Supplementary Materials

Ewald et al., A Combined NMR and Computational Approach to Investigate Peptide Binding to a Designed Armadillo Repeat Protein

Content

Overview of figures

Tables

1. Sequence details

Figure S1 The sequence of proteins based on $Y_IMR^1R^2R^3MA_{II}$ and associated modifications and mutations. The proteins were either produced with a non-cleavable his tag as encoded in the vector pPANK (first line) or a cleavable his tag as encoded in the vector pLIC (second line). After cleavage, only a single Gly is left. For example, the sequence of $Y_IMR^1R^2R^3MA_{II}$ includes the non-cleavable His_6-tag followed by the sequence of the capping and repeat sequences given in the subsequent lines.

Shorter versions, for example $Y_IMR^1R^2A_{II}$, in turn include the sequences of the pPANK-based His_6-tag, the Y_I N-cap, the first unrandomized M-type repeat, randomized repeats R^1 and R^2 and the A_{II} C-cap. N-terminally truncated proteins (not containing an N-cap) are headed by a TEV-cleavable His_6-tag (pLIC-based), which leaves a single N-terminal Gly after cleavage, and include the repeats indicated by their name, e.g. R^3MA_{II}. Due to the glycine residue left after TEV-cleavage (ENLYFQ/G) at the N-terminus of the fragment, N-terminally truncated proteins effectively start with the last glycine of the preceding repeat, which is naturally present in the sequence. To facilitate cloning, the $MR^1R^2R^3MA_{II}$ fragment starts two residues N-terminal of the glycine, adding "RD" to the sequence. C-terminally truncated proteins end at position 41 of their last respective internal repeat, e.g. $Y_IMR^1R^2$ ends with the glutamate in position 41 of the R^2 repeat.

Helices 1, 2 and 3 of each repeat and the C-cap are indicated by light blue boxes. the N-cap only contains two helices, H1 and H2. Randomized positions in H3 of internal repeats are indicated in red. Spontaneous point mutations acquired during ribosome display are indicated in green. Sites of mutations introduced in different cap generations to improve protein stability are indicated in bold and underlined: V34, R37 and R42 (continuous numbering including pPANK-based His_6-tag) are mutated to convert a Y_I N-cap into a Y_{II} N-cap (R34, S37, R42Δ); The LQ motif at the end of the A_{II} C-cap corresponds to QF in a A_I C-cap. Conserved asparagine positions in helix 3 of each repeat are N81, N123, N165, N207 and N249.

Table S1: Peptides, uncapped (unless indicated otherwise).

Peptide name	Sequence	Purpose
NT	pGlu-LYENKPRRPYIL	CSP and PRE-studies
NT7-13	Ac-PRRPYIL	CSP
NTQ1C	CLYENKPRRPYIL	PRE-studies
NTK6C	pGlu-LYENKPRRPYIC	PRE-studies
NTL13C	pGlu-LYENCPRRPYIL	PRE-studies
NT_{biotin}	Biotin-6ACA-βA-βA-NT	ELISA

Ac, acetyl; 6ACA, 6-amino-caproic acid; βA, beta-alanine; CSP, chemical shift perturbations; ELISA, enzyme linked immunosorbent assay; NT, neurotensin; pGlu, pyro-glutamate; PRE, paramagnetic relaxation enhancement.

Table S2 Details of Cloning: Oligonucleotides

Name	Sequence 5'-3' direction	Purpose
078_RMAm_FWD	GAA AAT TTA TAT TTT CAG GGT AAC GAA CAA CAA	Construction of R^3MA_{II} from pPANK_N_IM328MC_{II}
077_R23MAm_FWD	GAA AAT TTA TAT TTT CAG GGT AAC GAA CAA ACC	Construction of $R^2R^3MA_{II}$ and $R^1R^2R^3MA_{II}$ from pPANK_N_IM328MC_{II}
090_MRRRMAm_v10_FWD	GAA AAT TTA TAT TTT CAG GGT CGT GAT GGT AAC GAA CAA	Construction of $MR^1R^2R^3MA_{II}$ from pPANK_N_IM328MC_{II}
088_RxMAm_v3_REV	AGA TGA GAG TAA GGC TAT CAT TAG TGG GAC TGC AG	Reverse primer for 077 ($R^2R^3MA_{II}$), 078 (R^3MA_{II}) and 090 ($MR^1R^2R^3MA_{II}$)
053_B328_YMR_x_pPANK_Fwd	TAA TGA GGT ACC CCG GGT CGA CCT GCA GCC	Construction of $Y_IMR^1R^2R^3$ and Y_IMR^1 from pPANK_NyM328MCa encoding VG_328
054_B328_YMRRR_pPANK_REV	CAG TTC AGC GAT GTT AGT CAG AGC GTC CAG	Reverse primer for 053 ($Y_IMR^1R^2R^3$) and 101 ($Y_IMR^1R^2R^3A_{II}$)
056_B328_YMR_pPANK_REV	AGC GAA AGC GAT GTT GTT CAG AGC GAT AAG	Reverse primer for 101 ($Y_IMR^1A_{II}$) and 053 (Y_IMR^1)
063_YMRR_REV	TTC CAT AGC GAT GTT AGT C	Reverse primer for 101 ($Y_IMR^1R^2A_{II}$) and 65 ($Y_IMR^1R^2$)
065_YMRR_FWD	TAA TGA GGT ACC CCG GG	Construction of $Y_IMR^1R^2$ from pPANK_$Y_IMR^1R^2R^3$
101_AtypeC-cap_FW	GGT AAC GAA CAG AAA CAG GCT GTT AAA GAA G	Construction of $Y_IMR^1R^2R^3A_{II}$, $Y_IMR^1R^2A_{II}$, $Y_IMR^1A_{II}$ from pPANK_$Y_IMR^1R^2R^3MA_{II}$
118_delR42_FW	TGA TGG TAA CGA ACA AAT CC	Deletion of R42 from Y_I-cap
119_delR42_RV	GAC AGG ATC TGA CGG AAT TTA AC	Reverse primer for 118
121_R42A_FW	AGA TCC TGT CTG CTG ATG GTA ACG A	Introduction of R42A into Y_I-cap
122_R42A_RV	GAC GGA ATT TAA CGG TAG CAG	Reverse primer for 121
124_V34R_R37S_RV	AGC TGT TCC TGC ATG TCG TCG GAG TTC	Reverse primer for 139 and 140
139_Yf_V24R_FWD	GTC TGC TAC CCG TAA ATT CCG TCA GAT CCT G	Introduction of V34R into Y_I-cap and Y_I-cap_R42del
140_Yf_R27S_FWD	GTC TGC TAC CGT TAA ATT CTC TCA GAT CCT G	Introduction of R37S into Y_I-cap and Y_I-cap_R42del
141_M_E46A_FWD	GTC TCG TGA TGG TAA CGC ACA AAT CC	Introduction of E46A into Y_I-cap

142_M_E46A_REV	AGG ATC TGA CGG AAT TTA ACG GTA GC	Reverse primer for 141
152_V34R_FWD	GTC TGC TAC CCG TAA ATT CTC TCA G	Introduction of V34R into Y_I-cap_R37S and Y_I-cap_R37S_R42del
153_V34R_REV	AGC TGT TCC TGC ATG TCG TCG	Reverse primer for 152

2. Details of Cloning procedures

Polymerase chain reactions (PCR) were usually performed with 200 µM of total dNTPs, 1.5 µM of forward and reverse primers and 50 ng of template. *Taq* polymerase was used for colony-PCRs, *Pfu* polymerase for site-directed mutagenesis and general high fidelity DNA amplification. Point mutations were introduced by whole plasmid amplification using non-overlapping primers, one of which introduced the mutation, and blunt end ligation. All plasmids were verified by DNA sequencing.

DNA restriction, phosphorylation and ligation were performed according to the manufacturer's instructions (New England Biolabs, USA, Fermentas, Lithuania or Promega, USA). DNA quality and size was assessed on ethidium bromide stained (0.1 µg/ml) 1% agarose gels. PCR products were purified using standard DNA kits (Macherey-Nagel, Germany, or Roche, Switzerland). DNA concentrations were determined by UV absorption spectroscopy at 260 nm (ND-1000 spectrophotometer, Nanodrop Technologies, USA).

pLIC_CR [31] and pPANK, a pQE30 derivative lacking the *Bpi*I and *Bsa*I endonuclease restriction recognition sites (GenBank accession number AY327140) were used as cloning and protein expression vectors. All pPANK-based vectors introduce an $MRGSH_6$-tag at the N-terminus of the proteins. pLIC-based vectors carry a TEV protease cleavable N-terminal H_6-tag (see Figure S1).

Cloning of $Y_IMR^1R^2R^3MA_{II}$ Fragments

pPANK_$Y_IMR^1R^2R^3MA_{II}$ has previously been established from pPANK_NyM328MCa [23] by introducing two point mutations (Q292L, F293Q) [25] in the A_I-type C-cap.

