupled to multiangle light
751 scattering. Dotted lines indicate calculated molecular weights across the profile.

752

753 **Figure 6 — figure supplement 1. LL-37 binding motif.**

754 Depiction as in Fig. 6A. The first grouping contains M proteins that have the LL-37 binding motif
755 identified in M87 protein (experimentally verified for M87, M25, M58, and M68, and predicted for
756 M79, M82, M90, M96, M103, M107, M144, M168, M180, M209, and M231). M1 protein, which
757 binds LL-37 but does not have the motif identified in M87 protein, is next, followed by M proteins
758 that have been shown experimentally to not bind LL-37. Lastly, eleven M proteins randomly
759 selected from the remaining pool of M proteins are shown (M12, M15, M18, M42, M63, M71,
760 M73, M101, M114, M176, and M178).

Figure 1

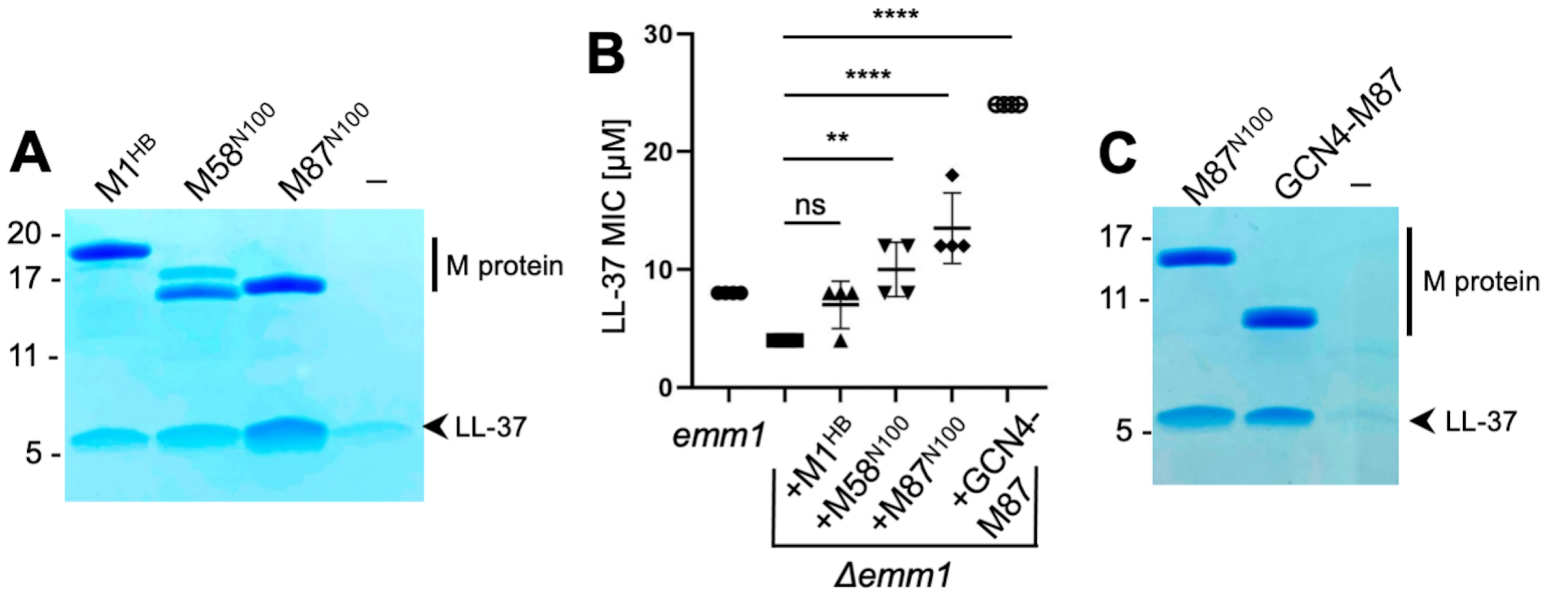

A
M1HB
M58N100
M87N100
–
20 -
17 -
11 -
5 -
M protein
LL-37
B
LL-37 MIC [μM]
30
20
10
0
****
****
**
ns
emm1
+M1HB
+M58N100
+M87N100
+GCN4-M87
Δemm1
C
M87N100
GCN4-M87
–
17 -
11 -
5 -
M protein
LL-37

