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, which results in its neutralization. The crystal structure of a complex of LL-37 with M87 protein, an antigenic 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.

INTRODUCTION

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

A particular challenge lies in understanding how M proteins specifically bind and thereby neutralize the human antimicrobial peptide LL-37. This is so because LL-37 also has a simple topography, consisting only of a short amphipathic α-helix. LL-37 is a member of the cathelicidin antimicrobial peptide family, and constitutes a major host immune defense against Strep A (7, 8). The 37-amino acid peptide is proteolytically generated from the precursor protein hCAP-18, which is produced by neutrophils, macrophages, mast cells, and keratinocytes along with other epithelial cell types (7, 9, 10). Like other amphipathic α-helical antimicrobial peptides, LL-37 functions by inserting into and lysing bacterial plasma membranes (11-13). Notably, M1 protein confers resistance against LL-37 action by sequestering it into a ‘protein trap’ on the Strep A surface, thereby preventing LL-37 from reaching its target of action, the bacterial inner membrane (8, 14). M1 protein released (in soluble form) from the Strep A surface, as occurs

during infection (15), shares this capacity (8). M1 protein also binds the LL-37 precursor hCAP-18, and consequently prevents the proteolytic generation of LL-37 (8).

To determine the mechanism of LL-37 binding by M proteins, we pursued co-crystallization. While M1 protein proved recalcitrant to co-crystallization, several new M protein types that bind and neutralize LL-37 were identified. LL-37 was co-crystallized with M87 protein, which is from a strain that is an emerging health threat (16). The structure revealed a remarkable and novel mode of interaction for a coiled coil, in which a portion of the M87 protein coiled coil unfurled and exposed its hydrophobic core for interaction with LL-37. Visualization of the LL-37-binding motif in M87 protein made it clear that other M protein types, which are prevalent in human populations, also possessed this motif for binding and neutralization of LL-37. Our results provide specific detail for inhibiting the interaction of M proteins with LL-37.

RESULTS

Structure of the M87/LL-37 Complex

To determine the mechanism of LL-37 by M proteins, we pursued co-crystallization beginning with M1 protein but were unable to obtain co-crystals despite various attempts, including fusion to segments of the canonical coiled-coil protein GCN4 (8, 14, 17). A number of other M protein constructs (M4, M5, M22, M28, M44, M58, M77, M87, and M89), which were available in the laboratory for other purposes, were tried next. These consisted of the N-terminal 100 amino acids (denoted below with superscripted “N100”) of these M proteins. $M58^{N100}$ and $M87^{N100}$ bound LL-37, with $M87^{N100}$ having the highest binding capacity, even higher than that of the $M1^{HB}$ fragment (Fig. 1A), while the others did not bind LL-37 above background level (Fig. S1).

To our knowledge, M58 and M87 proteins are the first M proteins besides M1 protein identified to bind LL-37. To determine whether this binding was functionally significant, $M58^{N100}$ or $M87^{N100}$ was exogenously added to an M1 Strep A strain in which *emm1* had been deleted

(*Δemm1*) (14). As previously shown (8), *Δemm1* is sensitive to the antimicrobial action of LL-37, but exogenously added $M1^{HB}$ provides resistance against LL-37 (Fig. 1B). Similarly, exogenous addition of $M58^{N100}$ or $M87^{N100}$ to *Δemm1* increased the LL-37 MIC, with $M87^{N100}$ having the greatest effect, consistent with its higher affinity for LL-37 (Fig. 1B).

A complex of LL-37 and a version of M87 protein (amino acids, aa 68-105) that had GCN4 (aa 250-278) fused in register to its N-terminus was co-crystallized. The structure of the GCN4-M87/LL-37 complex was determined through molecular replacement to 2.1 Å resolution limit (Table SI, Fig. S2A). The GCN4-M87 fusion protein was verified to bind and neutralize LL-37, as gauged by co-precipitation and MIC assays, respectively (Figs. 1B, C). GCN4-M87 bound LL-37 with an apparent affinity similar to that of $M87^{N100}$ and led to an increase in the LL-37 MIC to an even greater extent than $M87^{N100}$.

The structure revealed a single LL-37 molecule bound to a dimer of GCN4-M87 (Fig. 2A). Two nearly identical 1:2 complexes occupied the asymmetric unit of the crystal (Fig. S2B, RMSD 0.70 Å). Almost the entire length of LL-37, which was predominantly in α-helical conformation, was visible in the crystal structure. GCN4-M87 was likewise in α-helical conformation throughout, but strikingly, only the GCN4 portion formed a coiled coil while the M87 α-helices were unfurled and asymmetrically disposed. This was best appreciated by comparing the structures of the complexed and free form of GCN4-M87. The structure of the latter was determined to 2.4 Å resolution limit (Table SI, Fig. S2C). Free GCN4-M87 formed a dimeric coiled coil throughout (Fig. 2B), indicating that the unfurling of the M87 α-helices was unique to the bound form (Fig. 2C). The greatest extent of unfurling occurred at Phe91, in which the distance between Cα atoms was 8.7 Å in the free form but 15.6 Å in the bound form. Only the M87 portion contacted LL-37, and therefore we will refer only to M87 rather than GCN4-M87 hereafter. The two M87 α-helices and the LL-37 α-helix together formed a parallel, three-helix bundle. One of the M87 α-helices made more contact with LL-37 (1016 $Å^2$ total buried surface area, average of the two complexes in the asymmetric unit of the crystal) (Fig 3, dark blue) while

the other made less but still substantial contact (491 $Å^2$ average total buried surface area) (Fig. 3, cyan); to differentiate between the two M87 α-helices, the one making greater contact will be denoted M87α1 and the other M87α2.

Interface dominated by hydrophobic interactions

Overall, hydrophobic interactions dominated the interaction site. Most notably, a number of hydrophobic M87 amino acids that occupied the core *a* or *d* positions and contacted each other in the free form, as is typical of coiled coils, were exposed in the bound form and instead contacted the hydrophobic face of the amphipathic LL-37 α-helix. Nearly five helical turns of LL-37 (Phe5-Val21) and M87 protein (Leu74-Trp92) engaged one another. Near the N-terminus of LL-37, a consecutive pair of phenylalanines, Phe5 and Phe6, were surrounded by hydrophobic *a* and *d* position amino acids of M87 — Leu74 (*a*, usual position in heptad), Ala77 (*d*), and Tyr81 (*a*) from M87α along with Ala77 from M87α2 (Fig. 3A). The two LL-37 Phe's and M87 Tyr81 formed a series of π-stacks. Two helical turns later in LL-37 was Ile13 which engaged in a ring of isoleucines, with M87 contributing Ile84 (*d*) from each of its helices (Fig. 3B). A further helical turn later in LL-37 was Phe17 which packed against a pair of M87 leucines, Leu88 (*a*), one from each of the M87 helices, and π-stacked against M87α2 Phe91 (*d*) (Fig. 3C). Lastly, one more helical turn further in LL-37 were Ile20, which was surrounded by M87α1 Phe91 (*d*) and Trp92; Val21, which packed against M87α2 Phe91; and Arg23 whose aliphatic side chain atoms packed against M87α1 Trp92 (Fig. 3D). These hydrophobic contacts were supplemented by a sparingly few polar ones, which occurred all within a small polar break in the hydrophobic face of LL-37, near its N-terminus: LL-37 Ser9 and Lys12 formed a hydrogen bond and salt bridge, respectively, with M87α1 Glu85 (Fig. 3E), and in one of the two complexes, LL-37 Lys12 also formed a salt bridge with M87α1 Glu89.

Importance of Hydrophobic Interactions

The role of the contacts observed structurally were evaluated through site-directed mutagenesis of M87 protein. To ensure that our structural observations were not limited to a fragment of M87 protein, these experiments were carried out with intact M87 protein. Intact His-tagged M87 protein and LL-37 were incubated at 37 °C, and their interaction was determined through a Ni^{2+}-NTA agarose bead co-precipitation assay. Dual alanine-substitutions of M87 Tyr81 (*a*) and Ile84 (*d*), which interacted with LL-37 Phe5, Phe6, and Ile13 (Figs. 3A and B), significantly decreased LL-37 interaction (Figs. 4A and B). M87 Y81A/I84A was verified through circular dichroism (CD) to have a similar structure to wild-type M87 protein at 37 °C (Fig. S3A). Additionally, M87 Y81A/I84A had a temperature-induced unfolding profile similar to that of wild-type M87 protein (Fig. S3B), also as monitored by CD. These results indicated that alanine-substitutions of Tyr81 and Ile84 affected LL-37 binding directly rather than indirectly through compromised structure, stability, or both. Alanine-substitution of M87 Leu88 (*a*) and Phe91 (*d*), which interacted hydrophobically with LL-37 Phe17, Ile20, and Val21 (Figs. 3C and D), likewise markedly decreased LL-37 interaction (Figs. 4A and B). The secondary structure and stability of M87 L88A/F91A resembled that of wild-type M87 protein as well (Figs. S3A and B). A substantial amount of the surface area of M87 Trp92 was buried by contact with LL-37, but surprisingly, substitution of this amino acid with alanine increased LL-37 interaction (Figs. 3C and D). M87 Trp92 was adjacent to LL-37 Arg23 (Fig. 3D), and thus we asked whether Arg-substitution of M87 Trp92 would interfere with LL-37 binding. Indeed, M87 W92R was almost entirely deficient in LL-37 interaction (Figs. 4A and B), while showing no changes in secondary structure or stability (Figs. S3A and B).

The only M87 amino acid seen to make polar contacts in both complexes in the asymmetric unit of the crystal was Glu85 (Fig. 3E). Ala-substitution of M87 Glu85 had little effect on LL-37 binding (Figs. 4A and B). Since M87 Glu85 was adjacent to LL-37 Lys12, we asked whether Arg-substitution of this amino acid would decrease interaction with LL-37. Consistent

with our structural observations, M87 E85R had significantly decreased interaction with LL-37 (Figs. 4A and B). This decrease was a direct effect, as M87 E85R had greater α-helical content than wild-type M87 protein and similar stability (Figs. S3A and B).

