Supplementary Material to

Optimization of designed armadillo repeat proteins by molecular dynamics simulations and NMR spectroscopy

Pietro Alfarano[1¥], Gautham Varadamsetty[1¥], Christina Ewald[2], Fabio Parmeggiani[1], Riccardo Pellarin[1], Oliver Zerbe[2], Andreas Plückthun[1*], Amedeo Caflisch[1*]

[1] Department of Biochemistry, University of Zürich, Winterthurerstrasse 190, CH-8057 Zürich, Switzerland
[2] Department of Organic Chemistry, University of Zürich, Winterthurerstrasse 190, CH-8057 Zürich, Switzerland

Supplementary Materials and Methods

Design and synthesis of DNA encoding designed ArmRPs

The cloning, expression, and purification of designed ArmRPs was carried out essentially as described previously.[1] All primers were synthesized and purified by Microsynth GmbH (Balgach, Switzerland) (**Table S1**). The melting temperature (T_m) was calculated as described (http://www.stratagene.com/manuals/200519.pdf). First, single modules were assembled. In case of the KK-module assembly, in a first step partially overlapping primer pairs (1)-(2), (3)-(4) and (5)-(6) (**Table S1**) were annealed and the double strand was completed by PCR. Then, 2 µl from these PCR reaction mixtures were combined as templates for a second assembly reaction in the presence of oligonucleotides (1) and (6). All the oligonucleotides were used at final concentrations of 1 µM. The annealing temperature was 50°C for the first and second reaction. Thirty PCR cycles were performed with an extension time of 30 s. The same procedure was applied for the internal and other capping repeats. Four oligonucleotides were used for the N-terminal capping repeats.

BamHI and KpnI restriction sites were used for direct insertion of modules into plasmid pQE30 and the modules were sequenced. The single modules were PCR amplified from the vectors, using external primers pQE_f_1 and pQE_r_1. Neighboring

modules were digested with restriction enzymes BpiI and BsaI and directly ligated together. The genes coding for the whole proteins were assembled by stepwise ligation of the internal and capping modules. BamHI and KpnI restriction sites were used for insertion of whole genes into the vector pPANK. Proper assembly of constructs was validated by DNA sequencing.

Q and K mutations were introduced by QuikChange mutagenesis, carried out in 50 µl with 50-100 ng template, 0.4-2 µM primer pair, 200 µM dNTPs and 2 u of *Pfu* DNA polymerase (Stratagene, CA). The reaction was initiated by pre-heating to 94°C for 3 min; followed by 18 cycles of 94°C for 1 min, 52°C for 1 min and 68°C for 15 min, followed by incubation at 68°C for 1 h. Chemically competent E. coli XL1-Blue cells were transformed with an aliquot of 5 µl of DpnI-digested PCR products (to remove the original plasmid serving as PCR template) and inoculated on Luria–Bertani (LB) agar plates containing 100 µg/ml ampicillin. To further remove any residual wild-type plasmid from potentially doubly transformed cells a total of 5 colonies for each mutant were picked and their plasmids were isolated by mini-prep, digested, and the insert was ligated into fresh vector. The positive mutants were verified by sequencing analysis.

For the generation of N-cap and C-cap mutants, the pQE30-based plasmid pPANK[1] containing the designed ArmRP YM_4A gene fragment was used as the PCR mutagenesis template for introducing N- and C-cap mutations.

To assemble the entire YM_4A protein, the modules were digested with the type IIS restriction enzymes BpiI and BsaI and directly ligated together and this step was repeated to result in the quadruple M_4 pieces that were then ligated to the corresponding caps. BamHI and KpnI restriction sites were used for insertion of the whole genes into the vector pPANK and the plasmids were sequenced.

Reference

1. Parmeggiani F, Pellarin R, Larsen A P, Varadamsetty G, Stumpp M T, Zerbe O, Caflisch A Plückthun A (2008) Designed armadillo repeat proteins as general peptide-binding scaffolds: consensus design and computational optimization of the hydrophobic core. J. Mol. Biol. **376**: 1282-1304.

