Characterization and further stabilization of designed ankyrin repeat proteins by combining molecular dynamics simulations and experiments

Supplementary Material

Gianluca Interlandi[a,b], Svava K. Wetzel[a], Giovanni Settanni[c], Andreas Plückthun*, and Amedeo Caflisch*

Department of Biochemistry, University of Zürich,

CH-8057 Zürich, Switzerland

caflisch@bioc.unizh.ch, Tel +41 44 635 5521, Fax +41 44 635 6862

plueckthun@bioc.unizh.ch, Tel +41 44 635 5570, Fax +41 44 635 5712

*Corresponding authors

[a] *Present address: Dept. of Bioengineering, University of Washington, USA*

[b] *Present address: MRC Center for Protein Engineering, University of Cambridge, UK*

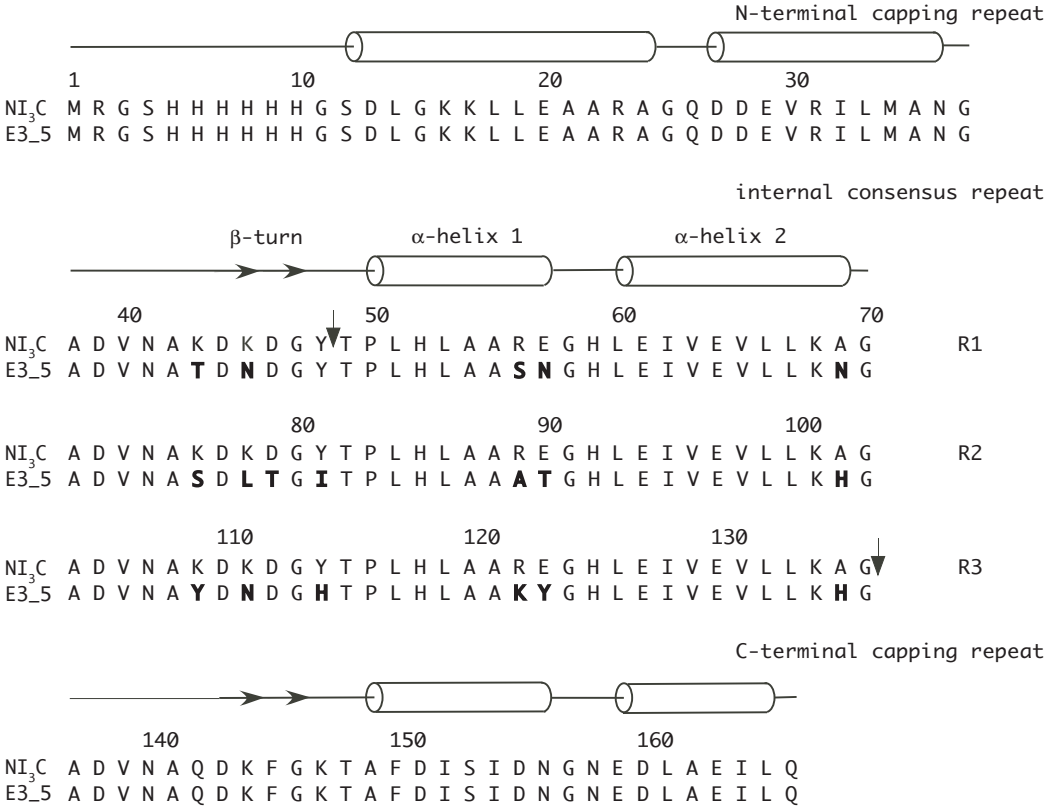

Figure 1: Sequence alignment of the full consensus ankyrin repeat protein NI_3C and the NX_3C library member E3_5 (PDB file 1MJ0). The different residues are in boldface. The "cut" of the capping repeats made in NI_3C in order to create I_3 is shown by vertical arrows after residue 48 and 136. In the experiments the proteins carry a MRGSHHHHHH-tag for purification. In the structures used for the simulations the tag was omitted.