These results validated the structural observations regarding the mode of interaction between M87 protein and LL-37, and indicated the importance of the M87 *a* and *d* heptad position amino acids to the interaction. They further indicated that the polar contact conferred by M87 Glu85 conferred specificity rather than binding affinity.

Stoichiometry of Interactions

For an LL-37 neutralization mechanism that involves an M protein trap (8), the 1:2 LL-37:M87 stoichiometry was puzzling. LL-37 forms variably sized oligomers in solution (18-20), and thus we undertook solution phase studies of the complex through size-exclusion chromatography (SEC) coupled to multiangle light scattering (MALS). The molecular weight of intact M87 protein alone was 73.6 ± 1.2 kDa (calc. 72.4 kDa) (Fig. 4C, black). A 10-fold excess of LL-37 to the intact M87 dimer was added and the sample was applied to SEC-MALS almost immediately after mixing. The complex had a mass of 80.4 ± 0.6 kDa, which corresponded to an M87 protein dimer bound to one or two molecules of LL-37 (calc. 4.5 kDa) (Fig. 4C, blue). Notably, after one hour of incubation of LL-37 with M87 protein, the mass of the complex was 97.1 ± 0.6 kDa (Fig. 4C, red), which corresponded to an M87 protein dimer bound to five or six molecules of LL-37. Incubation of four hours also resulted in the same increased mass, indicating self-limiting growth of the complex (Fig. S4). These results suggested that the single LL-37 molecule bound to an M87 dimer could recruit four or five additional LL-37 molecules.

***emm87* Confers Resistance to LL-37**

We asked whether the mode of interaction visualized through crystallography was applicable to M87 protein in its native conformation on the Strep A surface. An isogenic

Δemm87 strain was constructed, and complemented with a plasmid expressing either wild-type *emm87* or *emm87* containing the E85R substitution, the latter having greatly diminished LL-37 binding in solution (Fig. 4A). The wild-type M87 strain was resistant to LL-37 (Fig. 4D), and indeed had a greater LL-37 MIC than the M1 strain (Fig. 1C). Deletion of *emm87* led to significantly increased sensitivity to LL-37 (Fig. 4D). Notably, the *Δemm87* strain complemented with wild-type M87 was restored to the LL-37 resistance of the wild-type parental M87 strain, while complemented with M87 E85R remained sensitive to LL-37, similar to the level seen for the uncomplemented *Δemm87* strain. These LL-37 susceptibility results provide physiological validation for our structural observations.

Conservation of M87 motif in other M types

The LL-37 binding motif visualized in M87 protein was identified in the sequence variable N-terminal regions of 14 other M protein types (Fig. 5A). The motif consisted of two consecutive ideal heptads, that is, with *a* and *d* positions occupied by canonical hydrophobic amino acids (21-23), including a strictly conserved Tyr at the *a* position of the first heptad. These hydrophobic amino acids in M87 protein (Tyr81, Ile84, Leu88, and Phe91) were shown above to be crucial to interaction with LL-37 (Figs. 4A and B). Along with these, the motif had a hydrogen bond acceptor at the *e* position of the first heptad, which in M87 protein (Glu85) provided specificity by contacting LL-37 Lys12 (Fig. 3). In addition, a positively charged amino acid was excluded from the *e* position of the second heptad, which in M87 protein was Trp92 and proximal to LL-37 Arg23 (Fig. 3). Preceding these two ideal heptads was a nearly ideal heptad, with the *a* position occupied by canonical amino acids (i.e., Leu or Tyr) while the *d* position was tolerant of less than ideal amino acids (i.e., Ala77 as in M87, and also Gln) (22, 23).

Among the M proteins identified to have this motif was M58 protein. We showed above that $M58^{N100}$ binds LL-37 (Fig. 1A). Similar fragments consisting of the N-terminal 100 amino acids were constructed for M25 and M68. $M68^{N100}$ bound LL-37 with similar apparent affinity as

M87^{N100}, while M25^{N100} bound somewhat more weakly, similar to M58^{N100} (Fig. 5B). To test whether the LL-37 binding mode observed for M87 protein was conserved in these M proteins, Arg-substitutions of the equivalent of M87 E85 were constructed. As the E85R substitution had been evaluated only in intact M87 protein, it was introduced into the M87^{N100} fragment, which resulted in significantly decreased LL-37 binding as well (Figs. 5B and D). Arg-substitution of the equivalent Glu in M25^{N100} (E87), M58^{N100} (E97), and M68^{N100} (E79) also led to significantly decreased LL-37 binding (Figs. 5B and D). In the case of M25 and M58 proteins, which did not bind LL-37 as well as M87 or M68, we noticed that the amino acid preceding the strictly conserved Tyr was Gly, which is a helix-breaking amino acid. Therefore, we substituted the Gly in M25 and M58 proteins with Asp, as in M87 D80, and found that M25^{N100} G82D and M58^{N100} G92D had significantly higher levels of LL-37 binding as compared to their wild-type counterparts (Figs. 5C and D).

These results are consistent with M25, M58, and M68 binding LL-37 in the same mode as observed for M87 protein, and support the hypothesis that the M87 protein LL-37 binding motif is conserved in at least fourteen other M proteins, which belong to the E2 or E3 cluster of M proteins (24).

DISCUSSION

We sought to understand how M proteins achieve specific binding of LL-37, which is essential to neutralization of this human antimicrobial peptide by the M1 strain of Strep A and likely by other strains as well (7, 8, 14). This question was challenging as the fibrillar structure of M proteins excludes features commonly present in globular proteins that enable specific binding, such as pockets or cavities. Adding to the challenge, the structure of LL-37 lacks complexity, consisting simply of an amphipathic α-helix. We discovered a remarkable mode of interaction of M87 protein with LL-37 through structure determination. The two α-helices of M87 protein in the free state form a coiled coil, with amino acids in the core *a* and *d* positions

engaging in ‘knobs-into-holes’ packing. In the bound state, the two helices do not form coiled coils, and are instead unfurled and asymmetrically disposed. Most significantly, in the bound state, hydrophobic amino acids at the *a* and *d* positions were exposed and formed a continuous patch that contacted the hydrophobic face of the LL-37 α-helix. These hydrophobic contacts dominated the interaction interface and were shown to be essential by Ala-substitution mutagenesis. While coiled-coil asymmetry coupled to cognate partner binding has been noted in some proteins (25, 26), this has taken the form of a helical stagger but with the coiled coil maintained. To our knowledge, the unfurling and asymmetric exposure of consecutive *a* and *d* position amino acids to form a continuous interaction site is a novel binding mechanism for a coiled-coil protein.

The LL-37 binding site in M87 protein was formed by two ideal heptads preceded by a nearly ideal heptad (Ala and Gln being tolerated at the *a* position). M proteins identified or predicted to bind LL-37 in the same mode also had at least three such consecutive repeats, with the first heptad being ideal in most cases (Figs. 6A and S5). In contrast, M proteins identified not to bind LL-37 lacked consecutive ideal or near ideal heptads. Instead ideal heptads occurred as isolated singletons in the midst of non-ideal heptads. Notably, M1 protein has only an isolated ideal repeat (Fig. 6A), indicating that there are additional mechanisms for binding LL-37, different from those observed for M87 protein. In general, the occurrence of two or more consecutive ideal heptads is rare in M protein variable regions (Fig. S5), whose sequences frequently have coiled-coil destabilizing amino acids at the *a* and *d* positions (e.g., Glu or Lys) (3, 27). In the case of the M1 B-repeats region, non-ideal heptads have been shown to be functionally essential for creating protein dynamics, which are required for binding fibrinogen in a “capture-and-collapse” mechanism (28). Dynamics in the consecutive ideal repeats of M87 and related M proteins are likely to be much lower than in the non-ideal M1 B-repeats, but nevertheless, it appears that enough breathing motion exists even in these ideal coiled coils for the infiltration of LL-37 and disruption of coiled-coil structure.

The single molecule of LL-37 bound to M87 protein acted as a nucleator for the recruitment of additional LL-37 molecules. One or two molecules of LL-37 molecule bound to M87 protein at an initial time point, but grew to five or six molecules over time, as evidenced by SEC-MALS analysis and consistent with LL-37 oligomerizing in solution (18-20). Plausible mechanisms for the recruitment of additional LL-37 molecules are suggested by crystal structures of LL-37 alone. In one crystal structure, the polar faces of two LL-37 molecules contact one another in anti-parallel orientation (12), and in another, the hydrophobic faces do likewise (11). Thus, it is possible that the single LL-37 molecule bound through its hydrophobic face to M87 protein is able to recruit a second molecule of LL-37 through polar face-polar face interactions. This second LL-37 molecule is then able to recruit a third molecule of LL-37 through hydrophobic face-hydrophobic face interactions (Fig. 6B). While this sort of growth is not self-limiting, the LL-37/M87 complex was self-limiting at five or six molecules of LL-37 per M87 dimer, and thus the specific details of LL-37/LL-37 interactions are likely to differ. The excess of LL-37 bound by M87 protein is consistent with a 'protein trap' neutralization mechanism (8).

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

M87 Strep A strains are prevalent in human populations and are an emerging cause of human clinical disease (16, 29, 30). The genomes of almost all M87 strain isolates contain a recombination event that increases the expression of Strep A toxin genes (NADase and streptolysin O) (16). Importantly, the same recombination event is present in the M1 strain that

caused a global pandemic starting in the 1980's and an M89 strain that is responsible for an ongoing epidemic (31). A few isolates of the M82 strain also carry this recombination event (16). The M82 strain, along with a number of the other M types identified to have the M87 motif, are prevalent in human populations, including M25, M68, M90, and M103 (16, 29, 30). These observations provide motivation for pursuing therapeutic inhibition of LL-37 binding by M proteins, and our results provide specific detail to achieve this end.

