J Mol Model (2008) 14:747–759
DOI 10.1007/s00894-008-0300-5

ORIGINAL PAPER

Two C-terminal ankyrin repeats form the minimal stable unit of the ankyrin repeat protein p18^{INK4c}

Petr Sklenovský · Pavel Banáš · Michal Otyepka

Received: 21 November 2007 / Accepted: 11 March 2008 / Published online: 15 May 2008

Abstract Ankyrin repeat proteins (ARPs) appear to be abundant in organisms from all phyla, and play critical regulatory roles, mediating specific interactions with target biomolecules and thus ordering the sequence of events in diverse cellular processes. ARPs possess a non-globular scaffold consisting of repeating motifs named ankyrin (ANK) repeats, which stack on each other. The modular architecture of ARPs provides a new paradigm for understanding protein stability and folding mechanisms. In the present study, the stability of various C-terminal fragments of the ARP p18^{INK4c} was investigated by all-atomic 450 ns molecular dynamics (MD) simulations in explicit water solvent. Only motifs with at least two ANK repeats made stable systems in the available timescale. All smaller fragments were unstable, readily losing their native fold and α-helical content. Since each non-terminal ANK repeat has two hydrophobic sides, we may hypothesize that at least one hydrophobic side must be fully covered and shielded from the water as a necessary, but not sufficient, condition to maintain ANK repeat stability. Consequently, at least two ANK repeats are required to make a stable ARP.

Keywords Ankyrin repeat · p18^{INK4c} · Minimal stable unit · Fragmentation · Molecular dynamics

Introduction

Ankyrin repeat proteins (ARPs) are non-globular biomolecules containing repeating units named ankyrin (ANK) motifs arranged in linear, tandem arrays (Fig. 1). The ANK repeat is a protein structural unit that is ~33 residues long and consists of two β-strands and a pair of α-helices assembled in an antiparallel fashion and connected by a short loop (often formed by three residues). Two consensus sequences of the ANK repeat have been defined: the Consensus 1 ANK repeat comprising the β2α2 (ββαα) motif, and Consensus 2 ANK repeat comprising the βα2β (βααβ) motif [1] (Fig. 2). Members of the ARP family co-localize and interact with various membrane and cytoplasmic proteins. The specificity of ARP–protein interactions is likely to be conferred by non-conserved residues flanking each ANK repeat, located at the tips of exposed loops [2].

The number of ANK repeats in naturally occurring ARPs varies from four [3, 4] to 29 [5]. Until 2000, only ARPs containing at least four ANK motifs were considered to be stable proteins [6]. However, a pioneering determination of a minimal stable motif of ARPs by Zhang and coworkers [7], which was based on proteolytic experiments and prediction algorithms, showed that two C-terminal ANK repeats of the naturally occurring p16^{INK4a} [3] protein stacked on each other were stable and could fold independently of the rest of the protein. In addition, numerous artificial ARPs with enhanced thermostability and affinity have been designed recently [8–18], the smallest of which (2ANK) contained two identical ANK repeats [17].

Electronic supplementary material The online version of this article (doi:10.1007/s00894-008-0300-5) contains supplementary material, which is available to authorized users.

P. Sklenovský · P. Banáš · M. Otyepka (✉)
Department of Physical Chemistry and Center for Biomolecules and Complex Molecular Systems, Faculty of Science,
Palacký University,
tr. Svobody 26,
771 46 Olomouc, Czech Republic
e-mail: otyepka@aix.upol.cz

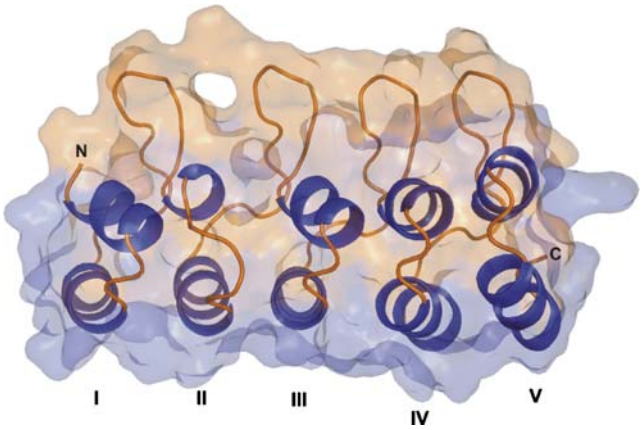


Fig. 1 Structure of the p18^{INK4c} protein (PDB entry 1IHB, chain B), which is a member of the cyclin-dependent kinase inhibitor (INK) tumor suppressor family with five ankyrin (ANK) repeat modules. The figure was generated by PyMol [30], http://pymol.sourceforge.net/

Furthermore, on the basis of proteolytic experiments and X-ray crystallographic studies, it was concluded that single ANK repeats cannot form stable native structures under physiological conditions [7, 8], so at least two ANK repeats appear to be required to form stable ARPs, and thus comprise a minimal folding unit. On the other hand, Ferreiro and coworkers [19] suggested from Gō-type simulations that the minimal topological folding module of ARPs may be even smaller, consisting of one fully folded ANK repeat followed by the C-terminal helix of the neighboring repeat.

Protein p18^{INK4c} [20] (p18) is a member of the cyclin-dependent kinase inhibitor (INK) tumor suppressor family and consists of five ANK repeat modules (Fig. 1). In the present study, the stability of various p18 C-terminal fragments was investigated using all-atomic molecular dynamics (MD) simulations in explicit water solvent at the tens of nanoseconds timescale, totaling 450 ns. Our results show that the minimal folding unit proposed by Ferreiro and coworkers [19] is unstable, and that the minimal stable unit of the p18 protein, at the studied timescale, consists of two C-terminal ANK repeats.

Methods

Studied fragments

The crystal structure of the p18 protein [20] (PBD entry 1IHB, chain B) with optimized positions of hydrogen atoms (obtained using the Sander module of AMBER [21]; *parm99* force field [22, 23]) was used as a template for studying all of the following p18 fragments in MD simulations (Table 1). The N-terminus of each fragment was acetylated and the C-terminus was capped with an *N*-methyl group:

- Two pairs of α-helices: α7+α8 (residues 106–112 + 116–125) and α9+α10 (residues 140–146 + 150–159)
- Two helix-turn-helix motifs: α7-turn-α8 and α9-turn-α10 (residues 106–125 and 140–159)
- Consensus 2 ANK repeat IV (residues 102–135)
- The hypothetical minimal folding unit suggested by Ferreiro et al. [19] (truncated form with six N-terminal β-hairpin residues 100–105 missing) α7-turn-α8-loop-α9 motif (residues 106–146)
- Consensus 2-based hypothetical minimal folding unit [19]: ANK IV-loop-α9 (residues 102–146)
- Pair of helix-turn-helix motifs: α7-turn-α8+α9-turn-α10 (residues 106–125 + 140–159)
- C-terminal fragment α7-turn-α8-loop-α9-turn-α10 (residues 106–159)

Molecular dynamics simulations

MD simulations of the p18 fragments were carried out using the AMBER [21] suite of programs with the *parm99* force field [22, 23] and the following simulation protocol [24–26]: each system was neutralized by adding counter ions (either Cl^- or Na^+ ions according to solute charge) and immersed in a rectangular water box (TIP3P [27]) with a minimum distance between the solute and the box wall of 10 Å. Then, each system was minimized prior to the production phase of the MD run, as follows. The protein

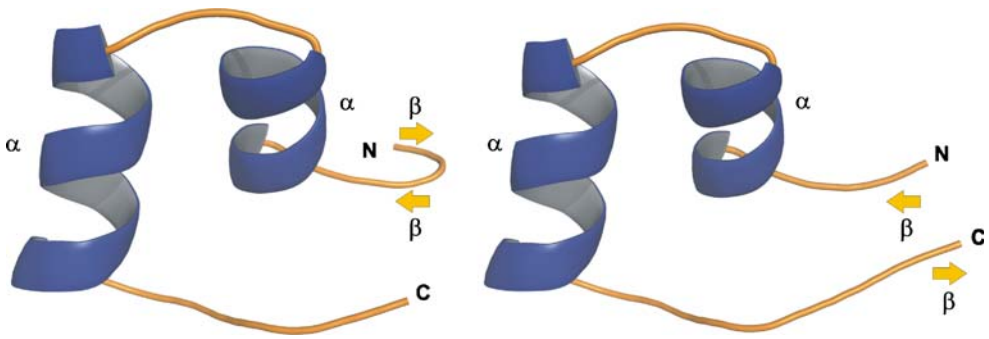


Fig. 2 Illustrations of the ankyrin (ANK) repeat motif: *left* Consensus 1, β2α2 motif (β-hairpin perpendicular to antiparallel α-helices linked by a short loop); *right* Consensus 2, β-strand-helix-loop-helix-β-strand sequence, for definitions, see [1]

