PAPER

Check for updates

Unfolding mechanism and free energy landscape of single, stable, alpha helices at low pull speeds†

Ana Elisa Bergues-Pupo, [ID] ‡ Reinhard Lipowsky [ID] and Ana Vila Verde [ID] *§

Single alpha helices (SAHs) stable in isolated form are often found in motor proteins where they bridge functional domains. Understanding the mechanical response of SAHs is thus critical to understand their function. The quasi-static force−extension relation of a small number of SAHs is known from single-molecule experiments. Unknown, or still controversial, are the molecular scale details behind those observations. We show that the deformation mechanism of SAHs pulled from the termini at pull speeds approaching the quasi-static limit differs from that of typical helices found in proteins, which are stable only when interacting with other protein domains. Using molecular dynamics simulations with atomistic resolution at low pull speeds previously inaccessible to simulation, we show that SAHs start unfolding from the termini at all pull speeds we investigated. Unfolding proceeds residue-by-residue and hydrogen bond breaking is not the main event determining the barrier to unfolding. We use the molecular simulation data to test the cooperative sticky chain model. This model yields excellent fits of the force−extension curves and quantifies the distance, $x_E = 0.13$ nm, to the transition state, the natural frequency of bond vibration, $\nu_0 = 0.82$ ns^{-1}, and the height, $V_0 = 2.9$ kcal mol^{-1}, of the free energy barrier associated with the deformation of single residues. Our results demonstrate that the sticky chain model could advantageously be used to analyze experimental force−extension curves of SAHs and other biopolymers.

1 Introduction

Single alpha helices (SAHs) differ from most helices found in proteins: whereas most protein helices are metastable and short-lived in the absence of intra- or intermolecular interactions with other protein domains, SAHs maintain a stable helical configuration in aqueous solution in the absence of interactions with other biomacromolecules. SAHs have been found in the proteomes of bacteria, archaea and eukaryotes.[1–4] They are typical components of motor proteins, where they act as spacer or connector segments between globular domains,[1,2,5,6] and could realize similar functions in artificial materials. Clarifying the properties of SAHs is thus important to understand the

functioning of motor proteins and for their targeted application in materials science.

Natural SAHs typically contain a large fraction of positively and negatively charged amino acids, and $(i, i + 3)$ or $(i, i + 4)$ salt bridges (where i denotes the position of an amino acid in the linear amino acid sequence) are thought to be particularly important for their thermodynamic stability.[7,8] The pattern of positively and negatively charged amino acids is not indispensable for SAH formation, however: polyglutamic acid[9] and polylysine[10,11] also form SAHs at pH levels where the side chain residues are uncharged.

The mechanical response to tensile forces has been investigated using single molecule force spectroscopy (SMFS) or optical tweezers experiments for only a few SAHs.[9–13] Current evidence suggests that both the magnitude of the unfolding force and the qualitative features of the force–extension curve depend on the amino acid sequence, for reasons that are not yet understood. Single molecule experiments show that the stretching of the six turn (≈ 22 amino acids) helical linker connecting the My12 and My13 domains of muscle protein myomesin occurs around 40 pN in a reversible way.[13] A comparable unfolding force ($F < 30$ pN) was also observed for the much longer SAH domain (97 residues) from myosin 10.[12] Substantially more intense forces, however, have been reported from atomic force microscopy for unfolding of homopolypeptide

Max Planck Institute of Colloids and Interfaces, Department of Theory & Bio-Systems, Am Mühlenberg 1, 14476 Potsdam, Germany.
E-mail: ana.vilaverde@mpikg.mpg.de; Fax: +49 (331) 567 9602;
Tel: +49 (331) 567 9601

† Electronic supplementary information (ESI) available: Individual force–strain traces at the lowest pull speed; hydrogen bonds and salt bridges *vs.* strain; Ramachandran plots as a function of time; secondary structure *vs.* time from umbrella sampling simulations. See DOI: 10.1039/d0sm01166e

‡ Present address: Max Delbrück Center for Molecular Medicine, Hannoversche Str. 28, 10115 Berlin, Germany.

§ Present address: University of Duisburg-Essen, Faculty of Physics, Lotharstr. 1, 47057 Duisburg, Germany.

SAHs: $150 < F/\text{pN} < 200$ for polylysine (in the form Cys$_3$–Lys$_{30}$–Cys)[10,11] and up to 1000 pN for polyglutamic acid (Glu$_{80}$–Cys).[9] For the natural SAHs, for polylysine and for coiled coils under tension, experimental evidence indicates that unfolding takes place at constant force, making these structures important force absorbers.[10–14] In contrast, for polyglutamic acid the force increases monotonically with extension.[9]

Steered molecular dynamics simulations can provide great mechanistic insight about the deformation of single helices under tension. To date, most simulations have focused on metastable helices (either homopolypeptides or natural sequences);[15–18] whether SAHs respond as metastable helices remains an open question. A constant force plateau was observed in atomistic, explicit solvent simulations of a 52-residue helix belonging to the vimentin AH1 domain simulated isolated in solution (i.e., in a metastable state).[18] Hydrogen bond breaking was proposed to be the key barrier to unfolding, in a velocity-dependent manner: at low pull speeds helices appeared to unfold turn-by-turn via the simultaneous breaking of 3-to-4 hydrogen bonds, whereas at large pull speeds (>1 nm ns^{-1}) hydrogen bonds appeared to break individually and helices unfold residue-by-residue. This velocity-dependent mechanism was thought to be behind the two regimes of dependence of the unfolding force on pulling speed that were observed in the simulations. Unfolding started randomly throughout the helix and not just at the pulled ends, i.e., cooperativity was low.[17,18] To date we are aware of only a single simulation study of a SAH, with 97 residues.[12] That study, with atomistic resolution and implicit solvent representation, indicated this helix unfolds at constant force, consistent with experiments on the same system and with simulations of metastable helices.[17,18] In contrast with simulations of metastable helices, unfolding started predominantly from the ends of the SAH suggesting that the cost of creating an interface between helical and unfolded domains is high, i.e., cooperativity in SAHs is high. Hydrogen bond breaking was identified as the key barrier to unfolding, but in a velocity-independent manner. The plateau force observed in simulations at the lowest simulated pull speed (0.01 nm ns^{-1}) was ≈ 50 pN, comparable to the experimentally detected force (<30 pN) for the same SAH[12] despite the large difference in pull speed between experiment and simulation.

There is a clear need to clarify the mechanism of unfolding of SAHs, the extent to which it is cooperative, its dependency on the pull speed and the dominant interactions behind the barrier to unfolding, among other open questions; here we address some of these issues using all-atom molecular simulation studies of a short SAH. Reliable and complete understanding, however, can only be obtained with a strong connection between experiment and simulation. But, how to facilitate this connection given that pull speeds rarely overlap? Both simulation and experimental studies suggest that, for helices as well as for many other polymers under tension, rupture events are well described by a 1-dimensional free energy landscape consisting of two stability wells separated by a barrier.[19–21] Coarse insight into this free energy landscape is often extracted, for experimental and for simulation data, via dynamic force spectroscopy models: most often the Bell–Evans,[22] Yoreo[23] and Dudko[24] models. These models yield the distance, x_E, from the first stable state to the transition state, and the zero-force off rate, k_{off}, of the rupture event; the Yoreo model in addition includes the equilibrium unfolding force. One way, albeit indirect, of connecting experiment and simulation is by comparing these parameter values. In addition, for a given estimate of the natural vibration frequency, the height, V_0, of the free energy barrier can be calculated from k_{off}. This procedure results in estimates of V_0 with substantial uncertainty, though. This uncertainty is problematic because our ability to predict the mechanical response of helices – and thus directly connect experiments and simulations – depends critically on this parameter. Bad estimates of V_0 lead to predictions of the mechanical response that are not even qualitatively correct.[25–27] Another issue with using these models to treat data is that both the Bell–Evans and the Dudko models assume refolding does not take place, an assumption that is not met at low pulling speeds. A final issue is that the models do not consider the case of cooperative helix unfolding.