Materials and methods

Bacterial strains and culture conditions. *Streptococcus pyogenes* (Strep A) strains M87 20161436 (NCBI SRA accession: SAMN07154152) and its isogenic *Δemm87* strain (32), and M1T1 5448 and its isogenic Δ*emm1* strain (14) were used. *S. pyogenes* was grown as standing cultures in Todd-Hewitt broth in ambient atmosphere at 37 ˚C. *Escherichia coli* was cultured in Lysogeny Broth at 37 ˚C with agitation. For selection and maintenance of strains, antibiotics were added to the medium at the following concentrations: erythromycin 500 µg/mL for *E. coli* and 2 µg/mL for *S. pyogenes*; chloramphenicol, 2 µg/mL for *S. pyogenes*.

DNA manipulation. Coding sequences for M proteins were cloned into a modified pET28b vector (Novagen) that contained sequences encoding an N-terminal His_6-tag followed by a PreScission protease cleavage site (5). Amino acid substitutions and deletions were introduced into pET28b vectors with the QuickChange II Site-Directed mutagenesis kit (Stratagene), according to the manufacturer's directions, and into pM87 E85R (32), which was used for expression in *S. pyogenes*, with the Phusion Site Directed Mutagenesis Kit (ThermoScientific, Waltham, MA, USA). The coding sequence for GCN4 250-278 was subcloned from *Saccharomyces cerevisiae*, and was fused to M87 68-105 through strand overlap extension PCR. An M87 protein-expressing vector (pM87) was constructed by insertion of *emm87* into pDCerm (32).

Peptides and Proteins. LL-37, which was chemically synthesized and lyophilized as a fluoride salt (Genscript, 95% purity), was solubilized at 5 mg/ml in sterile deionized water for MIC assays or 100 mM NaCl, 20 mM HEPES-NaOH, pH 7.5 (HS) for other experiments.

Expression and purification of M proteins constructs was carried out as previously described (3, 5), except for the following minor modifications. After PreScission protease digestion, GCN4-M87 was subjected to gel filtration chromatography using a Superdex 200 (GE Healthcare) column equilibrated with HS buffer. For formation of the GCN4-M87/LL-37 complex, GCN4-M87 (2 mg/ml) was mixed with a three-fold molar excess of LL-37 (3 mg/ml), both in HS buffer, and the complex was purified by gel filtration chromatography using a Superdex 200 column that had been equilibrated with 100 mM NaCl, 20 mM MES-NaOH, pH 6.5. Intact wild-type and mutant M87 proteins, following Ni^{2+}-NTA agarose bead purification, were subjected to gel filtration chromatography on a Superdex 200 column that had been equilibrated with HS buffer. For CD measurements, the His-tag on intact wild-type and mutant M87 proteins was removed by PreScission protease digestion, and the cleaved product was further purified by reverse Ni^{2+}-NTA chromatography. Protein concentrations of M proteins were determined by measuring A_{280} with the sample in 6 M guanidine hydrochloride, 20 mM Tris, pH 8.5, and using a calculated molar 280 nm extinction coefficient. The concentration of LL-37 and GCN4-M87/LL-37 complex was measured using the Bradford assay (Bio-Rad) with BSA as a standard.

Co-precipitation assay. One and half nmol of His_6-tagged M protein constructs (2-10 µl) were added to 50 µl of Ni^{2+}-NTA agarose beads that had been pre-equilibrated with HS buffer and incubated with gentle agitation for 10 min at RT. Beads were centrifuged (3,000 x *g*, 30 s, RT) and the supernatant was removed. Six nmol of LL-37 in 150 µl of HS buffer was added to the beads and incubated with gentle agitation for 30 min at 37 °C. The beads were washed three times each with 1 ml HS buffer containing 5 mM imidazole, pH 8.0. For the washes, the resin

was mixed with the wash solution by gentle agitation, incubated for 1 min at RT and then centrifuged (3,000 x *g*, 30 s, RT). Bound proteins were eluted with 50 µl HS containing 400 mM imidazole, pH 8.0. Protein samples were resolved by SDS-PAGE and stained with InstaBlue (APExBIO). The intensity of gel bands was quantified as previously described (5).

Minimal Inhibitory Concentration. *S. pyogenes* that had been grown overnight were inoculated into Todd-Hewitt broth at 1:100 dilution and grown at 37 °C to an OD_{600} of 0.4. The culture was diluted to an OD_{600} of 0.1, and 5 µl (~10^5 CFU) was mixed into 100 µl of RPMI 1640 medium with glutamine, which contained 0, 2, 4, 8, 12, 18, or 32 µM LL-37. In some experiments, the medium also contained 10 µM $M1^{HB}$, $M58^{N100}$, or $M87^{N100}$ protein. *S. pyogenes* were grown in individual wells of a 96-well plate for 24 h at 37 °C. *S. pyogenes* viability was assessed at this time point by the color of the RPMI medium, where yellow indicated bacterial growth and red no bacterial growth. The MIC was defined as the LL-37 concentration at which no growth was detectable at 24 h.

Molecular mass determination. Intact His-tagged M87 protein (2.5 mg/ml) alone or mixed with LL-37 (1.6 mg/ml; 10-fold molar excess over M87 dimer) in HS (100 mM NaCl, 20 mM HEPES-NaOH, pH 7.5) was centrifuged (10 min, 20,000 x *g*, 20 °C) to remove aggregates. Samples (100 µl) were then either immediately applied to a Superdex 200 10/300 column that had been pre-equilibrated in HS, or incubated 1-4 h at RT before application to the column. Samples eluting from the column were monitored with a light scattering detector (DAWN HELEOS II, Wyatt Technology, Santa Barbara, CA, USA) and a differential refractometer (Optilab T-rEX; Wyatt Technology). Data processing and molecular mass calculation were performed with ASTRA software (Wyatt Technology).

Crystallization and data collection. Crystallization trials were carried out at 293 K using the hanging drop vapor diffusion method. The GCN4-M87/LL-37 complex was concentrated by ultrafiltration using a 3,500 MWCO membrane (Millipore; 4,500 x *g*, 30 min, 15 °C) to 8 mg/ml. GCN4-M87 alone was concentrated to 10 mg/ml by ultrafiltration through 3,500 MWCO membrane (Millipore; 4,500 x *g*, 30 min, 4 °C) The complex was brought to RT before introduction into crystallization drops to overcome its low solubility at 4 °C.

The GCN4-M87/LL-37 complex (0.8 µl) was mixed in a 1:1 ratio with 5% (v/v) acetonitrile, 0.1 M MES-NaOH, pH 6.5. Microclusters of plates (ca. 20 x 20 x 5 μm^3) that had grown after 2-4 days were crushed, diluted 125-fold with the precipitant solution, and centrifuged (2,000 x *g*, 30 s, RT). The supernatant was collected and used as a seed stock. GCN4-M87/LL-37 (0.9 µl) was mixed with 0.9 µl of 10% (v/v) acetonitrile, 0.1 M MES-NaOH, pH 6.5 and 0.2 µl of the seed stock. Clusters of diffraction-sized crystals (50–150 µm in each dimension) which were obtained after 3-7 days were crushed once again and used for a subsequent round of seeding, carried out in the same manner as the first round. These two rounds of seeding yielded single crystals (200 x 200 x 10 μm^3) that were cryo-preserved by three serial transfers to the precipitant solution supplemented with 10, 20, and 30% ethylene glycol, respectively, and flash-cooled in liquid N_2.

GCN4-M87 (0.8 µl) was mixed in a 1:1 ratio with 0.6 M monobasic ammonium phosphate, 0.1 M Tris-HCl, pH 8.0,. Single crystals (300 x 150 x 150 μm^3), which were obtained after 4-7 days, were cryo-protected in the precipitant solution supplemented with 30% ethylene glycol and flash cooled in liquid N_2 prior to data collection.

Diffraction data were collected at SSRL (beamline 9-2) at 0.979 Å, integrated with Mosflm (GCN4-M87/LL-37) (33) or DIALS (GCN4-M87) (34) and scaled with Aimless (35) (Table S1). Because of the highly anisotropic nature of both datasets, the resolution cutoffs for both were determined using anisotropic $CC_{1/2}$.

Structure determination and refinement. Phases for the GCN4-M87/LL-37 complex and free GCN4-M87 were determined by molecular replacement using Phaser (36). The search model was generated from the coiled-coil dimer structure of GCN4 fused to the coiled-coil dimer structure of striated muscle α-tropomyosin (PDB 1kql) using Sculptor (37) with default settings. Extensive model modification and building were performed with Coot and guided by the inspection of σA-weighted 2mFo-DFc and mFo-DFc omit maps (Fig. S2).

LL-37 in the GCN4-M87/LL-37 complex was manually modelled into well-defined difference electron density that was visible after a few rounds of refinement of the search model. The asymmetric unit of the GCN4-M87/LL-37 crystal contained two heteromeric assemblies, each composed of two chains of GCN4-M87 and one of LL-37. Refinement was performed using Refine from the Phenix suite (38) with default settings. At the final stages of refinement, TLS parameters were applied. TLS groups were applied as follows: chain A; chain B; chain C aa 41-55 and 56-93' chain D aa 38-55 56-104; chain E; and chain F. Side chains with no corresponding electron density were truncated to Cβ. Interaction interfaces between M87 and LL-37 were analyzed with PISA (39).

The GCN4-M87 dimer was modeled into continuous electron density. The asymmetric unit of the GCN4-M87 crystal contained a single coiled-coil dimer. Refinement of GCN4-M87 was performed using Refine with default settings in addition to the application of two-fold NCS restraints and TLS parameters. Each chain constituted a single TLS group. Three and six amino acids at N- and C-termini, respectively, of chain A and a single N-terminal amino acid of chain B lacked electron density and were not modeled.

Molecular figures were generated with PyMol (http://pymol.sourceforge.net).

Structures have been deposited in the PDB: 7SAY for GCN4-M87/LL-37 and 7SAF for GCN4-M87.

CD spectroscopy. CD spectra were measured on an Aviv 215 Circular Dichroism Spectrometer using a quartz cell with 1 mm path length. Protein samples were ~0.125-0.250 mg/ml in 5 mM sodium phosphate, pH 7.9. Wavelength spectra were recorded in a range of 190-260 nm at 37 °C at 0.5 nm intervals with a 0.5 s averaging time per data point. Melting curves were determined at 222 nm between 20-75 °C with 1 °C increments and 30 s equilibration time for each temperature point. Two independent experiments were carried out for each sample, and the data were averaged and presented as a mean molar residue ellipticity.

