How strong are hydrogen bonds in the peptide model?

Jakub Dąbrowski[a], Wiesław Nowak[b] and Arkadiusz Ptak*[a]

[a]*Institute of Physics, Faculty of Technical Physics, Poznan University of Technology, Piotrowo 3, PL-60965 Poznan, Poland*

[b]*Institute of Physics, Faculty of Physics, Astronomy and Informatics, Nicolaus Copernicus University, Grudziadzka 5, PL-87100 Torun, Poland*

[*]*E-mail: arkadiusz.ptak@put.poznan.pl*

Electronic Supplementary Information

1. Simulation protocol

The AAKA(AEAAKA)$_5$AC model was assembled from 36 residues (A – alanine, E – glutamic acid, K – lysine, C – cysteine) in predefined α-helical conformation with the VMD package.[1] For solvation purposes, a 22 Å × 212.5 Å × 22.5 Å cell containing 2949 TIP3P water molecules was implemented in the periodic boundary conditions scheme, allowing an elongation by about 300 % of the initial peptide length in steered molecular dynamics (SMD) simulations. The elongations of the peptide were calculated with the VMD as an increase in the molecule length measured between C_α atoms of the terminal residues. The cutoff for nonbonded interactions was 12 Å.

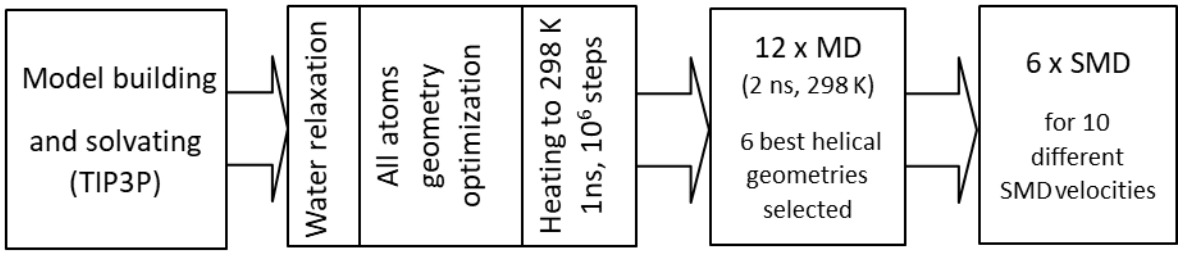

Fig. S1 A scheme of the simulation protocol. After the energy minimization and heating simulation, 12 MD simulations were performed, and 6 system configurations with the highest α-helix content were selected as starting points for successive SMD simulations.

The NAMD package[2] with the implemented CHARMM27[3,4] force field was used to perform energy minimization and all molecular dynamics (MD) calculations. The sequence of the calculations is schematically presented in Fig. S1. The optimization of the whole system was performed with the conjugate gradient method with standard parameters. At first, the positions and geometry of water molecules were optimized and then the positions of all atoms of the system (the peptide and water

molecules) were optimized in 50k steps. After the geometry optimization, a 1ns heating simulation was performed, from 0 K to 298 K. After the heating, 12 MD simulations were performed. For solving equations of motion, the Verlet-I /r-RESPA/ impulse MTS method was used with the integration step of 1 fs. Each of the MD simulations was for 2 ns, at a constant temperature 298 K, using the Langevin thermostat with a damping coefficient of 5 ps^{-1}. Then, 6 system configurations with the highest number of the peptide residues participating in the α-helical conformation were selected as starting points for SMD simulations.

2. SMD simulations

NAMD enables the use of an additional, harmonic potential (U_{SMD}) applied between an SMD atom (or a group of SMD atoms) in the molecule and the virtual atom (Fig. 2 in the paper):

$$U_{\mathrm{SMD}} = \frac{1}{2} k_{\mathrm{SMD}} [V_{\mathrm{SMD}} t - (R_t - R_{\mathrm{O}})]^2 \qquad \text{(S1)}$$

where: k_{SMD} is the spring constant of the SMD potential, V_{SMD} is the velocity of the virtual atom, R_t is the actual position of the SMD atom at time t, and R_0 is the initial position of the SMD atom.[5] The SMD force, i.e. the force applied by the harmonic potential and acting on the SMD atom, as well as the molecule elongation can be extracted from SMD simulation results. This enables plotting the force-versus-elongation curves. To reduce noise, the force-elongation curves were smoothed using the FFT filter (Origin 8.6, OriginLab, Northampton, MA). The bin values were set as 2 data points, because smoothing with bigger bins led to over-smoothed curves, which can lose local extrema.

3. Quantities determined from the SMD simulations and their statistical reliability

Values of the loading rate, the effective spring constant of the molecular system and the unbinding force of hydrogen bonds (HBs) were determined from the force versus elongation curves smoothed by the FFT filter for each SMD simulation. Average values and standard errors of the loading rate (r_{F}) and the unbinding force (F_{un}) were calculated for each of the SMD velocities (Table S1).

To test the reliability of the SMD results, in particular the unbinding force, we have performed calculations for 50 SMD simulations at 2 Å/ps, which is one of the middle SMD velocities used in present study. The average value of the unbinding force was 464 pN, which means that the difference between the average values for 6 (480 ± 61 pN) and 50 simulations is much smaller than the respective standard error.