Figure 2

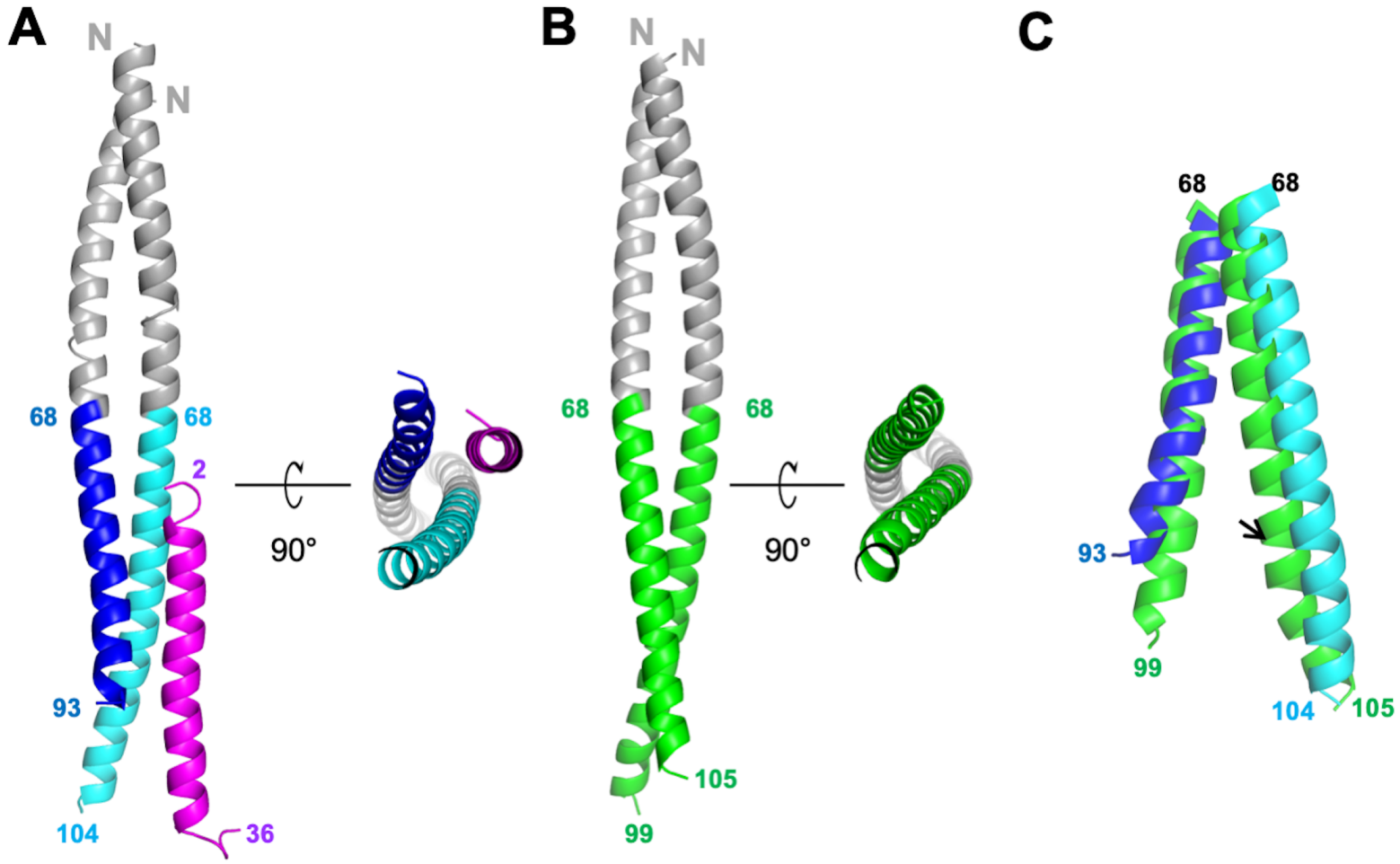


Figure 3

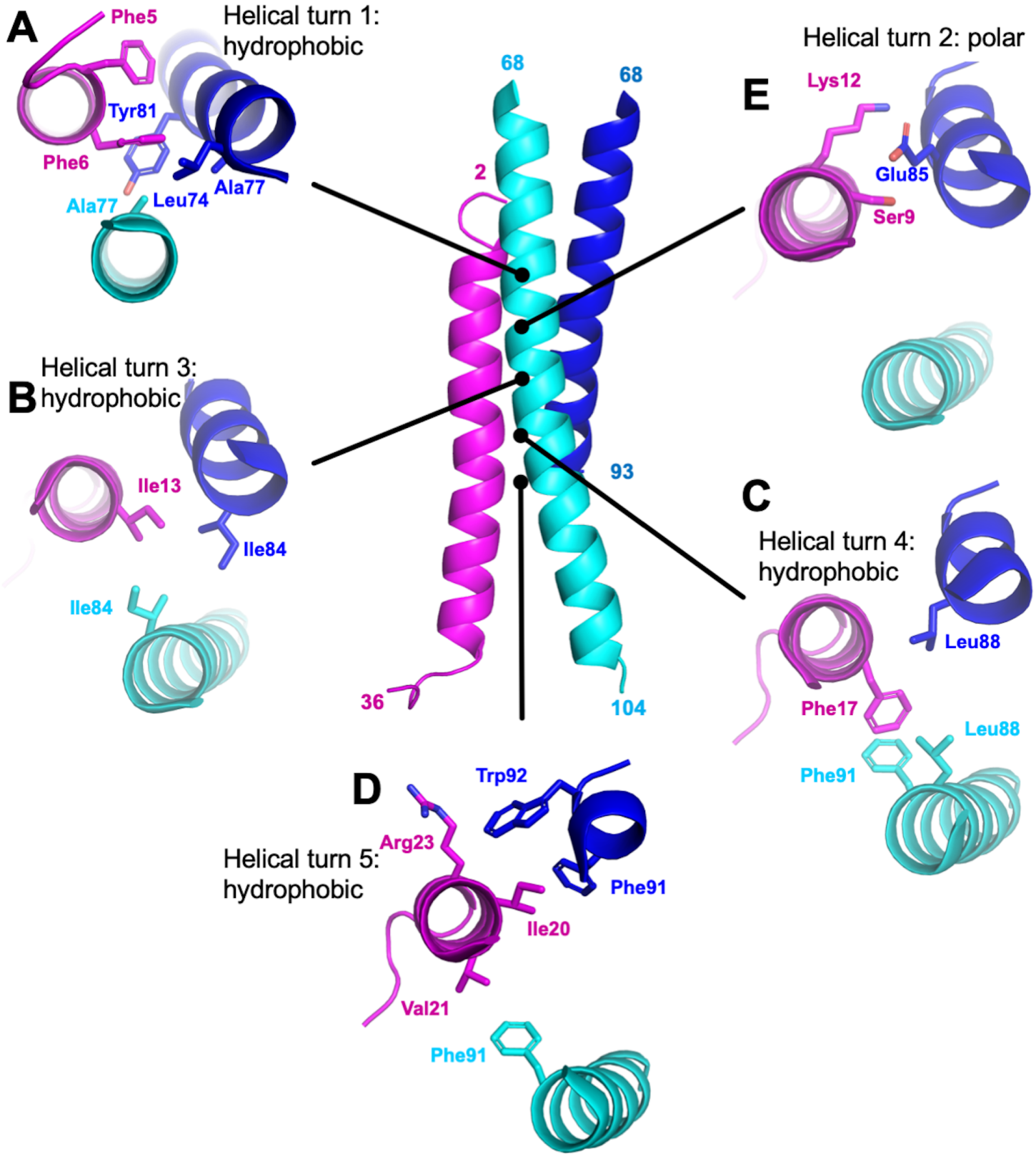

A
Phe5
Helical turn 1:
hydrophobic
Tyr81
Phe6
Ala77
Leu74
Ala77
68
68
2
E
Helical turn 2: polar
Lys12
Glu85
Ser9
B
Helical turn 3:
hydrophobic
Ile13
Ile84
Ile84
93
C
Helical turn 4:
hydrophobic
Leu88
Phe17
Leu88
Phe91
36
104
D
Trp92
Arg23
Helical turn 5:
hydrophobic
Phe91
Ile20
Val21
Phe91