Table 1 Basic features of studied fragments

Fragment	Charge (e)[a]	Number of residues	Number of heavy atoms	% of hydrophobic residues	Periodic box dimensions (Å)[b]
α7+α8	2	17	139	64.7	40.0×48.9×34.6
α9+α10	−1	17	126	58.8	41.2×44.6×37.4
α7-turn-α8	1	20	162	55.0	42.3×48.1×37.1
α9-turn-α10	0	20	153	50.0	42.4×46.4×38.2
ANK repeat IV	1	34	267	41.2	45.4×51.2×39.4
α7-turn-α8-loop-α9	2	41	317	43.9	45.1×54.3×43.6
ANK IV-loop-α9	1	45	346	42.2	45.7×54.6×45.1
α7-turn-α8+α9-turn-α10	1	40	315	52.5	44.2×49.7×44.6
α7-turn-α8-loop-α9-turn-α10	2	54	419	42.6	49.7×54.6×46.3

[a] The solute's total charge under physiological conditions
[b] Mean periodic box dimensions calculated from the last 5 ns of each molecular dynamic (MD) simulation

was constrained and the solvent molecules with counter ions were allowed to move during a 1,000-step minimization followed by a 10-ps-long MD runs under [*NpT*] conditions (p=1 atm, T=298.15 K). The side chains were then relaxed by several minimizations, with decreasing force constants applied to the backbone atoms. After the relaxation, each system was heated to 298.15 K as follows: from 10 K to 50 K for 20 ps, then from 50 K to 250 K for 20 ps, and finally from 250 K to 298.15 K for 30 ps. The particle-mesh Ewald (PME) method for treating electrostatic interactions was used, and all simulations were performed under periodic boundary conditions in the [*NpT*] ensemble at 298.15 K and 1 atm using a 2 fs integration step. The SHAKE algorithm with a tolerance of 10^{-5} Å was used to fix positions of all hydrogens, a 9.0 Å cutoff was applied to nonbonding interactions and coordinates were stored every picosecond. Totally, nine independent simulations each 50 ns long were performed. Thus, the cumulative production time amounted to 450 ns.

The stability and structural properties of the p18 fragments were evaluated by calculating root-mean-square-deviations (RMSDs) for backbone atoms from the initial structure, radius of gyration (R_g), secondary structure elements obtained from the DSSP program [28], and native contacts obtained using the in-house program RESDIST (P. B., unpublished software). The RESDIST program calculates a contact map (map of distances) among all residues represented by centers of masses of side chains. Contacts between $i...i$ + 4 residues and higher are considered as native (excluding contacts between $i...i$ + 1, $i...i$ + 2 and $i...i$ + 3 residues) if the distance between two residues is smaller than or equal to 6.0 Å. The program analyzes the number (percentage) of saved native contacts during the MD simulation. The mean values of all structure parameters are listed in Table 2.

Table 2 Summary of trajectories' characteristics

Fragment	t (ns)[a]	R_g (MD)[b]	RMSD (Å)[c]	Native contacts (%)[d]	α-helicity (%)[e]
α7+α8	50	10.1±3.4	6.1±2.7	5.0±6.7	14.0
α9+α10	50	8.5±2.6	5.6±1.9	0.7±2.6	5.8
α7-turn-α8	50	7.7±0.5	3.7±0.3	22.3±6.3	19.7
α9-turn-α10	50	9.0±0.4	5.2±0.6	12.6±5.9	33.6
ANK repeat IV	50	8.3±0.1	2.8±0.1	24.3±3.0	39.4
α7-turn-α8-loop-α9	50	11.9±0.5	10.0±0.7	10.7±3.1	34.6
ANK IV-loop-α9	50	9.1±0.1	3.2±0.2	31.6±3.4	54.0
α7-turn-α8+α9-turn-α10	50	8.5±0.1	1.2±0.3	75.3±6.2	89.1
α7-turn-α8-loop-α9-turn-α10	50	9.4±0.1	1.7±0.3	53.5±5.3	83.7

[a] The time evolution of presented structure characteristics are shown in the Supplementary material
[b] Mean radius of gyration ± SD of main chain atoms calculated from the last 5 ns of each simulation
[c] Root-mean-square-deviation of helices main chain atoms compared with the initial structure calculated as the mean ± SD from the last 5 ns of each MD simulation
[d] Mean number of native contacts ± SD calculated from the last 5 ns of each MD simulation
[e] Mean native α-helical content averaged along the MD simulation. The native α-helical contents are normalized to the initial structure, for which the α-helical content equals 100%

Results

Helix pair α7+α8

Several hydrophobic contacts form between residues in the α7+α8 pair of α-helices (Leu122 with Pro106 and Leu107; Leu107 with Val123; and Ala110 with Val118 and Val119, see Fig. 3a), but this fragment is highly unstable and both helices rapidly lose their initial α-helical structures. The unfolding of the α7 helix began from the N-terminus at ~200 ps. Shortly thereafter (~300 ps) unfolding of α8 started, again from the N-terminus. Then, at ~600 ps, the middle turn of α8 was reestablished and remained stable until Ala110 lost its contact with Val119 (~800 ps), which also initiated separation of the α7 C-terminus from the α8 N-terminus, accompanied by a significant lowering of the number of native contacts and α-helical content (cf. Fig. 4). The helices were then held together via hydrophobic contacts of Leu107 with Val119 and the hydrophobic portion of Glu120, and via an H-bond that formed between the Glu120 carboxyl oxygen and the His108 side chain (Fig. 3b). Those interactions were sufficiently strong to prevent dissociation of the two helices. Even at 1.5 ns, α-helices were still oriented in the antiparallel arrangement due to hydrophobic interaction among Leu107, Leu122 and Val119 (Fig. 3b). From this time until ~9.5 ns, the α7 helix refolded repeatedly, whereas the α8 secondary structure remained in the non-helical conformation. Thereafter, refolding of the α8 was initiated from the C-terminus and assisted by the formation of a hydrophobic cluster consisting of Leu107, Val118, Val119, Leu122 and Val123 (Fig. 3b). The reestablished α-helical structure of α8 was maintained for ~3 ns. Subsequently (at ~12.5 ns), both α-helices lost their α-helical content (cf. Fig. 4). The dissociation of both α-helices appeared at ~32 ns and no other refolding event was observed during the rest of the simulation.