Table S1 The average loading rate (r_F) and unbinding force (F_{un}) for each SMD velocity (V_{SMD})

V_{SMD}[Å/ps]	r_F [pN/ps]	F_{un} [pN]
0.01	1.3 ± 0.5	140 ± 43
0.05	3.2 ± 0.9	140 ± 32
0.1	6.5 ± 1.6	140 ± 34
0.5	40 ± 10	260 ± 60
1	160 ± 31	390 ± 50
2	260 ± 17	480 ± 61
3	420 ± 47	680 ± 98
5	890 ± 85	570 ± 129
7.5	1900 ± 510	800 ± 30
10	2800 ± 560	1000 ± 110

The uncertainties were calculated as standard errors.

4. Theoretical fits to the dynamic force spectrum

User defined non-linear fit functions, based on the theoretical models: Bell-Evans (BE),[6,7] Dudko-Hummer-Szabo (DHS),[8] and Friddle-Noy-De Yoero (FNDY)[9], were used in the Origin 8.6 software to fit the data obtained from SMD simulations – $F_{un}(r_F)$. From the fits unbinding characteristics of HB were determined.

To determine the free energy of HB unbinding, i.e. the free energy of the unbound state in relation to the bound one, the following formula obtained from the FNDY model (Equation 5 in the manuscript) was used:

$$\Delta G_{bu} = \frac{F_{eq}^{\ 2}}{2k_{eff}} \tag{S2}$$

Here, F_{eq} is the equilibrium force and k_{eff} is the effective spring constant of the system consisting of the α-helix and the SMD spring:

$$k_{\mathrm{eff}} = \frac{k_{\mathrm{helix}} k_{\mathrm{SMD}}}{k_{\mathrm{helix}} + k_{\mathrm{SMD}}} \tag{S3}$$

where: k_{helix} is the stiffness of the α-helical model calculated from the force-versus-elongation curve just before the first rupture point.

5. Hydrogen bond analysis

For HB characterization, the average number of broken HBs was calculated, assuming 4 Å as a threshold donor-acceptor distance for HB rupture. The distances between donors and acceptors for all 32 pairs of the AAKA(AEAAKA)$_5$AC backbone residues (with the "n...n+4" principle for α-helical HB formation) were measured in the VMD package[1] with a TCL script. The data are presented as a function of the molecule elongation for the lowest (0.01 Å/ps) and the highest (10 Å/ps) SMD velocities (Fig. S2).

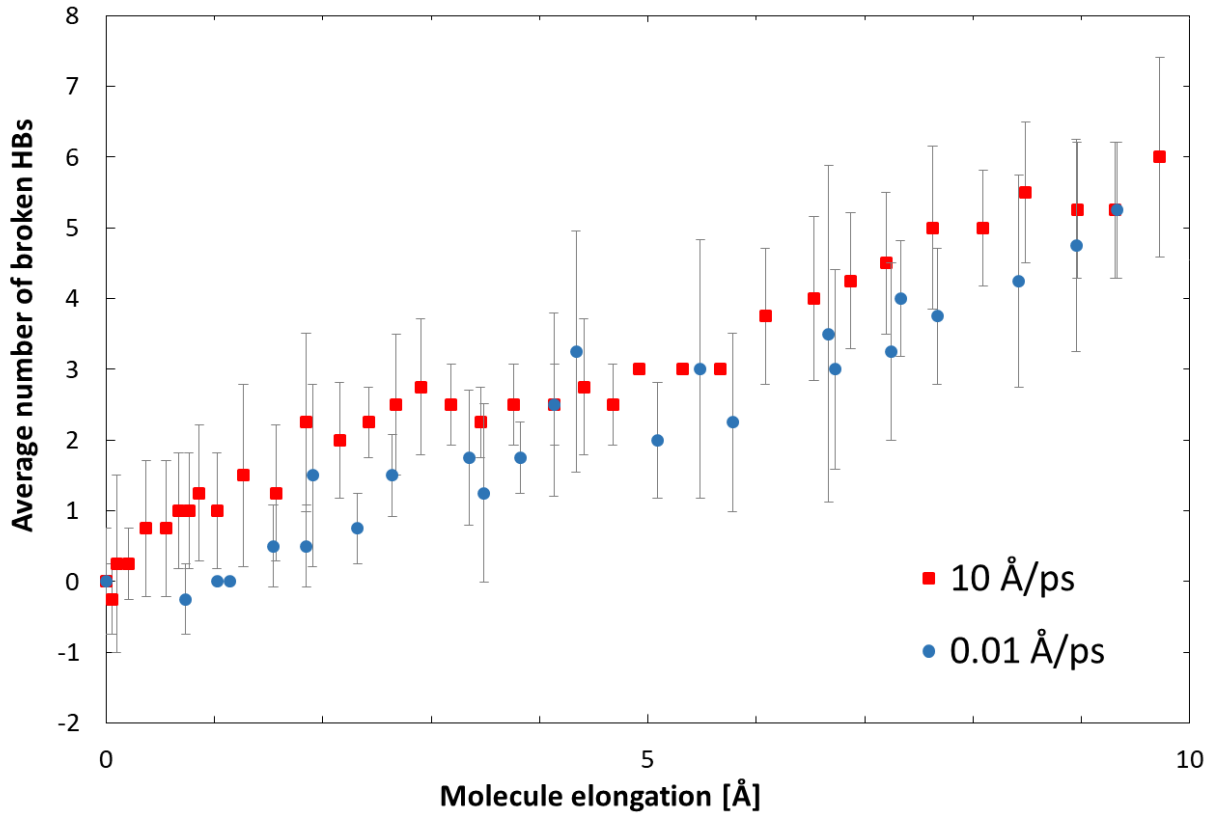

Fig. S2 The average number of broken HBs in function of the molecule elongation for 10 Å/ps (red squares) and 0.01 Å/ps (blue circles) SMD velocities