Identification of M87 LL-37-binding motif in other M proteins. The sequence of the N-terminal 250 amino acids of mature M87 was aligned against the sequence of the N-terminal 250 amino acids of the mature form of 179 M proteins using Clustal Omega (40). Alignments were manually curated for the presence of a Tyr (or Phe, although none were found) occupying the position equivalent to M87 Tyr81, the occurrence of hydrophobic amino acids at the *d*, *a,* and *d* positions following the Tyr, and a hydrogen bond acceptor at the *e* position following the Tyr.

Acknowledgements

We thank S. Rees for his help with crystallographic data collection. This work was supported by NIH 1R21AI144901 (P.G.).

Supplementary Information

Table S1.

Figures S1-S5

References

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

28. Stewart CM, Buffalo CZ, Valderrama JA, Henningham A, Cole JN, Nizet V, Ghosh P. Coiled-coil destabilizing residues in the group A Streptococcus M1 protein are required for functional interaction. Proc Natl Acad Sci U S A. 2016;113(34):9515-20. PMC5003295. PubMed 27512043

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

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

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

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

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

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

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

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

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

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

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

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

FIGURE LEGENDS

Figure 1. LL-37 binding and detoxification.

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

B. LL-37 MIC for wild-type Strep A *emm1* and isogenic *Δemm1* alone or supplemented with 10 μM $M1^{HB}$, $M58^{N100}$, $M87^{N100}$, or GCN4-M87. Means and standard deviations from four independent experiments are shown.

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

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

A. Cartoon representation of the GCN4-M87/LL-37 complex. LL-37 is in magenta. The GCN4 portion of GCN4-M87 is in gray, and the M87 portion in blue for the chain that makes more contacts to LL-37 (M87α1), and cyan for the chain that makes fewer contacts to LL-37 (M87α2). The N- and C-terminal amino acids for LL-37 and M87 are indicated.

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

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

Figure 3. M87/LL-37 Interface.

LL-37 is in magenta, and the M87 α-helix that forms a greater number of contacts with LL-37 is in blue (M87α1) and the one that forms fewer contacts in cyan (M87α2). Shown are contacts

formed with LL-37 (**A**) Phe5 and Phe6; (**B**) Ile 13; (**C**) Phe17; (**D**) Ile20, Val21, and Arg23; and (**E**) Ser9 and Lys12.

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

A. LL-37 binding by His-tagged, intact wild-type M87 or M87 Y81A/I84A, L88A/F91A, E85A, E85R, W92A, or W92R at 37 °C as determined by Ni^{2+}-NTA agarose bead co-precipitation. The last lane contains no M protein. Bound fractions were resolved by SDS-PAGE and visualized by Coomassie staining. The gel is representative of four experiments.

B. Quantification of LL-37 binding to intact wild-type and mutant M87 proteins. The ratio of band intensities of LL-37 co-precipitated by mutant M87 protein as compared to wild-type M87 protein is shown. Band intensities of LL-37 were first normalized by the quantity of co-precipitated M87 protein. Data from four independent experiments are presented as means and SD. Statistical significance was calculated using one-way ANOVA with Dunnett post-hoc test. P values are as follows: ns > 0.05, * ≤ 0.05, ** ≤ 0.01, *** ≤ 0.001 and **** ≤ 0.0001.

C. Molecular weight determination of intact M87 alone (black), M87 and LL-37 added together and immediately analyzed (blue), or M87 and LL-37 incubated for 1 h before being analyzed (red) by size-exclusion chromatography coupled to multiangle light scattering. Dotted lines indicate calculated molecular weights across the profile. Data are representative of three experiments.

D. LL-37 MIC for *emm87* and *Δemm87* alone or expressing wild-type M87 (pM87) or M87 E85R pM87(E85R) from a plasmid. Data from eight independent experiments are presented as mean ± SD. The statistical test and P values are the same as described for panel B.

Figure 5. Conservation of the M87LL-37-binding motif.

A. M87 amino acids that make hydrophobic contact to LL-37 are boxed in red and those that disrupt interaction with LL-37 through Arg-substitution are in blue boxes, as are amino acids that are similarly positioned in other M proteins. The arrow indicates the position of M87 D80.

B. LL-37 interaction with His-tagged wild-type or mutant $M87^{N100}$, $M25^{N100}$, $M58^{N100}$, $M68^{N100}$ proteins determined by Ni^{2+}-NTA agarose bead co-precipitation at 37 °C. Bound fractions were resolved by SDS-PAGE and visualized by Coomassie staining. Data are representative of six experiments.

C. LL-37 interaction with His-tagged $M25^{N100}$, $M25^{N10}$ G82D, $M58^{N100}$, and $M58^{N100}$ G02D, determined as in panel B. Data are representative of six experiments.

D. Quantification of the effect of mutagenesis on LL-37 binding. Data from six independent experiments are presented as means and SD. Statistical significance was calculated using Student's t-test to compare mutant and the corresponding wild-type protein. P values are as follows: ns > 0.05, * ≤ 0.05, ** ≤ 0.01, *** ≤ 0.001 and **** ≤ 0.0001.

Figure 6. LL-37 binding motif and higher-order complex formation.

A. Occurrence of ideal heptads (dark grey boxes; Ile, Leu, Met, Phe, Tyr, or Val at both *a* and *d* positions) and near ideal heptads (light grey boxes; ideal amino acid at *a* position and Ala at *d* position, with M25 being an exception with Gln at this position) within the N-terminal regions of various M protein types. Within these boxes, yellow circles indicate Tyr at an *a* position; green circles Asp, Asn, Glu, or Gln at an *e* position; and red circles Arg or Lys at an *e* position. Yellow and red circles favor LL-37 binding, while the red circles disfavor LL-37 binding.

B. Model of higher-order M87/LL-37 complex. M87α1 and M87α2 are in blue and cyan, respectively, and LL-37 is in magenta. Left, LL-37 bound to M87 protein has its polar face free to form polar contacts with a second LL-37 molecule (light gray, as seen in a structure of LL-37 alone, PDB 7PDC). Right, the light gray LL-37 has its hydrophobic face free to form hydrophobic

contacts with yet another LL-37 molecule (dark gray, as seen in another structure of LL-37 alone, PDB 5NNM). The light and dark gray LL-37 molecules do not contact M87 protein.

Figure 1

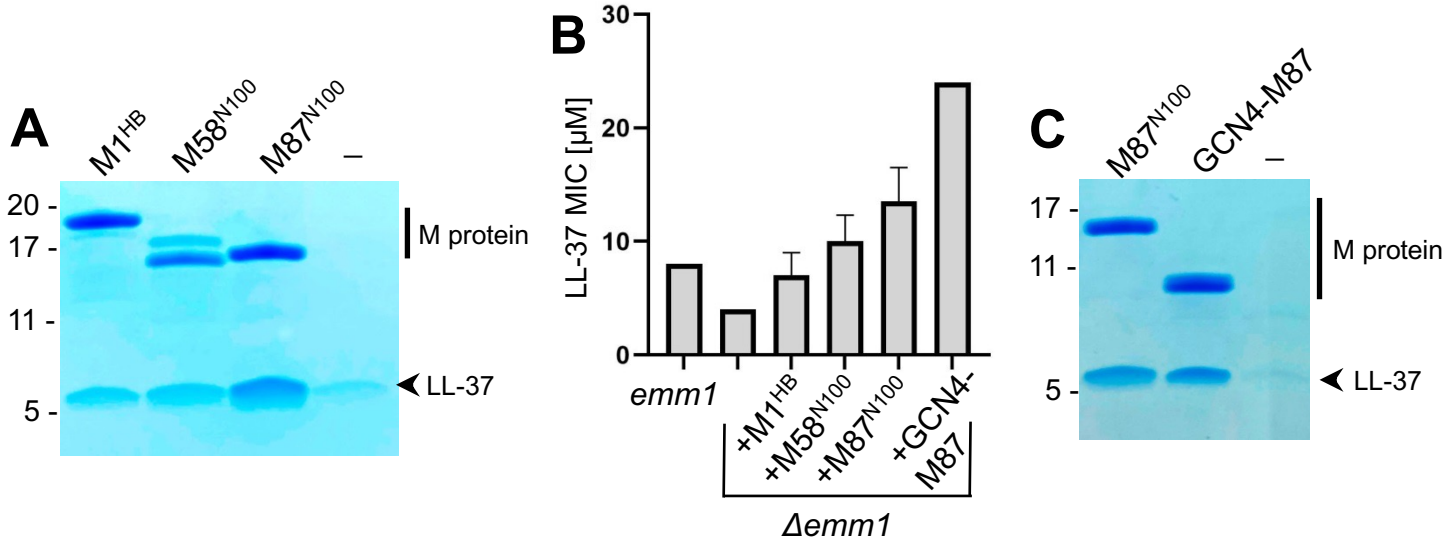

A
M1HB
M58N100
M87N100
–
20 -
17 -
11 -
5 -
M protein
LL-37
B
LL-37 MIC [μM]
30
20
10
0
emm1
+M1HB
+M58N100
+M87N100
+GCN4-M87
Δemm1
C
M87N100
GCN4-M87
–
17 -
11 -
5 -
M protein
LL-37