Table S1: List of oligonucleotides used for generating point mutants

Oligonucleotides		
name	**sequence 5'-3' direction**	**description (for=forward, rev=reverse)**
Y_{II}_f	GCTACCCGTAAATTCTCTCAGATCCTGTCTGATGGTAACGAACAAATC	for primer for introducing three mutations in N-cap
Y_{II} _r	CCATCAGACAGGATCTGAGAGAATTTACGGGTAGCAGACAGCTGTTCC	rev primer for introducing three mutations in N-cap
C Q240M_f	CTGGAGAAGATGTTCTCCCACTAATGAGGTACCCCGG	for primer for introducing Met at position 240 in C-cap
C Q240M_r	GTGGGAGAACATCTTCTCCAGAGCTTCCTGAGCTTCTTTC	rev primer for introducing Met at position 240 in C-cap
C Q240L_f	CTGGAGAAGCTGTTCTCCCACTAATGAGGTACCCCGG	for primer for introducing Leu at position 240 in C-cap
C Q240L_r	GTGGGAGAACAGCTTCTCCAGAGCTTCCTGAGCTTCTTTC	rev primer for introducing Leu at position 240 in C-cap
C Q240M,F24Q_f	CTGGAGAAGATGCAGTCCCACTAATGAGGTACCCCGGGTC	for primer for introducing Met at position 240 and Gln at position 241 in C-cap
C Q240M,F241Q_r	CATTAGTGGGACTGCATCTTCTCCAGAGCTTCCTGAGCTTCTTTC	rev primer for introducing Met at position 240 and Gln at position 241 C-cap
A_{II} cap_f	CTGGAGAAGCTGCAGTCCCACTAATGAGGTACCCCGGGTCG	for primer for introducing Leu at position 240 and Gln at position 241 in C-cap
A_{II} cap_r	CATTAGTGGGACTGCAGCTTCTCCAGAGCTTCCTGAGCTTCCTGAGCTTCTTTC	for primer for introducing Leu at position 240 and Gln at position 241 in C-cap
Cons_1_for (1)	CCAGGGATCCTAGGAAGACCTTGGTAACGAACAAATCC	for assembly consensus M module of KK type and amplification
2A_rev (2)	AGCCGGCAGAGCACCAGCATCGATAACAGCTTGGATTTGTTCGTTACCAAGG	rev assembly consensus M module of KK type
3A_for (3)	GGTGCTCTGCCGGCTCTGGTTCAACTGCTGTCCTCTCCGAACG	for assembly consensus M module of KK type
4L_rev (4)	CCACAGAGCTTCTTTCAGGATCTTCTCGTTCGGAGAGGACAGC	rev assembly consensus M module of KK type
5L-I_for (5)	CTGAAAGAAGCTCTGTGGGCTCTGTCTAACATCGCTTCTGGTGGTTGAG	for assembly consensus M module of KK type
6I_rev (6)	TTCCTGGTACCCTAAGGTCTCAACCACCAGAAGCGAT	rev assembly consensus M module of KK type and amplification
KQ_for	CCGAACGAGAAGATCCTGCAAGAAGCTCTGTGGGC	for mutation KK->KQ
KQ_rev	GCCCACAGAGCTTCTTGCAGGATCTTCTCGTTCGG	rev mutation KK->KQ
QK_for	CCTCTCCGAACGAGCAGATCCTGAAAGAAGC	for mutation KK->QK
QK_rev	GCTTCTTTCAGGATCTGCTCGTTCGGAGAGG	rev mutation KK->QK
QQ_for	CCTCTCCGAACGAGCAGATCCTGCAAGAAGCTCTGTGGGC	for mutation KK->QQ
QQ_rev	GCCCACAGAGCTTCTTGCAGGATCTGCTCGTTCGGAGAGG	rev mutation KK->QQ
pQE_f_1	CGGATAACAATTTCACACAG	forward primer for pQE vectors
pQE_r_1	GTTCTGAGGTCATTACTG	reverse primer for pQE vectors