The average number of broken HBs for the molecule elongation stage (0-1.5 Å) at which the unbinding force is determined, is approximately 1. The values below 1 can result from formation of new HBs between residues 34-36 and 30-31 (Fig. S3). Detailed inspection of the peptide C-terminus (pulled end) conformation during SMD simulations discloses two possible paths of the unfolding of this fragment: 1) with HBs reformation, and 2) without HBs reformation (Fig. S3 – blue and red conformations, respectively). This ambiguous behaviour could cause some uncertainties observed in the

SMD force and the average number of broken HBs, particularly at the beginning of the peptide lengthening (Fig. S2).

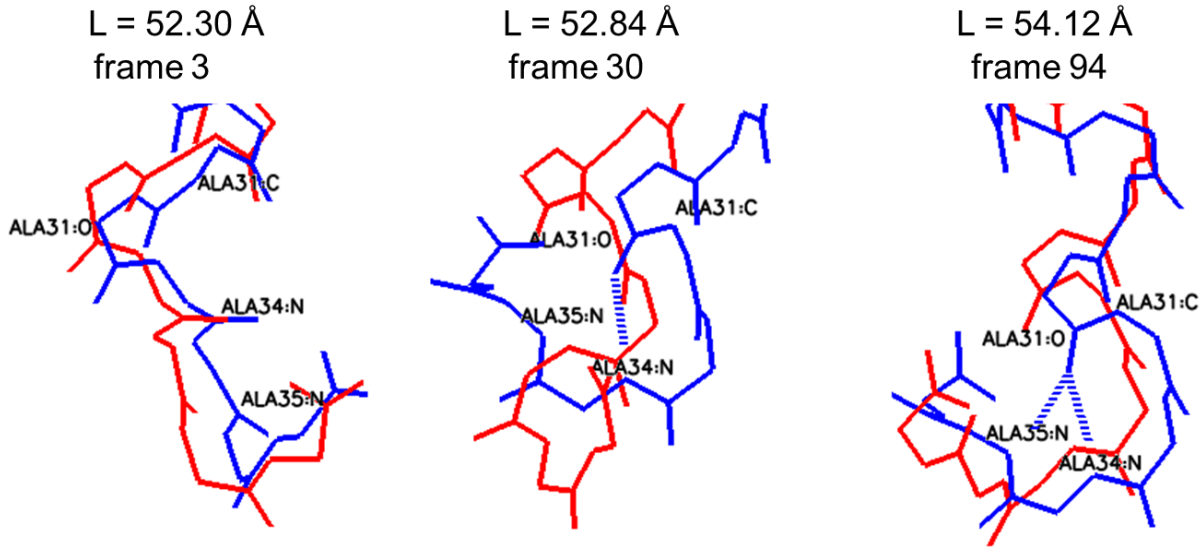

Fig. S3 Exemplary conformations of the pulled peptide terminus with two possible paths of unfolding: 1) with HBs formation between residues 34–36 and 30–31 (blue), and 2) without HBs formation (red) for three different molecule lengths (frames)