Figure 4

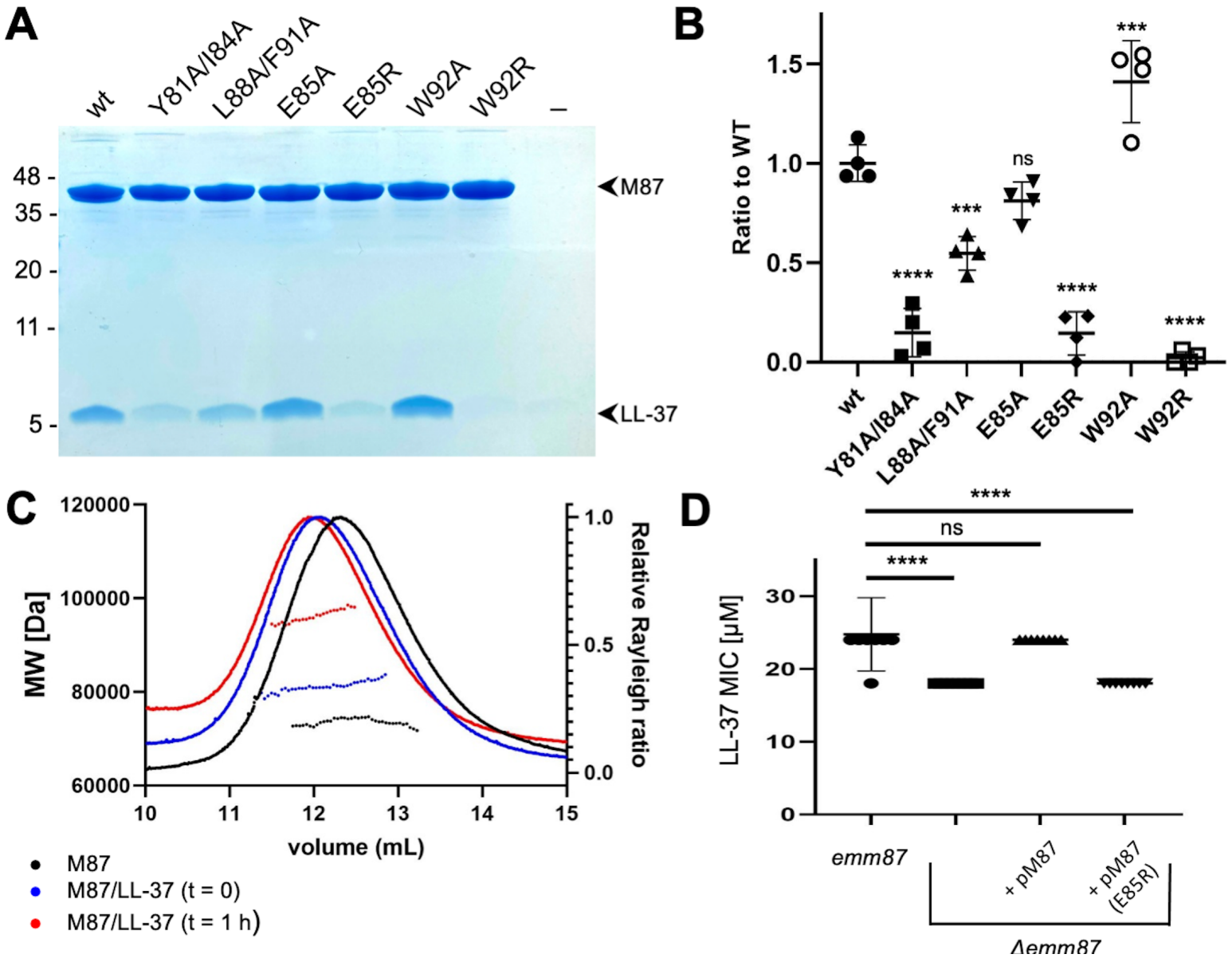
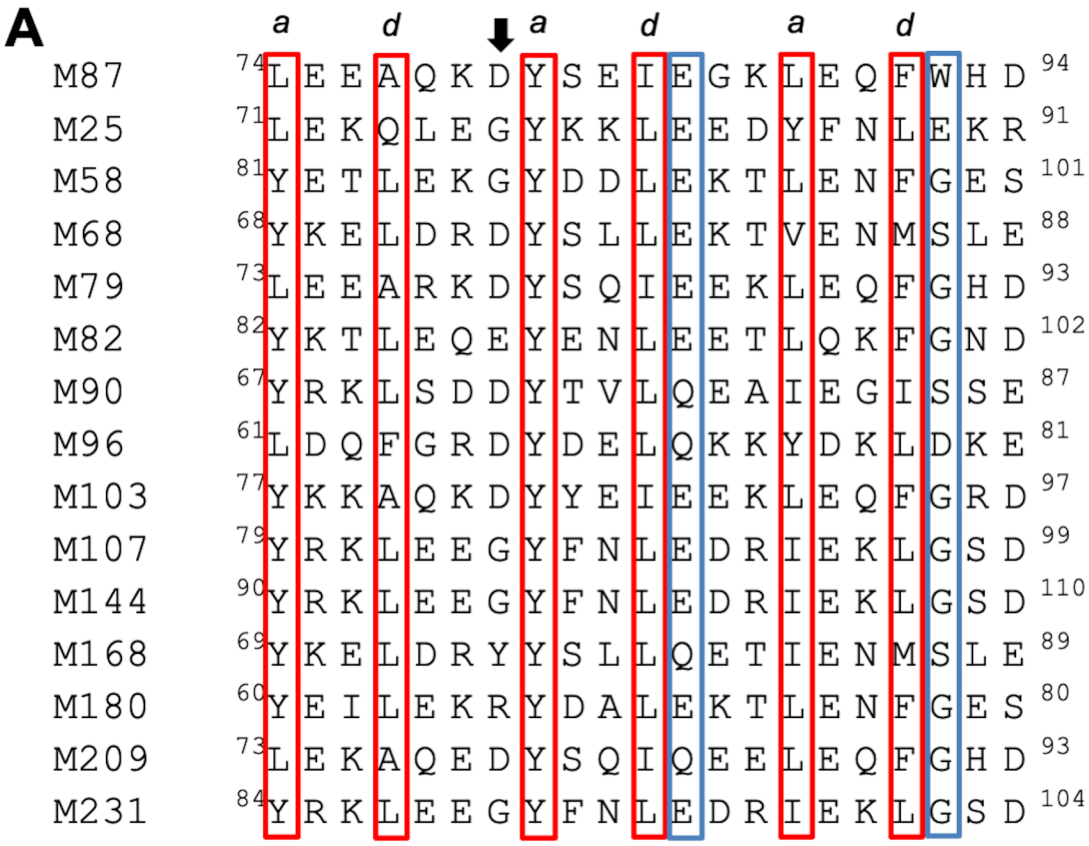
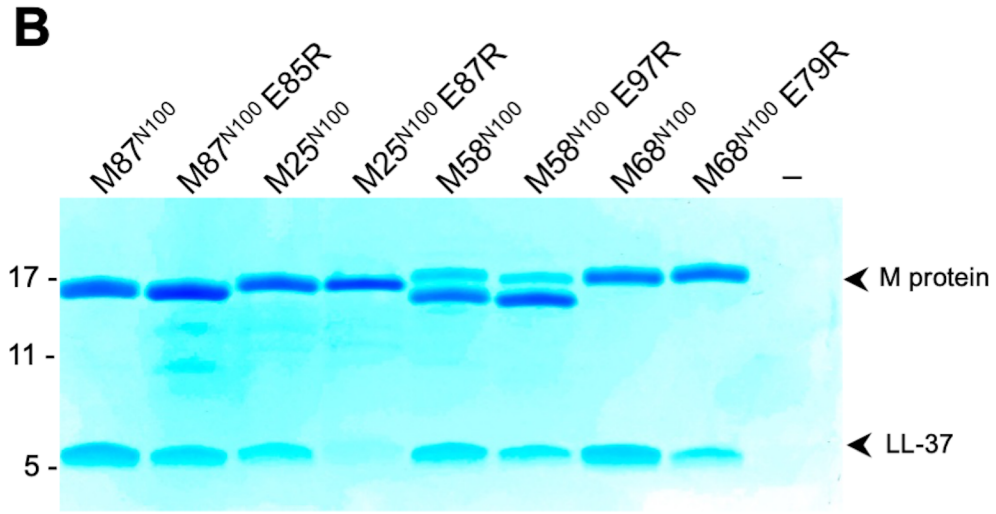
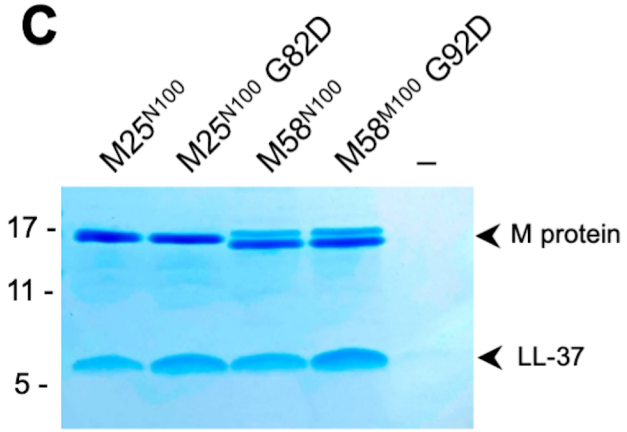