Figure 2

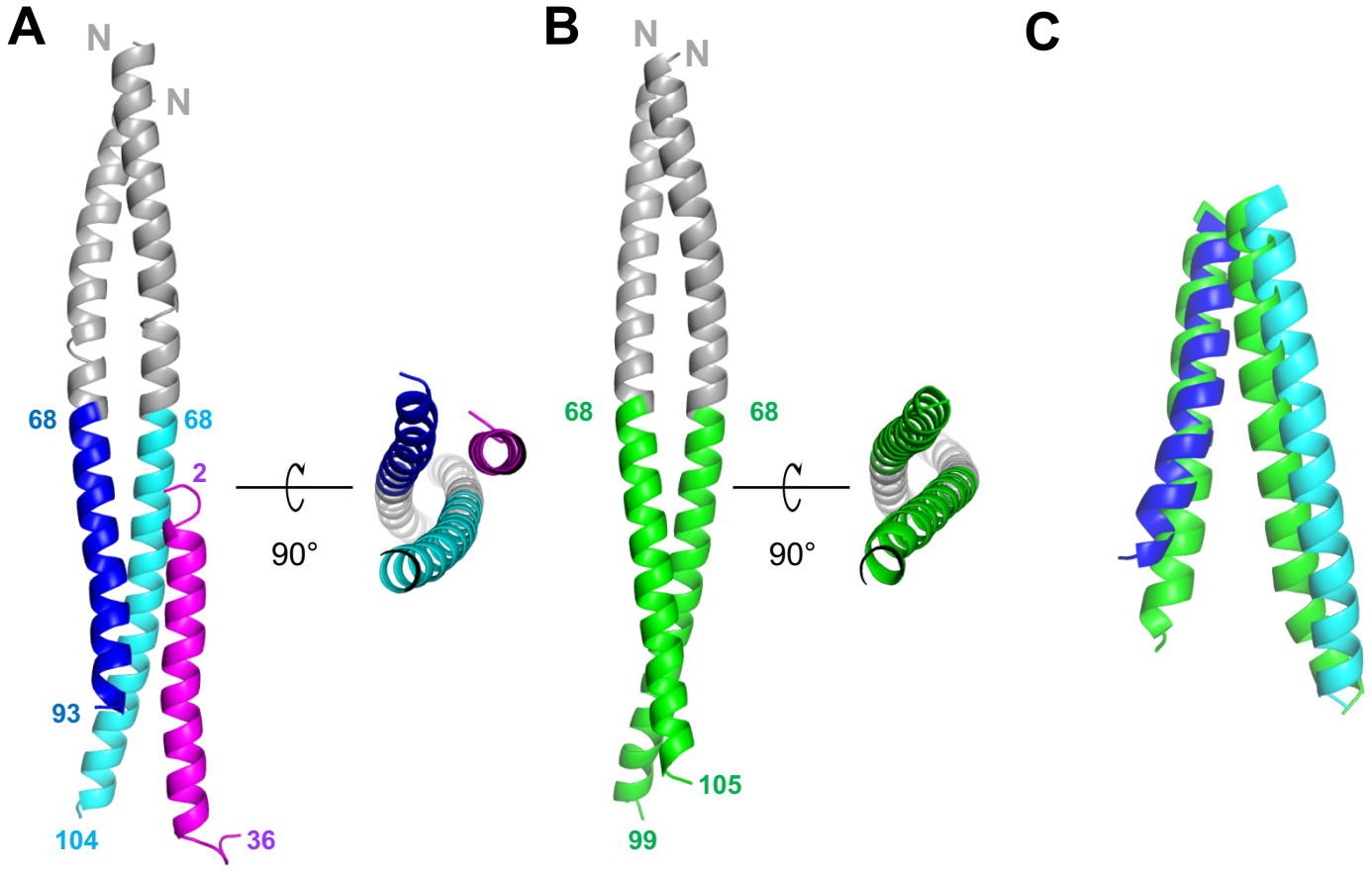


Figure 3

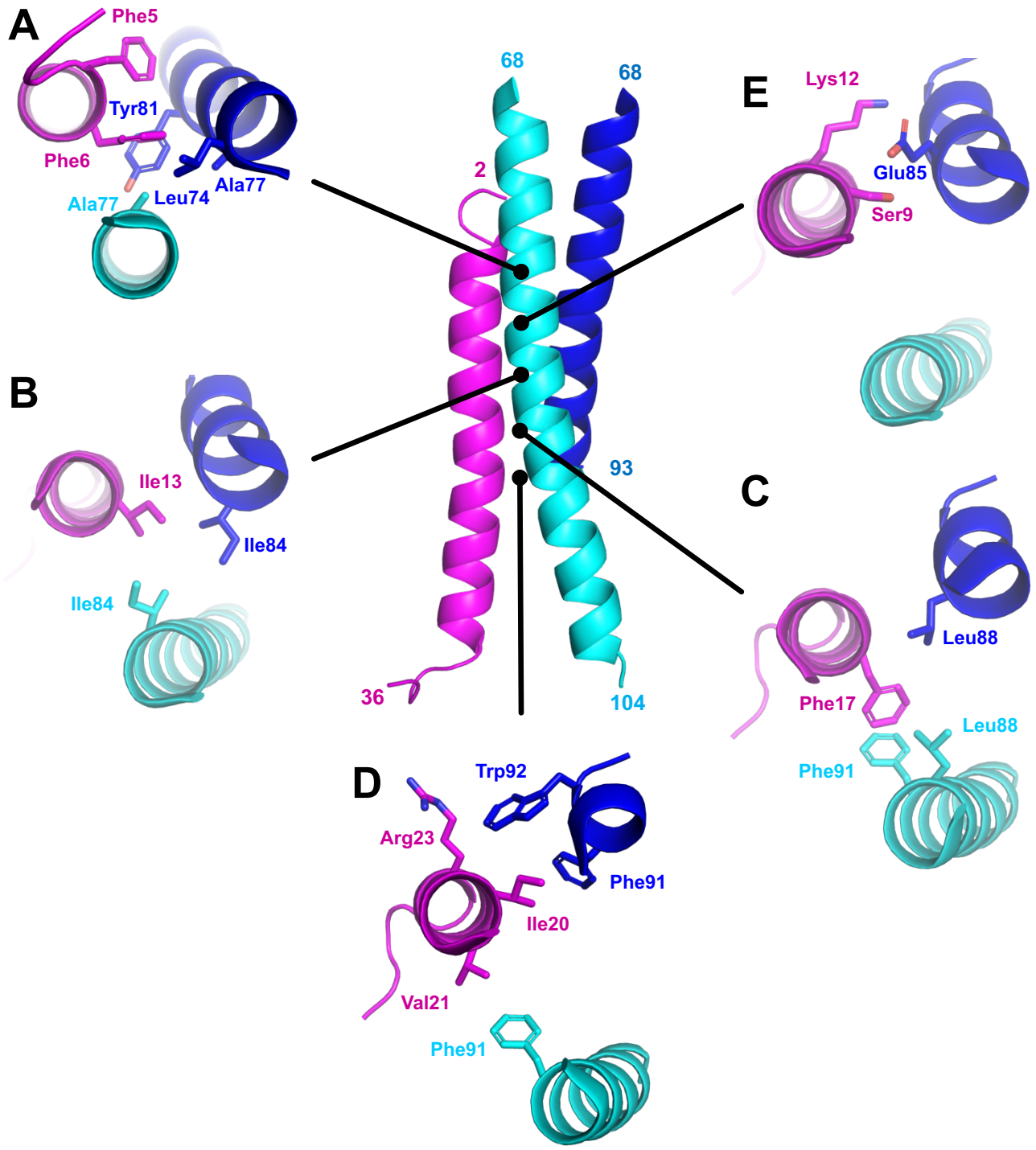

A
Phe5
Tyr81
Phe6
Ala77
Leu74
Ala77
B
Ile13
Ile84
Ile84
68
68
2
93
36
104
E
Lys12
Glu85
Ser9
C
Leu88
Phe17
Leu88
Phe91
D
Trp92
Arg23
Phe91
Ile20
Val21
Phe91