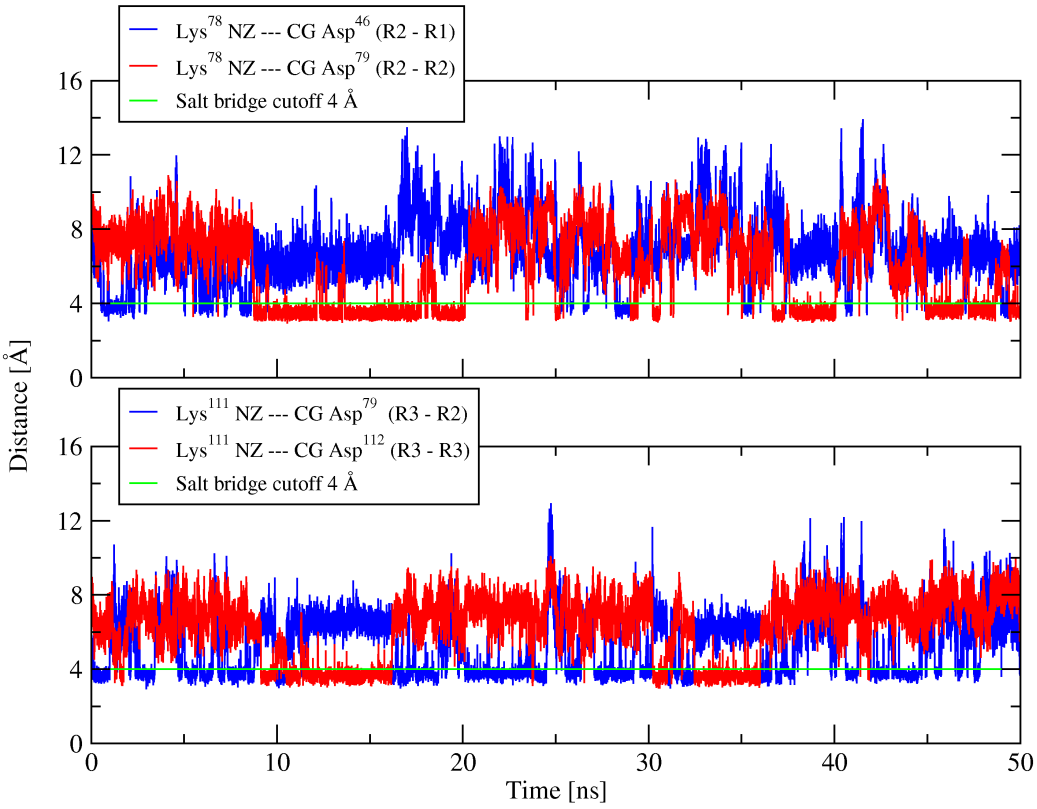

Figure 2: Time series of the salt bridge switch involving Lys^{78} in the second loop **(top)** and Lys^{111} in the third loop **(bottom)** (cf. Figure 1b in the paper) during the 50-ns simulation of NI_3C at 300 K. The side chain of lysine is within salt bridge distance from either the aspartate following in sequence, i.e. intrarepeat, or the aspartate in the loop of the preceding repeat (Figure 1 in the paper). R2, R3 and R4 denote repeat 2, 3 and 4, respectively.

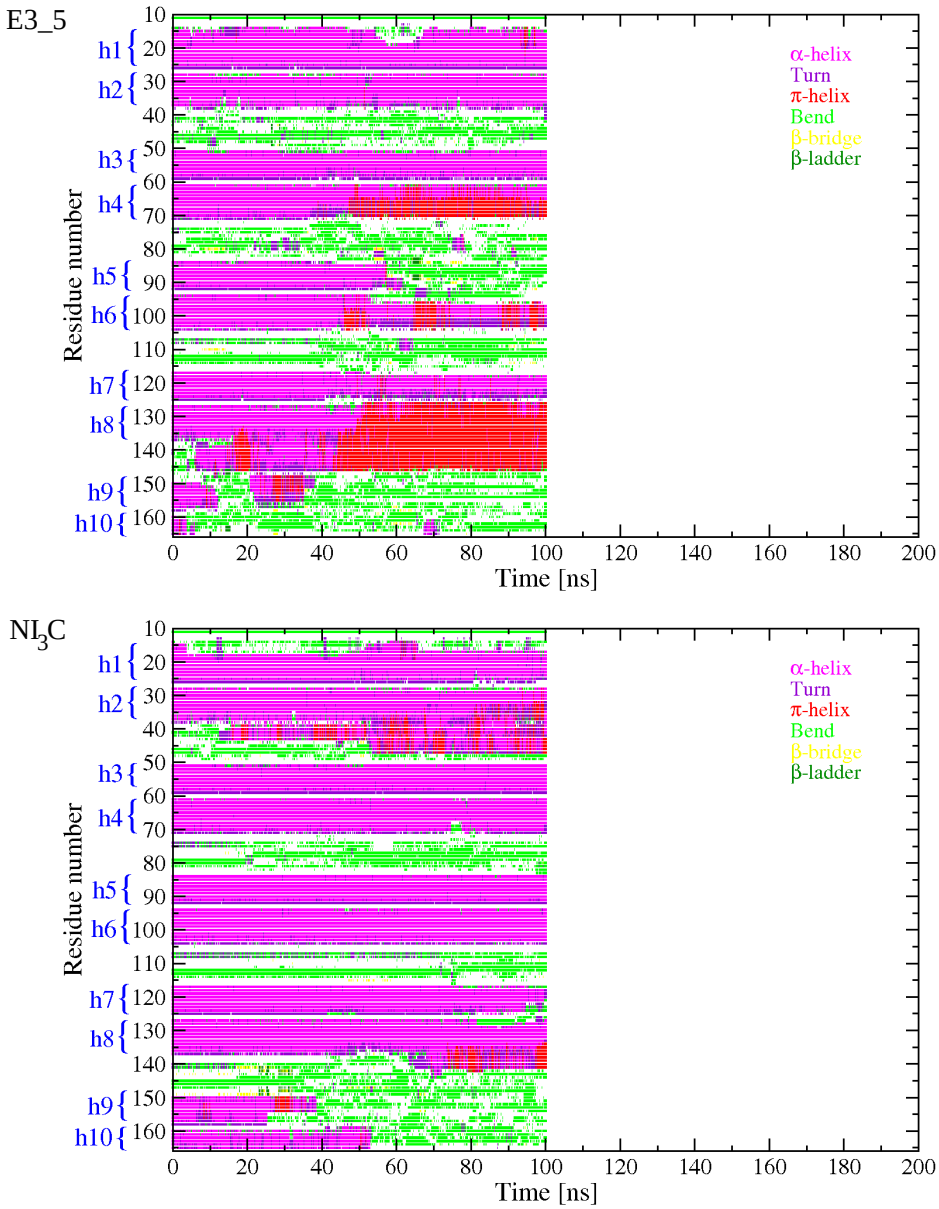

Figure 3: Time series of the secondary structure during the 400 K runs of all molecules investigated here (see also subsequent pages). The helices are labelled on the left side of the plots. The colors refer to different secondary structural elements calculated with DSSP [1].