The sticky chain model proposed by Jäger[28] can resolve these issues and thus provide a better connection between SMFS experiments and molecular simulations. This model, based on a 1-dimensional free energy landscape as the Bell and Yoreo models, can in principle be used to fit the full force–extension curve of a polymer under tension and yields, among other parameters, reliable values of V_0. It is a kinetic model that allows for unfolding and refolding and is for that reason valid also at low pull speeds, where the Bell–Evans model may fail.[20] It comes in two flavors – cooperative or not – enabling insight into this point. It is simple to implement, and thus advantageous relative to Monte Carlo models that could also be used for the same purpose.[19] The greater detail of the sticky chain model relative to Bell–Evans-like models thus offers a pathway to connect experimental and simulation force–extension curves. To our knowledge, however, this model has only been used to fit force–extension curves of bacterial pili;[29] its suitability for SAHs has not yet been tested.

In this work we combine atomistic, implicit solvent simulations of a SAH with the sticky chain model, to gain insight into the deformation mechanism of SAHs under tension and to test the potential of the sticky chain model to fit force–extension curves of protein helices. We use steered molecular dynamics and umbrella sampling simulations to systematically characterize the mechanical response and deformation mechanism of a short SAH in a range of pull speeds, starting from the high-speed, non-equilibrium unfolding regime all the way to quasi-static unfolding; our lowest pull speed (10^{-3} nm ns^{-1}) is one order of magnitude lower than reported all-atom simulation studies. We then use the simulation data to test the extent to and conditions under which the sticky chain model[28] can be used to fit force–extension curves of SAHs. We evaluate how the parameters of the 1-dimensional free energy landscape associated with residue unfolding obtained from the fit compare with expectations, and with values obtained from fits of the Bell–Evans and the Yoreo models.

2 Methods

2.1 Atomistic simulations of the alpha helix

We choose a *de novo* designed SAH of length 28 amino acids, consisting of four repeats of the basic sequence AEEEKRK, see Fig. 1. This sequence was designed synthetically with longer repeats, and contains numerous salt bridges between acidic and basic residues which likely contribute to its thermal stability.[30] All the simulations were conducted in implicit solvent in GROMACS 5.1.1, using the Amber99SB-ILDN force field.[31] The implicit solvent uses the Generalized Born (GB) and surface area algorithms. The Still approximation was used for calculating the GB part. Born radii were calculated at every time step, and a 1 nm cutoff was used for this calculation. The distance for the dielectric offset when calculating the Born radii was 0.009 nm. The non-polar surface area algorithm used an Ace-type approximation and the surface tension was set to 0.0049 kcal mol^{-1} Å^{-2}. The dielectric constant of the implicit solvent was set to 80 and the implicit salt concentration was 0.137 mol dm^{-3}. The initial helical structure was generated manually using Chimera,[32] and was minimized for 50 000 steps with the steepest descent algorithm. Equilibration was conducted in three steps of 200 ps each: the unrestrained system was heated from 10 K (step 1), to 150 K (step 2) and finally to 300 K (step 3). All subsequent simulations were conducted at $T = 300$ K. All simulations used a velocity rescaling thermostat that includes a stochastic term and thus samples the correct canonical ensemble; the coupling time constant was 2 ps. Constraints on all covalent bonds were applied with the LINCS algorithm, with 4th order expansion of the constraint coupling matrix and the number of iterative corrections set to 1. These settings enable using an integration time step of 2 fs. The cutoff distances for the Coulomb and the van der Waals interactions were 5 nm and 1 nm, respectively, and a charge group cutoff scheme was used.

Pulling simulations were conducted at constant speed, $v \in \{10^{-3}, 10^{-2}, 10^{-1}, 1\}$ nm ns^{-1}. For the three highest pull speeds, thirty independent runs were started from the same initial configuration, with velocity reassignment at $t = 0$. For the lowest pull speed, only four independent runs were performed

because of their high computational cost. Force–extension curves averaged over 4 runs already overlap with the 30-run average for each of the two intermediate pull speeds (comparison not shown). The 4 runs done for the lowest pull speed should thus be sufficient. Pulling was done *via* a virtual spring of constant k_s connected to the C_α carbon of glycine at the C-terminus of the helix; the C_α of the glycine at the N-terminus was restrained *via* a virtual spring of constant K, see Fig. 1. The virtual springs had values $K = 10 k_s = 1000$ kJ mol^{-1} nm^{-2}. Configurations were saved every $vt = 0.01$ nm unless otherwise noted, resulting in 1000 saved configurations per simulation run. For the analysis of the hydrogen bond breaking events, simulations at the lowest pull speed were repeated to save configurations every 10 ps (*i.e.*, the same as for the highest pull speed) but using a limited strain range, $0.5 < \varepsilon = vt/L_0 < 0.8$, to avoid excessively large trajectory files.

Umbrella sampling simulations were performed to build the potential of mean force as a function of the extension. Starting conformations were obtained from a pulling trajectory at $v = 10^{-2}$ nm ns^{-1}, by selecting configurations at equally spaced displacements ($vt = 0.2$ nm), yielding 30 separate umbrella simulations. The umbrella restraint was applied to the C_α of the terminal amino acids and the restraint constant was 100 kJ mol^{-1} nm^{-2}. Each umbrella was simulated for 50 ns.

2.2 Two-state chain model

The sticky chain model by Jäger[28] is a velocity-dependent mesoscopic model that describes the structural transition of a one-dimensional chain under a stretching force. In contrast with other helix–coil transition models (*e.g.*, the Zimm–Bragg model[33]), the sticky chain model describes non-equilibrium properties like the dependence of the force–extension curves on the pulling velocity. Each link of the chain can be in two states characterized by equilibrium lengths b_0 and $b_0 + \Delta L$, respectively, see Fig. 2. The deformation basin of each state is harmonic with elastic constants k_1 and k_2, respectively.