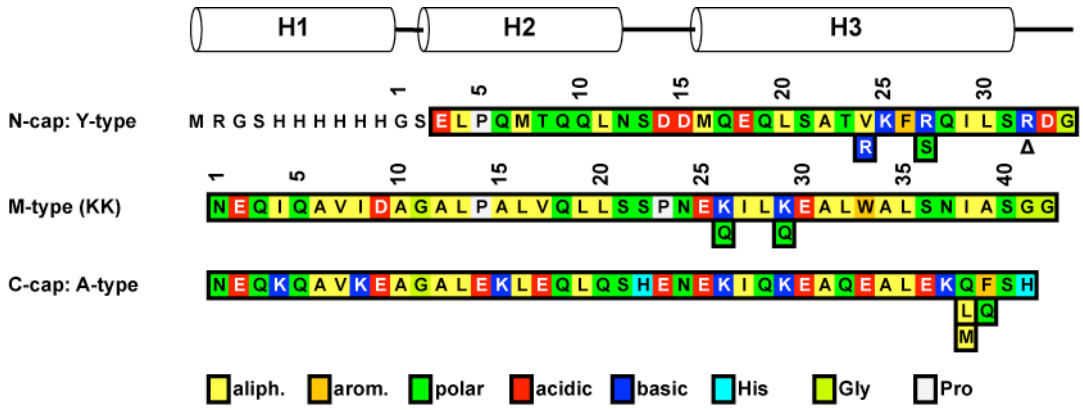


Fig. S1 Sequence of the armadillo repeats and mutants studied. In the proteins studied experimentally, the N-cap is preceded by the sequence MRGSHHHHHGS for purification and detection. Experimentally, YM_3A and YM_4A molecules have been tested. KK refers to the internal repeat as shown, QK with only a substitution at position 26, KQ with only a substitution at position 29, and QQ with both, as shown underneath. Y_{II} refers to an N-cap carrying all three mutations indicated underneath, A_{II} refers to a C-cap with the mutations L and Q indicated underneath.

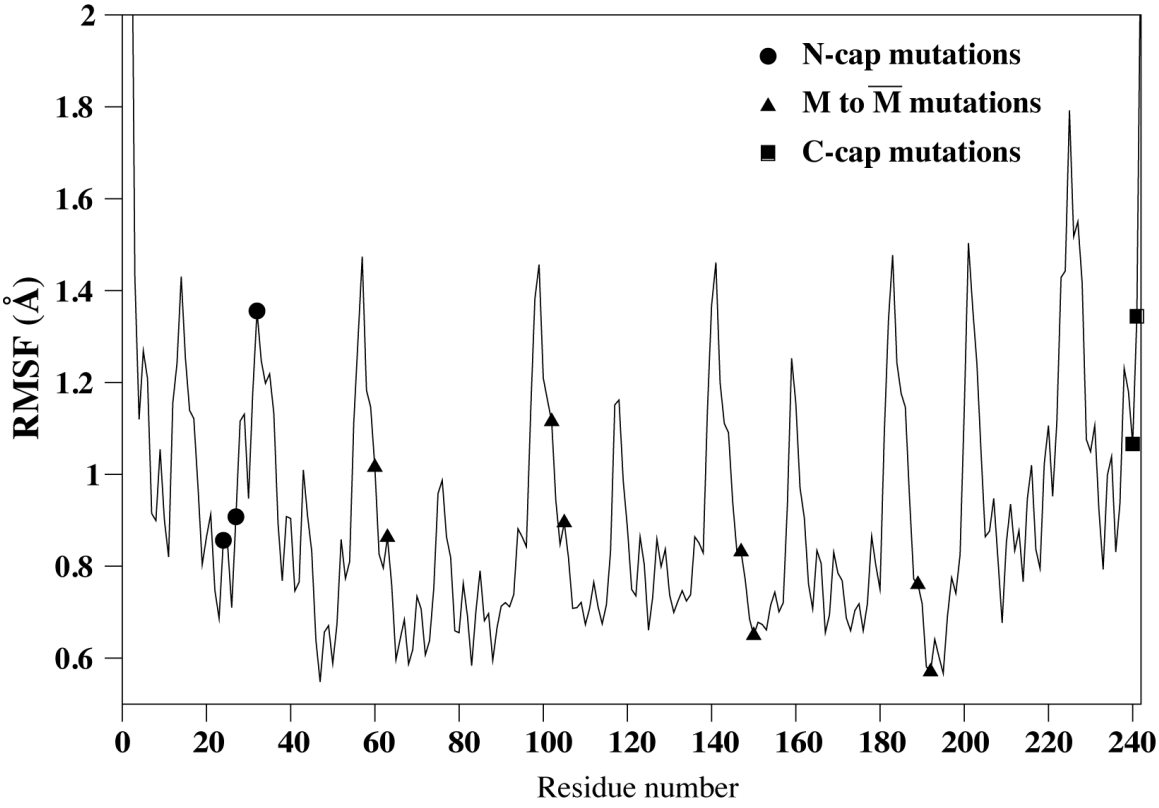


Fig. S2 RMSF plot of Cα atoms of YM_4A in the implicit solvent simulations. The circles show the position of the residues suggested for mutations at the N-cap: Val-24, Arg-27 and Arg-32. The triangles show the internal repeat positions Lys-26 and Lys-29; and the squares the residues suggested for mutations at the C-cap: Gln-240 and Phe-241.