N-terminal fragments (Y_IMR^1, $Y_IMR^1R^2$, $Y_IMR^1R^2R^3$) were cloned by deleting the unrequired C-terminal repeats from previously established pPANK_NyM328MCa encoding $Y_IMR^1R^2R^3MA_I$. C-cap primers 053 and 056 were used to create $YMR^1R^2R^3$ (template: pPANK_NyM328MCa), 054 and 056 for YMR^1 (template: pPANK_NyM328MCa) and 63 and 65 for YMR^1R^2 (template: pPANK_$Y_IMR^1R^2R^3$) (Table S2). Primers were designed to

allow amplification of the whole expression plasmid excluding the sequence to be deleted. The PCR product was then treated with *Dpn*I (Fermentas) to remove template DNA, treated with PNK T4-Kinase (Promega) and blunt end ligated to re-create a circular pPANK based plasmid. C-terminal fragments (R^3MA_{II}, $R^2R^3MA_{II}$, $R^1R^2R^3MA_{II}$ and $MR^1R^2R^3MA_{II}$) were amplified from pPANK_N_IM328MC_{II} using primers 078 and 088, (R^3MA_{II}), 077 and 088 ($R^2R^3MA_{II}$ and $R^1R^2R^3MA_{II}$) and primers 090 and 088 ($MR^1R^2R^3MA_{II}$) (Table S2). Primers were designed to allow ligation independent cloning (LIC) [Aslanidis and Jong, 1990, Nucl. Acids. Res. 18(1990), 6069-74] into the pLIC_CR vector [31] to introduce a cleavable N-terminal His_6-tag for IMAC purification and undisturbed fragment interaction. pLIC_CR contains an N-terminal $MRGSH_6$-tag followed by the TEV protease recognition site and the *sacB* gene as an additional negative selection marker. pLIC_CR_MA for the expression the C-terminal MA has been previously established by Watson *et al*. [31]. For LIC cloning 100 ng of pLIC_CR was *Bsa*I digested, and treated with T4-DNA-polymerase in the presence of dCTP. The target PCR fragment was treated with T4-DNA-polymerase and dGTP. Prepared plasmid and insert DNA were column purified (High Pure PCR Product Purification kit, Roche), mixed at approximately 1:1 molar ratio and incubated for 10 min at room temperature and then transformed into *E. coli* XL1-blue cells. Plasmids obtained from colonies on LB agar plates containing 100 µg/ml ampicillin and 7 % sucrose were sequenced for correctness.

Cloning of $Y_IMR^1R^2R^3A_{II}$, $Y_IMR^1R^2A_{II}$ and $Y_IMR^1A_{II}$

Vectors for $Y_IMR^1R^2R^3A_{II}$, $Y_IMR^1R^2A_{II}$ and $Y_IMR^1A_{II}$ were cloned by deleting the unrequired internal repeats from $Y_IMR^1R^2R^3MA_{II}$ via a blunt end ligation strategy. Primers were designed to allow amplification of the whole expression plasmid excluding the sequence to be deleted. The PCR product was then treated with *Dpn*I (Fermentas) to remove template DNA, treated with PNK T4-Kinase (Promega) and blunt end ligated to re-create a circular pPANK based plasmid. pPANK_$Y_IMR^1R^2R^3A_{II}$, was amplified from pPANK_$Y_IMR^1R^2R^3A_{II}$ using primers 101/054. pPANK_$Y_IMR^1R^2A_{II}$ and pPANK_$Y_IMR^1A_{II}$ were amplified from pPANK_$Y_IMR^1R^2R^3A_{II}$ using primers 101/063 and 101/056 respectively (See Table S2).

Cloning of $Y_IMR^1R^2A_{II}$-Mutants

The following single, double and triple point mutations in pPANK_$Y_IMR^1R^2A_{II}$ were introduced by whole plasmid amplification and blunt end ligation as described above using primers indicated in Table S2: V34R, R37S, R42Δ, R42A, E46A, V34R/R37S, V34R/R42Δ, R37S/R42Δ and V34R/R37S/R42Δ (= $Y_{Ii}MR^1R^2A_{II}$).

3. Spectroscopic details

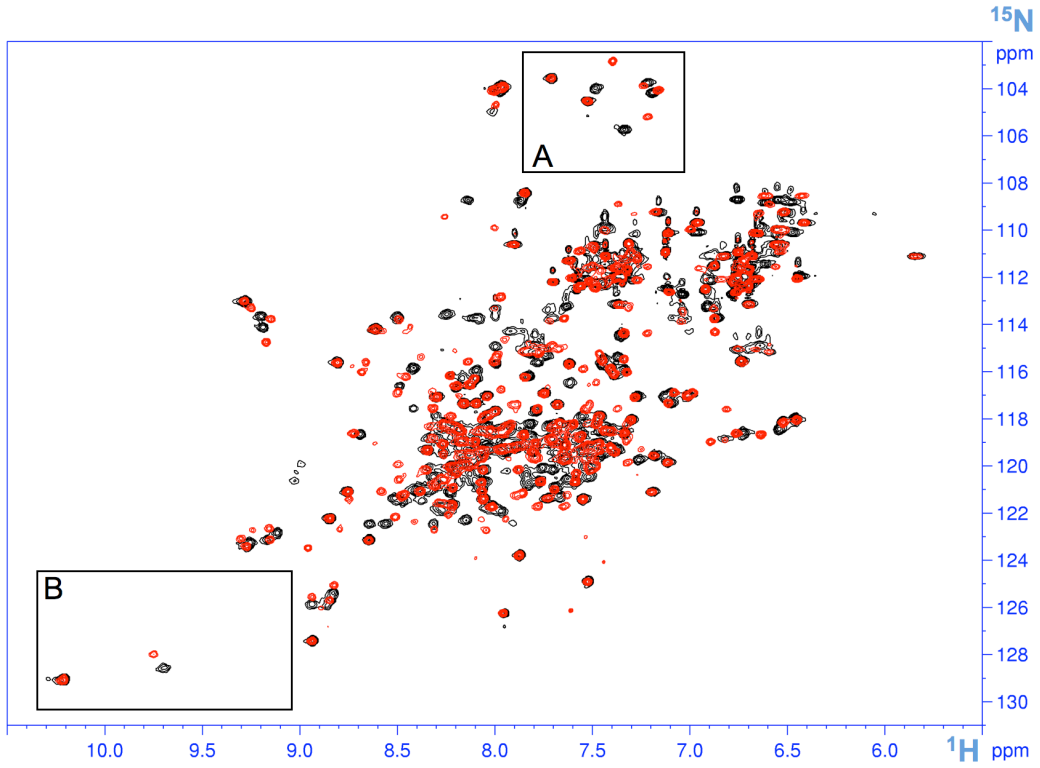

Figure S2: [^{15}N,^{1}H]-HSQC spectra showing chemical shift perturbations of 400 µM ^{15}N-labelled $Y_IMR^1R^2R^3MA_{II}$ (VG_328) without (black) and with (red) two equivalents of NT. Details for the Gly (A) and Trp indole (B) regions are shown in Figure 3A and B.

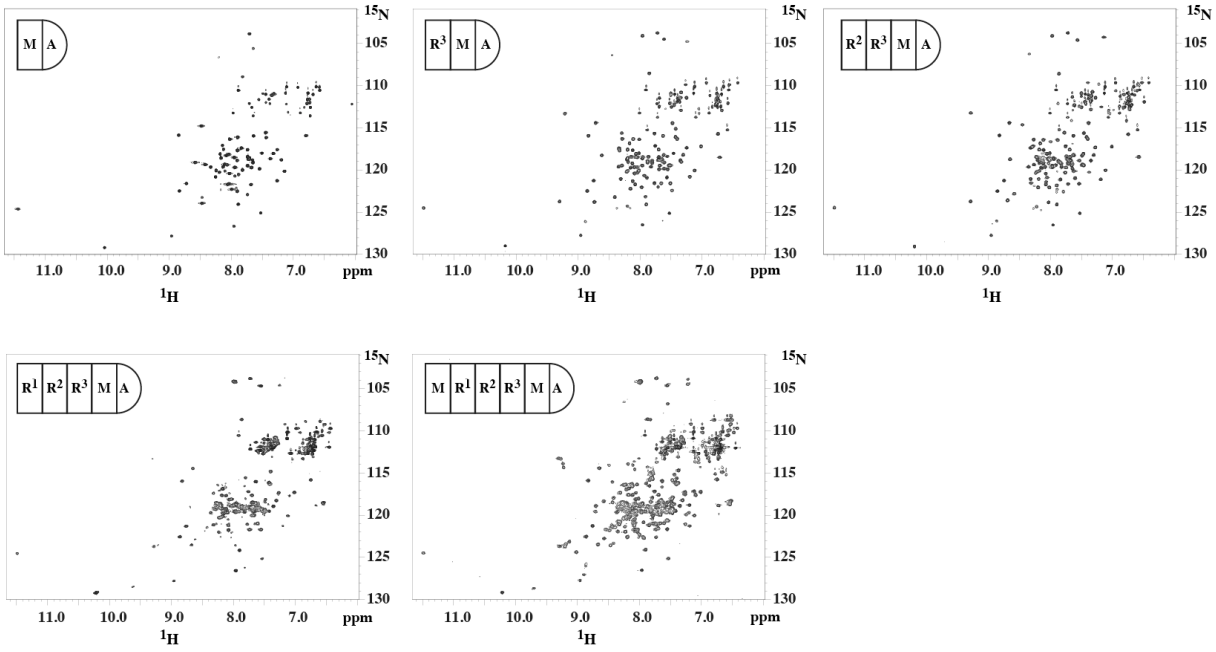

Figure S3: [^{15}N,^{1}H]-HSQC spectra of ^{15}N-labelled N-terminally truncated $Y_IMR^1R^2R^3MA_{II}$ fragments. All spectra were recorded at 32 °C, 300 µM concentration.