Figure 4

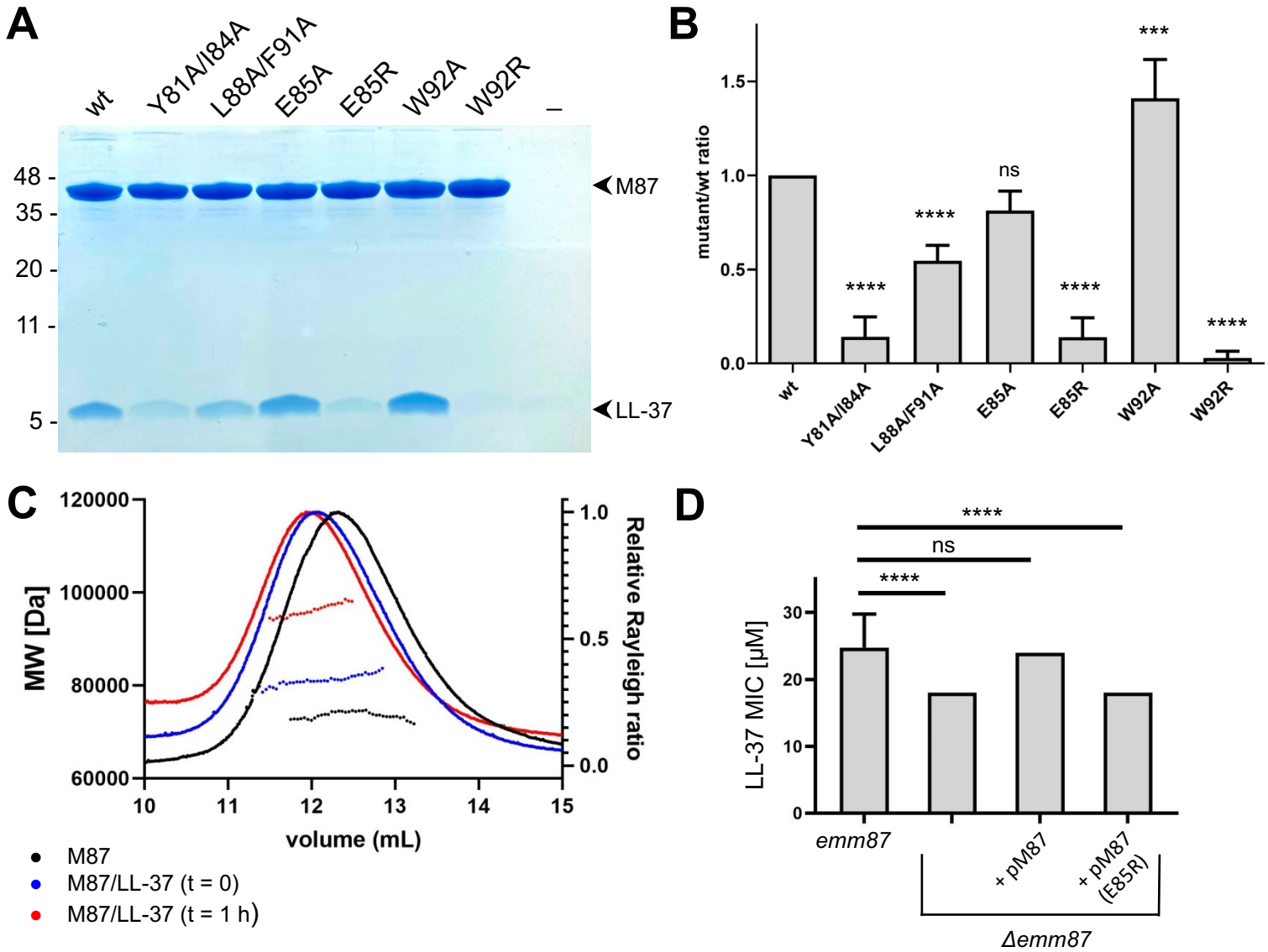
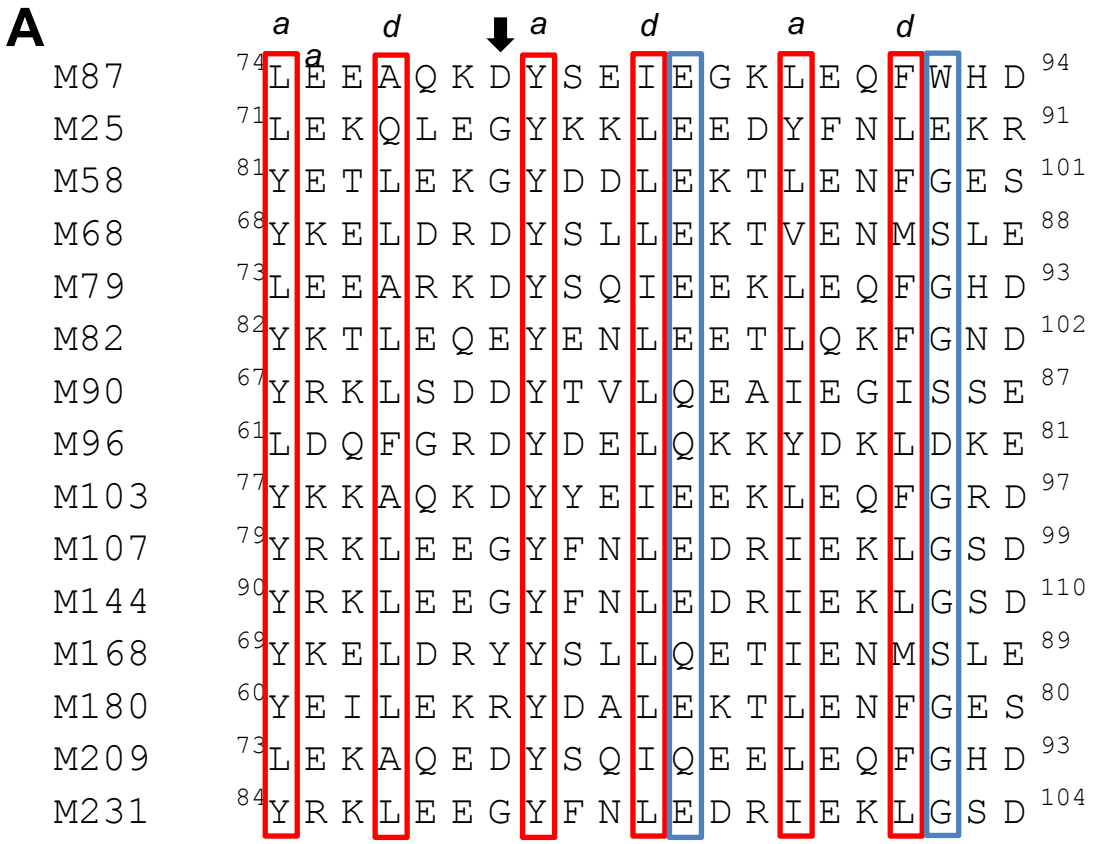
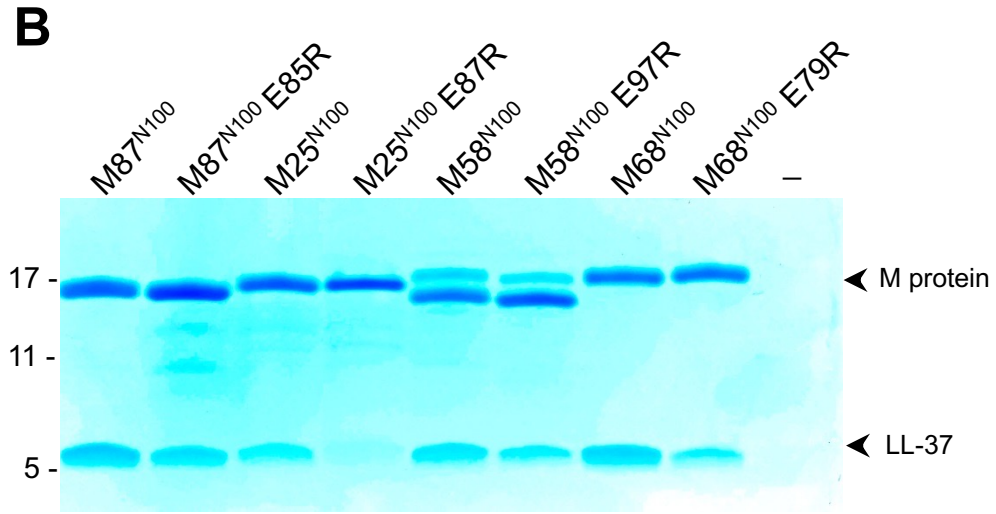
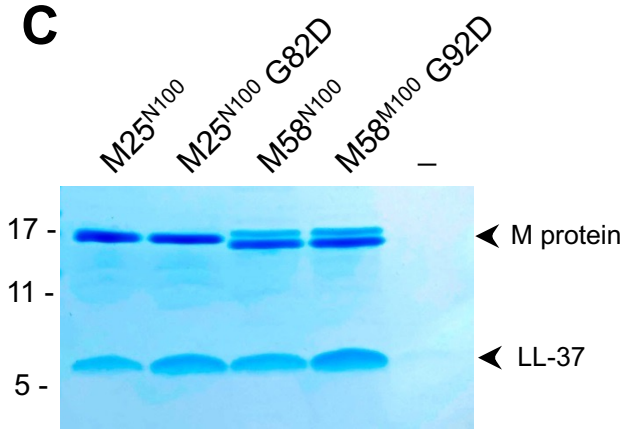

A
wt
Y81A/I84A
L88A/F91A
E85A
E85R
W92A
W92R
–
48
35
20
11
5
M87
LL-37
B
mutant/wt ratio
1.5
1.0
0.5
0.0
***
ns
****
****
****
****
wt
Y81A/I84A
L88A/F91A
E85A
E85R
W92A
W92R
C
MW [Da]
120000
100000
80000
60000
Relative Rayleigh ratio
1.0
0.5
0.0
10
11
12
13
14
15
volume (mL)
M87
M87/LL-37 (t = 0)
M87/LL-37 (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