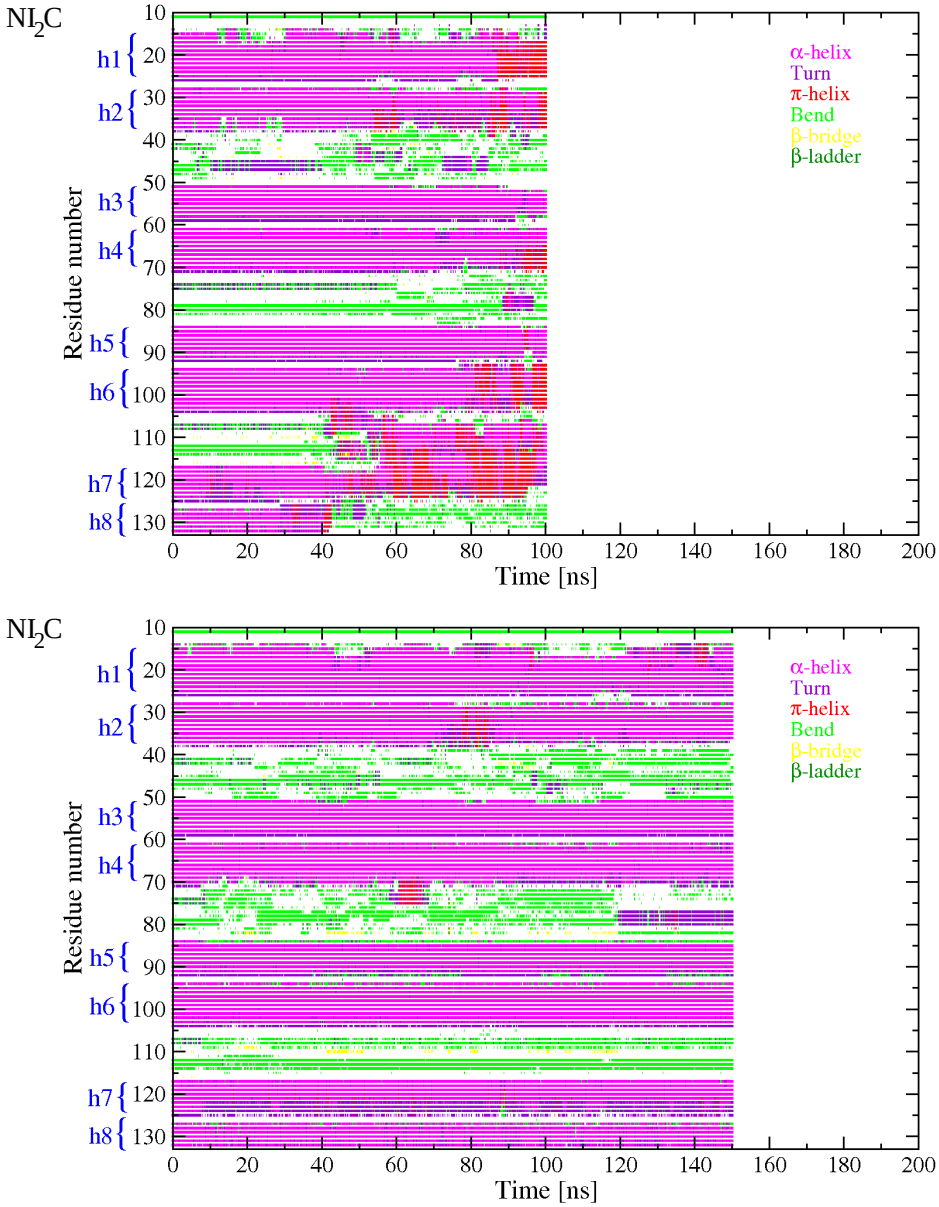

Figure 3 (continued)

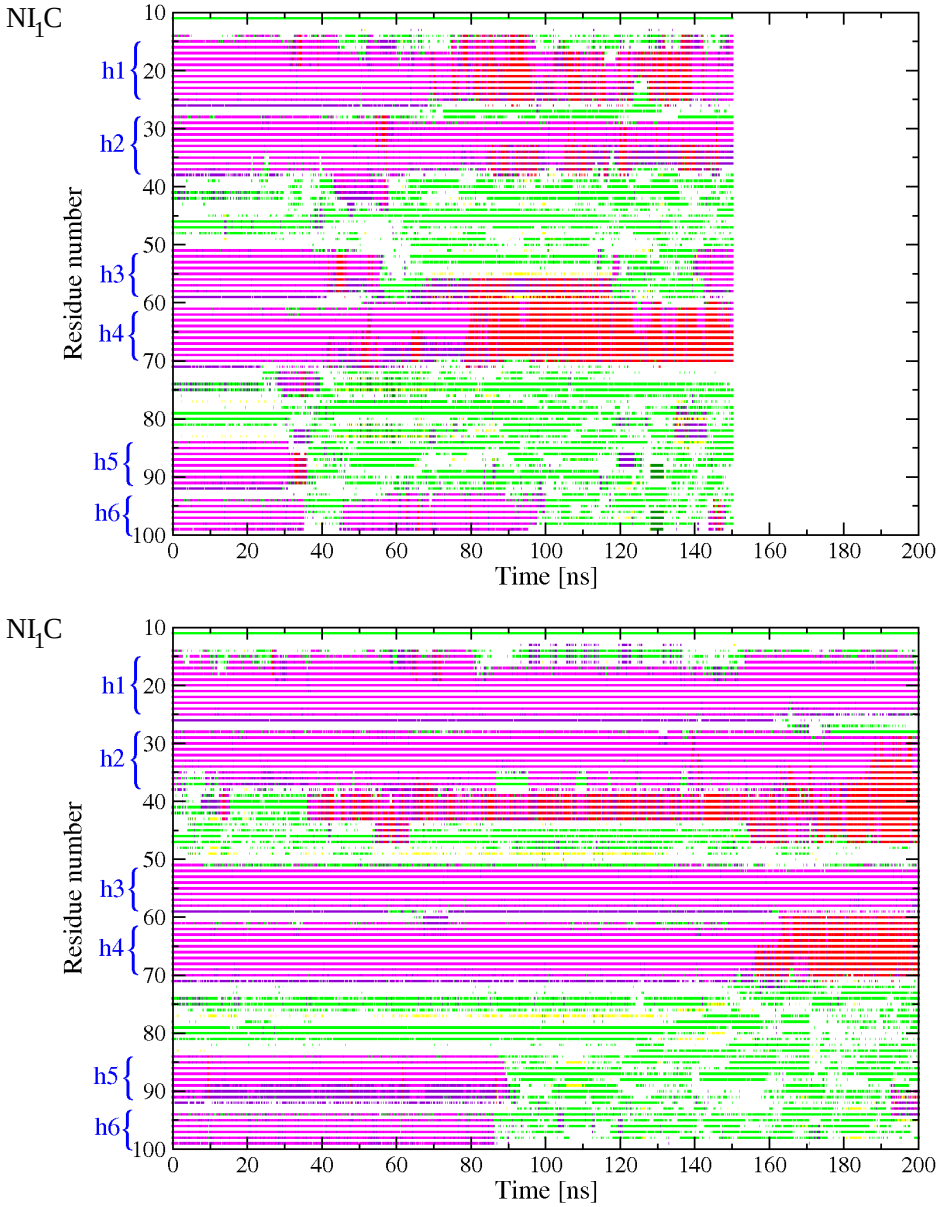

Figure 3 (continued)