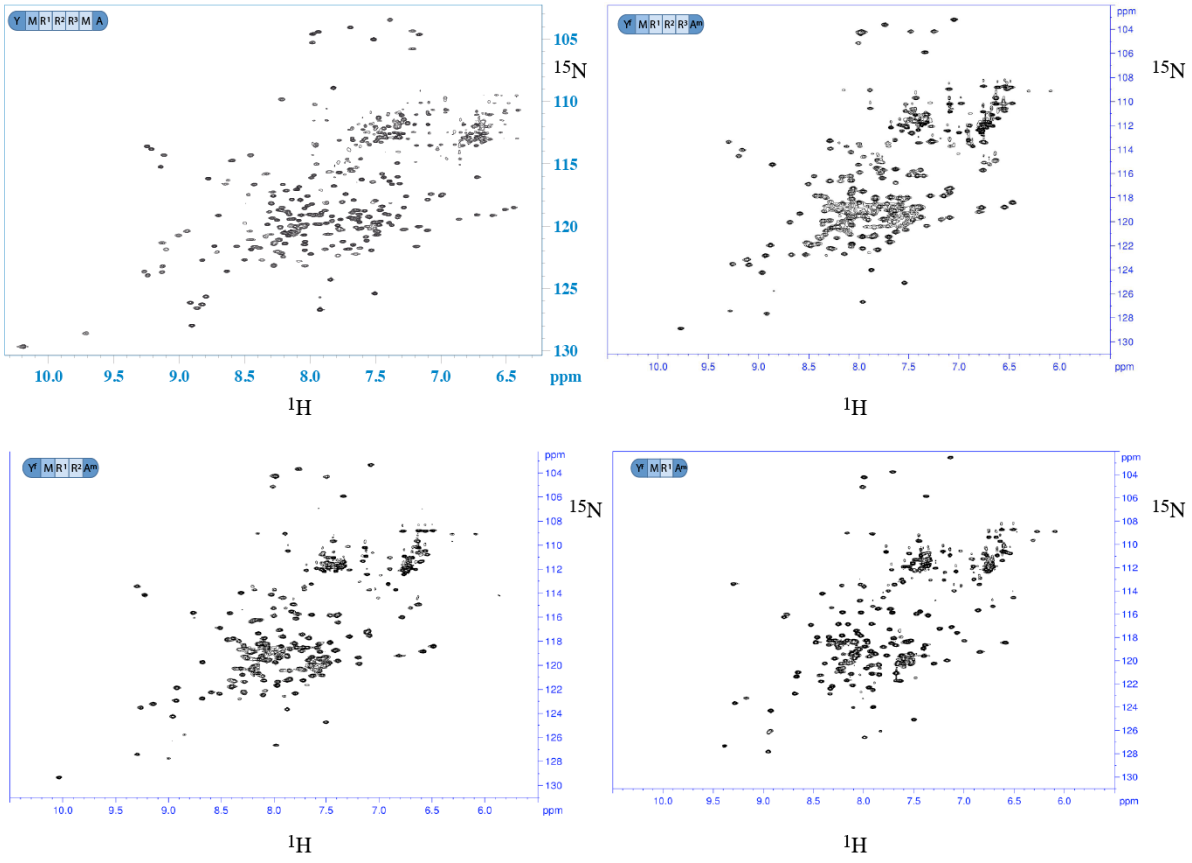

Figure S4 [^{15}N,^{1}H]-HSQC spectra of $Y_IMR^1R^2R^3MA^{II}$ (top left), $Y_IMR^1R^2R^3A^{II}$ (top right), $Y_IMR^1R^2A^{II}$ (bottom left) and $Y_IMR^1A^{II}$ (bottom right).

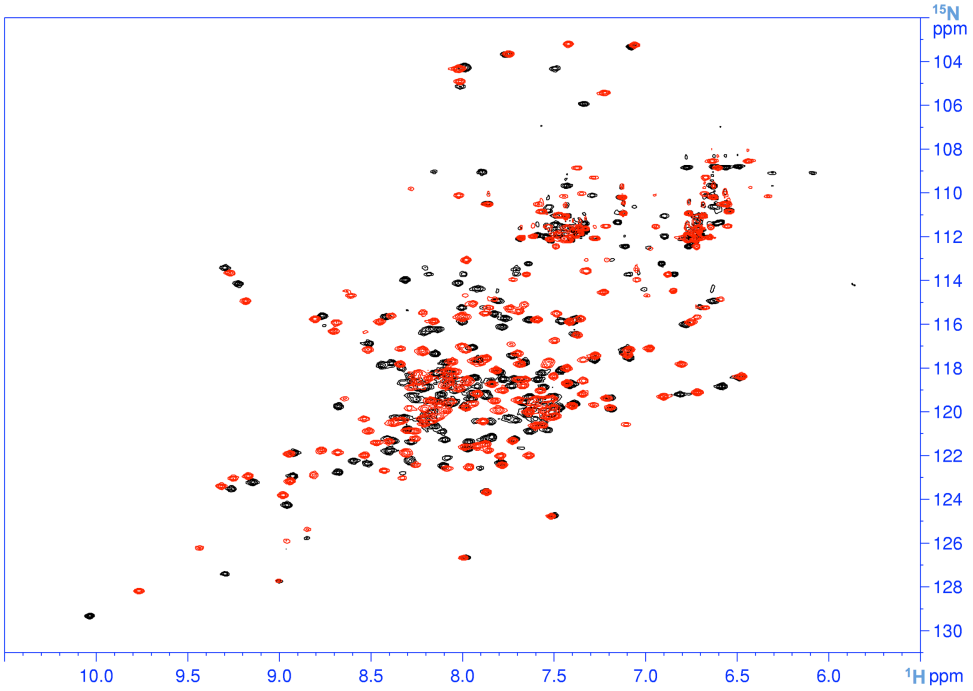

Figure S5: [^{15}N,^{1}H]-HSQC spectra showing chemical shift perturbations of 400 µM ^{15}N labelled $Y_IMR^1R^2A_{II}$ without (black) and with (red) two equivalents of NT.

4. SPR Data:

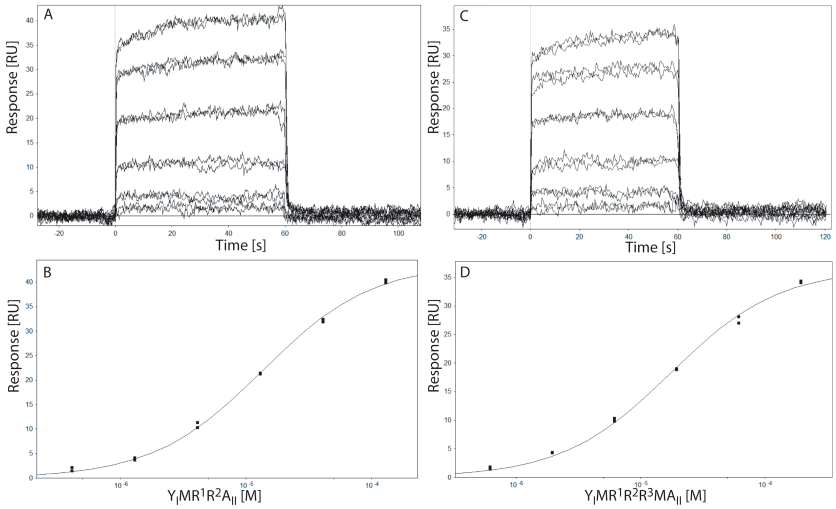

Figure S6: SPR response curves (top) and fitted data (bottom) for $Y_IMR^1R^2A_{II}$ (A, B, 14 µM) and $Y_IMR^1R^2R^3MA_{II}$ (C, D, 18 µM) at 8 °C.