Helix pair α9+α10

The hydrophobic face of the α9 helix consists of Cys141 and Ala144, while the hydrophobic face of α10 consists of

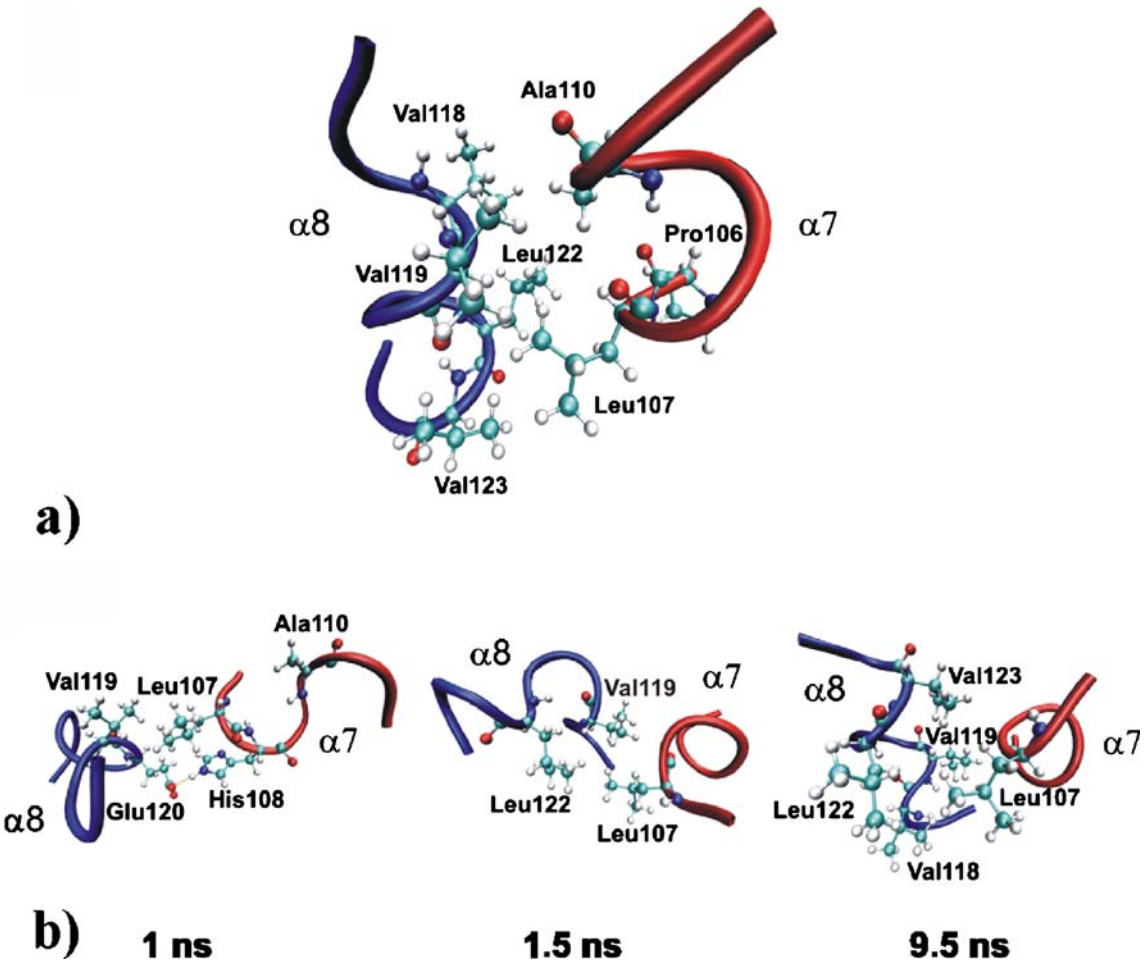
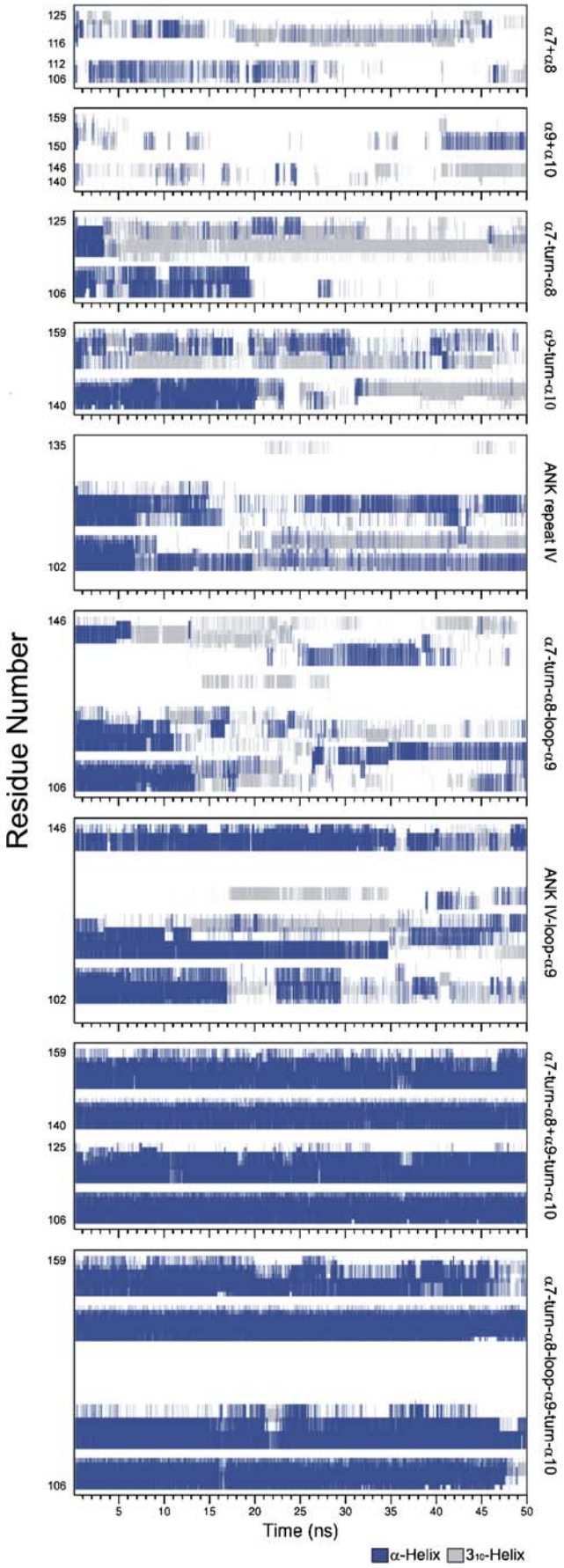


Fig. 3 **a** The α7 (*red*) and α8 (*blue*) helix pair in its native conformation. **b** Snapshots from the α7+α8 fragment molecular dynamics (MD) simulation taken at various times



◀ **Fig. 4** Plot of the secondary structure elements calculated by DSSP software as a function of time for all simulated p18 fragments. *Blue bars* residues with α-helical conformation, *gray bars* residues with 3_{10}-helical conformation. For the sake of simplicity, other secondary structure elements are not depicted

Val152, Val153 and Met156 (Fig. 5a). As in the previous case, this structural motif is highly unstable when taken out of its native context, readily losing its initial α-helical structure. Firstly, the α9 unfolded simultaneously (from both termini) at ~200 ps (Fig. 5b). This process was accompanied by a ~20% reduction in native contacts. Then, at ~350 ps, α10 unfolded from the C-terminus and an additional ~20% decrease in native contacts was observed (Fig. 5b). The α9+α10 pair dissociated at ~6 ns (Fig. 5b) and, simultaneously, the α-helical content significantly decreased (cf. Fig. 4). Notably, except for a reformation of the N-terminal turn of the α10 helix, which appeared at ~40 ns, no other refolding events were observed.

α7-turn-α8 motif

Three residues (Glu113, Gly114 and His115) with side chains exposed to water make the turn connecting α7 with α8. The α7-turn-α8 fragment proved to be less dynamic than the unconnected variant (i.e., the pair of helices); the presence of the turn significantly restricts the conformational space of the fragment, thereby slowing kinetic processes such as unfolding and structural rearrangements.