The transitions between the two states are thermally activated and are described by Bell-like kinetics, namely the transition rate depends exponentially on the external force. The equations describing the model are for the force–strain curve:

$$F = \frac{\varepsilon N b_0 - (N_0 - N_\Lambda)\Delta L}{N_\Lambda/k_1 + (N - N_\Lambda)/k_2 + 1/k_s + 1/K}, \quad (1)$$

and for the unfolding kinetics of each link, assuming cooperative unfolding:

$$\frac{dN_\Lambda}{dt} = -2\nu_0 e^{\beta F x_E}\left\{e^{-\beta V_0} - e^{-\beta F \Delta L}\right\}. \quad (2)$$

The $1/k_s + 1/K$ term takes into account the two external springs used in the atomistic simulations to stretch the single helix. N is the total number of links of the chain and N_Λ the number of links in the folded state. N_0 denotes the equilibrium number of folded links at $F = 0$. The strain ε can be expressed in terms of the pulling speed v:

$$\varepsilon = vt/L_0, \quad (3)$$
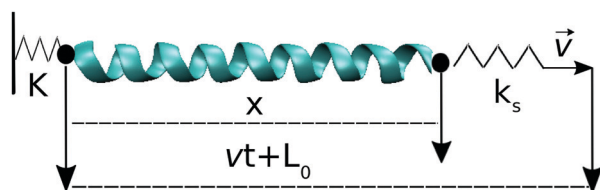
G AEEEKRK AEEEKRK AEEEKRK AEEEKRK G

Fig. 1 (top) Pulling geometry of the alpha helix. The distance x is the end-to-end distance of the chain measured from the terminal C_α atoms, projected along the pull axis; its value at $t = 0$ is $L_0 = 3.46$ nm. (bottom) Amino acid sequence in one-letter code. Two glycine residues (G) are added at either end of the helix. Positively and negatively charged amino acids are colored in blue and red respectively.
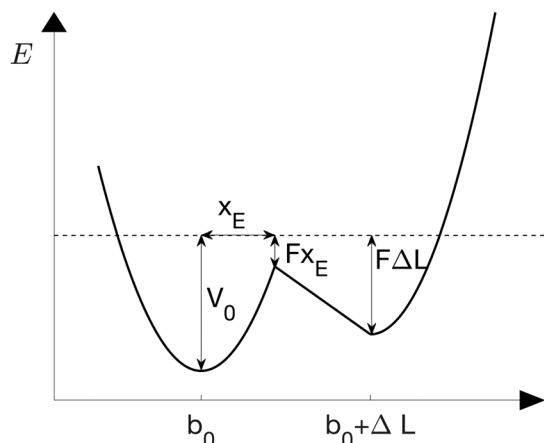
Fig. 2 Double well potential of the chain model under a constant force F.[28] The folded and unfolded states have equilibrium distances b_0 and $b_0 + \Delta L$, respectively and elastic constants k_1 and k_2. The position of the unfolding barrier is x_E and is force-independent.

with L_0 the initial length of the chain. In eqn (2), ν_0 is the natural vibration frequency and $\beta = 1/(k_B T)$ with k_B the Boltzmann constant.

A particularity of this double-well potential model is that at zero force, the energy of the unfolding barrier V_0 coincides with the energy of the unfolded state, *i.e.*, the barrier to refolding at zero force is zero. This choice has negligible impact on the dependence of the unfolding force on the pulling speed[20] but has the advantage of allowing a direct determination of the attempt frequency of opening of each link:

$$k_{off} = \nu_0 \times \exp^{-\beta V_0}. \qquad (4)$$

The three parameters defining the stability basin of the folded state are related as:

$$V_0 = \frac{1}{2}k_1 x_E^2 \qquad (5)$$

For the conditions of our simulations, only 3 parameters of the two state model (k_1, k_2 and ν_0) are obtained from fitting the force–extension curves obtained in the atomistic simulations, because the remaining parameters can be inferred from the atomistic model as explained below.

3 Results and discussion

3.1 Atomistic simulations in implicit solvent

SAHs are constant or quasi-constant force springs under quasi-static pulling and up to $v = 1$ nm ns^{-1}. The average force–strain curves at four different pull speeds are shown in Fig. 3. All curves show an initial rise phase, followed by a force plateau up to $\varepsilon \approx 2$, and a steep rise in the force beyond that point. Faster pull speeds push the onset of the force plateau to larger values of strain. The term "force plateau" is used loosely in this work: for the highest pull speed, statistical noise makes it impossible to ascertain whether it is a true force plateau or only a pseudo force plateau – a region where the force increases
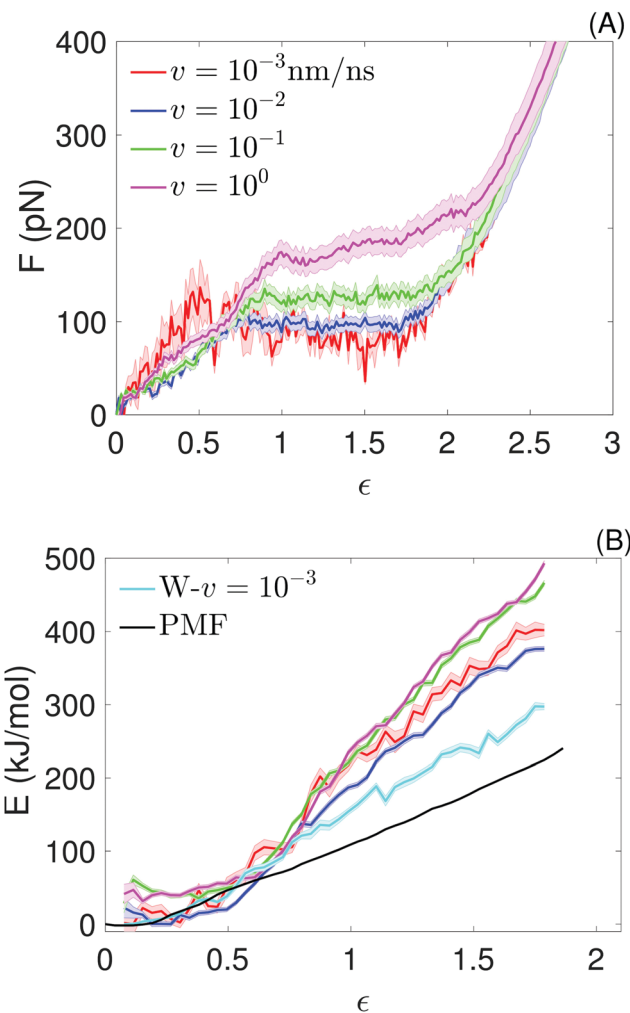

Fig. 3 Atomistic simulations: (A) average force–strain curves at the indicated pull speeds, obtained from the molecular dynamics simulations. For the lowest pull speed, the average is over 4 force traces; for the other pull speeds, it is over 30 force traces. The strain is defined as $\varepsilon = vt/L_0$. Average values of the plateau forces are shown in Table 1. (B) Potential of mean force from umbrella sampling (black), work (W–$v = 10^{-3}$; cyan) realized at the lowest pull speed, and potential energy for the 4 pull velocities (same colors as in (A)), as a function of the chain strain ($\varepsilon = x/L_0$), obtained from the molecular dynamics simulations. The different definition of strain used here is necessary to directly compare the PMF with the quantities extracted from the pull simulations. Every (E,ε) point of the potential energy curves is obtained by binning $E(t)$ values of all simulation runs performed for each v in the corresponding strain bin and averaging over the number of E-points per bin (this number is not constant); the shadow show the standard error of the mean. An analogous procedure is used to obtain the W–$v = 10^{-3}$ curve.

more slowly with strain than in the initial or final parts of the curve. Irrespective of whether the average response of this helix under tension results in a true or pseudo force plateau, these terms are useful only in terms of ensemble averages: individual force traces have substantial fluctuations (see ESI,† Fig. S1) reflecting the thermal nature of the system and also the existence of stick-slip events,[34] discussed in more detail below. The presence of a force plateau agrees with the only reported single molecule force spectroscopy experiment of a natural SAH.[12]

The SAH studied experimentally was derived from myosin-10 and was longer (97 amino acids) than the one simulated here, but was similarly rich in positively and negatively charged amino acids.