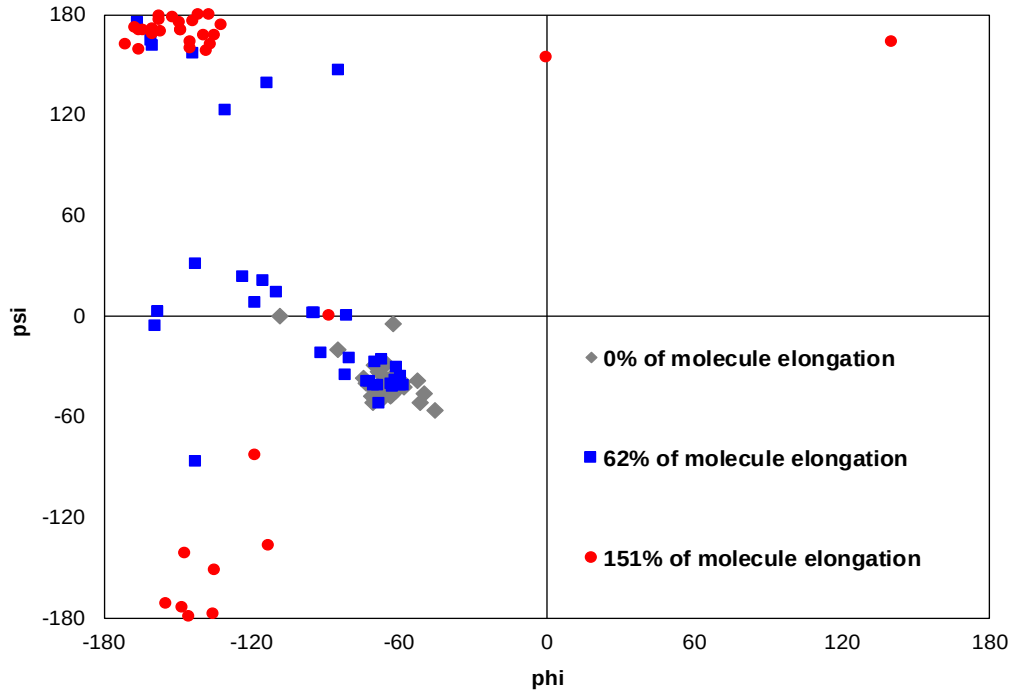

Fig. S4 Ramachandran plot for 35 pairs of psi-phi angles in the AAKA(AEAAKA)$_5$AC peptide model at three stages of SMD simulations: the beginning of simulations (grey diamonds), 62 % of the relative molecule elongation (blue squares), and 151 % of the relative molecule elongation (red circles)

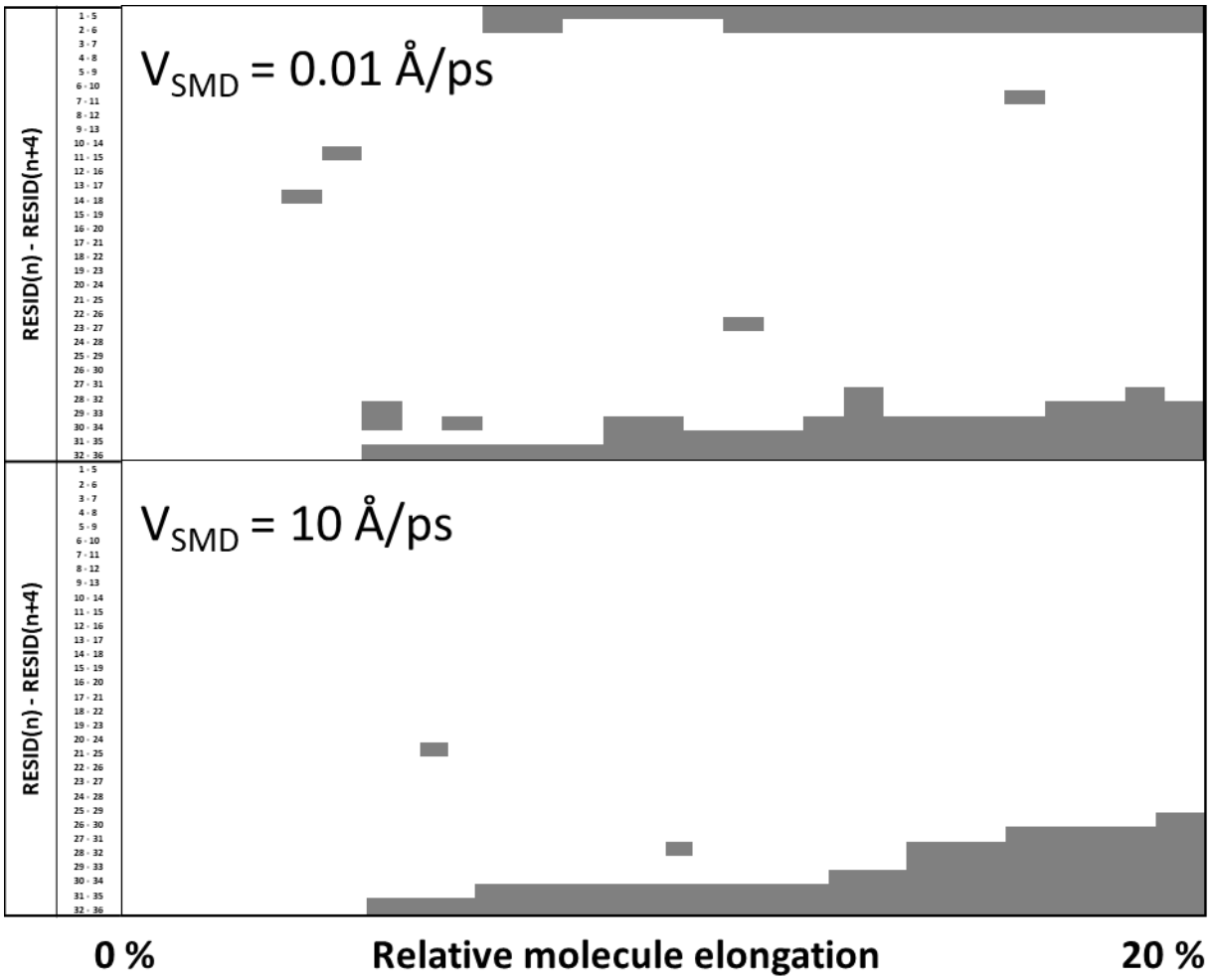

Fig. S5 Schematic representation of broken HBs (dark rectangles) in relation to the relative molecule elongation (from 0 % to 20 %) for two SMD velocities: 0.01 Å/ps (top panel) and 10 Å/ps (bottom panel)