5. Chemical shift perturbation (CSP) data

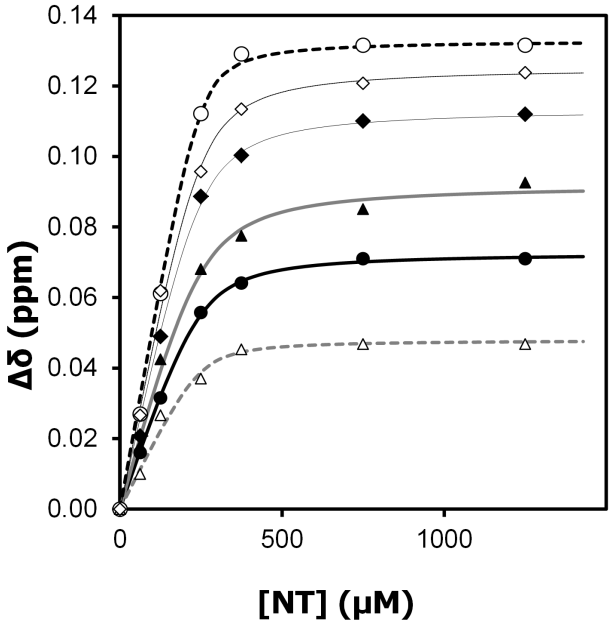

Figure S7: Representative fits from CSP raw data of $Y_IMR^1R^2A_{II}$ residues used for k_d determination; G86 (empty circle), G128 (empty diamond), L76 (filled diamond), S66 (filled triangle), A59 (filled circle), A92 (empty triangle).

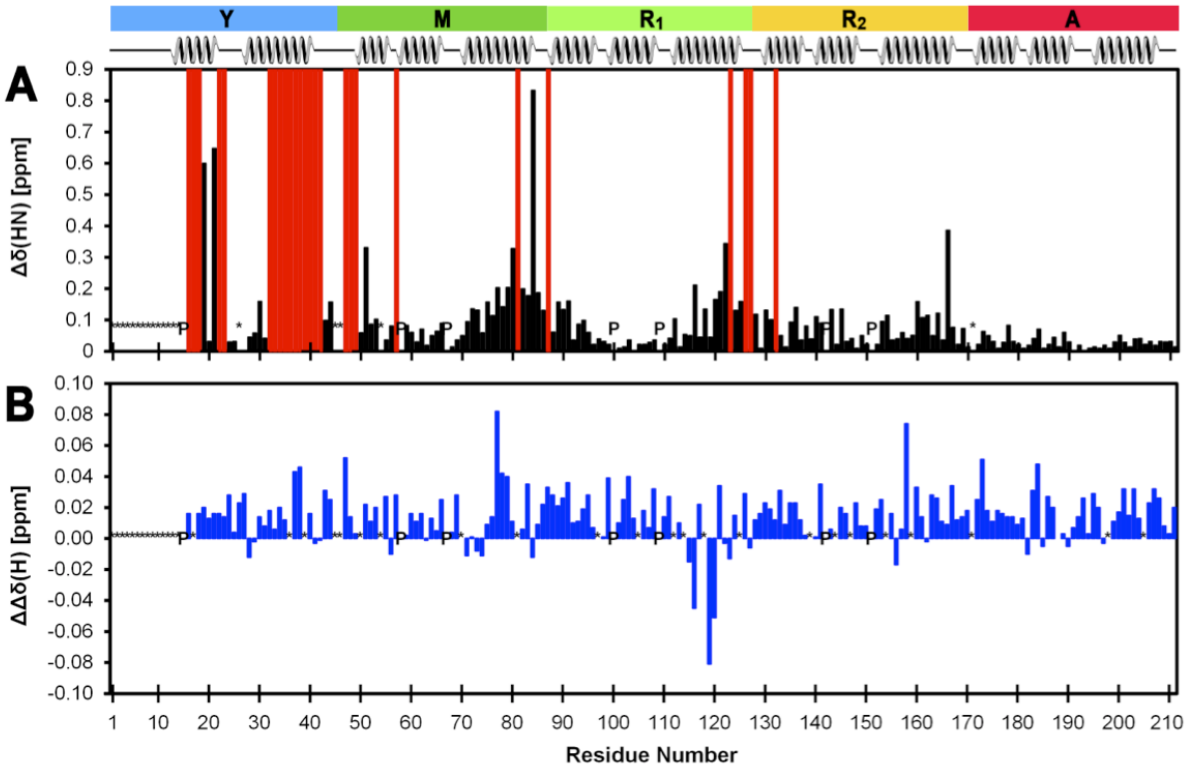

Figure S8: A Appropriately weighted combined chemical shift changes Δδ(HN) in ^{15}N and ^{1}H chemical shift (see Methods and Materials) upon addition of 2 equiv. of NT to $Y_IMR^1R^2A_{II}$. Red bars indicate unquantifiable CSPs when assignments are only available in the presence of NT. Prolines are indicated by "P", unassigned residues by "*" **B** Difference between ^{1}H chemical shift changes ΔΔδ(H) of $Y_IMR^1R^2A_{II}$ in presence of 2 equiv. NT and in presence of 2 equiv. NTK6C-MTSL

6. Molecular Docking results

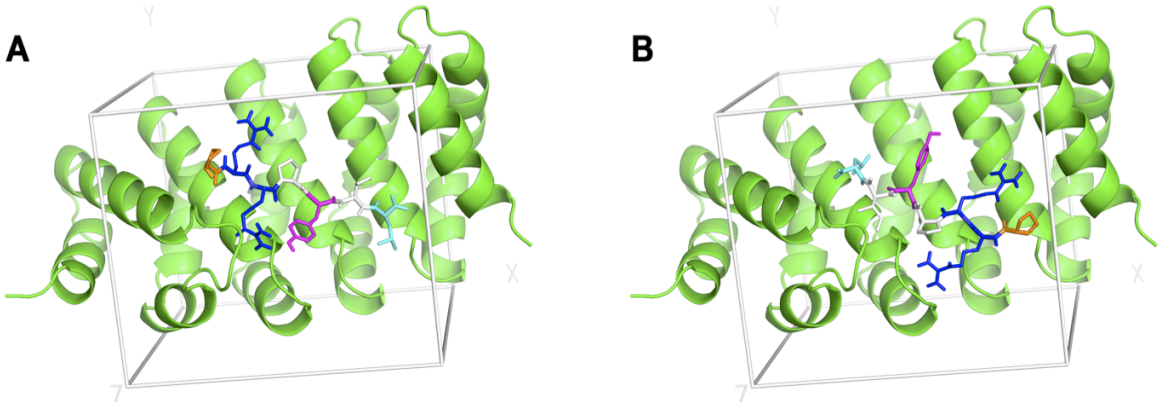

Figure S9: Parallel (A) and antiparallel (B) binding poses of NT7-13 (sticks) from automatic docking to $Y_IMR^1R^2A_{II}$ (green ribbon). The peptide was initially placed arbitrarily within the indicated box (see methods).

7. PRE results from additional spin-labeled NT forms

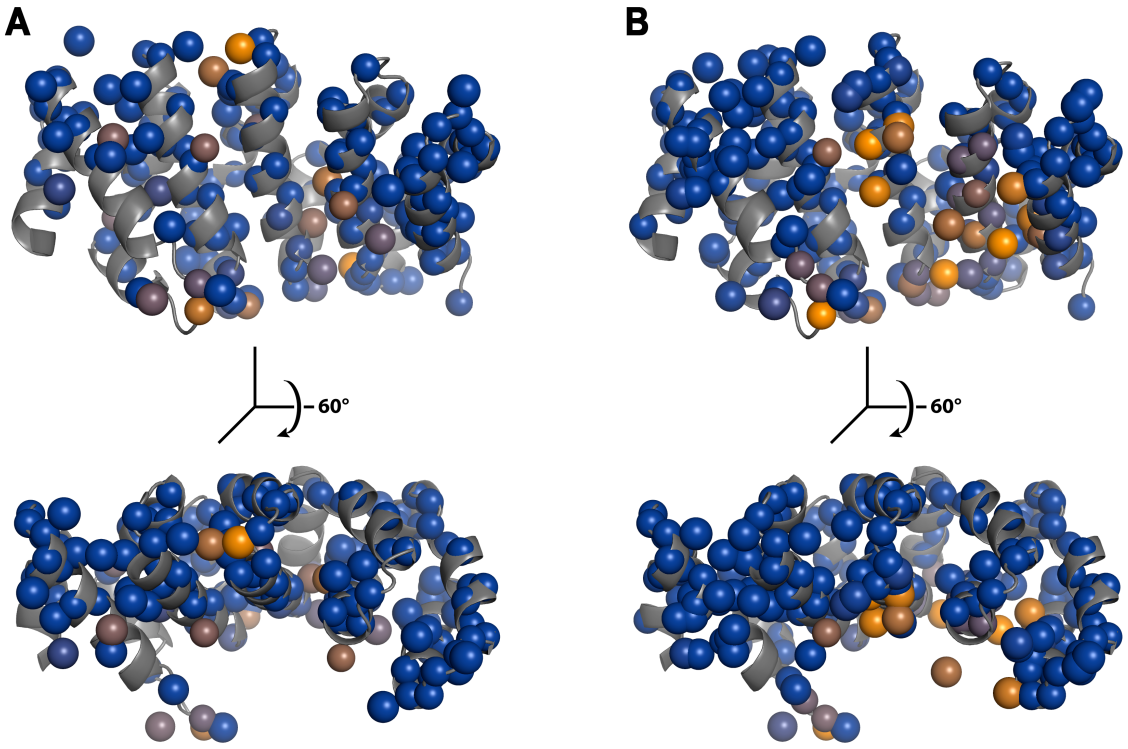

Figure S10: Interaction of $Y_IMR^1R^2A_{II}$ with NTQ1C-MTSL (A) and with NTL13C-MTSL (B) as probed by PREs. For further details see the caption of Figure 5 in the main paper.