At ~2.2 ns, α7 unfolded from the C-terminus, an event initiated by loss of the H-bond between the backbone amide hydrogen of His108 and the backbone carbonyl group of Lys112. By ~2.5 ns, all native α-helical H-bonds were broken and the entire α7 helix was fully unfolded. After that, the main chain axis of the unfolded α7 flipped out of the helix-turn-helix plane by ~90° (Fig. 6). The α8 helix then unfolded from the C-terminus at ~3.2 ns. Simultaneously, the distance between the N-terminus of α7 and C-terminus of α8 increased while the number of native contacts decreased (Fig. 6). Loss of the H-bond between the carbonyl group of Leu116 and the amide hydrogen of Glu120 probably initiated α8 melting. The middle turn of α8 adopted a 3_{10}-helix conformation, which was well maintained during the following simulation run. At ~3.5 ns, the N-terminal helix turn again reformed, and was maintained for ~1 ns, then, at 5.2 ns, a second refolding of the α7 helix occurred, which was propagated from the C-terminus and maintained for ~4 ns. Reestablishment of the α7 helical structure was initiated by the formation of hydrophobic contacts among the side chain of Leu109, the hydrophobic moiety of Lys112 and the hydrophobic moiety of Glu113 (Fig. 6). At ~6 ns, both helices again rearranged into the antiparallel conformation and the number of native

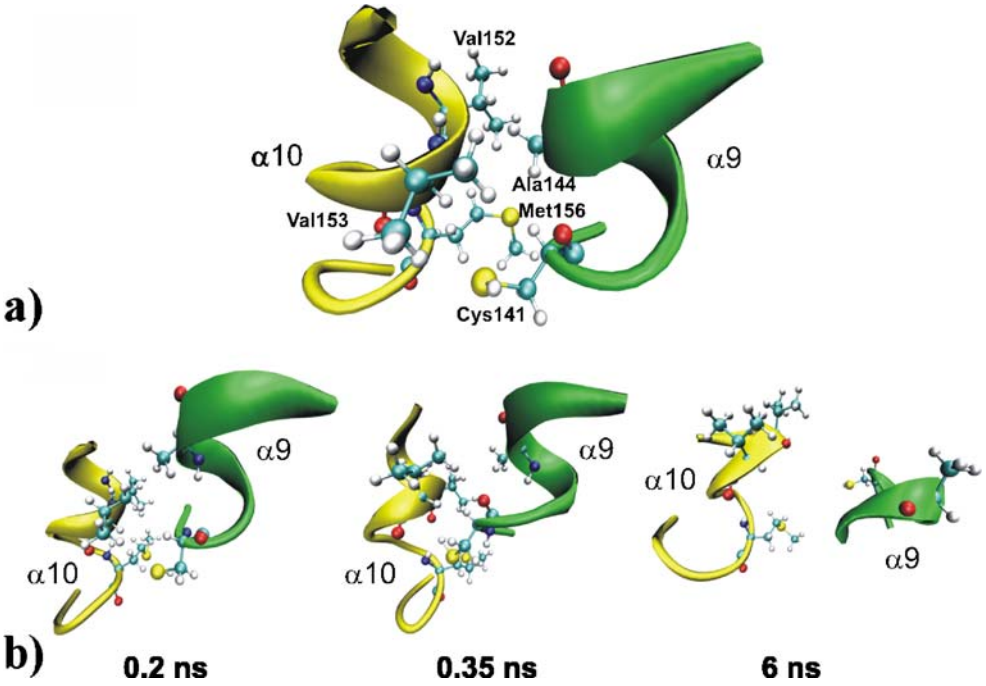


Fig. 5 **a** The α9 (*green*) and α10 (*yellow*) helix pair in its native conformation. **b** Snapshots from the α9+α10 fragment MD simulation taken at various times

contacts increased. At ~11.5 ns, the last refolding event of α7 occurred, beginning from the N-terminus, and the helical structure was maintained until ~19.5 ns. Then, the C-terminal turn of α8 refolded twice (at ~20 ns and ~23.5 ns), for ~2 ns on each occasion. No other refolding events were observed during the rest of the simulation. Finally, at ~46 ns, the opposite termini of individual helices moved away (Fig. 6). This extended structure of the α7-turn-α8 fragment was retained until the end of the simulation.

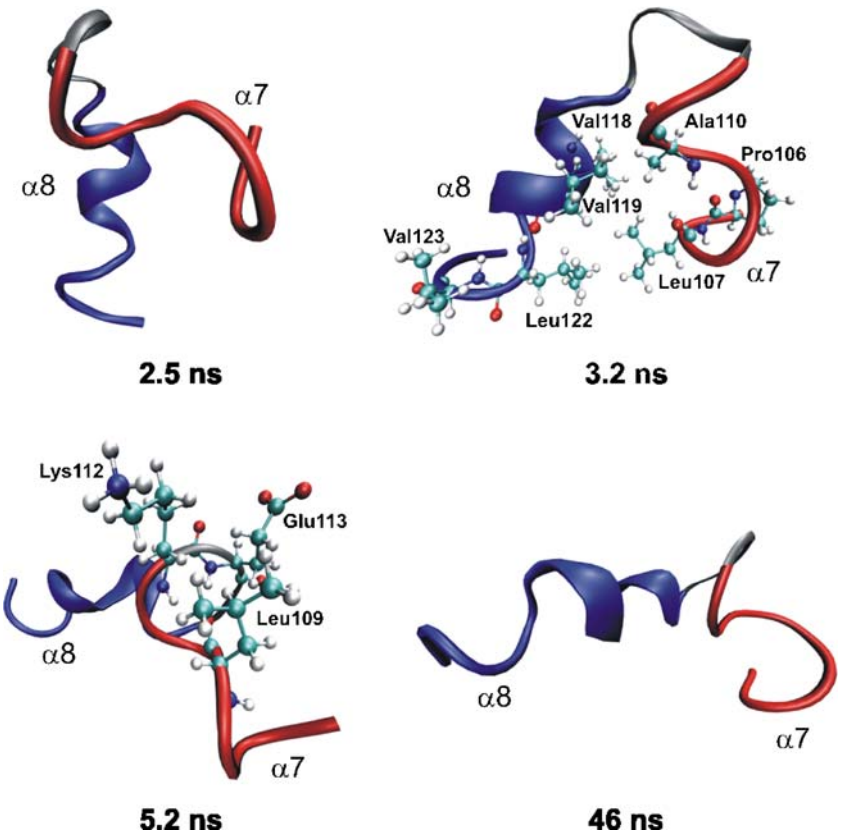


Fig. 6 Snapshots taken from the α7-turn-α8 fragment MD simulation at various times; *red* α7, *gray* turn, *blue* α8

α9-turn-α10 motif

The turn here consists of Tyr147, Gly148 and Arg149 residues, all of which have water-exposed side chains. The mean α-helical content of the α9-turn-α10 fragment is almost six times higher than that of the α9+α10 pair (cf. Table 2), i.e., the two helices without the covalent linker. In the case of both α7+α8 vs α7-turn-α8 and α9+α10 vs α9-turn-α10, the turns restrict the conformational space and significantly enhance the α-helical propensity of the respective fragments.

Unfolding of the α10 helix started from the N-terminus at ~0.7 ns and was initiated by formation of an H-bond between the backbone carbonyl group of Asn150 and the side chain hydroxyl of Ser154, thereby disrupting the H-bond between the backbone carbonyl of Asn150 and the amide hydrogen of Ser154. This was followed by loss of the H-bond between the backbone carbonyl moiety of Leu155 and backbone amide hydrogen of Asn159, and consequent unfolding of the C-terminal and middle turn of α10. After that, formation of an H-bond between the carbonyl group of Val153 and side chain of Gln157 mediated disruption of the H-bond between the carbonyl group of Val153 and backbone amide hydrogen of Gln157 (Fig. 7). At ~6.2 ns, the α9 helix flipped out of the helix-turn-helix plane by almost ~90° (T-shape arrangement, see Fig. 7). Then, at ~15 ns, the most significant refolding of the α10 middle turn appeared, assisted by formation of an H-bond between the carbonyl group of Val153 and the side chain of Gln157. The reformed middle turn of the α10 helix was stable for ~2 ns. Then, at ~17 ns, the hydroxyl group of Tyr147 (turn residue) formed an H-bond with the backbone carbonyl oxygen of Asn159 (Fig. 7). Simultaneously, the middle turn of the α10 again refolded due to formation of an H-bond between the backbone carbonyl of Ser154 and side chain of Gln157. After that (at ~20 ns), the distance between the two helices increased (the opposite helical termini moved away), accompanied by a reduction in native contacts. Simultaneously, α9 unfolded from its N-terminus. It is likely that the loss of hydrophobic contacts among Leu143, Ala145 and Val153 initiated α9 unfolding. At ~22.5 ns, the α9 helix briefly (for ~1 ns) refolded, mediated by reformation of the hydrophobic cluster consisting of Leu143, Ala145 and Val153. Between ~24.3 and ~24.8 ns, the fragment adopted a parallel arrangement (Fig. 7). After that, the T-shape arrangement of the