During the stretching of the peptide model, subsequent HBs break, leading to unfolding of the α-helix, which can be illustrated in the Ramachandran plot (Fig. S4). The unfolding process is quite random at low loading rates, when HB ruptures appear in various fragments of the molecule (the left edge of dark rectangles – Fig. S5), and some of the broken HBs are rebound (the right edge of dark rectangles – Fig. S5). With increasing the loading rate, the unfolding process becomes more sequential, with weaker influence of the rebinding. It is noticeable, that after about 10 % of the relative molecule elongation, at low SMD velocity (0.01 Å/ps) both ends of the α-helix unfold simultaneously, while at high SMD velocity (10 Å/ps) the unfolding propagates from the pulled end (C-terminus, residue 36 in Fig. S5) towards the fixed end of the peptide model (residue 1 in Fig. S5). A reverse experiment, i.e. pulling the opposite terminus, has been performed for two velocities: 0.2 Å/ps and 2 Å/ps. At higher velocity (Fig. S6), HBs start to brake from both sides of the peptide: the pulled N-terminus (residue 1) and the C-terminus

(residue 36). However, after the first ruptures of HBs, further unfolding occurs from the pulled N-terminus, whereas at lower velocities (not shown) the rupture events are more random, which is consistent with our observation for the C-terminus pulling.

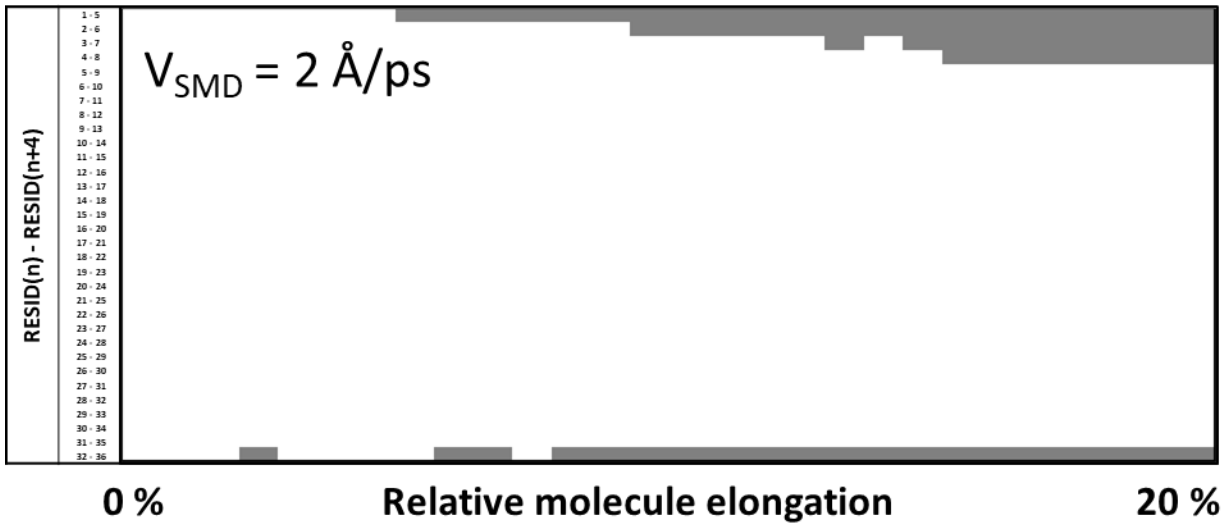

Fig. S6 Schematic representation of broken HBs (dark rectangles) in relation to the relative molecule elongation (from 0 % to 20 %) in the reverse experiment, i.e. pulling the N-terminus, for SMD velocity of 2 Å/ps

The HB rebinding probability, calculated as the ratio of the rebound HBs per the total number of HBs in the α-helix (32 for the AAKA(AEAAKA)$_5$AC peptide), is much higher at slow pulling than at fast one (Table S2).

Table S2 The HB rebinding probability for two ranges of the relative molecule elongation (RME) and for two SMD velocities (V_{SMD})

V_{SMD}	HB rebinding probability	
[Å/ps]	RME: 0–10 %	RME: 10–20 %
0.01	16 %	13 %
10	6 %	0 %

6. α-helix stability

The AAKA(AEAAKA)$_5$AC model was built with a predefined secondary structure by assigning the α-helix torsion angles values (ϕ = -57°, ψ = -47°) for each residue. The Molefacture Tool in VMD package has been used for this purpose. The stability of the α-helix was tested with two methods:

a) monitoring the donor (D)–acceptors (A) distances and D-H-A angles of HBs (Table S3),

b) monitoring the secondary structure with STRIDE algorithm[10] implemented in VMD package (Fig. S7).

Six final structures with the highest content of α-helical residues were selected for further analysis (Table S3). Based on the D–A distance (4 Å) and D-H-A angle (110°) criteria,[11] one can conclude that the HB co-formed by cysteine (residue 32–residue 36) is the weakest HB in the α-helix.