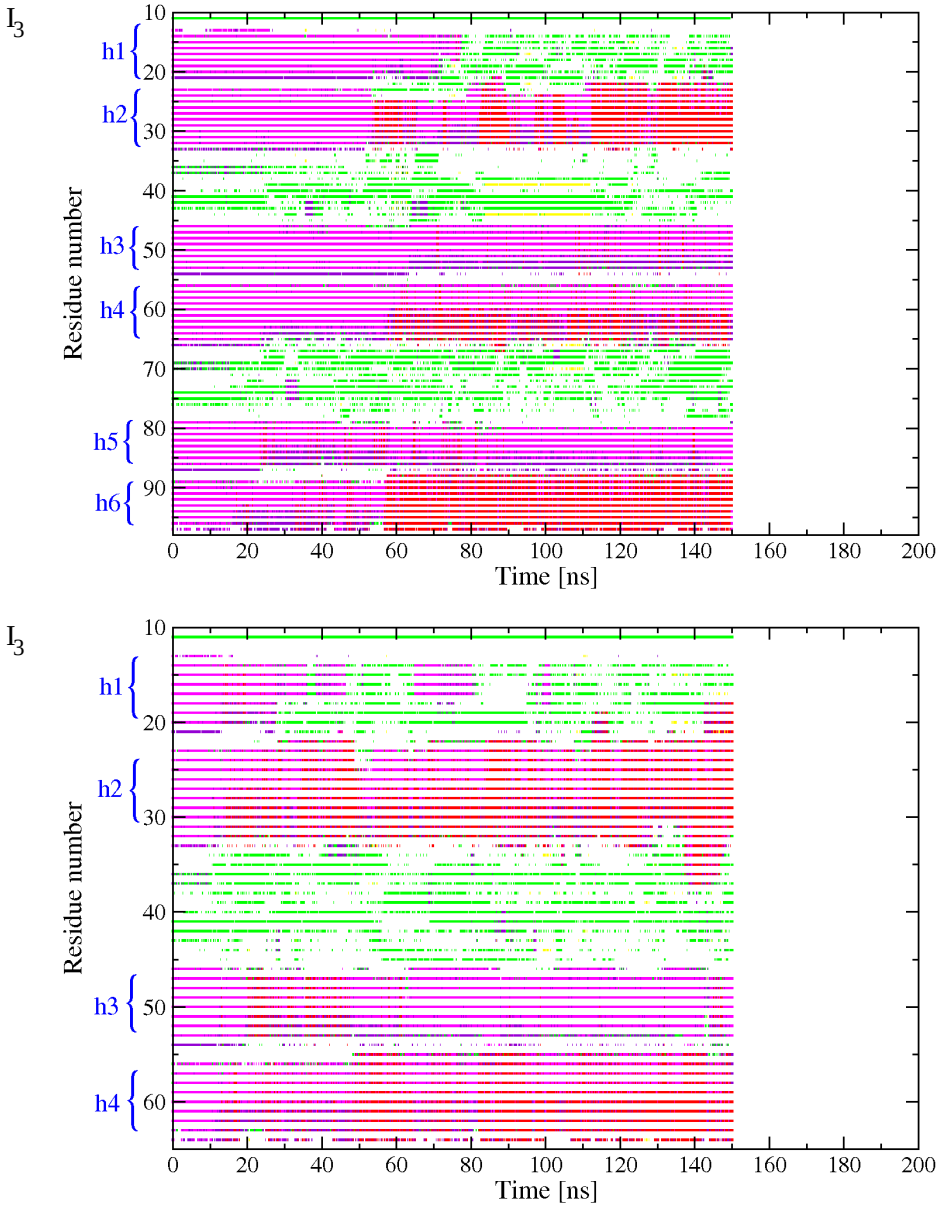

Figure 3 (continued)

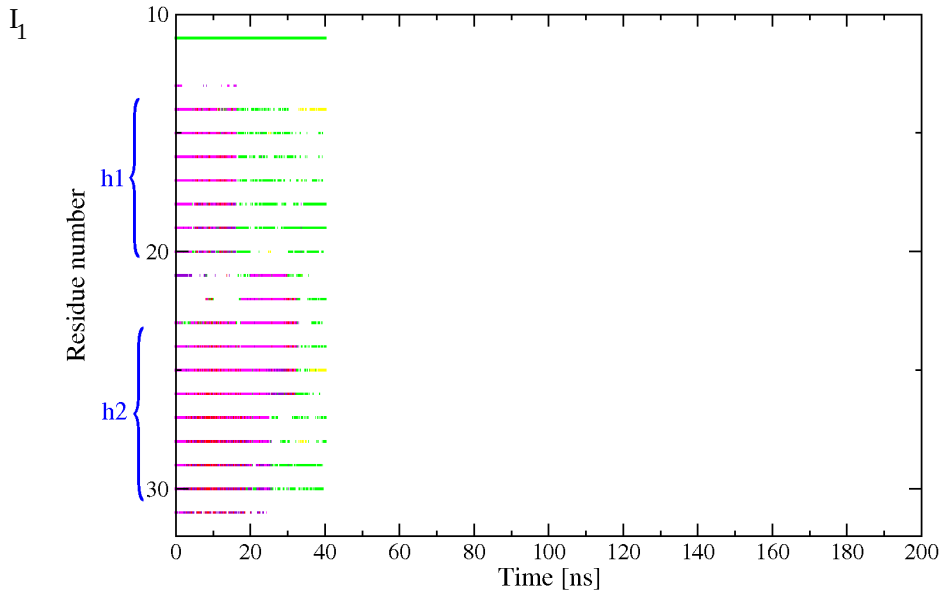

Figure 3 (continued)