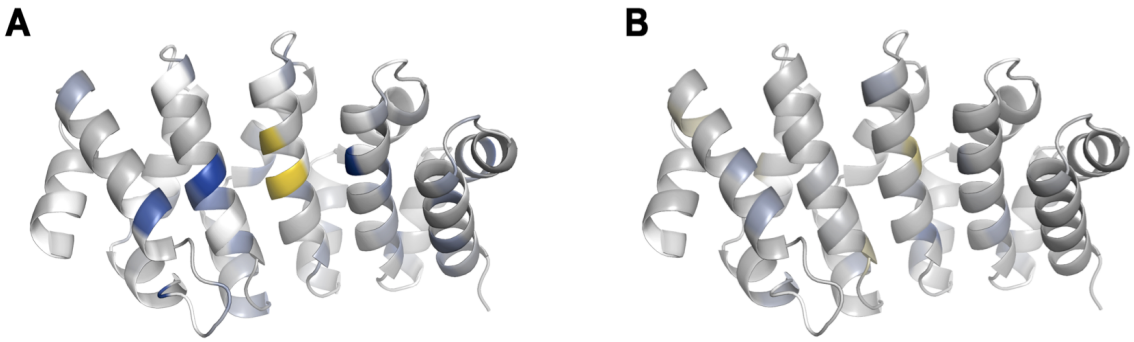

Figure S11: Location of bound MTSL by *differential* chemical shift perturbation ($\Delta\Delta\delta(^1H)$, see Methods) mapped onto $Y_IMR^1R^2A_{II}$ complexed with either unlabeled NT or with NT coupled to quenched MTSL. Positive and negative values of $\Delta\Delta\delta$ are depicted in blue and yellow, respectively for MTSL attached at position 6 (A) or at position 13 (B).

8. MD Simulations

Multiple MD runs were carried out in the absence and presence of the peptide. The MD simulations for determining the binding mode of NT(7-13) to $Y_IMR^1R^2A^{II}$ were started from the parallel or the antiparallel orientation of NT(7-13). The same simulation protocol was used in both cases (see Molecular Dynamics Simulations section in Materials & Methods). The two independent simulations of 2 μs each for the parallel orientation with the wild type $Y_IMR^1R^2A^{II}$ used different random seeds for the initial assignment of the velocities.

Table S3: MD simulations performed to probe the intrinsic plasticity of armadillo proteins in the absence of peptide (apo), and to determine the binding mode of the NT7-13 peptide.

system	Binding mode	Time (μs)	Temperature (K)	Force field	vdW cutoff (nm)
$Y_IMR^1R^2A_{II}$	apo	2	310	Charmm36	1.0
		1	330	OPLS	0.9
$Y_IMR^1R^2A_{II}$_V34R	apo	2	310	Charmm36	1.0
$Y_IMR^1R^2A_{II}$_R42Δ	apo	2	310	Charmm36	1.0
$Y_{III}M_3A_{II}$	apo	2	310	Charmm36	1.0
		1	330	OPLS	0.9
$Y_IMR^1R^2A_{II}$_R37S	apo	2	310	Charmm36	1.0
$Y_IMR^1R^2A_{II}$ – NT7-13	parallel	2 x 2	310	Charmm36	1.0
$Y_IMR^1R^2A_{II}$ – NT7-13	antiparallel	2	310	Charmm36	1.0
$Y_IMR^1R^2A_{II}$_R37S – NT7-13	parallel	2	310	Charmm36	1.0

Table S4: Average interaction energy between residues of NT(7-13) and their main interacting partners of $Y_IMR^1R^2A_{II}$ (parallel orientation). The interaction energy is the sum of the Coulomb energy and the van der Waals energy. The energies are averaged over a 2-μs trajectory and the values larger than 4 kJ/mol are shown; the errors are the standard deviation. Values from two independent simulations are shown in plain and italics letters, respectively. Energies larger than 12 kJ/mol (~5kT) are shown in bold letters.

NT-$Y_IMR^1R^2A^{II}$	res-res[1] (kJ/mol)	bb-bb[2] (kJ/mol)	bb-sc[3] (kJ/mol)	sc-bb[4] (kJ/mol)	sc-sc[5] (kJ/mol)
P7-W77	**-14 ± 6**	**-1 ± 1**	**-4 ± 4**	**-2 ± 2**	**-7 ± 4**
	-13 ± 4	***0 ± 0***	***-2 ± 1***	***-4 ± 2***	***-7 ± 3***
P7-S80	**-13 ± 11**	**-2 ± 2**	**-10 ± 10**	**-1 ± 1**	**0 ± 2**
	-22 ± 11	***-2 ± 1***	***-19 ± 10***	***-1 ± 1***	***0 ± 1***
P7-N81	-4 ± 3	0 ± 0	-1 ± 2	-1 ± 1	-2 ± 2
	-7 ± 2	*-1 ± 0*	*-1 ± 1*	*-2 ± 1*	*-4 ± 2*
P7-Y116	-4 ± 3	0 ± 0	-1 ± 2	0 ± 0	-2 ± 2
	--	--	--	--	--
P7-I119	-5 ± 2	0 ± 0	-3 ± 2	0 ± 0	-1 ± 1
	-4 ± 2	*0 ± 0*	*-3 ± 1*	*0 ± 0*	*0 ± 0*
P7-N123	-7 ± 6	0 ± 0	-7 ± 6	0 ± 0	0 ± 0
	-10 ± 4	*0 ± 0*	*-9 ± 4*	*0 ± 0*	*0 ± 0*
R8-S84	-4 ± 4	0 ± 0	-3 ± 3	0 ± 0	-1 ± 1
	-8 ± 3	*0 ± 0*	*-7 ± 3*	*0 ± 0*	*0 ± 0*
R8-I119	-9 ± 3	0 ± 0	-1 ± 1	-2 ± 1	-6 ± 2
	-9 ± 3	*0 ± 0*	*-1 ± 0*	*-2 ± 1*	*-5 ± 3*
R8-N122	8 ± 4	0 ± 0	0 ± 0	0 ± 0	8 ± 4
	6 ± 4	*0 ± 0*	*0 ± 0*	*0 ± 0*	*6 ± 4*
R8-N123	-11 ± 6	0 ± 0	-4 ± 5	0 ± 0	-7 ± 5
	-12 ± 4	***0 ± 0***	***-5 ± 3***	***0 ± 0***	***-7 ± 3***
R8-E158	**-147 ± 7**	**0 ± 0**	**0 ± 0**	**0 ± 0**	**-147 ± 7**
	-138 ± 19	***0 ± 0***	***0 ± 0***	***0 ± 0***	***-138 ± 19***
R8-W161	--	--	--	--	--
	-4 ± 4	*0 ± 0*	*0 ± 0*	*0 ± 0*	*-4 ± 4*
R9-S41	-6 ± 9	0 ± 0	0 ± 0	-6 ± 9	-1 ± 2
	--	--	--	--	--
R9-R42	-4 ± 8	0 ± 0	0 ± 0	-3 ± 8	0 ± 0
	-	-	-	-	-
R9-D43	**-40 ± 52**	**0 ± 0**	**0 ± 0**	**0 ± 2**	**-41 ± 52**
	--	--	--	--	--
R9-N81	-8 ± 12	0 ± 0	-1 ± 2	-2 ± 2	-6 ± 10
	--	--	--	--	--
R9-A83	--	--	--	--	--
	-6 ± 7	*0 ± 0*	*0 ± 0*	*-5 ± 7*	*0 ± 0*