Table S3 HB donor–acceptor (D-A) distances (Å) and donor–hydrogen–acceptor (D-H-A) angles (°) for the final structures obtained from 12 MD simulations (2 ns long) for the AAKA(AEAAKA)$_5$AC model. Red boxes indicate HBs that do not meet the distance (< 4 Å) or angle (< 110°) criteria. Six peptide conformations (grey-blue columns) have been selected for SMD simulations.

		D-A distance												D-H-A angle											
D	A	1	2	3	4	5	6	7	8	9	10	11	12	1	2	3	4	5	6	7	8	9	10	11	12
1	5	3.0	3.5	2.7	4.7	2.7	3.3	3.4	2.9	2.8	3.4	3.0	2.9	162	148	157	149	149	132	143	155	165	159	169	162
2	6	2.8	2.9	2.9	3.9	2.8	3.2	3.5	3.2	3.1	3.0	3.0	4.4	166	165	170	154	169	148	160	171	140	144	148	135
3	7	3.2	3.0	3.0	2.9	3.1	2.9	3.1	3.4	3.4	3.2	2.8	3.5	172	159	155	160	170	169	164	163	158	152	142	154
4	8	3.1	2.7	3.3	2.8	3.0	2.9	3.0	3.1	3.3	3.3	3.0	3.2	153	162	159	174	171	151	174	162	160	151	154	145
5	9	2.8	2.8	3.3	2.8	3.0	2.9	3.0	2.9	3.0	2.9	2.9	3.4	155	171	155	155	155	145	167	138	138	167	157	170
6	10	3.0	3.1	2.8	2.9	3.0	3.0	3.1	2.9	3.4	3.2	3.2	3.5	152	157	148	172	163	162	172	156	149	147	146	145
7	11	3.3	3.1	3.1	2.9	3.1	3.5	2.8	3.1	3.4	3.4	3.1	3.2	145	130	172	168	152	161	175	166	159	154	151	152
8	12	3.0	2.9	3.2	2.7	2.7	3.7	3.0	3.0	2.9	3.4	2.9	3.0	170	163	167	154	168	167	154	154	141	153	143	147
9	13	2.9	3.0	3.1	2.8	2.8	3.6	3.0	2.7	3.1	3.8	3.0	2.7	151	159	136	159	153	158	155	146	164	153	171	163
10	14	3.5	3.0	3.3	2.8	2.9	3.1	2.8	3.0	3.4	3.0	3.3	3.1	158	133	151	148	149	158	147	155	152	139	136	141
11	15	3.0	2.8	3.4	3.0	3.0	3.1	2.9	3.0	3.0	2.9	3.2	2.7	178	158	146	174	165	159	164	158	134	159	136	137
12	16	2.7	2.9	3.2	3.0	2.8	2.8	2.8	2.8	3.0	3.6	2.9	2.9	141	158	172	164	157	133	162	160	163	154	170	168
13	17	2.8	3.0	2.7	2.9	3.1	3.1	3.4	2.9	3.5	3.4	2.9	3.1	160	150	149	162	157	167	161	164	154	163	166	167
14	18	3.1	3.1	3.1	2.9	2.9	3.5	3.5	3.0	3.3	3.1	3.8	3.3	166	160	176	142	156	156	157	155	126	147	141	172
15	19	3.0	3.0	2.9	3.1	3.3	2.9	3.0	2.7	3.3	3.3	3.4	3.3	161	163	168	164	168	170	153	149	172	157	136	160
16	20	2.8	3.1	2.9	3.0	3.2	3.5	3.3	2.9	2.8	3.4	3.3	3.0	162	149	157	163	159	154	129	169	152	160	148	151
17	21	3.0	3.5	2.9	3.7	3.0	4.1	3.0	3.3	3.0	2.9	2.9	3.2	159	167	153	153	151	155	164	131	158	161	161	171
18	22	2.7	2.9	3.0	3.3	3.6	3.2	2.9	2.8	3.1	3.1	3.4	3.0	173	164	143	152	159	137	146	146	153	150	151	164
19	23	3.2	2.9	2.8	3.0	2.8	2.8	2.8	3.1	2.9	3.0	2.8	2.8	155	138	134	176	150	148	133	153	162	174	133	172
20	24	3.4	3.2	2.8	3.0	2.9	3.1	3.0	3.3	3.0	3.0	2.9	2.9	150	157	170	169	177	165	175	145	172	163	159	146
21	25	2.9	3.3	2.9	3.0	3.3	2.9	3.2	3.8	2.8	2.8	3.3	3.2	176	174	165	145	156	142	170	166	167	174	175	158
22	26	2.9	2.8	3.1	3.1	3.0	3.1	3.0	3.2	2.9	3.4	2.8	2.9	157	145	161	151	154	160	147	149	144	149	154	155
23	27	2.9	2.9	2.8	2.8	2.8	3.0	2.8	3.2	3.0	3.4	3.0	3.1	152	135	156	138	160	149	149	172	152	159	132	150
24	28	3.5	2.8	2.9	2.9	2.8	3.1	3.3	3.0	2.8	2.9	3.1	3.3	154	174	156	160	157	129	148	147	165	160	149	153
25	29	3.2	2.9	3.0	3.0	2.9	2.9	3.4	3.4	3.1	2.9	3.1	3.3	167	158	173	144	150	159	153	168	153	141	151	139
26	30	2.9	3.6	2.9	3.2	2.9	2.9	2.9	3.3	3.2	2.9	2.8	3.5	168	153	156	159	175	160	161	160	142	144	155	167
27	31	3.0	3.5	2.7	2.8	2.7	3.0	2.8	2.8	3.3	3.0	2.8	2.7	156	162	167	165	163	160	158	166	165	147	173	166
28	32	3.1	3.0	2.9	2.9	3.0	3.3	3.0	3.5	3.3	3.6	3.0	3.2	151	158	165	158	174	154	162	136	159	137	142	127
29	33	2.8	3.4	3.2	3.1	3.1	3.3	2.7	3.2	3.1	4.3	3.1	3.2	173	156	158	135	168	145	149	146	174	130	168	155
30	34	3.1	4.4	2.9	2.7	2.8	3.1	3.0	3.5	3.1	3.6	2.9	3.0	129	150	149	151	178	140	168	154	125	116	170	160
31	35	3.2	4.2	3.4	3.1	3.4	4.1	3.1	3.2	3.3	3.5	3.2	3.2	139	153	165	153	131	148	144	138	138	91	164	136
32	36	3.3	3.7	3.6	3.1	3.1	3.5	3.6	4.0	3.9	6.8	3.4	3.1	106	98	164	142	90	143	131	124	80	127	98	125