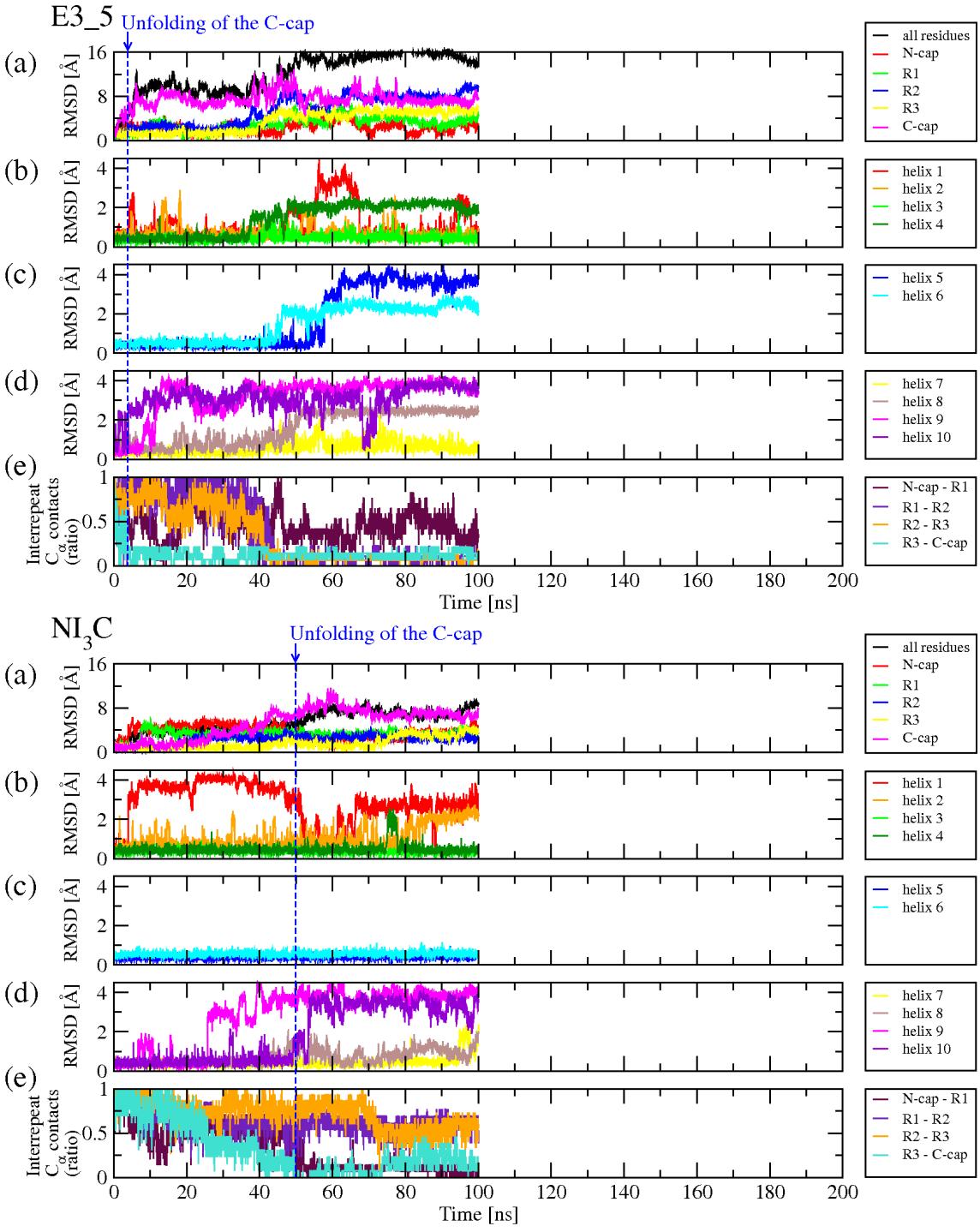

Figure 4: Time series of the 400 K simulations with all molecules investigated here (see also subsequent pages). C_α RMSD from the initial conformation for **(a)** whole protein (black line), single repeats (colored lines), **(b)-(d)** single helices. **(e)** Ratio of native interrepeat C_α contacts. R1 to R3 denote repeat 1 to repeat 3.