A
wt
Y81A/I84A
L88A/F91A
E85A
E85R
W92A
W92R
–
48
35
20
11
5
M87
LL-37
B
Ratio to WT
1.5
1.0
0.5
0.0
***
ns
***
****
****
****
wt
Y81A/I84A
L88A/F91A
E85A
E85R
W92A
W92R
C
MW [Da]
120000
100000
80000
60000
Relative Rayleigh ratio
1.0
0.5
0.0
10
11
12
13
14
15
volume (mL)
M87
M87/LL-37 (t = 0)
M87/LL-37 (t = 1 h)
D
LL-37 MIC [µM]
30
20
10
0
****
ns
****
emm87
+ pM87
+ pM87 (E85R)
Δemm87

Figure 5


A
a d a d a d
M87 74 L E E A Q K D Y S E I E G K L E Q F W H D 94
M25 71 L E K Q L E G Y K K L E E D Y F N L E K R 91
M58 81 Y E T L E K G Y D D L E K T L E N F G E S 101
M68 68 Y K E L D R D Y S L L E K T V E N M S L E 88
M79 73 L E E A R K D Y S Q I E E K L E Q F G H D 93
M82 82 Y K T L E Q E Y E N L E E T L Q K F G N D 102
M90 67 Y R K L S D D Y T V L Q E A I E G I S S E 87
M96 61 L D Q F G R D Y D E L Q K K Y D K L D K E 81
M103 77 Y K K A Q K D Y Y E I E E K L E Q F G R D 97
M107 79 Y R K L E E G Y F N L E D R I E K L G S D 99
M144 90 Y R K L E E G Y F N L E D R I E K L G S D 110
M168 69 Y K E L D R Y Y S L L Q E T I E N M S L E 89
M180 60 Y E I L E K R Y D A L E K T L E N F G E S 80
M209 73 L E K A Q E D Y S Q I Q E E L E Q F G H D 93
M231 84 Y R K L E E G Y F N L E D R I E K L G S D 104


B
M87N100
M87N100 E85R
M25N100
M25N100 E87R
M58N100
M58N100 E97R
M68N100
M68N100 E79R
–
17 -
11 -
5 -
M protein
LL-37