Analytical and numerical models have demonstrated that force plateaus naturally arise from a two-state free energy landscape associated with extension of links in a chain; the extent of cooperativity determines whether the force remains truly constant or increases slowly with extension in the plateau or quasi-plateau region.[21,28] Force plateaus have also been seen in simulations and experiments of different molecules under tension and in a wide range of pulling velocities, suggesting two-state free energy landscapes associated with stretching are common to many (bio)polymers. For example, force plateaus have been observed in simulations of helices in tension, even at higher pulling velocities[16,18,35] or under adiabatic conditions;[34,36,37] in single molecule measurements of polylysine SAHs,[10] of helical segments connected by joints,[37,38] of coiled coils,[14,39] of dextran[19] and of double stranded DNA;[40,41] in steered molecular dynamics simulations of duplex DNA,[42] and simulations of coiled coils in tension[39] and shear.[43,44]

The average plateau force (F_p; see Table 1), increases substantially with increasing pull speed, varying between $F_p^{v=10^{-3}} = (87 \pm 1)$ pN at the lowest pull speed up to ≈ 200 pN at the highest pull speed. The plateaus at the two lowest pull speeds, $v \in \{10^{-3}, 10^{-2}\}$ nm ns^{-1}, almost overlap, suggesting that unfolding at the lowest pull speed may take place under quasi-static conditions. To test this possibility, we calculated the work realized on the helix at the lowest pull speed and compare it with the potential of mean force (PMF) as a function of strain calculated using the umbrella sampling simulations. Both curves (light blue and black) are shown in Fig. 3(B) and clearly do not overlap, indicating that unfolding at the lowest velocity in fact does not proceed under quasi-static conditions.

The difference between the PMF and the work realized at $v = 10^{-3}$ nm ns^{-1} (Fig. 3(B)) together with the fact that the plateau forces at the two lowest pulling velocities are very similar, suggest that there is a timescale separation between the processes playing a significant role at equilibrium and those important for $v \geq 10^{-3}$ nm ns^{-1}. We attempted to gain insight into these processes by plotting the secondary structure as a function of time for two fixed values of strain of the umbrella simulations. The results, shown in ESI,† Fig. S8, indicate that the secondary structure substantially changes as a function of time in some parts of the peptide and that some of the changes occur only once within the simulated 50 ns. We speculate that these infrequent but substantial structural

Table 1 Average plateau forces for the three lowest pull speeds. The average is over $0.86 < \varepsilon < 1.7$ for $v = 10^{-3}$ nm ns^{-1} and $0.9 < \varepsilon < 1.8$ for the other speeds; the uncertainty is the standard error of the mean. The average for the highest pull speed is not calculated because a true plateau cannot be identified

v (nm ns^{-1})		
10^{-1}	10^{-2}	10^{-3}
126 ± 0.5	97 ± 0.5	87 ± 1

changes increase the chain entropy in the PMF calculations and thus soften the helix. These rare events are not well-sampled even at the slowest pull speed used in the pull simulations, which would explain why the work to extend the helix realized at this pull speed is larger than the PMF.

A plateau force < 30 pN was reported from experiment for a SAH rich in negative and positive charged amino acids;[12] quasi-static experiments done on a SAH of comparable composition reported a similar unfolding force.[13] Those forces are significantly lower than $F_p^{v=10^{-3}}$. This difference between experiment and simulation may be an artifact of the simulations, such as the limited timescales accessed as discussed above, or shortcomings of the force field. Alternatively, it may reflect true differences between the helix simulated here and those studied experimentally. Clarifying this point requires closer connection between experiment and simulation setups and is beyond the scope of this work. Investigating the non-equilibrium mechanical response of SAHs to applied tension, which will likely require high-speed SMFS setups,[45] will be an important step towards this aim.

Entropy contributes significantly to the mechanical response. To gain insight into the enthalpic/entropic origin of the features of the force–extension curve, we examine the potential energy, E, as a function of the strain for the four pulling velocities shown in Fig. 3(B). E includes all the potential energy terms from the atomistic model (angle, dihedral, Lennard-Jones, electrostatics, and implicit solvent). The potential energy curves at the two highest pull speeds are similar between them, but are clearly distinct from the other pull speeds. For a given value of strain beyond $\varepsilon > 0.8$, the potential energy increases with pull speed. This suggests that at higher pull speeds, folded amino acids explore steeper regions of their free energy basin before opening. This behavior is also observed in the sticky chain model (results not shown). Our analysis of hydrogen bonds and salt bridges (ESI,† Section S2) supports this scenario. For $\varepsilon > 0.6$, all potential energy curves clearly have a steeper slope than the PMF curve indicating that entropy substantially softens helical stretching at equilibrium. The substantial contribution of entropy to the mechanical response of helices under tension appears to be general: e.g., it has also been observed in decaalanine in the vacuum.[46]

The large contribution of entropy to the mechanical response of helices under tension may have substantial implications: if unfolding is cooperative – i.e., the price to create multiple interfaces between folded and unfolded sections throughout the helix is high, so the helix unfolds residue-by-residue starting from the termini – then the entropic contribution arises at the level of individual amino acids. This contribution is by definition included in 1-dimensional models of the free energy landscape of amino acid stretching, and so these models should be sufficient to model the mechanical response of helices under tension. On the contrary, if unfolding under tension is non-cooperative – i.e., multiple unfolding/refolding events occur randomly throughout the length of the helix – then chain entropy plays a role, and models of the mechanical response of helices under tension must include it. We note that the random opening version

of the sticky chain model indeed allows for unfolding/refolding throughout the helix but does not correctly account for chain entropy.[28]

Unfolding mechanism is cooperative and occurs residue-by-residue at all pull speeds. The changes in secondary structure upon unfolding are shown in Fig. 4 for the individual simulations done at the highest and lowest velocities. The unfolding mechanism does not change noticeably in the velocity range investigated. The initial stretching of the helix occurs without loss of helical structure, in agreement with prior simulations of (metastable) helices[15,17] and similarly to the response of coiled coils in tension[47] and shear.[43,44] Larger strain involves loss of helical structure starting predominantly from the ends of the helix, *i.e.*, in the range of pull velocities investigated, unfolding is a cooperative mechanism. Cooperative unfolding was also observed in simulations of a 97 amino acid natural SAH[12] but not in simulations of some metastable helices[17] and is consistent with the higher thermodynamic stability of SAHs relative to metastable helices. Refolding (stick-slip) events are frequent, and consistent with the force peaks in individual force–strain curves as discussed above. The strain beyond which unfolding starts varies widely between different realizations, and unfolding may start from either end of the helix.

Fig. 4 gives the impression that unfolding occurs in clusters of 3 or 4 consecutive amino acids. This impression is misleading and results from the manner in which STRIDE assigns secondary structures. For example, if amino acids $i + 1$, $i + 2$, $i + 3$ and $i + 4$ initially meet all the criteria used to define a α-helix, but the hydrogen bond between i and $i + 4$ subsequently breaks, all four amino acids will lose their α-helix status. To ascertain whether unfolding occurs in clusters or residue-by-residue, we examined the backbone ϕ and ψ angles of each amino acid as a function of time. Results for selected residues are shown in ESI,† Section S3. The results confirm that at both speeds, unfolding occurs predominantly residue-by-residue.