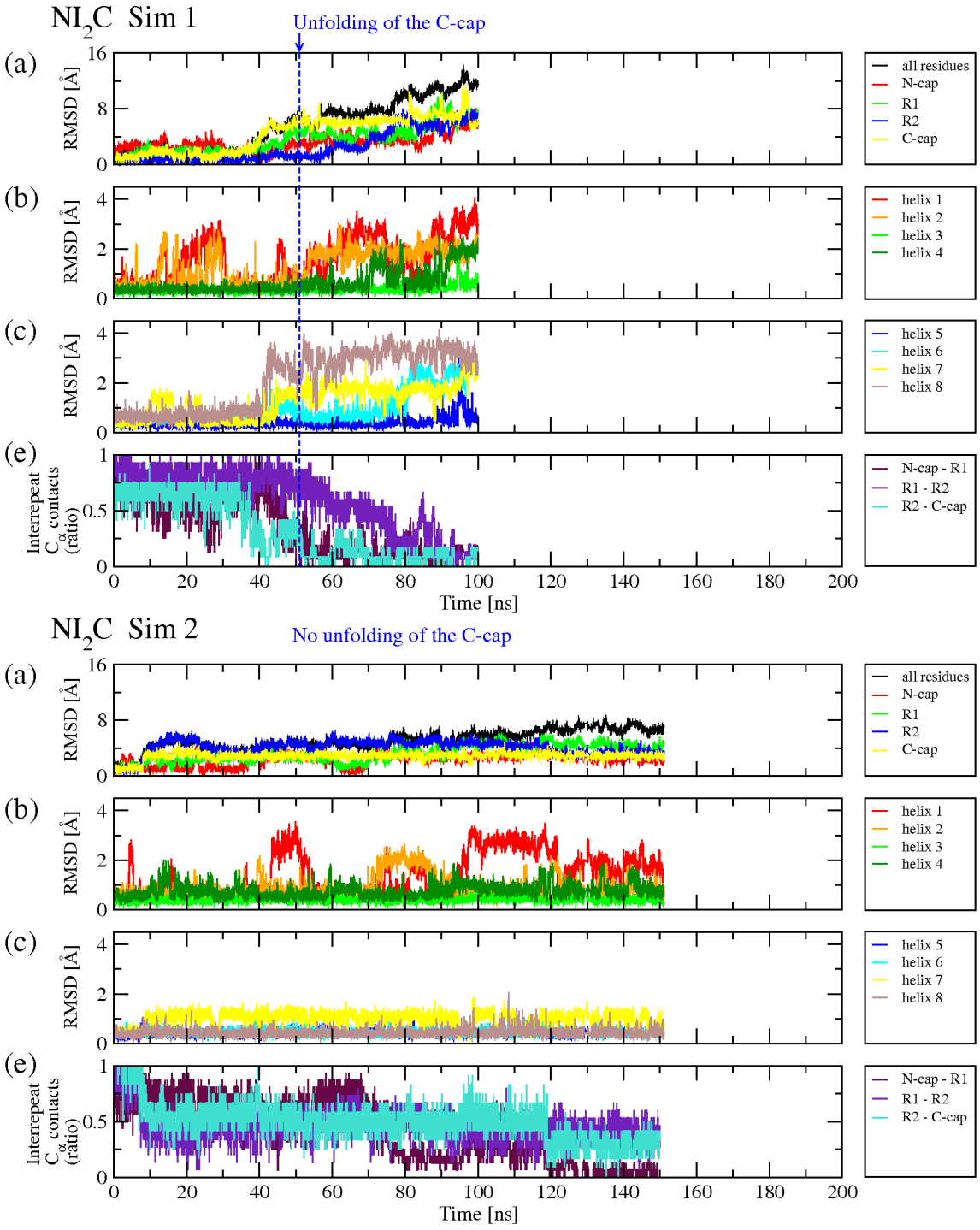

Figure 4 (continued): Two simulations were run with NI_2C: Sim 1 (100 ns) and Sim 2 (150 ns).

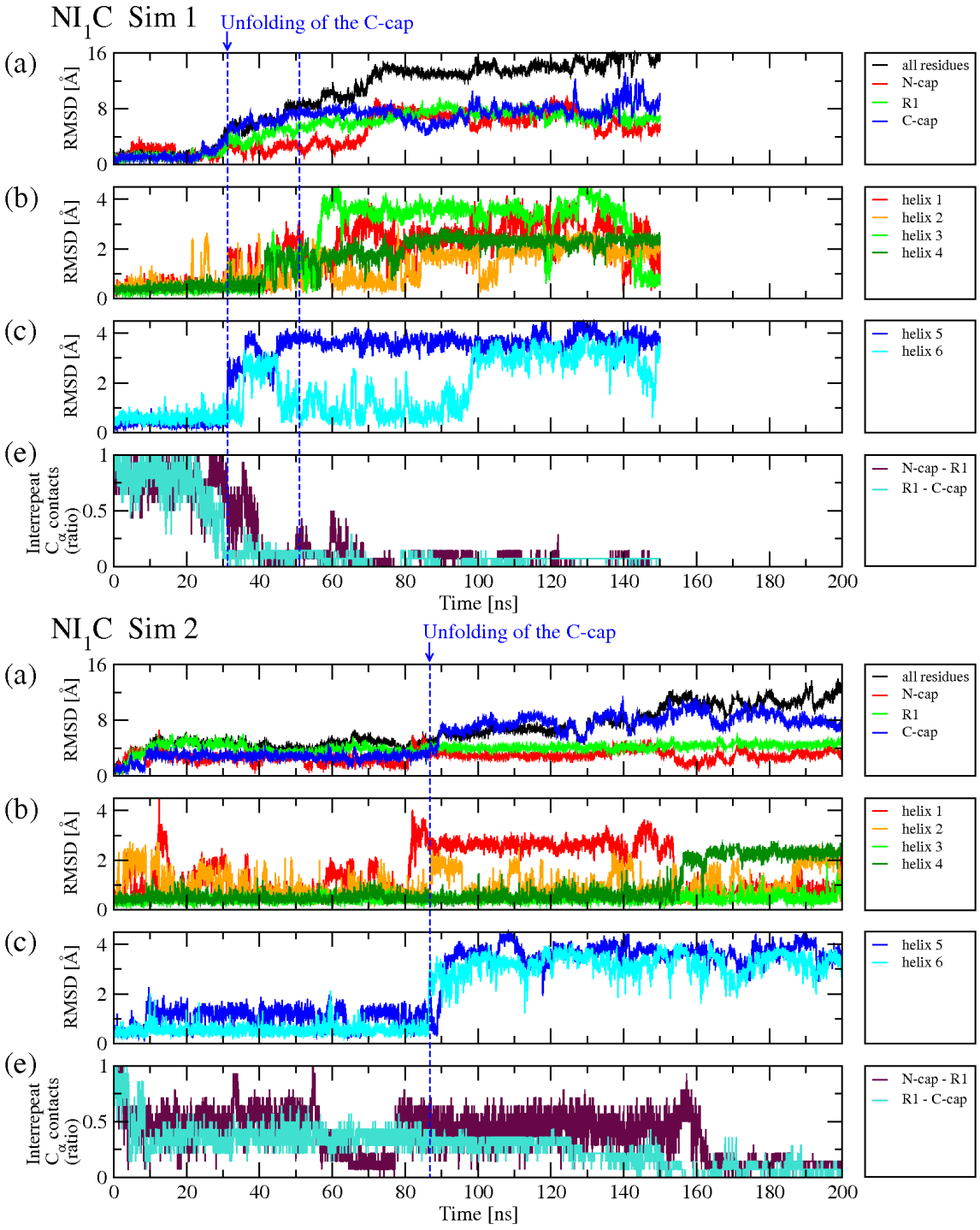

Figure 4 (continued): Two simulations were run with NI_1C: Sim 1 (150 ns) and Sim 2 (200 ns).