The graphical representation of the secondary structures identified by the STRIDE algorithm (Fig. S7) allows following the conformational changes during the peptide unfolding. At low SMD velocities more random unfolding proceeds, with HB ruptures occurring at both termini of the peptide model as well as in the middle at further stages of the unfolding (left panel). In contrast, at high SMD velocities, sequential unfolding is observed, which starts from the pulled terminus of the peptide (right panel). At further stages of the peptide unfolding, in particular for slow pulling, helical conformations other than α-helix appear, e.g. 3_{10} helix (amino acids 8-12 in Fig. S7). Similar conformational changes were observed earlier.[12,13] However, this is out of the scope of our research, the main purpose of which was to analyze the initial stage of the peptide unfolding, in which the first ruptures of HBs occur.

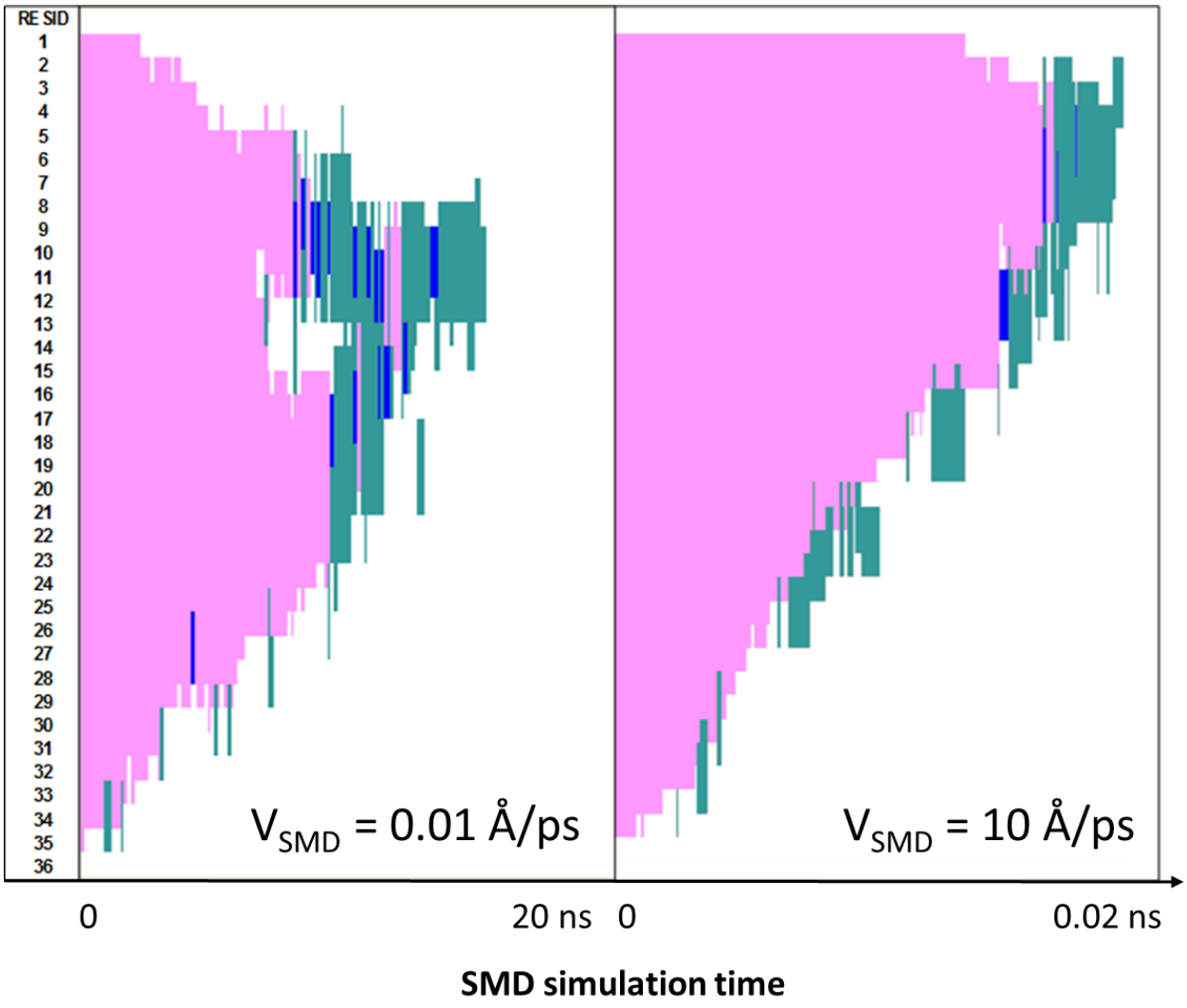

Fig. S7 Graphical representation of the secondary structures identified by STRIDE algorithm during the stretching of the AAKA(AEAAKA)$_5$AC model at low (0.01 Å/ps, left panel) and high (10 Å/ps, right panel) SMD velocities. Colours representing secondary structure type: pink – α-helix, dark turquoise – turn, blue – 3_{10} helix, white – coil or unfolded geometry.

In our research, we applied the TIP3P water model because it is compatible with the CHARMM27 force field (used in our MD and SMD simulations). TIP4P is another model that could be employed with CHARMM27. In order to check the sensitivity of HBs to the water model used, we have conducted test simulations for the two water models: TIP3P and TIP4P. Both simulations were carried out using the same procedure: 2 ns MD simulations were conducted at a constant temperature of 298 K following geometry optimization and 0.9 ns heating simulation from 0 K to 298 K. The average donor-acceptor distance and donor-hydrogen-acceptor angle (Table S4) are almost identical for both water models.

Table S4 The average HB donor–acceptor (D-A) distance and donor–hydrogen–acceptor (D-H-A) angle obtained from 2 ns-long MD simulations for the AAKA$(AEAAKA)_5$AC model for two water models: TIP3P and TIP4P

	TIP3P	TIP4P
D-A distance [Å]	3.13 ± 0.34	3.16 ± 0.51
D-H-A angle [°]	154 ± 15	154 ± 15

The uncertainties were calculated as standard deviations.

REFERENCES