Unfolding pathways are similar at all pull speeds. Fig. 4 suggests that amino acid unfolding may proceed *via* three pathways: pathway I includes a transition to a 3_{10}-helix (α-helix → ⋯ → 3_{10}-helix → ⋯ → coil); pathway II excludes that transition and the only intermediate configuration is a turn (α-helix → turn → coil); pathway III is a direct transition to the coil state (α-helix → coil). We quantify the fraction, f_i, of pathways of each type (i = I, II, III) at each pulling speed by averaging over all amino acids (a) as:

$$f_i = \frac{\langle n_i \rangle_a}{\langle n_I \rangle_a + \langle n_{II} \rangle_a + \langle n_{III} \rangle_a} \tag{6}$$

Here n_i is the number of times an amino acid unfolds *via* pathway i. The unfolding pathway(s) for each amino acid is (are) identified by monitoring its secondary structure changes between the instant(s), t_1, it assumes a α-helical configuration, and the instant t_2 it assumes a coil configuration for the first time after t_1. Incomplete unfolding events followed by refolding to the α-helix state are excluded from the statistics by resetting t_1 every time refolding occurs. The fraction of pathways of each type for the highest and lowest pulling speed are
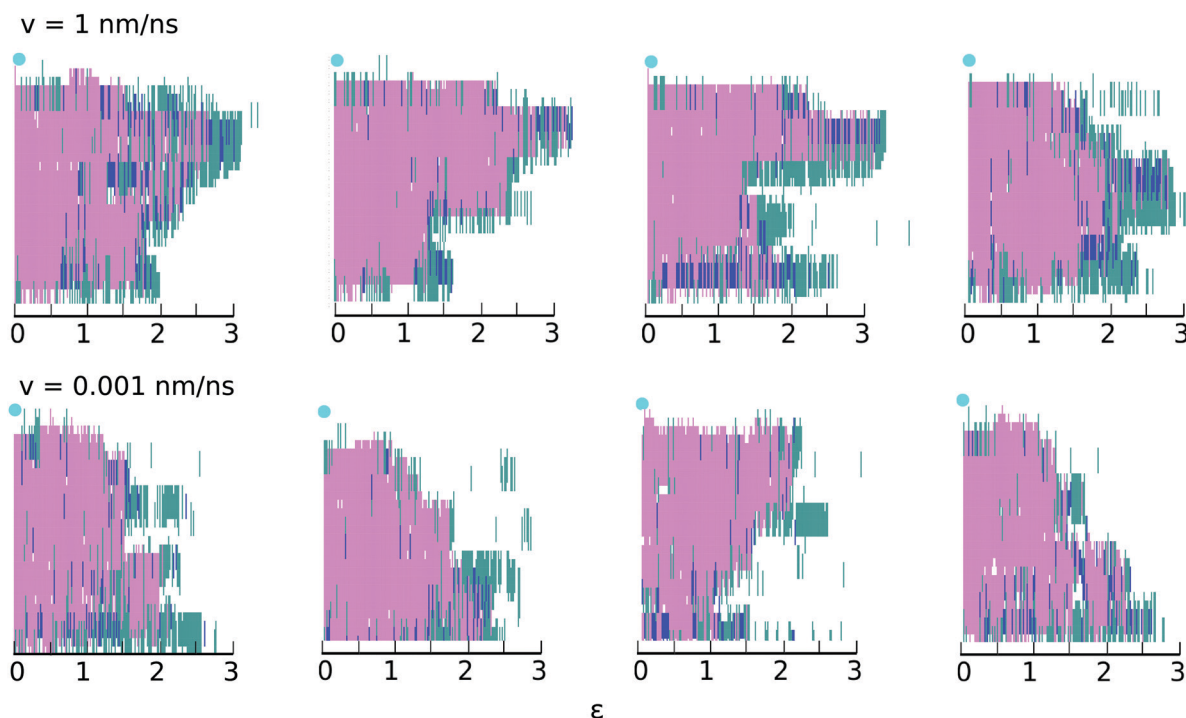

Fig. 4 Secondary structure changes as a function of strain at the highest (top) and lowest (bottom) velocities for four different realizations. The *y*-axis shows the secondary structure of each amino acid in the chain, determined using STRIDE:[48] turn (turquoise); α-helix (pink); 3_{10}-helix (dark blue) and coil (white). The light blue circle indicates the restrained end of the helix (amino terminus); the acid terminus, connected to the traveling virtual spring, is at *y* = 0.

Table 2 Fraction of unfolding pathways of each type (I, II and III) observed at the highest and lowest pulling speed, v. The uncertainty is estimated as the standard error of the mean over 4 independent molecular dynamics simulations

v (nm ns^{-1})	I	II	III
1	0.10 ± 0.02	0.26 ± 0.02	0.64 ± 0.03
10^{-3}	0.13 ± 0.02	0.28 ± 0.03	0.59 ± 0.04

shown in Table 2. The amino acid unfolding pathways are independent of the pull speed. The large majority of unfolding events takes place without intermediate configurations, at least at the resolution with which configurations were saved (every $vt = 0.01$ nm). The second most important pathway involves an intermediate turn configuration. Pathways with 3_{10}-helix intermediates represent only 10% of the total. Simulations of alanine-rich helical peptides also detected the presence of 3_{10}-helices in their unfolding pathways, both in water and in the vacuum.[15,36,37,49] This pathway, while non-dominant, appears to nevertheless be ubiquitous.

Backbone hydrogen bond breaking events are not the main barrier to loss of helical structure. Prior molecular dynamics studies of α-helices in tension have been interpreted to suggest that the barrier to α-helix stretching is primarily determined by the barrier to break hydrogen bonds, and that this occurs in a velocity-dependent manner: at low pulling velocities, consecutive backbone hydrogen bonds preferentially break simultaneously in clusters of 3-to-4; in contrast, at high pull speeds they break individually.[17,18] To determine if SAHs respond similarly, we calculated time series of hydrogen bond breaking events for all $(i, i + 4)$ hydrogen bonds along the helical backbone. The hydrogen bond criteria used is lax, *i.e.*, weak interactions are also identified as hydrogen bonds. The top graph (HB-break) in each panel of Fig. 5 shows representative time series for consecutive hydrogen bonds $(i, i + 4)$ and $(i + 1, i + 5)$ along the backbone. Each point (drawn with drop lines) indicates a hydrogen bond breaking event. At both high and low pull speeds, hydrogen bond breaking events are extremely frequent, in agreement with prior reports.[7,8,50,51]

α-Helical structure is determined by backbone angles ϕ and ψ. To assess whether the frequent hydrogen bond breaking events seen in the HB-break time series in Fig. 5 are accompanied by loss of helical structure, we plotted these angles as Ramachandran plots for all residues forming the loop stabilized by the backbone hydrogen bonds $(i, i + 4)$ indicated in that figure. These results are shown in ESI,† Fig. S4–S7. The Ramachandran plots confirm that transient hydrogen bond breaking is not accompanied loss of α-helical structure. Permanent destruction of backbone $(i, i + 4)$ hydrogen bonds only occurs after at least one of the relevant ϕ and ψ transitions out of α-helical configuration.