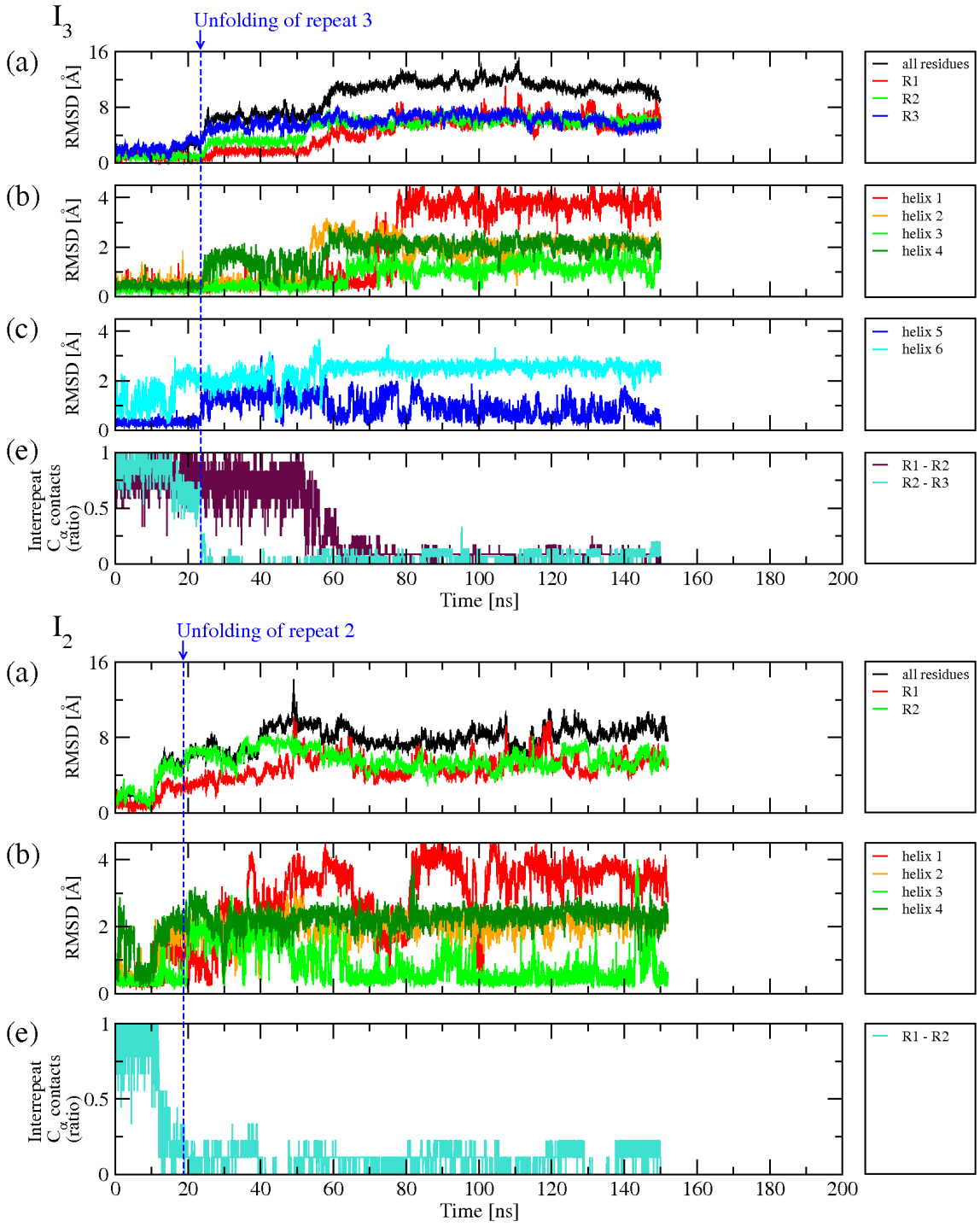

Figure 4 (continued)