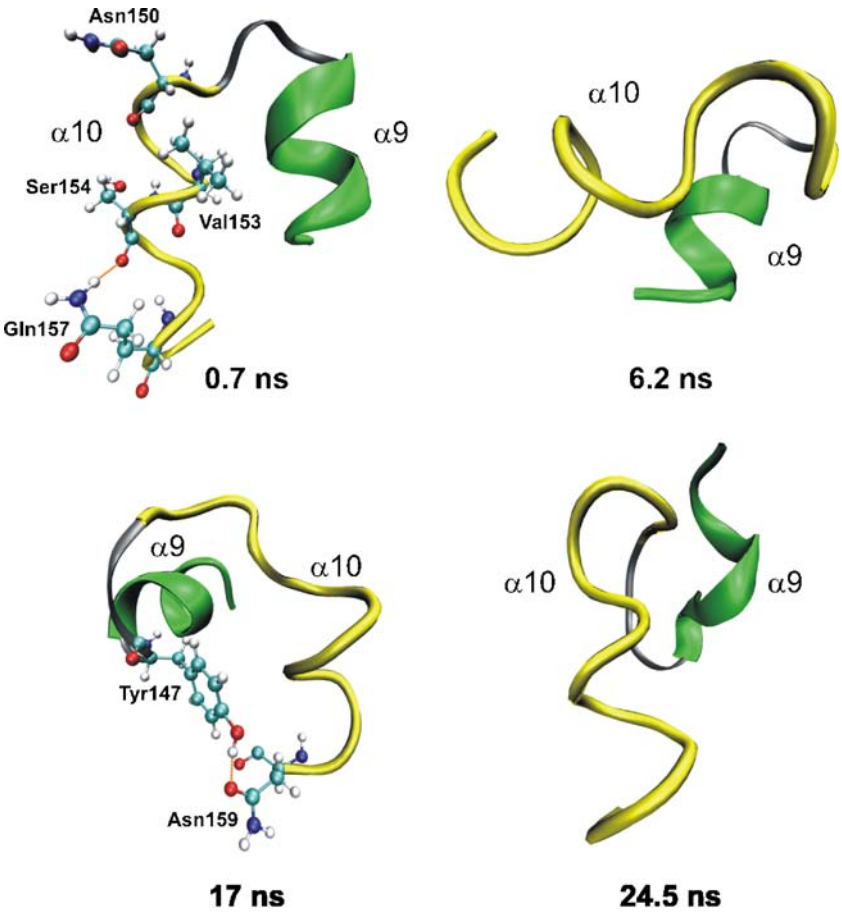


Fig. 7 Snapshots taken from the α9-turn-α10 fragment simulation at various times; *green* α9, *gray* turn, *yellow* α10

fragment again appeared. At ~31.1 ns, the last refolding of α9 was observed, and the helical structure was maintained for ~0.9 ns. From this point onwards, no other refolding of the α9 helical structure was observed, although α10 refolded several times in the remaining simulation time. At ~31.3 ns, the antiparallel arrangement of both helices was reestablished and maintained until ~38.5 ns. Thereafter, the T-shape conformation again appeared. Finally, at ~42 ns, opposite termini of the two helices moved away and the fragment adopted an extended structure, which was maintained until the end of the simulation.

ANK repeat IV

The four C-terminal residues of α8 quickly (at ~0.4 ns) adopted a flexible loop conformation. Then, at ~3.4 ns, all H-bonds between the N-terminal loop (residues 102–105) and the C-terminal loop (residues 126–135) disrupted, and the flexibility of the C-terminal loop then rapidly increased (Fig. 8). After that, at ~6.5 ns, losses of H-bonds between the carbonyl group of Ala110 and the backbone amide hydrogen of His115, and between the carbonyl moiety of Ala111 and backbone amide hydrogen of Gly114 initiated unfolding of α7 from the C-terminus (Fig. 8). The α8 helix unfolded at ~8.5 ns from its N-terminus. During the following simulation run, the N-terminal turn of α7 and the middle turn of α8 refolded several times. At ~9 ns, the N-terminus of α7 was reestablished and remained stable until ~13.5 ns. Between ~11 ns and 14.5 ns, the middle turn of α8 reformed, and this process was followed by propagation of the helical structure to the N-terminus. The two helices moved away at ~17.4 ns and the helix-turn-helix motif adopted a "semicircular" arrangement (Fig. 8). At ~18.8 ns, an H-bond between the His108 side chain and Asn129 carbonyl group initiated attachment of the C-terminal loop to the helix-turn-helix motif. At 27.5 ns, the ANK repeat IV adopted a globular arrangement, which remained stable until the end of the simulation. In this arrangement, the α7 was kinked and connected via an H-bond between the Ala111 amide hydrogen and the Val130 carbonyl group (Fig. 8) to the C-terminal loop. The core of this globular fold consisted of a hydrophobic cluster of residues Ala111, Val119, Val123 and Val130.

α7-turn-α8-loop-α9 motif

This fragment corresponds to the hypothetical folding unit suggested by Ferreiro et al. [19], the hydrophobic core of

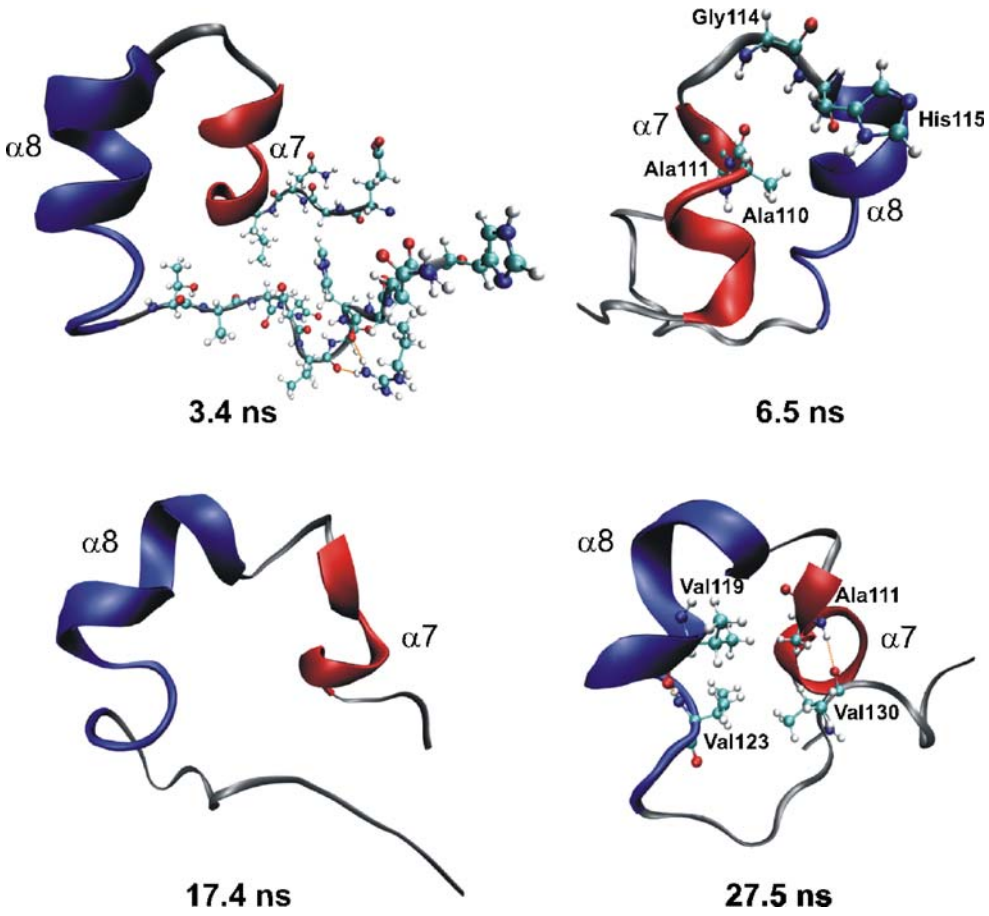


Fig. 8 Snapshots taken from the ANK repeat IV fragment simulation at various times; *red* α7, *gray* turn and loop, *blue* α8

which consists of the following residues: Leu107, Ala110, and Ala111 in α7; Val118, Val119, Leu122, and Val123 in α8; and Ala140, and Ala144 in α9. There is a salt bridge between the side chains of Lys112 and Glu113, which presumably stabilizes the C-terminal turn of α7, and a salt bridge between Glu120 and Lys124 that stabilizes the helical structure of α8 in the native conformation. The unstructured loop (residues 126–139) involves two H-bonds: one formed between the carbonyl oxygen of Asn129 and amide hydrogen of His132, and another between the amide hydrogen of Asn134 and carbonyl oxygen of Asp138 (Fig. 9a). The entire system is positively charged (+2e), since it includes three Lys, three Arg, two Asp and two Glu residues.