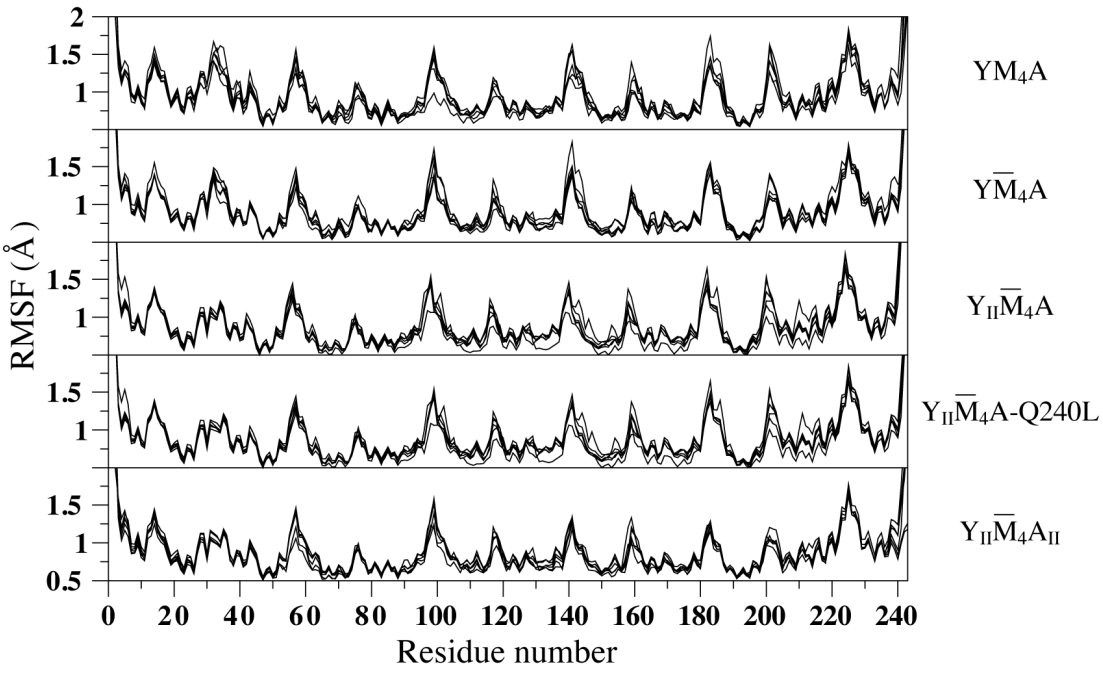


Fig. S3 RMSF plot superposition of all implicit solvent simulations.