1 W. Humphrey, A. Dalke and K. Schulten, VMD: Visual Molecular Dynamics, *J. Mol. Graphics,* 1996, 14, 33–38.

2 J. C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, Ch. Chipot, R. D. Skeel, L. Kale and K. Schulten, Scalable molecular dynamics with NAMD, *J. Comput. Chem.,* 2005, 26, 1781–1802.

3 A. D. MacKerell Jr., D. Bashford, M. Bellott, R. L. Dunbrack Jr., J. D. Evanseck, M. J. Field, S. Fischer, J. Gao, H. S. Guo et al. , All-atom empirical potential for molecular modeling and dynamics studies of proteins, *J. Phys. Chem. B,* 1998, 102, 3586–3616.

4 A. D. MacKerell Jr., M. Feig and C. L. Brooks III, Extending the treatment of backbone energetics in protein force fields: Limitations of gas-phase quantum mechanics in reproducing protein conformational distributions in molecular dynamics simulations, *J. Comput. Chem*., 2004, 25, 1400–1415.

5 S. Izrailev, S. Stepaniants, B. Isralewitz, D. Kosztin, H. Lu, F. Molnar, W. Wriggers and K. Schulte, *Steered Molecular Dynamics. Computational Molecular Dynamics: Challenges, Methods, Ideas*, Lecture Notes in Computational Science and Engineering, Springer-Verlag, Berlin, 1999; pp. 39–65.

6 G. I. Bell, Models for specific adhesion of cells to cells, *Science,* 1978, 200, 618–627.

7 E. Evans and K. Ritchie, Dynamic strength of molecular adhesion bonds, *Biophys. J.,* 1997, 72, 1541–1555.

8 O. Dudko, G. Hummer and A. Szabo, Intrinsic rates and activation free energies from single-molecule pulling experiments, *Phys. Rev. Lett*., 2006, 96, 108101.

9 R. W. Friddle, A. Noy and J. J. De Yoreo, Interpreting the widespread nonlinear force spectra of intermolecular bonds, *Proc. Natl. Acad. Sci. U. S. A.,* 2012, 109, 13573–13578.

10 D. Frishman and P. Argos, Knowledge-based protein secondary structure assignment, *Proteins: Struct., Funct., Genet.,* 1995, 23, 566–579.

11 E. Arunan, G. R. Desiraju, R. A. Klein, J. Sadlej, S. Scheiner, I. Alkorta, D. C. Clary, R. H. Crabtree, J. J. Dannenberg et al., Definition of the hydrogen bond (IUPAC Recommendations 2011), *Pure Appl. Chem.*, 2011, 83, 1619–1641.

12 S. M. Kreuzer, R. Elber and T. J. Moon, Early events in helix unfolding under external forces: A milestone analysis, *J. Phys. Chem. B,* 2012, 116, 8662–8691.

13 G. S. Jas and K. Kuczera, Helix–coil transition courses through multiple pathways and intermediates: Fast kinetic measurements and dimensionality reduction*, J. Phys. Chem. B,* 2018, 122, 10806–10816.