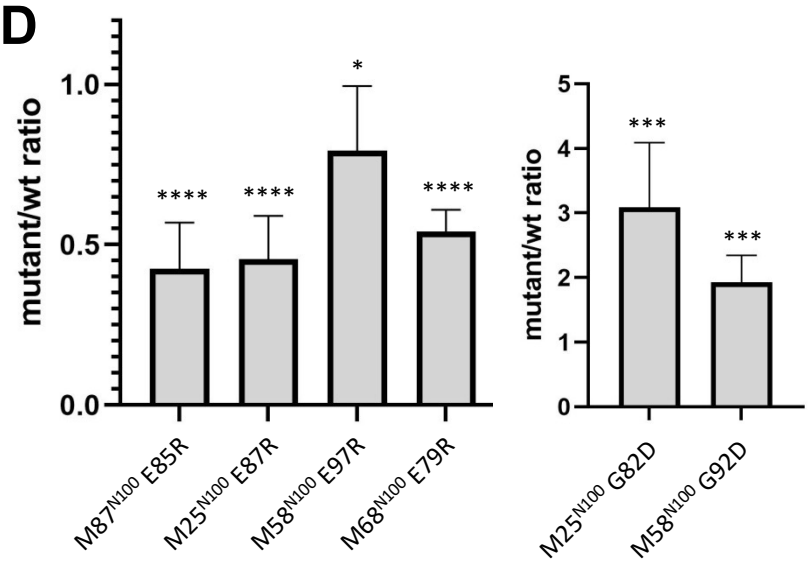
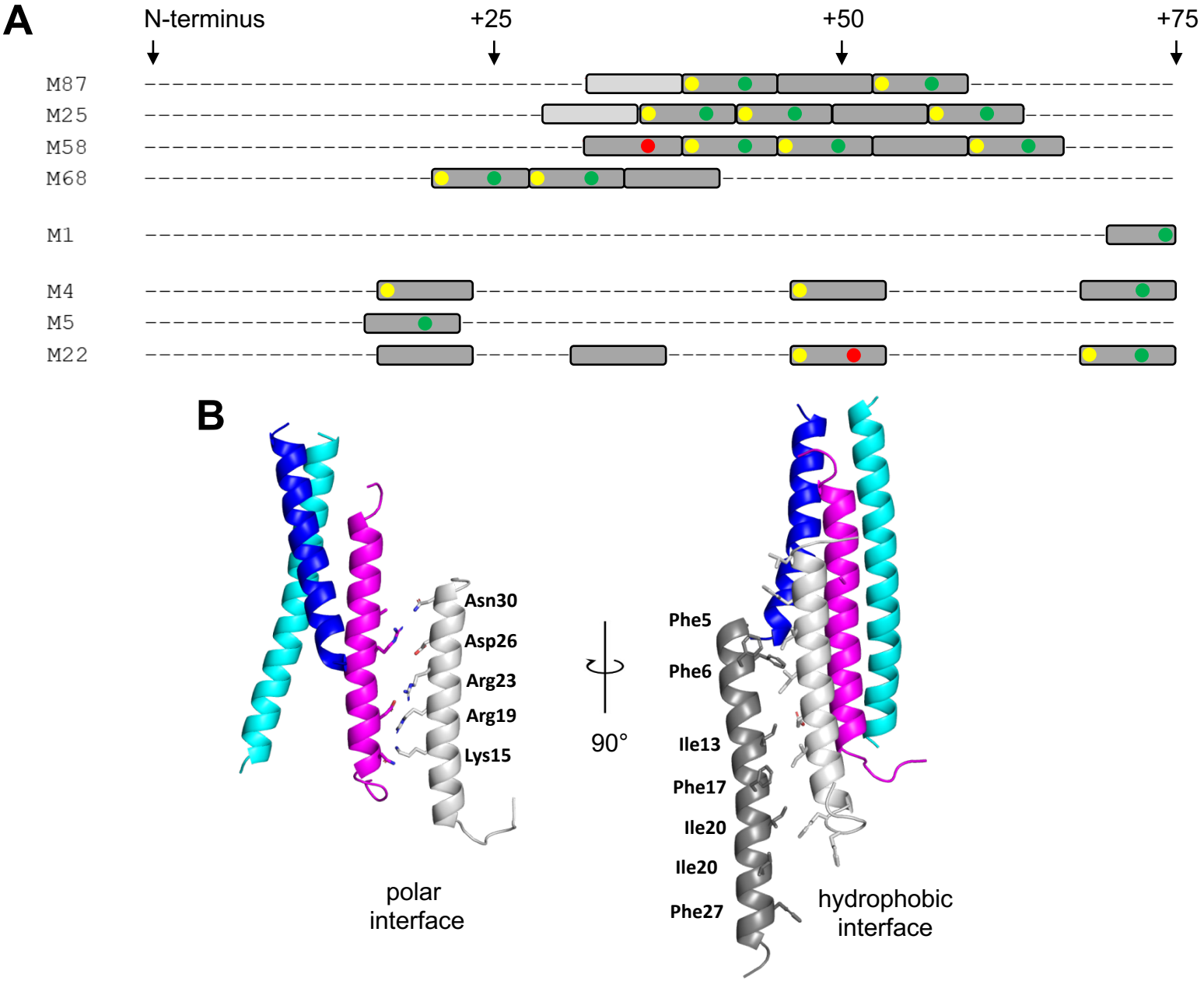

D
mutant/wt ratio
1.0
0.5
0.0
****
****
*
****
M87N100 E85R
M25N100 E87R
M58N100 E97R
M68N100 E79R
mutant/wt ratio
5
4
3
2
1
0
***
***
M25N100 G82D
M58N100 G92D

Figure 6


A
N-terminus
+25
+50
+75
M87
M25
M58
M68
M1
M4
M5
M22
B
Asn30
Asp26
Arg23
Arg19
Lys15
90°
Phe5
Phe6
Ile13
Phe17
Ile20
Ile20
Phe27
polar interface
hydrophobic interface

SUPPLEMENTARY FIGURE LEGENDS

Figure S1. LL-37 Binding.

LL-37 binding to His-tagged $M1^{HB}$, $M4^{N100}$, $M5^{N100}$, $M22^{N100}$, $M28^{N100}$, $M44^{N100}$, $M77^{N100}$, or $M89^{N100}$ as determined by a Ni^{2+}-NTA agarose bead co-precipitation assay at 37 °C. The last lane contains no M protein. Bound fractions were resolved by SDS-PAGE and visualized by Coomassie staining. The gel is representative of three experiments.

Figure S2. Structures of the GCN4-M87/LL-37 complex and free GCN4-M87.

A. Electron density for GCN4-M87 from a simulated annealing σA-weighted 2mFo-DFc composite omit map contoured at 1.4σ. Part of the interface between M87 (left) and LL-37 (right) is shown.

B. Superposition of the two GCN4-M87/LL-37 complexes present in the asymmetric unit of the crystal. LL-37 is in different shades of magenta. The GCN4 portions are in shades of gray, and the M87 portions are in blue and cyan.

C. Electron density for free GCN-M87 from a simulated annealing σA-weighted 2mFo-DFc composite omit map contoured at 1.4σ.

Figure S3. Structure and Stability of M87 mutant proteins.

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

B. Melting curves of intact wild-type and mutant M87 proteins, as monitored by the CD mean molar residue ellipticity at 222 nm. The data are an average of two independent measurements.

Figure S4. Self-limiting growth of the M87/LL-37 complex.

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

Figure S5. LL-37 binding motif.

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

Table S1. 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

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

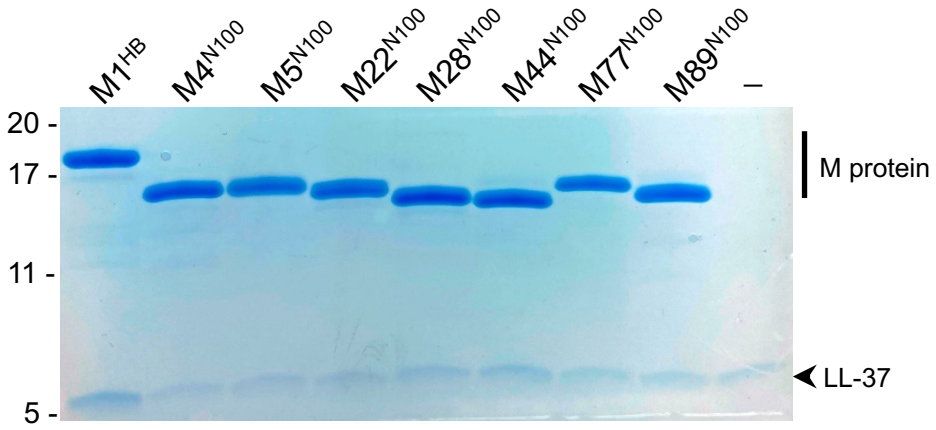
[b]$R_{meas} = \sum_{hkl} \sqrt{\frac{n}{n-1}} \sum_{j=1}^{n} |I_{hkl\,j} - \langle I_{hkl} \rangle| / \sum_{hkl} \sum_{j} I_{hkl\,j}$

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

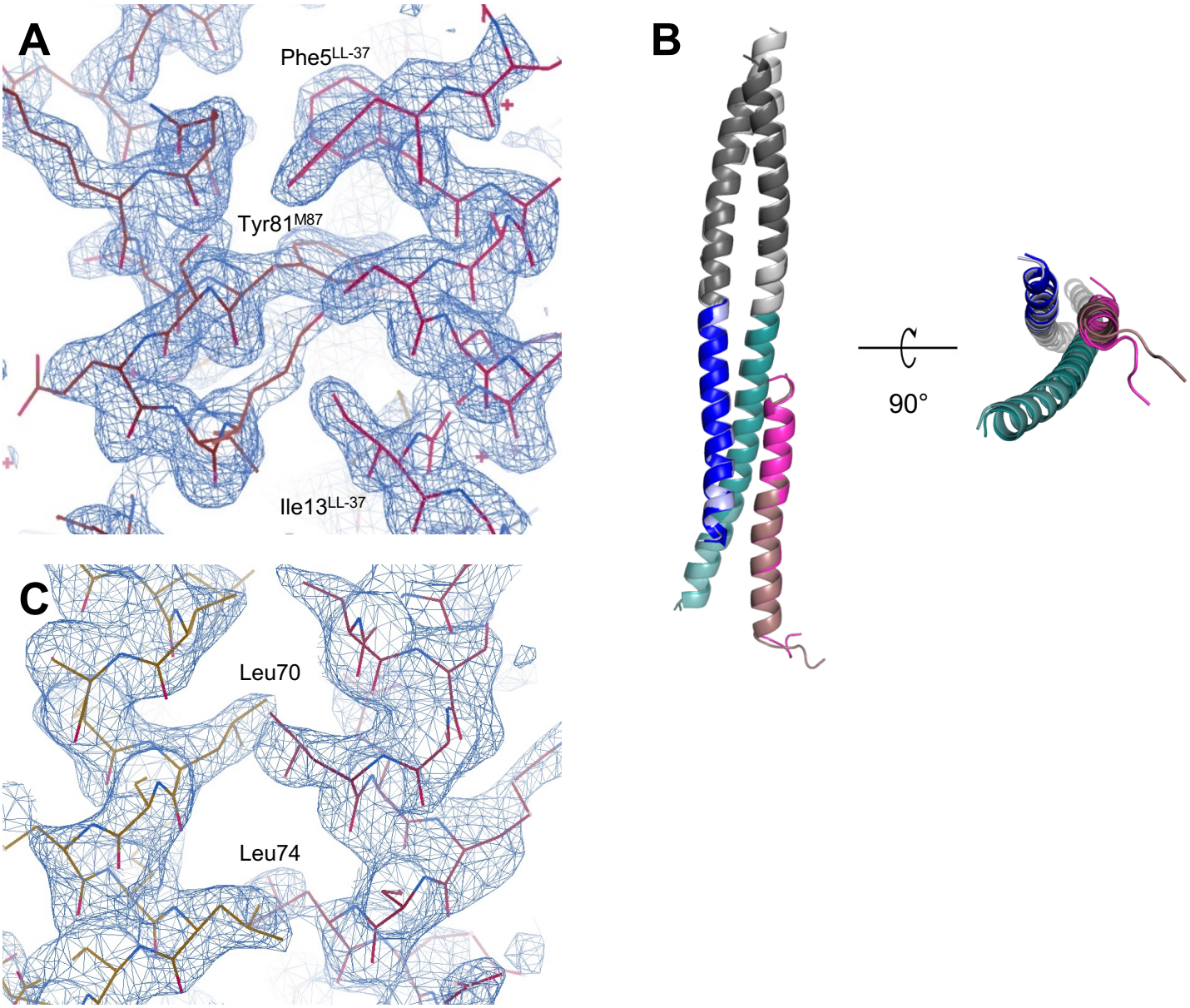
[d]$R_{work} = \sum_{hkl} |F_{obs} - F_{calc}| / \sum_{hkl} F_{obs}$

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}.