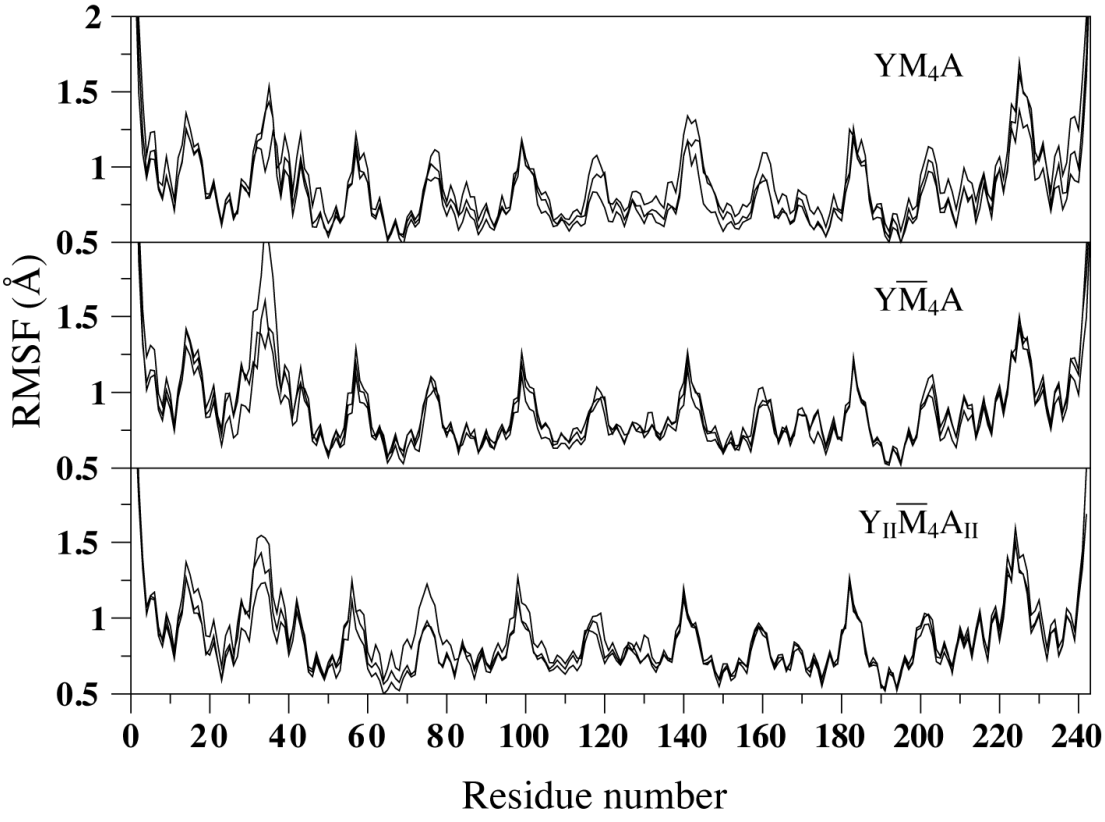


Fig. S4 RMSF plot superposition of all explicit solvent simulations.