C
M25N100
M25N100 G82D
M58N100
M58M100 G92D
–
17 -
11 -
5 -
M protein
LL-37

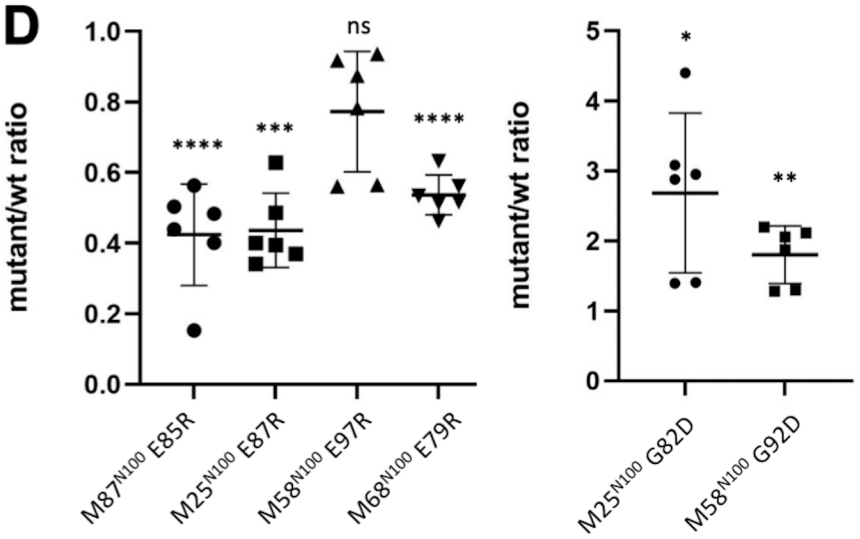
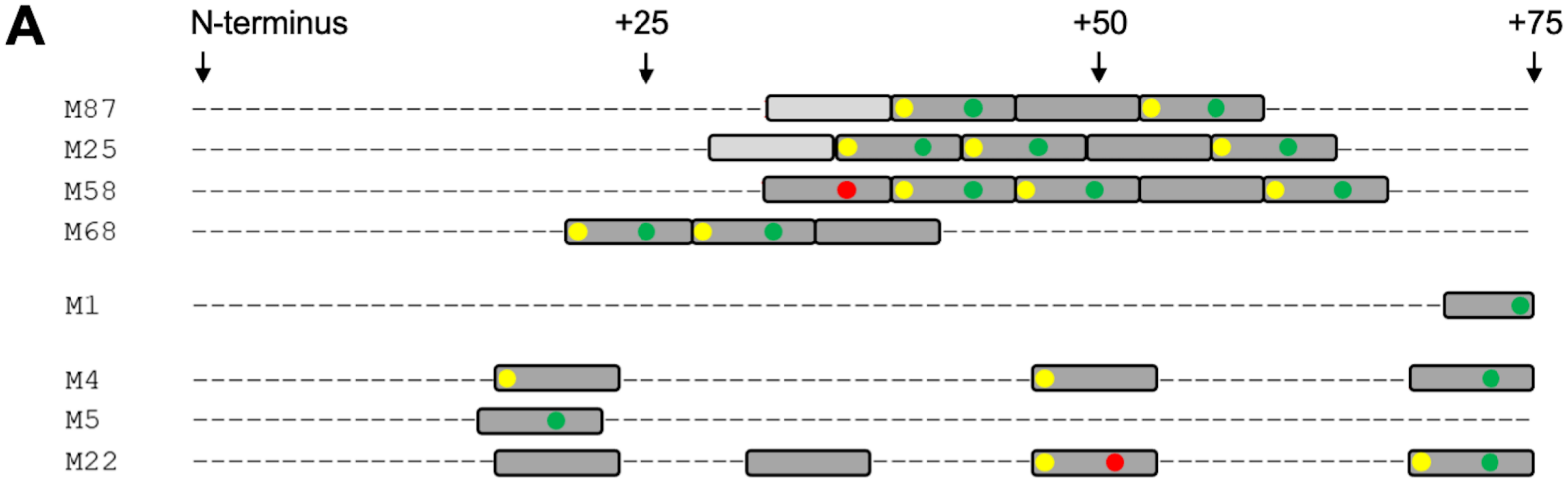

D
mutant/wt ratio
1.0
0.8
0.6
0.4
0.2
0.0
****
***
ns
****
M87N100 E85R
M25N100 E87R
M58N100 E97R
M68N100 E79R
mutant/wt ratio
5
4
3
2
1
0
*
**
M25N100 G82D
M58N100 G92D

Figure 6


A
N-terminus
+25
+50
+75
M87
M25
M58
M68
M1
M4
M5
M22

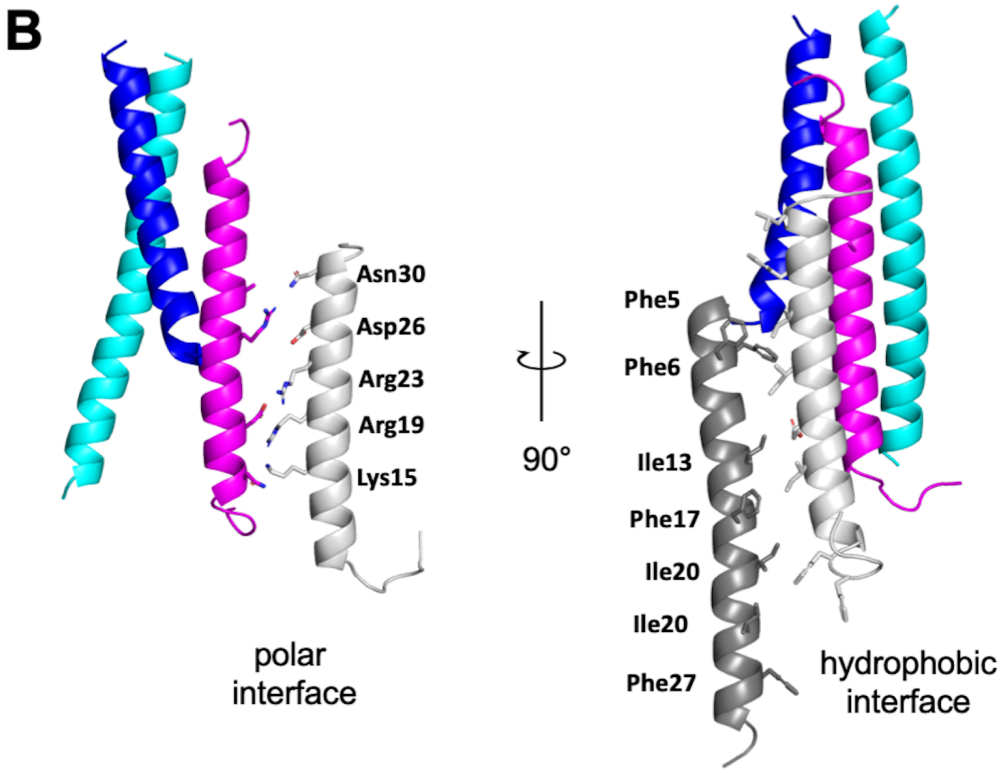

B
Asn30
Asp26
Arg23
Arg19
Lys15
polar interface
90°
Phe5
Phe6
Ile13
Phe17
Ile20
Ile20
Phe27
hydrophobic interface