Apart from the increased flexibility of the α8 and α9 C-terminal turns, no other structural changes were observed in the first ~6.5 ns, during which the α7-turn-α8-loop-α9 fragment fold remained stable. Then, formation of a salt bridge among Arg145, Asp142 and Asp138 caused α9 to adopt a 3_{10}-helix conformation (Fig. 9b). After a further 1 ns (at ~7.5 ns), α9 shifted towards the solvent-exposed side of α7. This motion was presumably directed by the formation of a salt bridge between Lys112 and Asp142 and a hydrophobic contact between Leu143 and the nonpolar moiety of Lys112 (Fig. 9b). After several nanoseconds (at ~11 ns), α8 unfolded from the C-terminus and it is assumed that the bifurcated H-bond between the backbone carbonyl oxygen of His115 and the amide hydrogens of Val118 and Val119 initiated the unfolding process. The assembly of α7, α8 and α9 formed at ~7.5 ns was broken at ~12 ns because the interaction between Lys112 and Asp142 vanished, due to the formation of a new salt bridge between Lys112 and Glu113 (Fig. 9b). After that (~13 ns), α9 moved away from α7 and α8, and losses of native hydrophobic contacts between α7 and α9 as well as between α8 and α9 were observed. Consequently, α7 unfolded from the C-terminus. The three-dimensional structure of this fragment was then extended and its core residues were significantly exposed to the water solvent. At ~24.5 ns, the antiparallel arrangement of the α7-turn-α8 motif was broken. Some hydrophobic residues became exposed to the solvent and the others formed non-native contacts that were not observed in the native fold (Fig. 9b), some of which assisted reformation of non-native α-helices spanning residues 113–117 and 139–143 (cf. Fig. 4).

ANK IV-loop-α9 motif

The Consensus 2 hypothetical minimal folding unit of the p18 protein consists of one ANK repeat motif (Consensus 2) followed by a C-terminal α-helix (Fig. 10a). The residues forming the hydrophobic core of the α7-turn-α8-loop-α9 fragment are listed above, and the H-bond network is similar to that described for the α7-turn-α8-loop-α9 fragment, except that additional H-bonds are formed between the backbone carbonyl oxygen and amide hydrogen of Glu102 and His135, and between the Asn104 side chain carbonyl oxygen and side chain amide hydrogen of Asn134 (Fig. 10a). The formal charge of the respective fragment is +1e. Three Lys, three Arg, three Glu and two Asp are involved in the fragment structure.

The conformation of the C-terminus of the α9 changed to a turn, stabilized by H-bonds formed among the main chain carbonyls of Asp142 or Leu143 and backbone amide hydrogens of Arg145, Leu146 and *N*-methyl. At ~2.5 ns, the C-terminus of α8 lost its helical conformation and also adopted a turn conformation. Loss of the α8 C-terminal helical turn then presumably initiated a shift (lasting ~5 ns) of α8 towards α9, to the position where α10 (which is missing in this fragment) occurs in the p18 structure (Fig. 10b). The C-terminal turn of α8 adopted a 3_{10}-helix conformation at ~13 ns. Then, the α7 unfolded from both termini at the same time, at ~17.3 ns, and the C-terminus of the α7 changed to a turn conformation, which was stabilized mainly by the formation of a bifurcated H-bond between the backbone carbonyl of Ala110 and both the amide hydrogen of Glu113 and His115 (Fig. 10b). At ~35 ns, α8 completely unfolded, while α9 unfolded from the C-terminus (Fig. 10b). From this time onward, all of the helices were able to refold to α-helical conformations, but these secondary structure elements were short-lived, appearing for periods ranging from approximately 1 to 8 ns. The major refolding events appeared in the α8 and α9 helices. The α8 helix reestablished its middle turn at ~37 ns and maintained it until ~45 ns. After that, the C-terminal turn of the α8 was quickly reformed and remained stable during the next simulation run. At ~48 ns, the helical structure of α9 fully reformed, and was maintained until the end of the simulation. It is likely that the hydrophobic contacts among Val119, Leu122, Val123, together with the hydrophobic moiety of Arg133, Cys141 and Ala144 stabilized the C-terminus of α8 as well as the helical structure of the α9 helix (Fig. 10b).

Four-helical bundles

Four-helical bundles (i.e., α7-turn-α8+α9-turn-α10 and α7-turn-α8-loop-α9-turn-α10 fragments) were significantly more stable than the smaller fragments, and their α-helical contents were well maintained in the simulation timescale (cf. Fig. 4). The α7 and α9 helices were generally more rigid than the α8 and α10 helices (Fig. 11), and the average α-helical contents of the α7-turn-α8+α9-turn-α10 and α7-turn-α8-loop-α9-turn-α10 fragment were 89% and 84%, respectively. Thus, the four-helical bundles are

slightly less stable when the loop connecting the two helix-turn-helix motifs is present than when it is absent (i.e., in the α7-turn-α8+α9-turn-α10 fragment, see Fig. 11). This is because of the greater flexibility of the α8–α9 loop (the loop connecting α8 with α9 in the α7-turn-α8-loop-α9-turn-α10 fragment), which readily makes contacts with the helices. It has substantially higher flexibility in simulations of these bundles, since it lacks