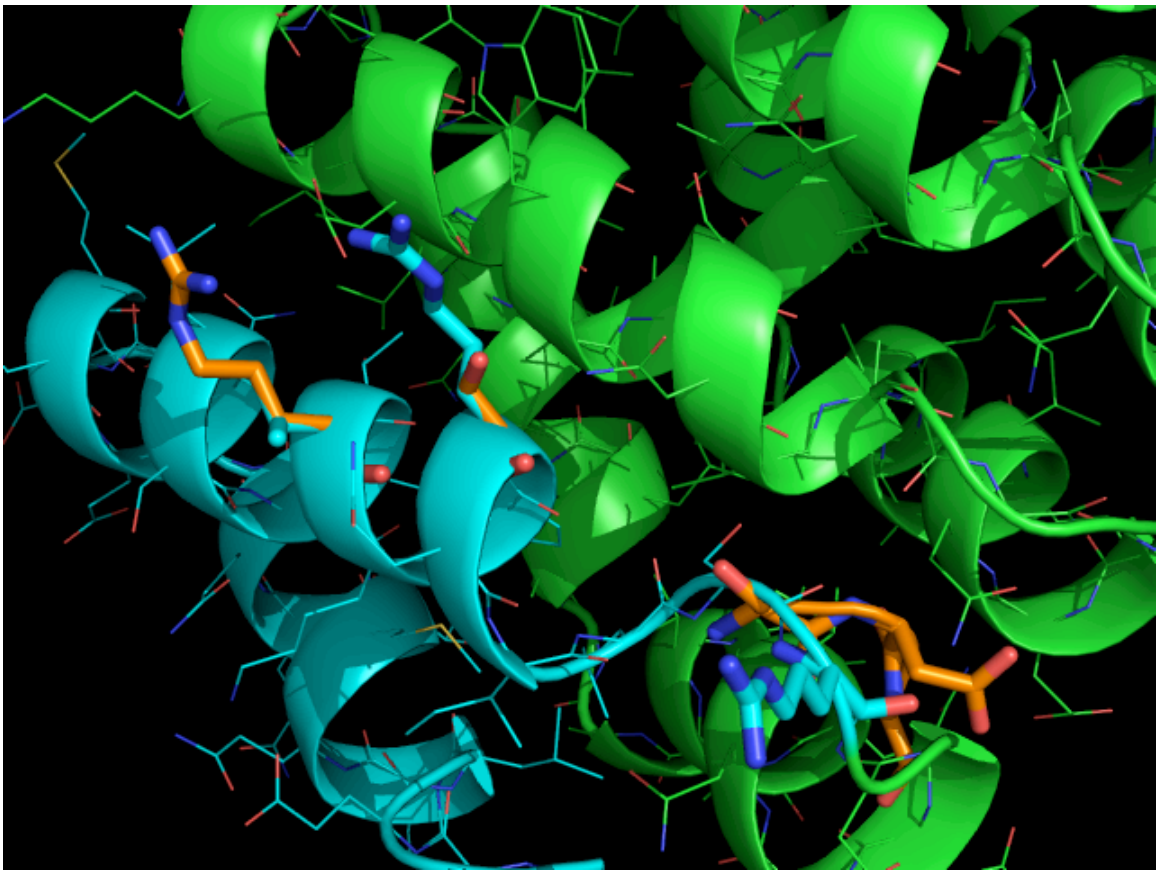
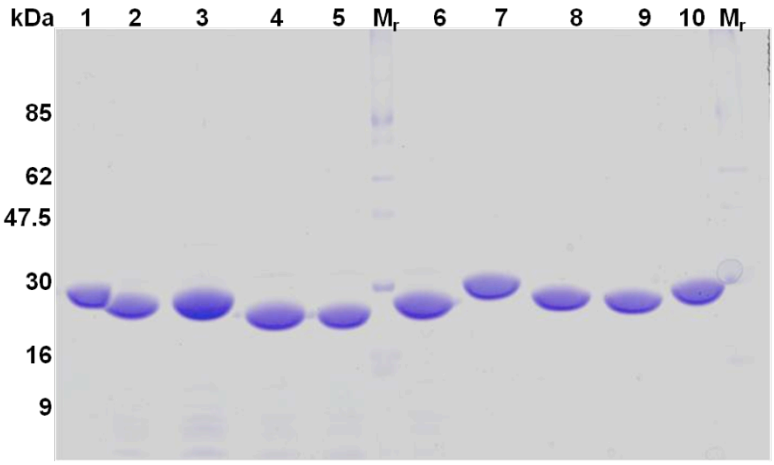


Fig. S5 Mutations introduced into the N_{II}-cap compared to wild-type. Wild-type residues (V24, R27 and R32) are in depicted in cyan while the three mutations introduced in the N_{II}-cap (R24, S27 and ΔR32 (deletion of arginine) are shown in orange. Helices colored in cyan refer to the N_{II}-cap while helices in green to the internal repeats.



Lane 1: Y$\overline{M}_4$A
Lane 2: $Y_{II}\overline{M}_4$A
Lane 3: Y$\overline{M}_4$A-Q240M
Lane 4: Y$\overline{M}_4$A-Q240L
Lane 5: $Y_{II}\overline{M}_4A_{II}$
Lane 6: $Y_{II}M_4A_{II}$
Lane 7: $Y_{II}M_4$A
Lane 8: YM_4A-Q240M
Lane 9: YM_4A-Q240L
Lane 10: YM_4A

Figure S6: Single-step IMAC purification of YM/$\overline{M}_4$A N-and C-cap variants. The expected size is approximately 27 kDa. Proteins were analyzed by SDS-PAGE (15%). The size marker is indicated in kDa.