Figure 1 — figure supplement 1

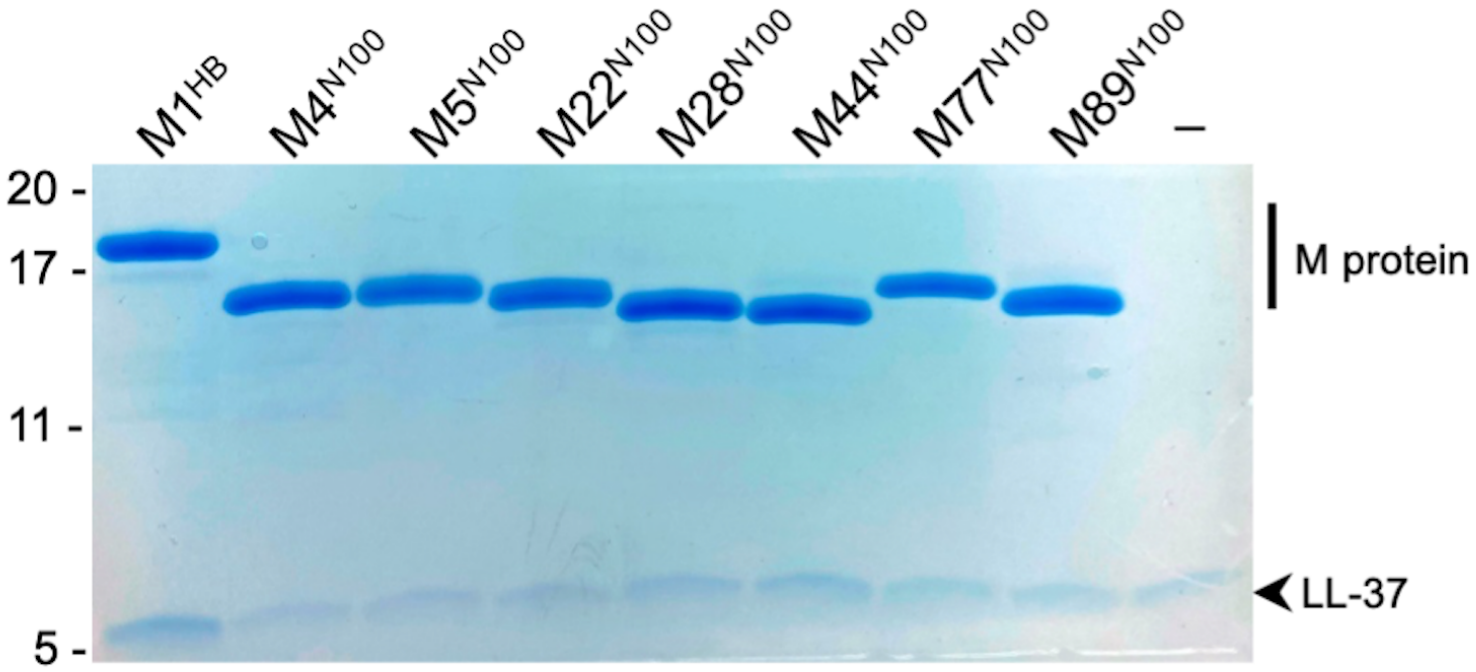


Figure 2 — figure supplement 1