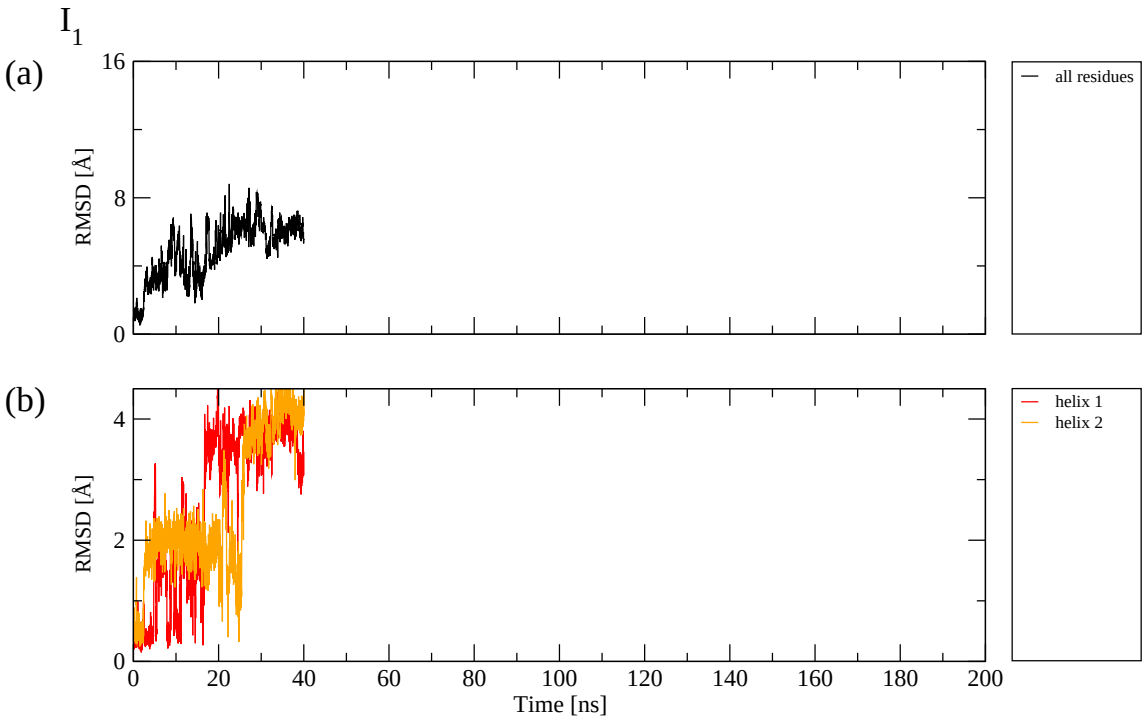

Figure 4 (continued)

Synthesis of DNA Encoding DARPins with C-cap mutations

Oligonucleotides were obtained from Microsynth (Balgach, Switzerland).

forMu2 (forward):
5'-CCATCGACAACGGTAACGAGGACATTGCTGAAGTGC-3'
revMu2 (reverse):
5'- GGAGTCCAAGCTCAGCTAATTAAGTCATTATTGCAGCACTTCAGCAATGTCC-3'
forMu4 (forward):
5'-GGACCTGGCTGAAATCCTGCAAAAAGCGGCGTAATGACTGAGC-3'
revMu4 (reverse):
5'-CAACAGGAGTCCAAGCTCAGTCATTACGCCGCTTTTTGCAGG-3'
pQE_f_1 (forward):
5'-CGGATAACAATTTCACACAG-3'
C1rev (reverse):
5'-GGTCTTACCGAATTTGTCCT-3'
CMu15a (forward):
5'-AGGACAAATTCGGTAAGACCCCGTTCGACTTAGCGATCGACAACGGTAACGAGG-3'
CMu1b (reverse):
5'- ATTATTGCAGGATTTCAGCCAGGTCCTCGTTACCGTTGTC-3'
CMu1rev (reverse):
5'-ATAATTAAGCTTTCATTATTGCAGGATTTCAGCC-3'
CMu3b (reverse):
5'-ATTATTGCAGCACTTCAGCAATGTCCTCGTTACCGTTGTC-3'
CMu3rev (reverse):
5'-CTAATTAAGCTTTCATTATTGCAGCACTTCAGC-3'
CMu5b (reverse):
5'- CTTTTTGCAGCACTTCAGCAATGTCCTCGTTACCGTTGTC-3'

CMu56rev (reverse):
5'- ATAATTAAGCTTTCATTACGCCGCTTTTTGCAGCACTTCAGC-3'
CMu6a (forward):
5'-AGGACAAATTCGGTAAGACCCCGTTCGACTTAGCGATCCGCGAAGGTCATGAGG-3'
CMu6b (reverse):
5'- CTTTTTGCAGCACTTCAGCAATGTCCTCATGACCTTCGCG-3'

The NI_1C mutants Mut 2 and Mut 4 were obtained by site-directed mutagenesis using the oligonucleotides forMu2 and revMu2 or forMu4 and revMu4, respectively, the NI_1C plasmid as template and Pfu Turbo® polymerase (1 min at 95 °C; followed by 18 cycles of 30 s at 95 °C, 1 min at 59 °C (Mut 2) or 52 °C (Mut 4), 20 min at 68 °C; followed by 5 min 68 °C; standard Pfu® polymerase buffer). The template plasmid was *Dpn*I digested, *E. coli* was transformed with the PCR product plasmid and positive clones were sequenced using standard techniques.

Mut 1, Mut 3, Mut 5 and Mut 6 were generated in two steps by assembly PCR. In the first step a constant DNA fragment NI_1 (or NI_3) and a fragment with the C-cap mutation were amplified. For the constant fragment NI_1 (or NI_3) the oligonucleotides pQE_f_1, C1rev, NI_1C plasmid (or NI_3C plasmid) as template and Vent® polymerase were used (5 min at 95 °C; followed by 25 cycles of 30 s at 95 °C, 1 min at 50 °C, 45 s at 72 °C; followed by 5 min 72 °C; standard Vent® polymerase buffer with a final concentration of 5% DMSO). The mutated C-cap fragments for Mut 1, Mut 3 and Mut 5 were amplified with Vent® polymerase using the oligonucleotides CMu15a and CMu1b, CMu3b, CMu5b, respectively, while CMu6a and CMu6b were used for Mut 6. In the second step these PCR products served for assembly PCR including the oligonucleotides pQE_f_1 and CMu1rev, CMu3rev or

CMu56rev, respectively, to generate the whole molecules of Mut 1, Mut 3, Mut 5 and Mut 6. The resulting DNA was cloned via *Bam*HI/*Hin*dIII into pPANK[2], a pQE30 (QIAgen, Germany) derivative lacking the *Bbs*I and *Bsa*I sites, and sequenced using standard techniques.

References

[1] Kabsch, W. & Sander, C. (1983). Dictionary of protein secondary structure - pattern-recognition of hydrogen-bonded and geometrical features. *Biopolymers,* **22**, 2577–2637.

[2] Binz, H. K., Stumpp, M. T., Forrer, P., Amstutz, P. & Plückthun, A. (2003). Designing repeat proteins: well-expressed, soluble and stable proteins from combinatorial libraries of consensus ankyrin repeat proteins. *J. Mol. Biol.* **332**, 489–503.