These results demonstrate that hydrogen bonds, while clearly contributing to the thermodynamic stability of helices, are not the main barrier to helix stretching. We (and others)[52] propose that this conclusion is general because experiments do not suggest a substantial dependence between the strength of
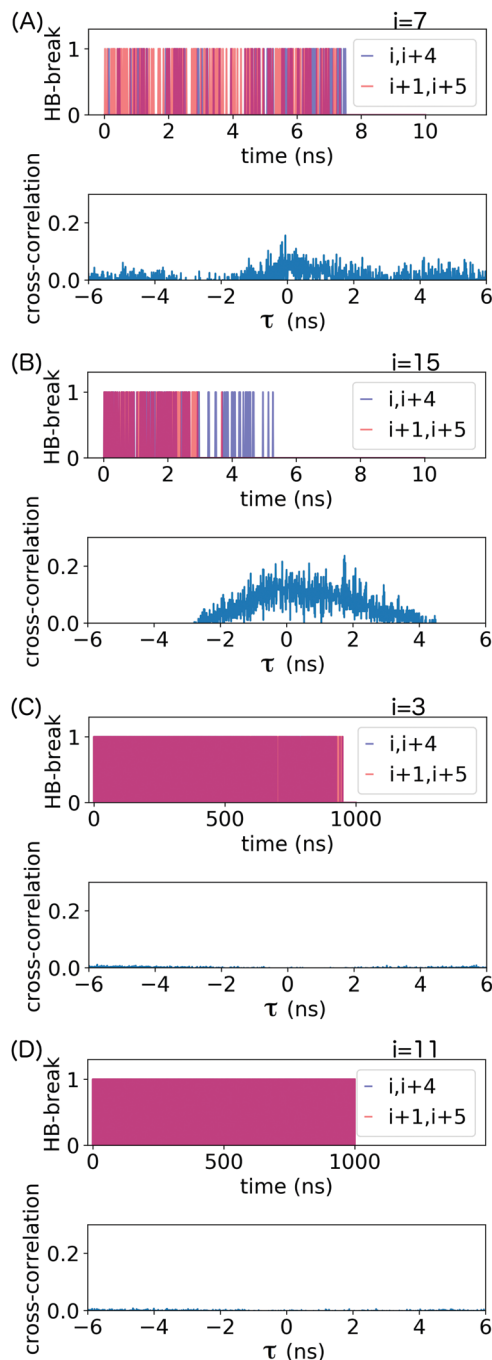

Fig. 5 Hydrogen bond breaking events (HB-break) for selected pairs of consecutive backbone hydrogen bonds, identified by the residue number i, and their respective time cross-correlation, from molecular dynamics simulations at (A and B) $v = 1$ nm ns^{-1} ($t = [0,10]$ ns corresponds to $\varepsilon = [0,2.9]$) and (C and D) $v = 10^{-3}$ nm ns^{-1} ($t = [0,1000]$ ns correspond to $\varepsilon = [0.5,0.8]$). The time resolution for analysis was 10 ps for both pull speeds. A hydrogen-bond breaking event occurs if the donor–acceptor distance becomes larger than 0.35 nm or the donor–hydrogen–acceptor angle becomes smaller than 140°. The HB-break time series represents each event using points with drop lines for ease of viewing; purple indicates simultaneous events within image resolution.

intrahelical hydrogen bonds and amino acid identity,[53] but polylysine, polyglutamate and SAHs have very different mechanical

response under tension as discussed in the introduction. Which interaction(s) contribute the most to the free energy barrier to unfolding of helices under tension remains for now an open question and is being investigated by our group.

Backbone hydrogen bonds break individually at all pull speeds. To assess whether hydrogen bonds preferentially break individually or in clusters, we analyzed the correlation between hydrogen bond breaking for all pairs of consecutive hydrogen bonds along the helical backbone. The normalized cross-covariance function is calculated as

$$R(\tau) = \frac{E[(y_1(t+\tau) - E(y_1))(y_2(t) - E(y_2))]}{\sigma_{y_1}\sigma_{y_2}} \quad (7)$$

where y_1 and y_2 are time series of hydrogen bond breaking events, σ_{y_1} and σ_{y_2} are their standard deviations, $E(y_1)$ and $E(y_2)$ are their means, and τ is the lag time.[54] The maximum of $R(\tau)$ indicates the time interval τ at which the two time series best align. If pairs of hydrogen bonds preferentially break simultaneously, there should be a large peak at $t = 0$ in this function. We focused on pairs of consecutive hydrogen bonds because, if hydrogen bonds indeed simultaneously break in clusters, consecutive pairs should show the strongest correlation. Our time resolution is 10 ps, so hydrogen bond pairs that break and remain broken within this time interval are considered to have broken simultaneously.

The cross-covariance functions are shown in the bottom plot of each panel in Fig. 5. A large peak at $t = 0$ is never observed, which demonstrates that consecutive hydrogen bonds along the helical backbone do not preferentially break simultaneously at any pull speed. Instead, all cross-covariance functions show distributions of time lags of several nanoseconds. These distributions are, as expected, narrower for the low pull speed because strain values leading to substantial loss of helical structure and thus permanent breakage of hydrogen bonds (see also discussion in ESI,† Section S2) are reached faster.

3.2 Parameters of the chain model for the cooperative opening

Analysis of the simulation trajectories indicate that the SAH unfolds cooperatively. The cooperative version of the sticky chain model proposed by Jäger should thus be the most appropriate to fit the force–strain curves as a function of pull speed. Before doing the actual fit, we determine several of the parameter values for the mesoscopic model directly from the atomistic simulations, and from equilibrium conditions derived within the mesoscopic model; these parameters are show in Table 3. The atomistic simulations show that the unfolding units are the individual amino acids rather than the helical turns. The number of links in the mesoscopic model is set at $N = 30 - 2 - 1 = 27$. The two terminal glycines are excluded from the number of links; each link in the mesoscopic model is defined by the presence of a peptide bond, so 28 non-glycine amino acids define 27 links. The parameter b_0 was estimated from known geometrical parameters for α-helices: the helical pitch is 0.54 nm and comprises 3.6 amino acids,

Table 3 Values of the parameters calculated from the atomistic simulations and used as input to the fits of the sticky chain model

Parameter	Value
N	27
b_0	0.15 nm
ΔL	0.23 nm
V_0	20 pN nm ($4.8k_BT$)

corresponding to a distance $b_0 = 0.15$ nm between consecutive C_α atoms projected along the helical axis. We calculated the parameter ΔL from the average distance between consecutive C_α atoms from multiple simulations at constant velocity, using configurations beyond the force plateau. Under these conditions the helix is unfolded and the C_α atoms are aligned with the pull axis so the distance between them is equal to the projected distance, $b_0 + \Delta L = 0.38$ nm, corresponding to $\Delta L = 0.23$ nm. V_0 was calculated from the force plateau $F_p^{v=10^{-3}} = 87$ pN value as $V_0 = F_p^{v=10^{-3}} \times \Delta L$ (eqn (A3) in ref. 28). This equality assumes that $F_p^{v=10^{-3}}$ is in the quasi-static limit which, as discussed above, is not true. However, the timescale separation between processes important under quasi-static conditions and those important for $v > 10^{-3}$ nm ns^{-1} and the fact that the plateaus at $v \in \{10^{-3}, 10^{-2}\}$ nm ns^{-1} are very close mean that, for the purpose of applying the mesoscopic model in the range $10^{-3} < v/(\text{nm ns}^{-1}) < 1$, the lowest v can be considered in the quasi-static limit. We obtain $V_0 = 4.8k_BT = 2.9$ kcal mol^{-1}, a value substantially lower than the 11.1 kcal mol^{-1} reported by Buehler *et al.*[18] for a metastable helix; below we discuss possible origins of this discrepancy. Because $x_E = \sqrt{2V_0/k_1}$ (eqn (5)), the parameters that remain to be obtained from the fit are the natural vibration frequency, ν_0, and the bond constants k_1 and k_2 associated with the potential basins of the folded and unfolded states.