R9-S84	**-13 ± 10**	**-1 ± 1**	**-5 ± 7**	**-4 ± 6**	**-2 ± 4**
	-14 ± 6	***-1 ± 1***	***-3 ± 3***	***-7 ± 4***	***-3 ± 3***
R9-G85	-5 ± 9	0 ± 0	0 ± 0	-6 ± 9	0 ± 1
	-11 ± 8	*0 ± 0*	*0 ± 0*	*-11 ± 8*	*0 ± 0*
R9-N87	-10 ± 20	0 ± 0	0 ± 0	0 ± 1	-10 ± 19
	--	--	--	--	--
R9-N123	-10 ± 13	0 ± 0	-5 ± 7	-3 ± 5	-2 ± 5
	-20 ± 8	***0 ± 0***	***-8 ± 5***	***-7 ± 4***	***-5 ± 3***
R9-F126	-4 ± 6	0 ± 0	-1 ± 1	-1 ± 2	-2 ± 4
	-8 ± 4	*0 ± 0*	*-1 ± 1*	*-3 ± 4*	*-4 ± 3*
Y11-S84	-6 ± 8	-2 ± 4	-1 ± 1	-2 ± 4	-2 ± 3
	--	--	--	--	--
Y11-I119	--	--	--	--	--
	-4 ± 2	*0 ± 0*	*0 ± 0*	*-3 ± 1*	*-1 ± 1*
Y11–N122	--	--	--	--	--
	-10 ± 4	*0 ± 0*	*0 ± 2*	*-3 ± 2*	*-7 ± 3*
Y11-N123	-9 ± 13	0 ± 0	0 ± 0	-2 ± 5	-7 ± 12
	-31 ± 11	***0 ± 0***	***0 ± 0***	***-2 ± 1***	***-29 ± 11***
Y11-F126	-10 ± 6	0 ± 0	-2 ± 3	-1 ± 1	-7 ± 5
	-15 ± 4	***0 ± 0***	***-6 ± 3***	***-1 ± 0***	***-9 ± 3***
Y11-W161	--	--	--	--	--
	-9 ± 4	*0 ± 0*	*-1 ± 1*	*-1 ± 0*	*-8 ± 4*
Y11-N165	--	--	--	--	--
	-6 ± 4	*0 ± 0*	*0 ± 1*	*0 ± 0*	*-5 ± 4*
I12-S84	-5 ± 9	-4 ± 7	0 ± 1	-1 ± 1	0 ± 0
	--	--	--	--	--
L13-E203	--	--	--	--	--
	5 ± 14	*0 ± 0*	*5 ± 14*	*0 ± 0*	*0 ± 1*

[1] interaction energy between an NT residue and a $Y_IMR^1R^2A^{II}$ residue

[2] interaction energy between the backbone atoms of an NT residue and the backbone atoms of a $Y_IMR^1R^2A^{II}$ residue

[3] interaction energy between the backbone atoms of an NT residue and the side chain atoms of a $Y_IMR^1R^2A^{II}$ residue

[4] interaction energy between the side chain atoms of an NT residue and the backbone atoms of a $Y_IMR^1R^2A^{II}$ residue

[5] interaction energy between the side chain atoms of an NT residue and the side chain atoms of a $Y_IMR^1R^2A^{II}$ residue

Table S5: Average interaction energy between residues of NT(7-13) and their main interacting partners of $Y_IMR^1R^2A_{II}$ (antiparallel orientation). The interaction energy is the sum of the Coulomb energy and the van der Waals energy. The energies are averaged over a 2-μs trajectory and the values larger than 4 kJ/mol are shown; the errors are the standard deviation. Energies larger than 12 kJ/mol (~5kT) are shown in bold letters.

NT-$Y_IMR^1R^2A^{II}$	res-res (kJ/mol)[1]	bb-bb (kJ/mol)[2]	bb-sc (kJ/mol)[3]	sc-bb (kJ/mol)[4]	sc-sc (kJ/mol)[5]
P7 - S84	-6 ± 7	-3 ± 3	-3 ± 3	-1 ± 1	0 ± 1
P7 - N122	-4 ± 5	0 ± 0	-1 ± 2	-1 ± 1	-2 ± 2
P7 - N123	-7 ± 7	0 ± 0	-4 ± 5	-1 ± 1	-1 ± 2
P7 - F126	-4 ± 4	0 ± 0	-1 ± 1	0 ± 0	-3 ± 3
R8 - D43	**-15 ± 42**	**0 ± 0**	**0 ± 0**	**0 ± 0**	**-15 ± 42**
R8 - A83	**-13 ± 15**	**0 ± 0**	**0 ± 0**	**-13 ± 15**	**0 ± 0**
R8 - S84	**-26 ± 19**	**-5 ± 7**	**-5 ± 4**	**-13 ± 11**	**-3 ± 3**
R8 - G85	-11 ± 11	-1 ± 1	0 ± 0	-11 ± 11	1 ± 1
R8 - G86	-4 ± 4	0 ± 0	0 ± 0	-5 ± 4	2 ± 2
R8 - N87	-11 ± 15	0 ± 0	0 ± 0	-1 ± 1	-10 ± 14
R8 - N123	**-16 ± 12**	**0 ± 0**	**-7 ± 9**	**-7 ± 8**	**-2 ± 3**
R8 - F126	-7 ± 6	0 ± 0	-2 ± 2	-1 ± 3	-5 ± 4
R8 - E169	-5 ± 24	0 ± 0	0 ± 0	0 ± 0	-5 ± 24
R9 - S84	-8 ± 9	-1 ± 1	-6 ± 8	0 ± 0	-1 ± 1
R9 - I119	-6 ± 3	0 ± 0	0 ± 0	-2 ± 2	-4 ± 2
R9 - N122	6 ± 7	0 ± 0	0 ± 0	0 ± 0	6 ± 7
R9 - N123	-9 ± 9	0 ± 0	-1 ± 1	0 ± 0	-8 ± 9
R9 - E158	**-143 ± 22**	**0 ± 0**	**0 ± 0**	**0 ± 1**	**-143 ± 23**
Y11 - W77	-10 ± 9	0 ± 0	-2 ± 3	-1 ± 2	-7 ± 6
Y11 - S80	-6 ± 8	0 ± 0	0 ± 0	-2 ± 3	-4 ± 7
Y11 - N81	-10 ± 10	0 ± 0	0 ± 1	-2 ± 2	-8 ± 9
Y11 - S84	-4 ± 8	0 ± 0	0 ± 0	0 ± 1	-4 ± 8

[1] interaction energy between an NT residue and a $Y_IMR^1R^2A^{II}$ residue

[2] interaction energy between the backbone atoms of an NT residue and the backbone atoms of a $Y_IMR^1R^2A^{II}$ residue

[3] interaction energy between the backbone atoms of an NT residue and the side chain atoms of a $Y_IMR^1R^2A^{II}$ residue

[4] interaction energy between the side chain atoms of an NT residue and the backbone atoms of a $Y_IMR^1R^2A^{II}$ residue

[5] interaction energy between the side chain atoms of an NT residue and the side chain atoms of a $Y_IMR^1R^2A^{II}$ residue