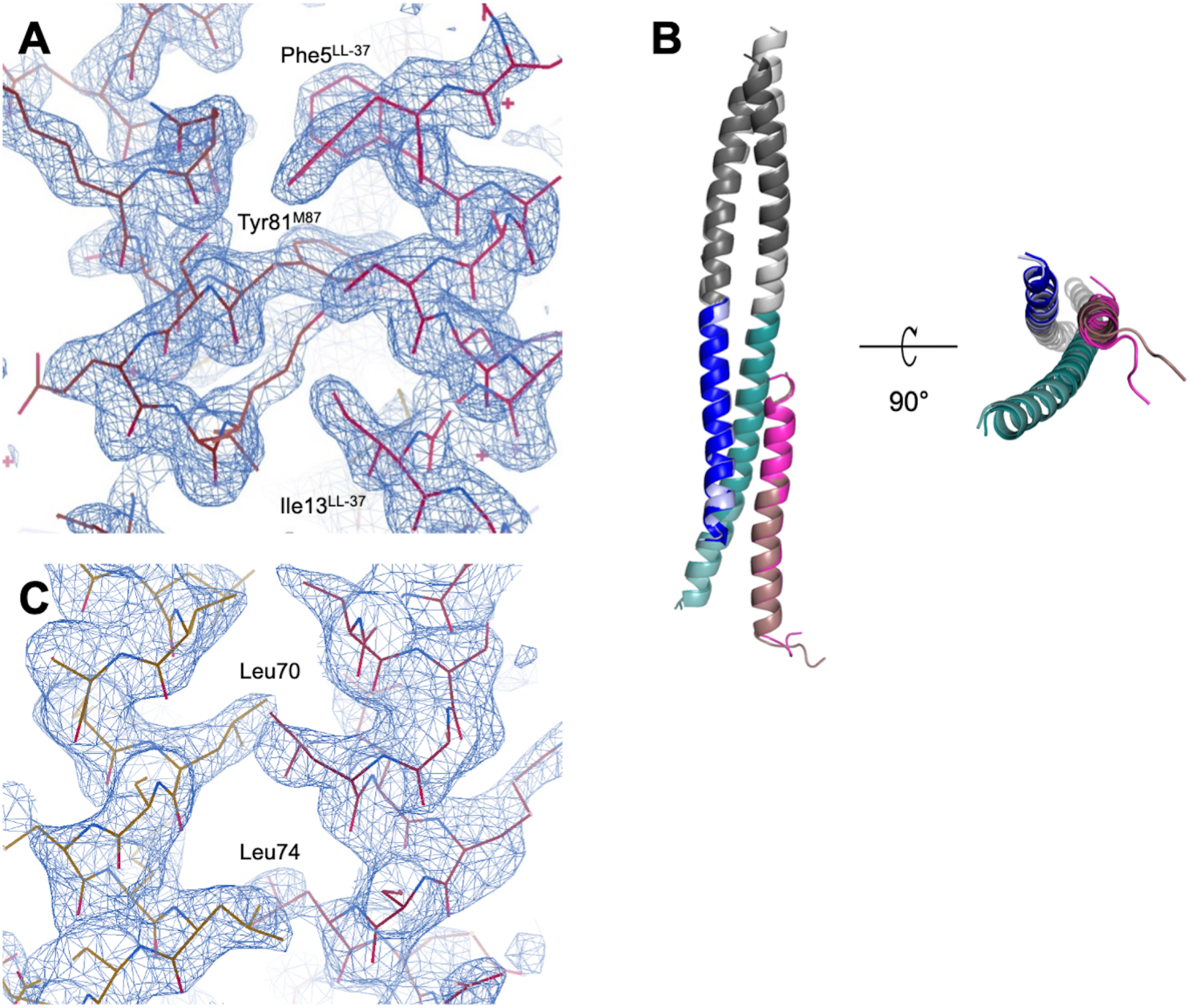


Figure 4 — figure supplement 1

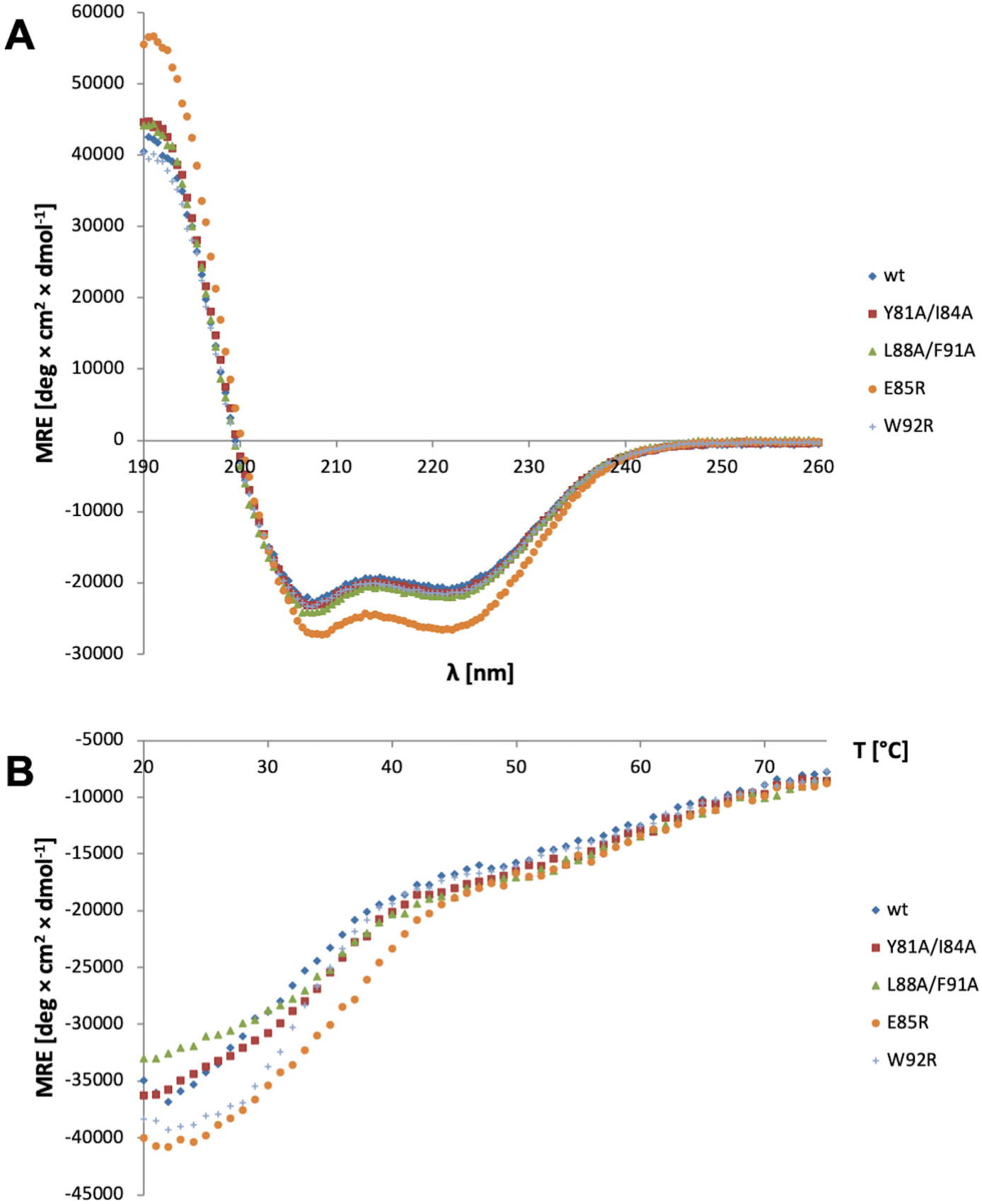


Figure 4 — figure supplement 2