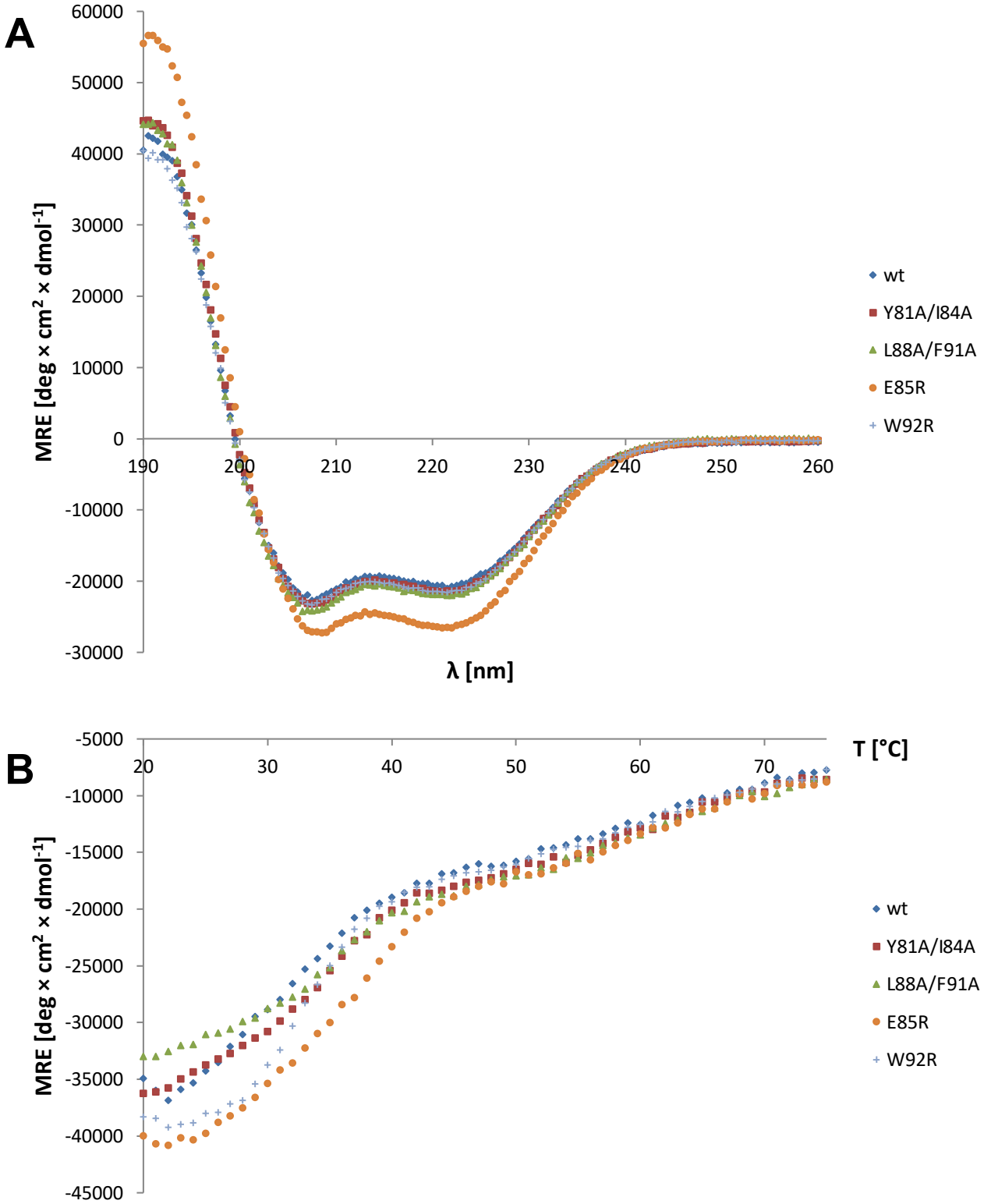
Supplementary Figure S1



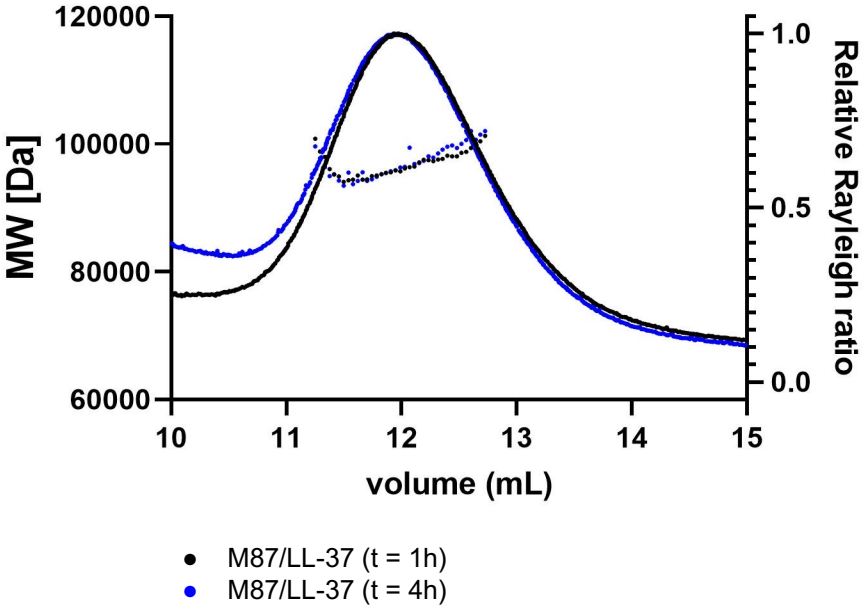
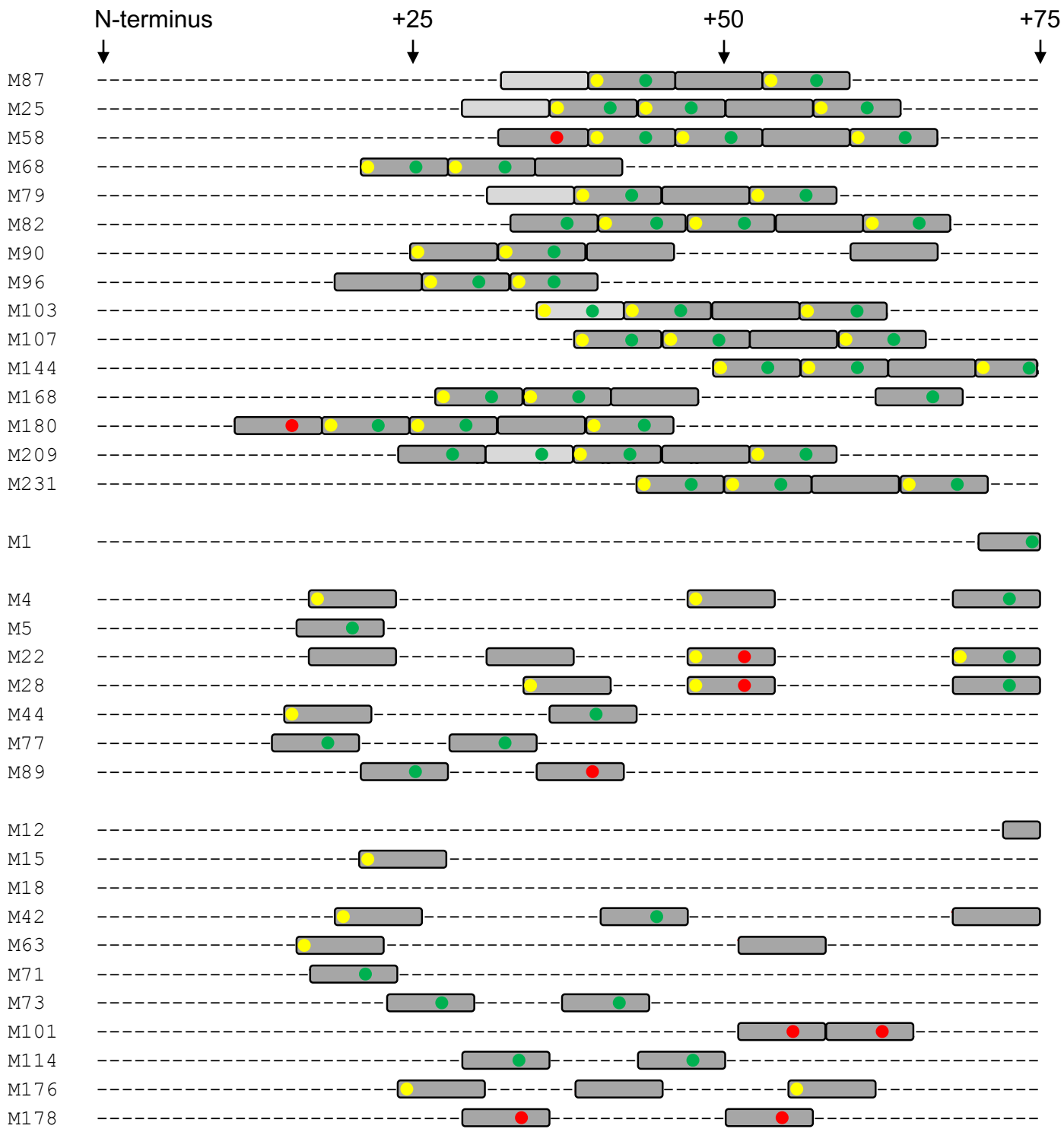
Supplementary Figure S2



Supplementary Figure S3



Supplementary Figure S4



Supplementary Figure S5