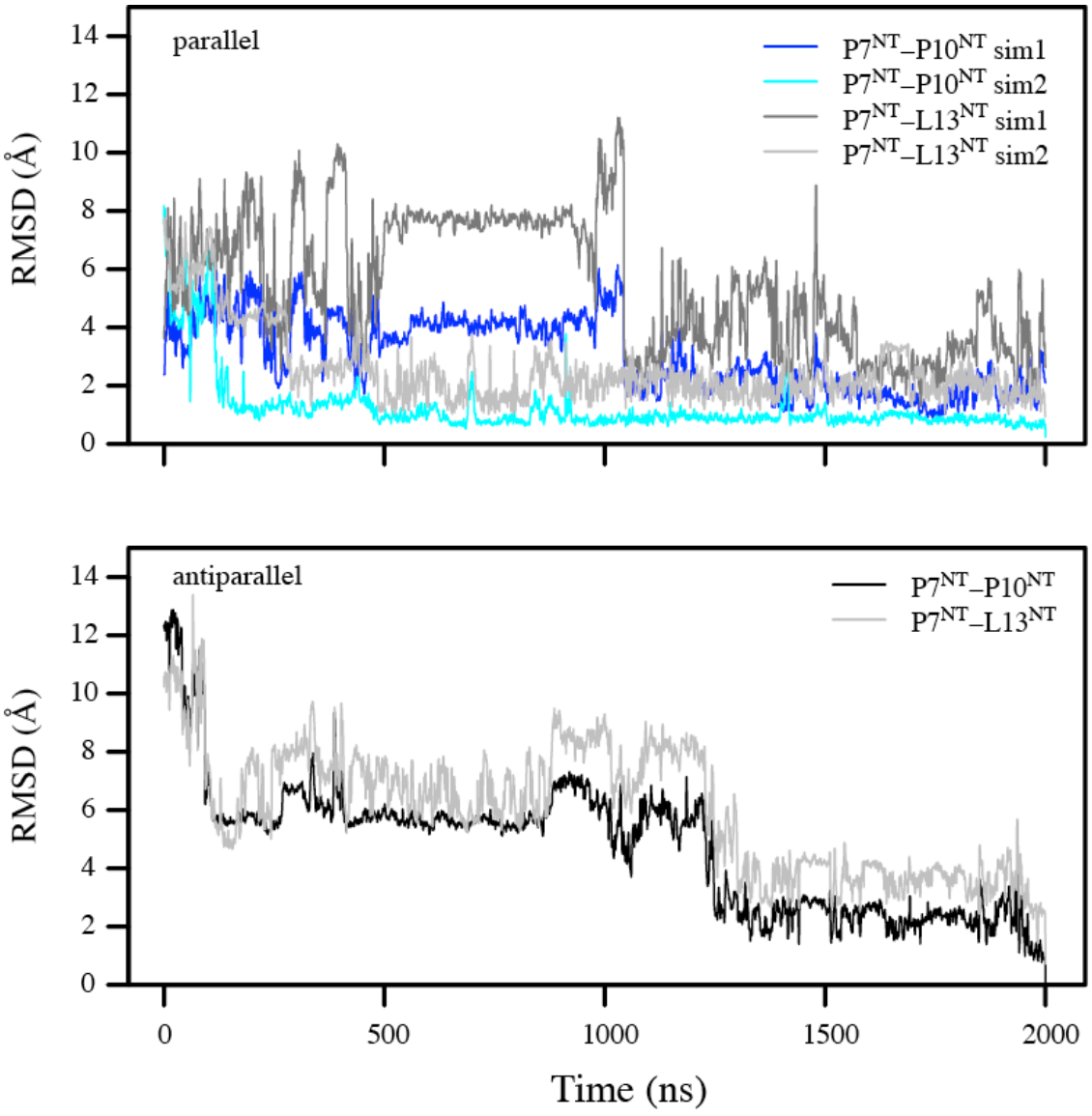

Figure S12: **Time series of root-mean-square deviations (RMSD) of the Cα atoms of NT(7-10) or NT(7-13) from two independent MD simulations.** Data are shown with NT(7-13) in the parallel orientation, calculated upon fitting MR^1R^2 repeats of $Y_IMR^1R^2A_{II}$ to the MR^1R^2 repeats of the structure obtained at the end of the second simulation ("sim2") (upper panel), or from a MD simulation with NT(7-13) in the antiparallel orientation, calculated upon fitting MR^1R^2 to that of the structure at the end of the simulation (lower panel).

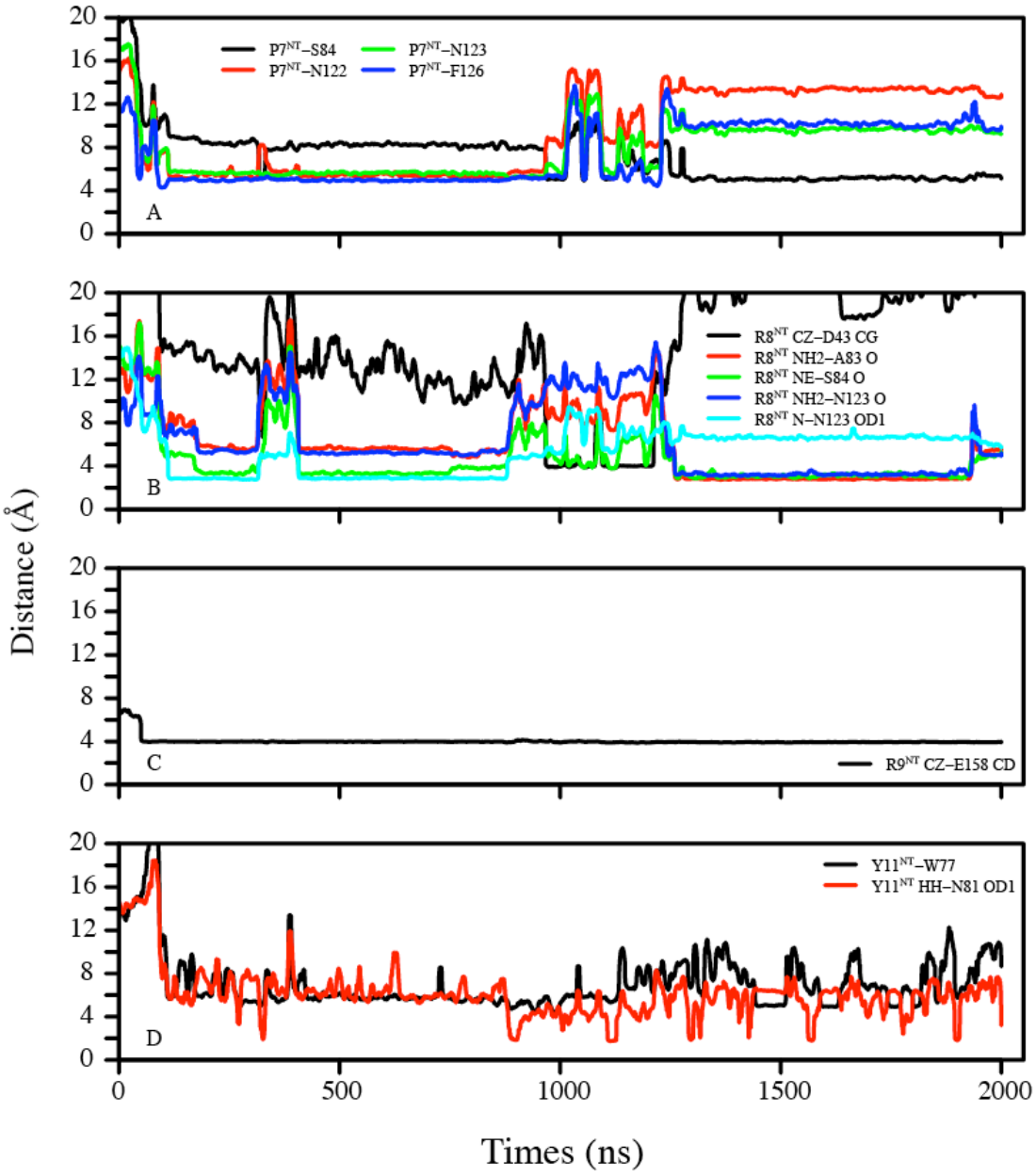

Figure S13: **Time evolution of distances between Pro7^{NT}, Arg8^{NT}, Arg9^{NT} and Tyr11^{NT} and main interacting residues of $Y_IMR^1R^2A_{II}$**. Data are shown when NT(7-13) is in the *antiparallel* orientation. In (A) the distances are calculated between the center of mass of the two residues. In (B) and (C), explicit distances between the indicated atoms are given. In (D) the distance between Tyr11^{NT} and Trp77 is the distance between the center of masses of the aromatic rings.

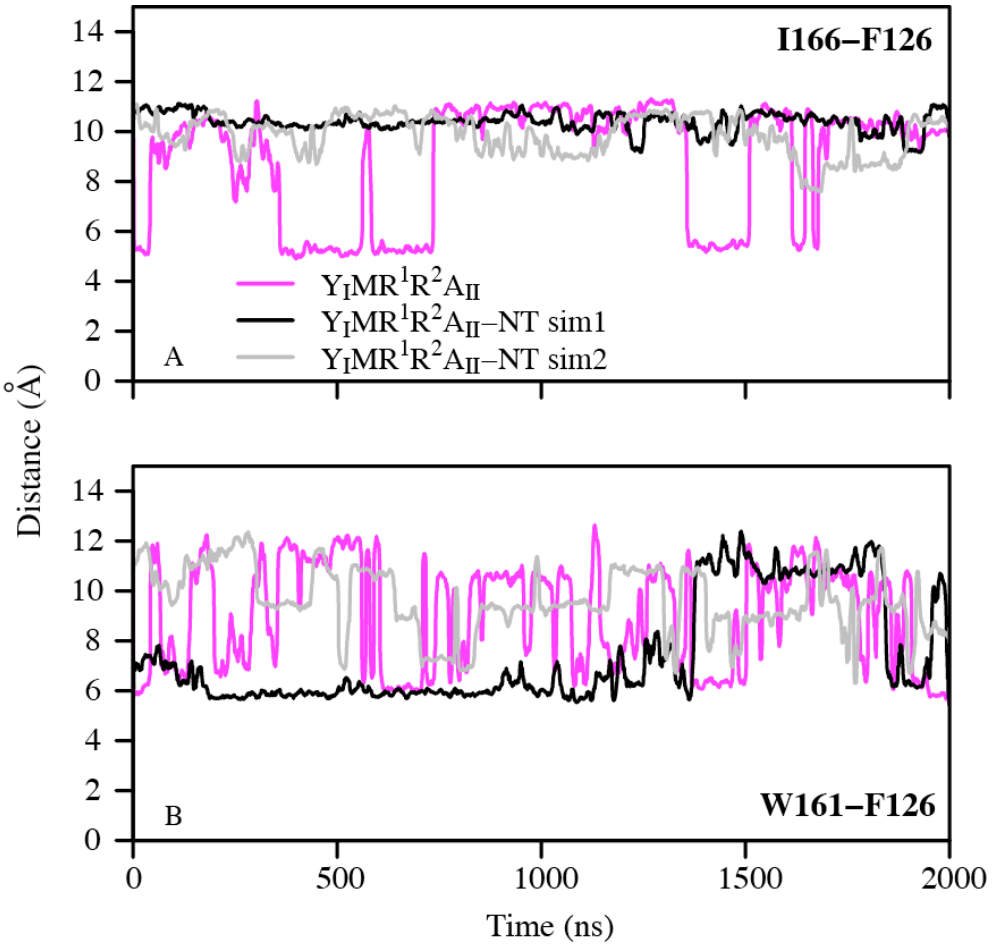

Figure S14: Evolution of atomic distances during MD simulations: (A) Phe126 (center of mass of ring) to Cγ1 of Ile166 and (B) Phe126 (center of mass) to Trp161 (center of mass).