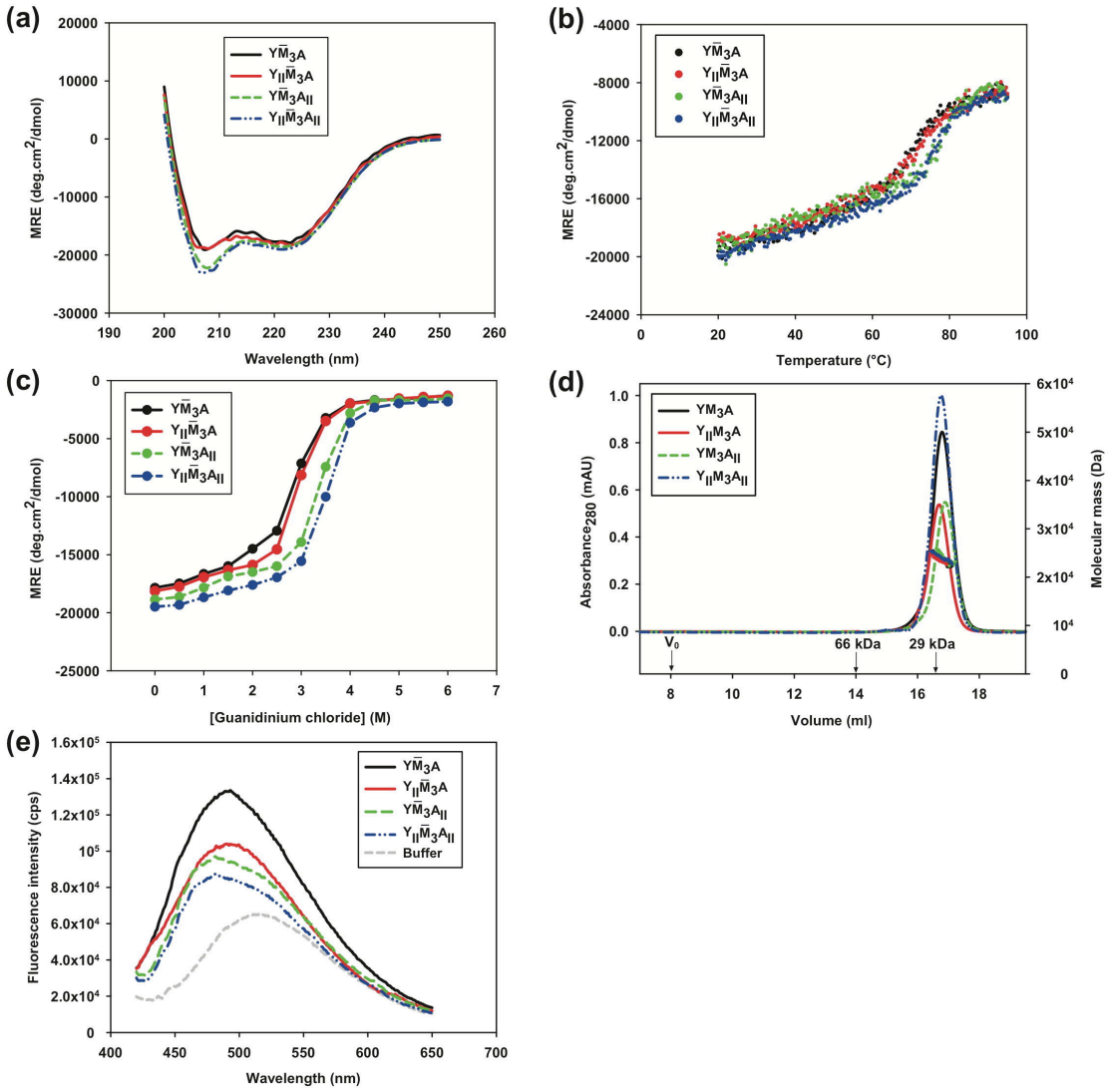


Figure S7: Biophysical characterization of designed ArmRPs $Y\bar{M}_3A$ and its cap variants ($Y_{II}\bar{M}_3A$, $Y\bar{M}_3A_{II}$ and $Y_{II}\bar{M}_3A_{II}$). **(a)** CD spectra, **(b)** thermal denaturation curves and (**c)** GdnCl-induced denaturation curves. The denaturation experiments were followed by CD. The values of MRE at 222 nm are reported. **(d)** SEC and MALS of designed ArmRPs. The absorbance at 280 nm from SEC is shown on the left y-axis, the calculated MW from MALS on the right y-axis. V_0 indicates the void volume of the column. Bovine serum albumin (MW 66 kDa), and carbonic anhydrase (MW 29 kDa) were used as molecular weight markers, and the corresponding elution volumes are indicated by the arrows. **(e)** ANS binding. The values without buffer subtractions are shown. The protein concentration used was 10 µM for a,b,c,e and 30 µM for d.