3.3 Fitting the force–strain curves from all-atom simulations

We fit the average force–strain curves from the all-atom simulations using eqn (1) and (2) and the parameters in Table 3. Fitting is done by substituting eqn (1) in eqn (2) and numerically integrating eqn (2) to get the number, N_Λ, of folded links as a function of the time. The fitting parameters are varied automatically with the function fmincon of MATLAB2016b in order to minimize $|F_m - F_s|$, where F_m and F_s are the values of the force obtained by the model and the force obtained in the all-atom simulations, respectively. To test the accuracy of the fitting procedure, we make two implementation attempts: (1) fitting one curve ($v = 10^{-1}$ nm ns^{-1}), or (2) simultaneously fitting two curves ($v \in \{10^{-1}, 10^{-3}\}$ nm ns^{-1}). In both cases we use the parameters obtained from the fit to generate the force extension curves at the remaining pull velocities of the atomistic simulations.

The fitting parameters obtained in each case are shown in Table 4, and the corresponding curves are shown in Fig. 6. Both fitting attempts yield comparable fitting parameters and reasonable force–strain curves for the 4 pull speeds, confirming that the model and fitting procedure are robust. These results

Table 4 Values of the parameters k_1, k_2 and ν_0 obtained from fitting the sticky chain model to the all-atom force–strain curves at the indicated pull speed(s), and values of derived parameters k_{off} (from eqn (4)) and x_E (from eqn (5))

Parameter	ν (nm ns^{-1})	
	10^{-1}	$\{10^{-1}, 10^{-13}\}$
k_1 (pN nm^{-1})	1321	2118
k_2 (pN nm^{-1})	5188	4792
ν_0 (ns^{-1})	0.23	0.82
k_{off} (s^{-1})	0.18×10^7	0.63×10^7
x_E (nm)	0.17	0.13

indicate that reliable insight into the 1-dimensional free energy landscape associated with helix unfolding under tension may be obtained already with one quasi-static and one non-equilibrium force–strain curve. Fits using curves at more than one pull speed are nevertheless expected to yield better estimates of the fit parameters, so the following discussion focuses on the results of the two-curve fit.

The bond constant, $k_2 = 4792$ pN nm^{-1}, characterizing the free energy basin of the unfolded state is substantially larger

than $k_1 = 2118$ pN nm^{-1}, which characterizes the free energy basin of the folded state. This difference is consistent with the fact that stretching an extended amino acid chain involves deforming angle potentials (bonds are constrained in our model), which are substantially stiffer than any stabilizing contribution to the folded state (dihedrals, hydrophobic effect, hydrogen bonds and salt bridges).

The natural vibration frequency, $\nu_0 = 0.82$ ns^{-1}, obtained from the fit is substantially lower than the typical value of 10^4 ns^{-1} used as estimate of this quantity.[55] This difference demonstrates that the commonly used estimate of ν_0 is in fact far from suitable in this case, where each link has much higher mass and much lower bond stiffness (k_1) than covalent bonds between 2 atoms, for which the estimate was originally made. A reliable value of ν_0 is indispensable to calculate V_0 from values of k_{off} (eqn (4)), which can be easily obtained from Bell–Evans fits of rupture force vs. velocity data. This was the procedure followed by Buehler et al.[18] to analyze force–strain curves from simulation; from a $k_{off}^{Buehler} = 0.807 \times 10^5$ s^{-1} they obtain $V_0^{Buehler} = 11.1$ kcal mol^{-1} for a metastable helix and the same pull speed range investigated here. Using our estimate of
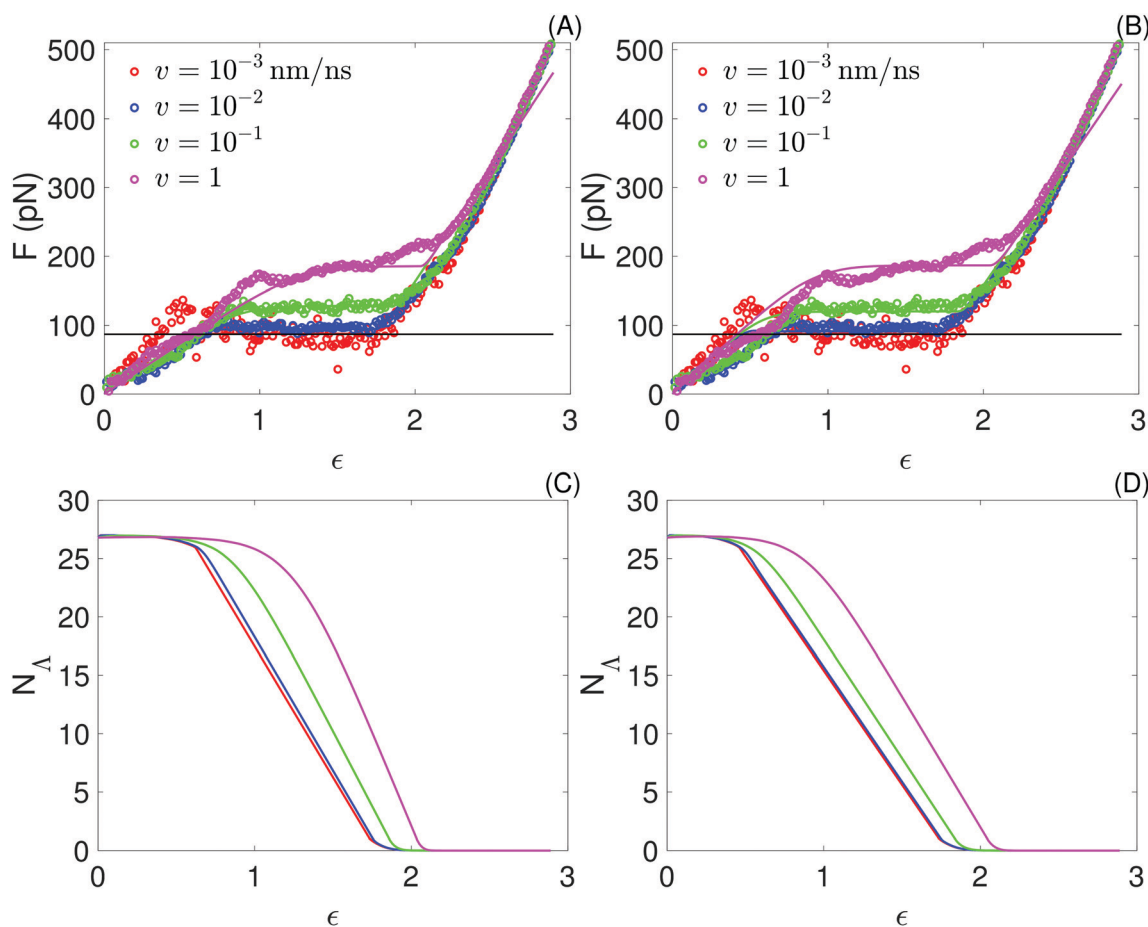

Fig. 6 Top: Force as a function of strain obtained with the sticky chain model (lines) using the two set of parameters in Table 3; to facilitate comparisons, the force–strain curves from all-atom simulations, already presented in Fig. 3, are repeated here (points). (A) Fitting one curve, $v = 10^{-1}$ nm ns^{-1}. (B) Fitting 2 curves, at $v \in \{10^{-1}, 10^{-3}\}$ nm ns^{-1}. Bottom: Number of folded links as a function of strain obtained from the sticky chain model. (C) Fitting one curve. (D) Fitting 2 curves.