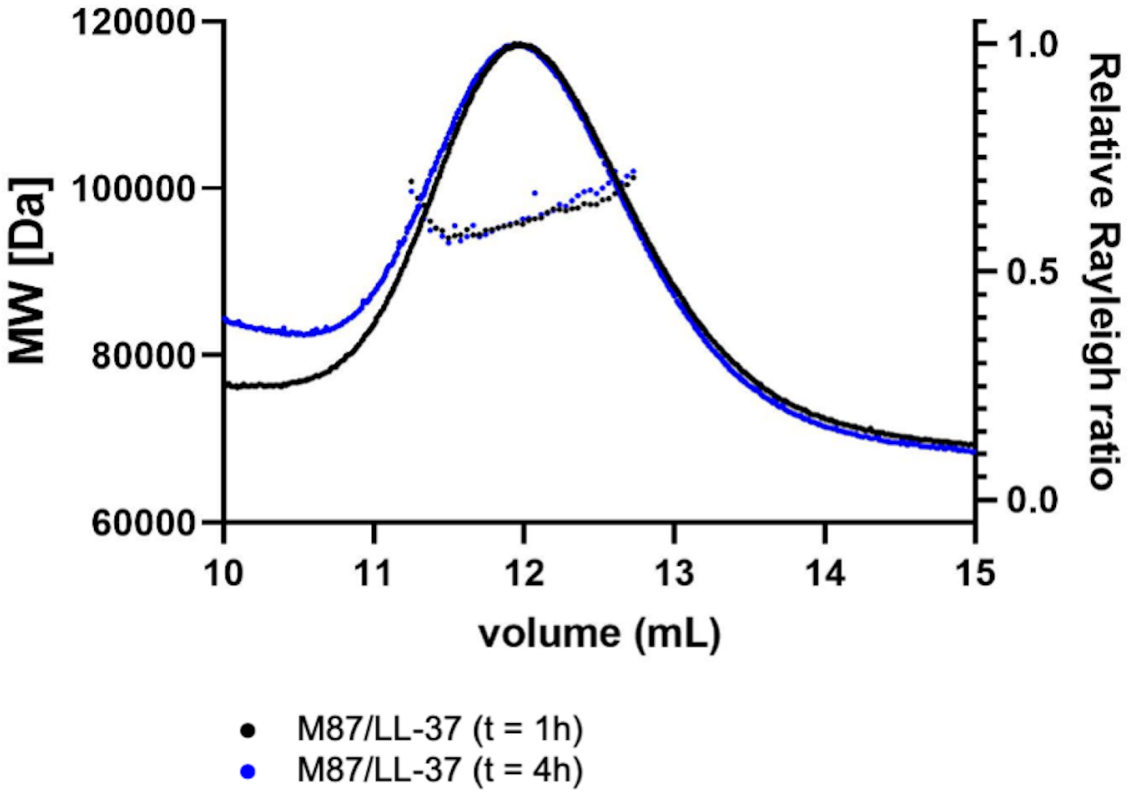
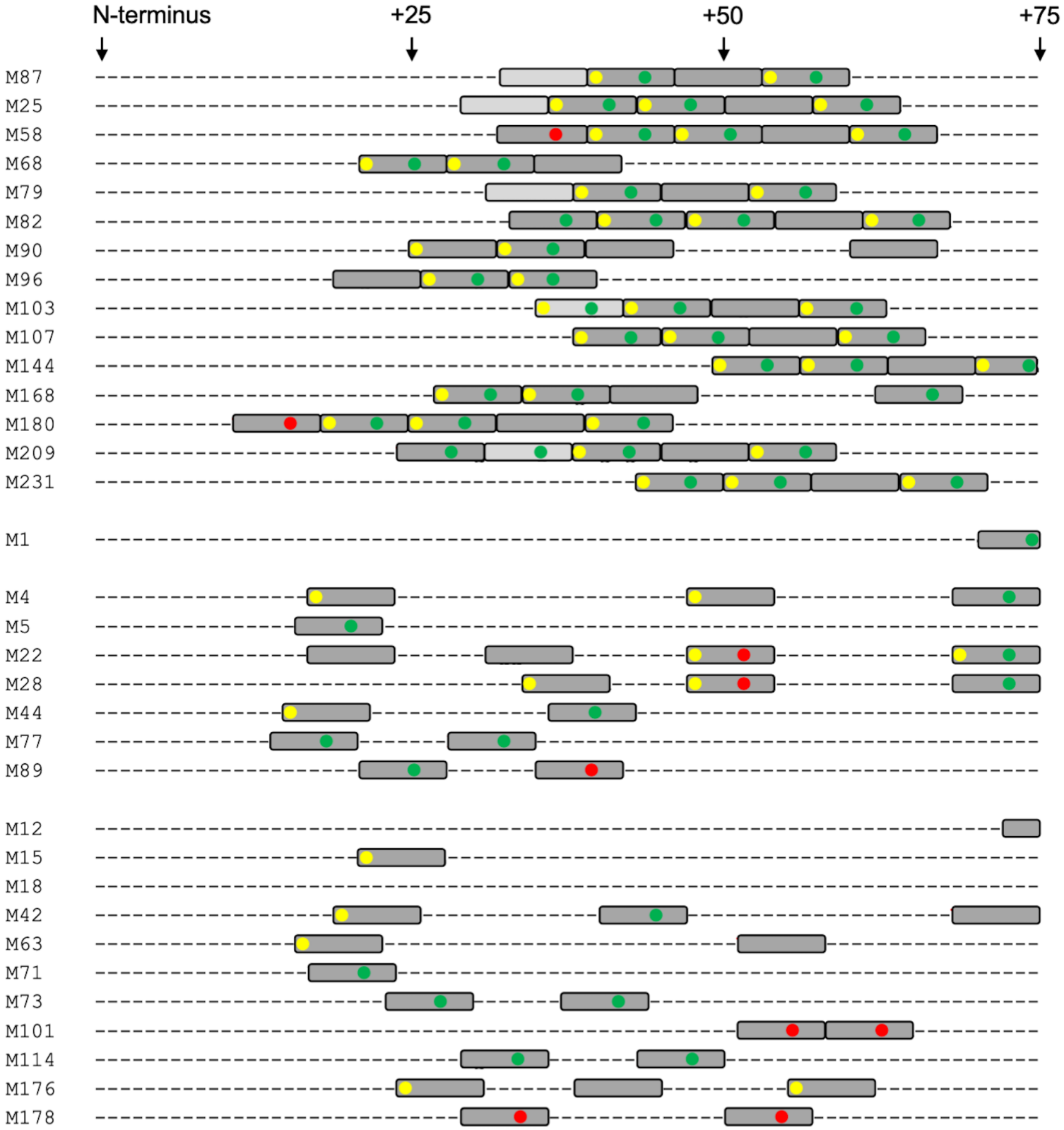



Figure 6 — figure supplement 1