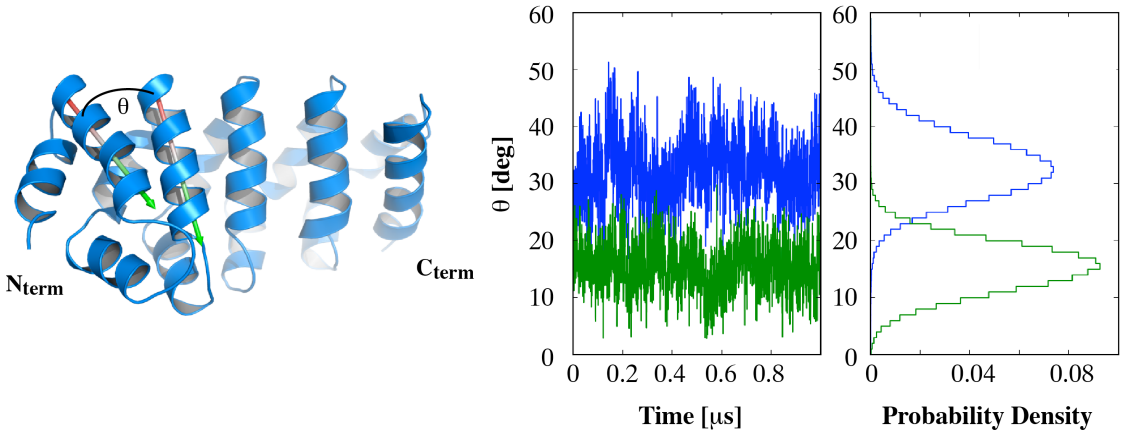

Figure S15: Inter-helical angle between helix 2 of the N-cap and helix 3 of the first internal repeat. The angle θ reflects the orientation of the N-cap relative to the first internal repeat (left). (Middle and right) Time series and probability density, respectively, of the inter-helical angle θ for $Y_IMR^1R^2A_{II}$ (blue lines) and $Y_{III}M_3A_{II}$ (green lines) at 330 K. The simulations were run using OPLS force-field, with the same simulation set-up as described in Molecular Dynamics Simulations section in Materials & Methods, except for the differences summarized in the Table S3.

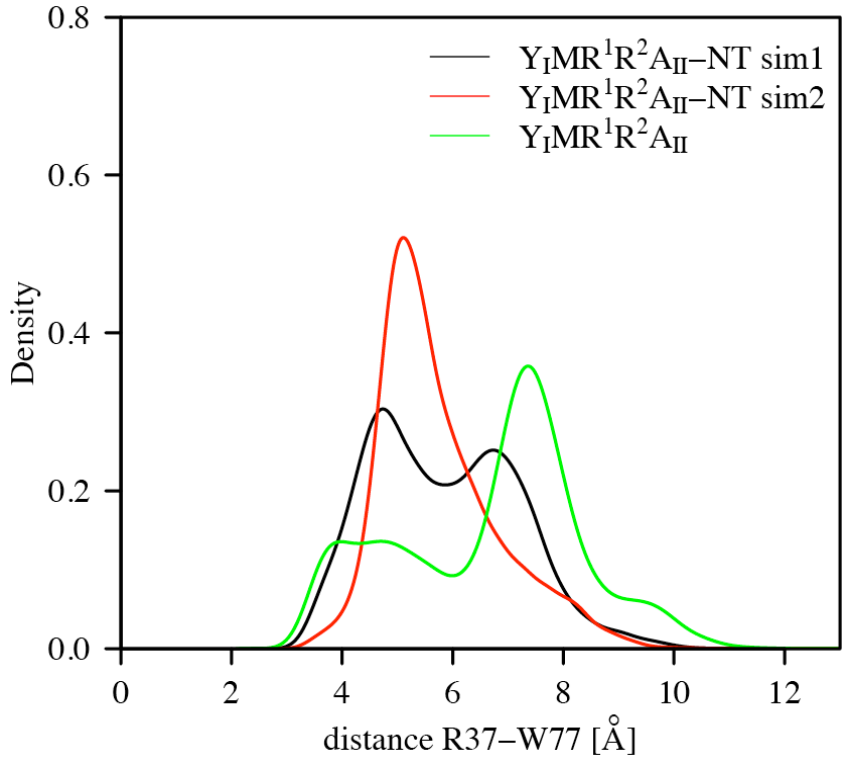

Figure S16: Distance between residues R37 and Trp77 in the two simulations in presence of NT (black and red line) as well as in absence of NT (green line). The distance was measured between the guanidinium group of Arg37 and the center of mass of the aromatic ring of Trp77.

9. ELISA Data:

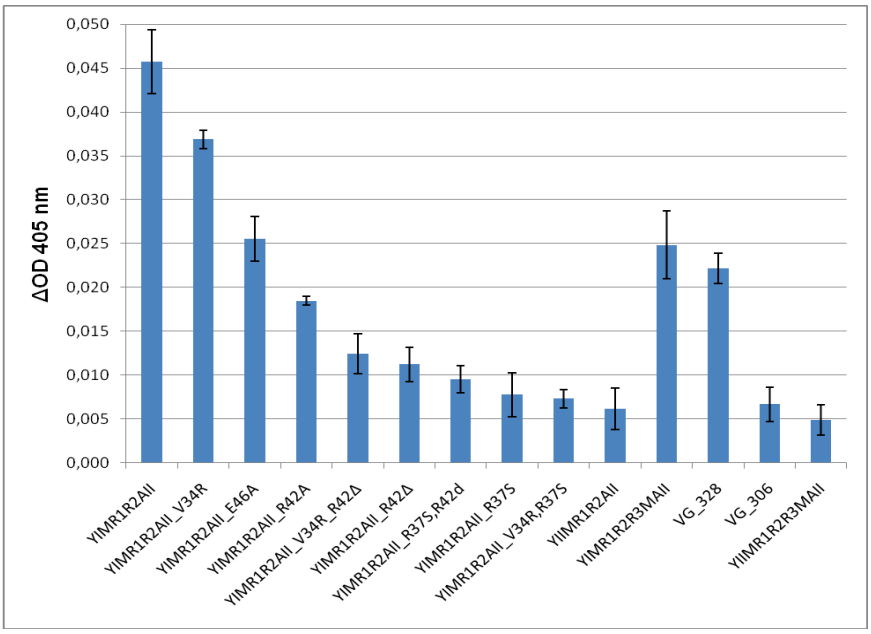

Figure S17: ELISA signal when adding $Y_IMR^1R^2A_{II}$ mutants and VG_328 based reference proteins (all at 200 nM protein) to NeutrAvidin-coated surface, to which biotinylated target peptide NT1-13 (200 nM) had been immobilized. Binding was detected by anti-$RGSH_6$ mouse antibody followed by goat anti mouse IgG alkaline phosphatase conjugate, and signals were read out at 405 nm after 2 h, after signals at the ,reference wavelength

and the background have been subtracted. Since the protein concentrations are very high and the affinities are rather weak, the signals correspond to affinities. We found that $Y_IMR^1R^2A_{II}$ gave a significantly higher signal than the original VG_328 and $Y_IMR^1R^2R^3MA_{II}$ indicating a size-dependent effect between the surface bound ELISA method and in-solution NMR method. For proteins of the same size a similar intensity of the ELISA signal translates reliably into a range of similar K_d values in NMR experiments. VG_306 is a previously published and characterized variant of VG_328 ($Y_IMR^1R^2$ R^3MA_I) carrying the point mutation Y116H. The mutation is located in the first randomized repeat R^1 in a randomized position on helix H3 at position 30 of the repeat that was observed to have a negative effect on peptide binding [23].

10. MEXICO results:

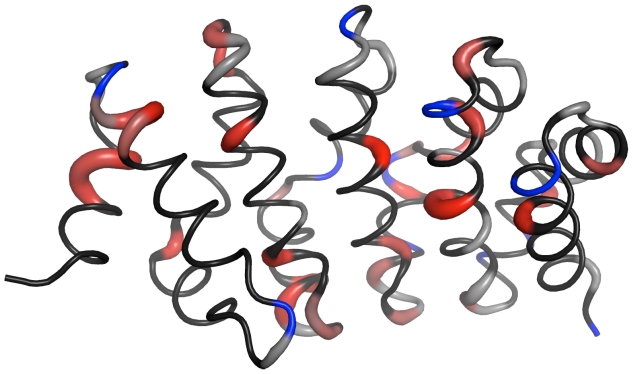

Figure S18: Effect of NT on $Y_IMR^1R^2A_{II}$ solvent accessibility (MEXICO data). The thickness of the backbone denotes the amplitude of normalized ratio of the proton exchange rates in presence to those in absence of NT. The precision of the data is indicated by gradients from gray to red (decreased exchange rate in presence of NT) and gray to blue (increased exchange rate in presence of NT), respectively; maximum color saturation reflects minimal fitting error. Prolines and residues that could not be analyzed are shown in thin black trace.