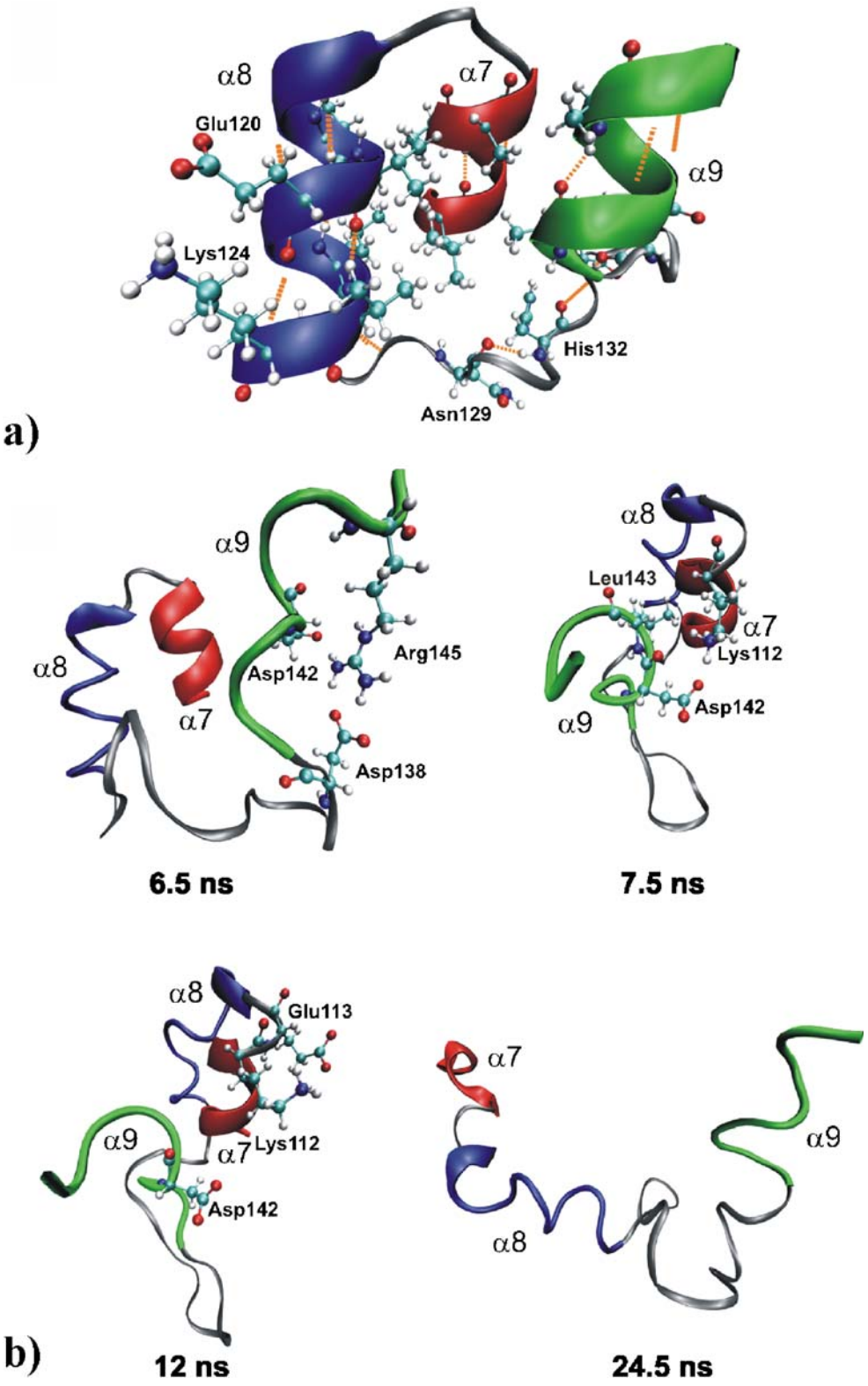


Fig. 9 **a** Native structure of the α7-turn-α8-loop-α9 fragment (*red* α7, *gray* turn and loop, *blue* α8, *green* α9). **b** Snapshots from the α7-turn-α8-loop-α9 fragment MD simulation taken at various times

inter-strand hydrogen bonds to the N-terminal hairpin of ANK IV, than in MD simulations of the ANK IV-loop-α9 fragment, where it does make these inter-strand hydrogen bonds.

Discussion

Here, we present all-atomic MD simulations, 450 ns long in total, in explicit solvent of various different C-terminal

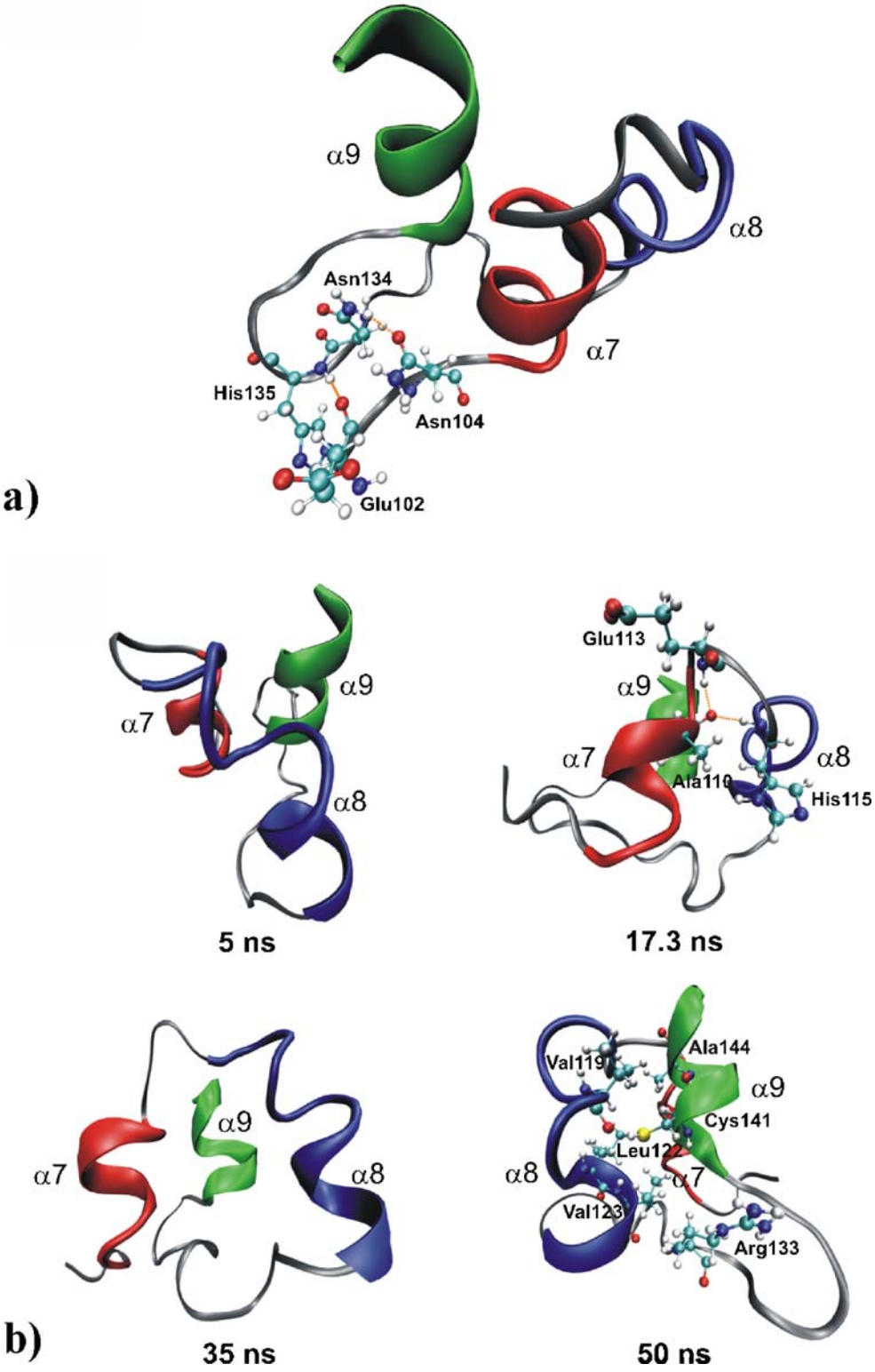


Fig. 10 **a** Native structure of the ANK IV-loop-α9 fragment (*red* α7, *gray* turn and loop, *blue* α8, *green* α9). **b** Snapshots from the ANK IV-loop-α9 fragment MD simulation taken at various times

fragments of the p18 protein (Table 1). The main aim of this study was to test the hypothesis of Ferreiro et al. [19], who proposed from Gō type simulations that the hypothetical minimal folding module of ARP consists of one complete ANK repeat motif and the first helix of the following ANK repeat. According to this hypothesis, the minimal folding module of ARP contains three helices connected by a turn and loop. We selected nine C-terminal fragments of p18 to test this hypothesis. The smallest studied systems consisted of pairs of helices (two) while the largest contained two ANK repeats, coinciding with the minimal folding unit of ARP suggested by Zhang and Peng based on results obtained from proteolytic analyses and prediction algorithm [7].

Separations of the helices, and resulting elongation of the structures, were observed after ~3.2 ns and ~20 ns in simulations of the helix-turn-helix motifs of the α7-turn-α8 and α9-turn-α10 fragments, respectively. Both of these fragments were unstable, with mean α-helical contents equal to 19.7% and 33.6%, respectively, and the contacts among hydrophobic residues in them were insufficiently strong to maintain their native folds. These observations are in agreement with findings of Du and Gai [29], who concluded that stability of the helix-turn-helix (α-hairpin) motif is enhanced by strong inter-strand hydrophobic clusters. The two helix pairs (α7+α8 and α9+α10) were the most unstable systems of all studied fragments, and both displayed similar behavior. Their native α-helical content quickly decreased due to unfolding of individual helices, accompanied by the loss of many native contacts, then several refolding events were observed until the individual helices dissociated, after which no further refolding of the α-helices was observed.

The C-terminal ANK repeat IV did not maintain its native fold and, after several nanoseconds, collapsed to a globular structure (cf. Fig. 8), which was likely stabilized by hydrophobic contacts and H-bonds between α-helical residues and residues located in the C-terminal loop. The mean α-helical content of the ANK repeat IV amounted to 39% and the mean percentage of preserved native contacts of the final structures was ~24%.