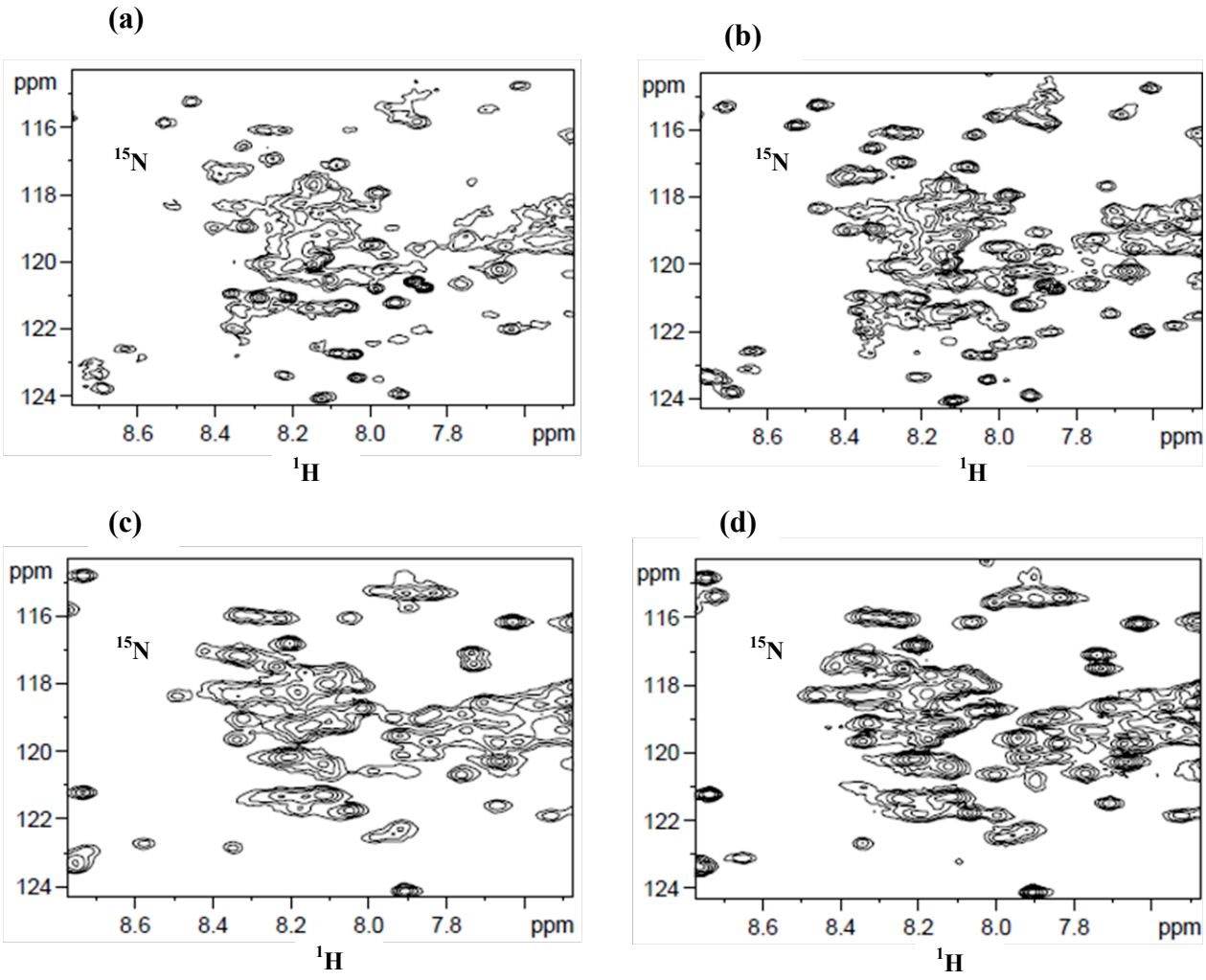


Figure S8: [^{15}N, ^{1}H]-HSQC spectra of designed ArmRP Y$\overline{\text{M}}_4$A **(a)** and its cap variants $Y_{II}\overline{M}_4A$ **(b),** $Y\overline{M}_4A_{II}$ **(c)** and $Y_{II}\overline{M}_4A_{II}$ **(d)** at pH 7.4. All spectra were recorded at a temperature of 310 K in 50 mM phosphate buffer, and 150 mM NaCl. The protein concentration was 0.5 mM. All proteins are of the QQ ($\overline{\text{M}}$) – type.