The three helical bundles occurring in the ANK IV-loop-α9 and α7-turn-α8-loop-α9 fragments were not stable during the observed timescale. The mean α-helical content of the ANK IV-loop-α9 motif was 54%, and the content of the α7-turn-α8-loop-α9 motif was just 35%. The C-terminal region of the α7-turn-α8-loop-α9 fragment was more flexible than the ANK IV-loop-α9 system. Lack of the four N-terminal β-hairpin residues in the α7-turn-α8-loop-α9 fragment prevents H-bonding of this moiety with the loop connecting α8 with α9. Therefore, the C-terminal region of the α7-turn-α8-loop-α9 fragment gained higher flexibility and, early in the simulation (at 24.5 ns), the fragment adopted an extended structure.

The largest fragments, containing two pairs of helix-turn-helix motifs (i.e., four helical bundles), were the most stable systems among those simulated (Fig. 11), as indicated by their high mean α-helical contents of 89%

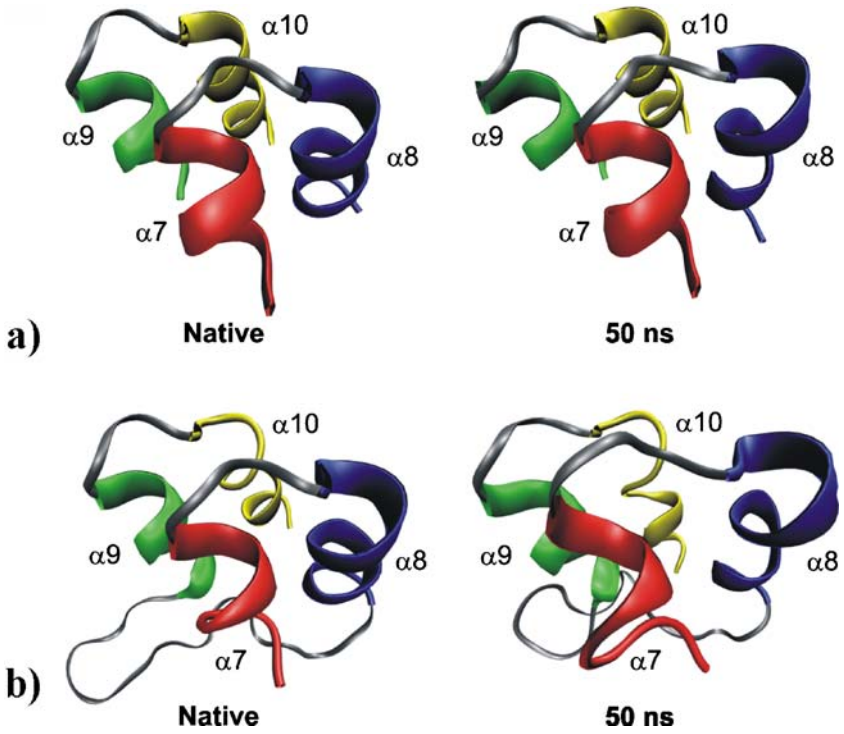


Fig. 11 **a** Initial and last structures from the MD simulation of α7-turn-α8+α9-turn-α10 fragment (*red* α7, *gray* turns, *blue* α8, *green* α9, *yellow* α10). **b** Initial and last structures from the simulation of α7-turn-α8-loop-α9-turn-α10 fragment (*red* α7, *gray* turns and loop, *blue* α8, *green* α9, *yellow* α10)

(α7-turn-α8+α9-turn-α10) and 84% (α7-turn-α8-loop-α9-turn-α10). These values, together with the relatively low associated and convergent RMSDs and R_gs values, and high percentage of preserved native contacts (see Table 2), indicate that these fragments are stable in the 50 ns long simulation timescale applied here.

Conclusions

The all-atomic MD simulations in explicit solvents presented here show that the hypothetical minimal folding module of ARP p18 comprising one ANK repeat and the first helix of the following ANK repeat (containing three connected helices) proposed by Ferreiro et al. [19], is not stable at a 50 ns timescale and should not be regarded as the minimal folding unit of p18. On the other hand, the minimal stable unit of ARP, consisting of two ANK repeats stacked with each other (containing four connected helices), as suggested by Zhang and Peng [7], was considerably more stable at the same timescale. The lower stability of the hypothetical folding module is attributable to the larger exposure of the hydrophobic core residues to water. Since each non-terminal ANK repeat has two hydrophobic sides, we may further hypothesize that at least one hydrophobic side must be fully covered, shielded from water, to maintain ANK repeat stability. Thus, at least two ANK repeats are required to make a stable ARP.

Acknowledgments Support through the MSMT (Ministry of Youths, Sports and Education, Czech Republic) grants LC512 and MSM6198959216 is gratefully acknowledged. We thank Sees-Editing, Ltd., (UK) for language revision.

References

1. Sedgwick SG, Smerdon SJ (1999) Trends Biochem Sci 24:311–316
2. Mohler PJ, Gramolini AO, Bennett V (2002) J Cell Sci 115:1565–1566
3. Serrano M, Hannon GJ, Beach D (1993) Nature 366:704–707
4. Mosavi LK, Williams S, Peng ZY (2002) J Mol Biol 320:165–170
5. Walker RG, Willingham AT, Zuker CS (2000) Science 287:2229–2234
6. Michaely P, Bennett V (1993) J Biol Chem 268:22703–22709
7. Zhang B, Peng ZY (2000) J Mol Biol 299:1121–1132
8. Mosavi LK, Minor DL, Peng ZY (2002) Proc Nat Acad Sci USA 99:16029–16034
9. Tripp KW, Barrick D (2007) J Mol Biol 365:1187–1200
10. Ferreiro DU, Cervantes CF, Truhlar SME, Cho SS, Wolynes PG, Komives EA (2007) J Mol Biol 365:1201–1216
11. Binz HK, Kohl A, Pluckthun A, Grutter MG (2006) Proteins: Struct Funct Bioinf 65:280–284
12. Binz HK, Amstutz P, Kohl A, Stumpp MT, Briand C, Forrer P, Grutter MG, Pluckthun A (2004) Nat Biotechnol 22:575–582
13. Binz HK, Stumpp MT, Forrer P, Amstutz P, Pluckthun A (2003) J Mol Biol 332:489–503
14. Devi VS, Binz HK, Stumpp MT, Pluckthun A, Bosshard HR, Jelesarov I (2004) Protein Sci 13:2864–2870
15. Zahnd C, Wyler E, Schwenk JM, Steiner D, Lawrence MC, McKern NM, Pecorari F, Ward CW, Joos TO, Pluckthun A (2007) J Mol Biol 369:1015–1028
16. Kohl A, Binz HK, Forrer P, Stumpp MT, Pluckthun A, Grutter MG (2003) Proc Nat Acad Sci USA 100:1700–1705
17. Mosavi LK, Peng ZY (2003) Protein Eng 16:739–745
18. Main ERG, Lowe AR, Mochrie SGJ, Jackson SE, Regan L (2005) Curr Opin Struct Biol 15:464–471
19. Ferreiro DU, Cho SS, Komives EA, Wolynes PG (2005) J Mol Biol 354:679–692
20. Venkataramani R, Swaminathan K, Marmorstein R (1998) Nat Struct Biol 5:74–81
21. Pearlman DA, Case DA, Caldwell JW, Ross WS, Cheatham TE, Debolt S, Ferguson D, Seibel G, Kollman P (1995) Comput Phys Commun 91:1–41
22. Cornell WD, Cieplak P, Bayly CI, Gould IR, Merz KM, Ferguson DM, Spellmeyer DC, Fox T, Caldwell JW, Kollman PA (1995) J Am Chem Soc 117:5179–5197
23. Wang JM, Cieplak P, Kollman PA (2000) J Comput Chem 21:1049–1074
24. Bartova I, Otyepka M, Kriz Z, Koca J (2004) Protein Sci 13:1449–1457
25. Otyepka M, Bartova I, Kriz Z, Koca J (2006) J Biol Chem 281:7271–7281
26. Bartova I, Otyepka M, Kriz Z, Koca J (2005) Protein Sci 14:445–451
27. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML (1983) J Chem Phys 79:926–935
28. Kabsch W, Sander C (1983) Biopolymers 22:2577–2637
29. Du DG, Gai F (2006) Biochemistry 45:13131–13139
30. DeLano WL (2002) DeLano Scientific. http://pymol.sourceforge.net/