ν_0 would result in $V_0^{\text{Buehler}} = 5.5$ kcal mol^{-1}. This value is still substantially higher than our own estimate, $V_0 = 4.8k_BT = 2.9$ kcal mol^{-1}, because $k_{\text{off}}^{\text{Buehler}}$ is 2 orders of magnitude lower than what we obtain (see Table 4). Whether these differences reflect true differences between the SAH investigated here and the AH1 helix from the vimentin intermediate filament investigated by Buehler *et al.*, or are instead artifacts arising from the different force fields used, remains unclear.

3.4 How do typical analysis procedures used in dynamic force spectroscopy compare with the sticky chain model?

To answer this question, we simulate force–strain curves using the sticky chain model and the parameters in Tables 3 and 4. Subsequently we analyze the value of the plateau force as a function of the loading rate with two of the models most often used in dynamic force spectroscopy: the Bell–Evans model, which is valid for velocities at which refolding can be neglected, and the Yoreo model, which is an extension of the first that

Fig. 7 (A) Force–strain curves generated from the sticky chain model using the parameters in Tables 3 and 4 (for fits using the curves at $v \in \{10^{-1}, 10^{-3}\}$ nm ns^{-1}); the black crosses indicate the values taken as the plateau force shown in panel (B). The pull speed is indicated as a multiple of the quasi-static velocity, $v_{\text{eq}} = 10^{-3}$ nm ns^{-1}. (B) Plateau force *vs.* loading rate, $k_s v$, from the sticky chain model (points) and fits using the Bell–Evans and Friddle–Yoreo model (lines).

Table 5 Parameters of the Bell–Evans and Yoreo models resulting from the fits in the bottom panel of Fig. 7. The reported uncertainty for the Yoreo parameters is their 95% confidence interval, obtained directly from the fit. The uncertainty for the Bell–Evans parameters was obtained using error propagation from the 95% confidence interval of the fit parameters, neglecting correlations between them

	Bell–Evans	Yoreo
Δx (nm)	0.14 ± 0.06	0.11 ± 0.01
k_{off} (s^{-1})	$(0.09 \pm 0.2) \times 10^8$	$(0.19 \pm 0.04) \times 10^8$

captures the quasi-static region where both unfolding and refolding take place. The Bell–Evans model[56] relates the most probable value F^* of the rupture force and the loading rate $r = \mathrm{d}F/\mathrm{d}t$ as:

$$F^* = \frac{k_B T}{\Delta x} \ln\left(\frac{r\Delta x}{k_{\text{off}}k_B T}\right), \tag{8}$$

where Δx is the distance between the folded state and the barrier to unfolding, and k_{off} is the unfolding rate at zero force.

The Yoreo model[23] relates the same two observables,

$$F^* = f_{\text{eq}} + \frac{k_B T}{\Delta x}\exp\left(\frac{k(f_{\text{eq}})k_B T}{r\Delta x}\right)E_1\left(\frac{k(f_{\text{eq}})k_B T}{r\Delta x}\right), \tag{9}$$

but is expected to be valid for the whole range of loading rates. This model includes as additional parameter the equilibrium unfolding force, f_{eq}. $k(f_{\text{eq}}) = k_{\text{off}}\exp\left(\frac{1}{k_B T}\left(f_{\text{eq}}\Delta x - 0.5k_s\Delta x^2\right)\right)$ and $E_1(z) = \int_z^\infty \frac{e^{-s}}{s}\mathrm{d}s$ is the exponential integral. We note that Δx has the same physical significance as x_E.

In the top panel of Fig. 7 we show force–strain curves obtained with the sticky chain model and the parameters in Tables 3 and 4 (for fits done at $v \in \{10^{-1}, 10^{-3}\}$ nm ns^{-1}) for a wide range of pull velocities. Values of the plateau force at each pull speed, indicated by the black markers, are re-plotted as a function of the loading rate in the bottom panel of the same figure. We fit the Bell–Evans model to the 3 points with higher loading rate, for which re-binding is expected to be minimal, and fit the Yoreo model to the entire curve. The Bell–Evans model is linearized for the fit ($F = p_1 \times x + p_2$), with Δx and k_{off} obtained as $\Delta x = k_B T/p_1$ and $k_{\text{off}} = (p_1 \exp(-p_2/p_1))^{-1}$.

The parameter values obtained from the fits are shown in Table 5. Parameters extracted with both models are comparable, but not identical, to those used as input (Table 4) to the sticky chain model, indicating that both the Yoreo and the Bell–Evans models are reasonable choices for initial analyses, but yield only coarse estimates of Δx and k_{off}. These results illustrate that, for helical systems where a (quasi-)force plateau can be observed as a function of strain, fits using the sticky chain model are preferable.

4 Concluding remarks

Our simulations indicate that the deformation mechanism of SAHs under applied tension is highly cooperative at all pull speeds tested: because the free energy penalty for creating
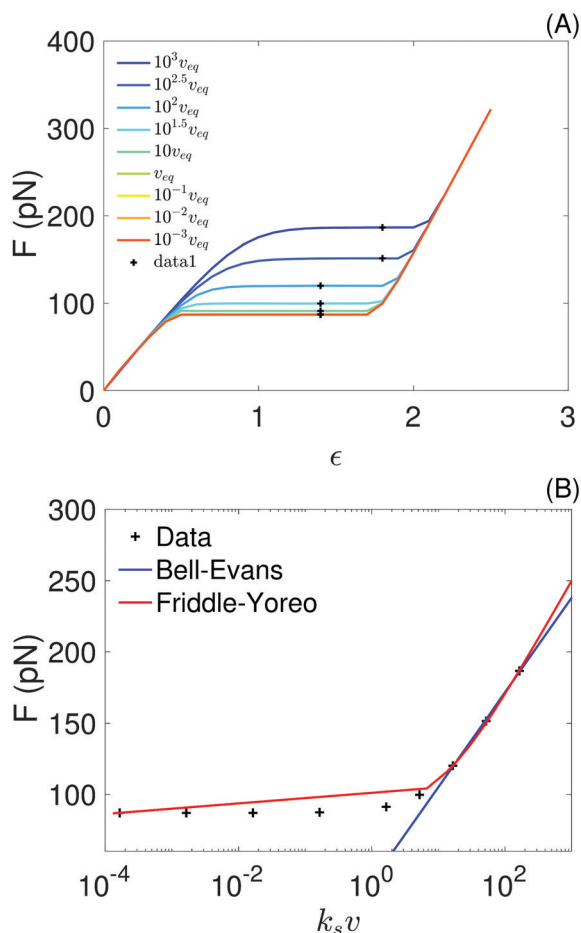
multiple interfaces between folded and unfolded domains is high, helix unraveling starts from the termini. This response is consistent with the high thermodynamic stability of SAHs. In contrast, the typical α-helices found in proteins are stabilized partly by interactions with other protein domains, *i.e.*, in the absence of those interactions the helical conformation is only metastable. Simulation studies indicate that, when metastable α-helices are stretched in solution, unfolding starts randomly along the helix.[17,18] The main barrier to unfolding of residues cannot be hydrogen bond breaking as several literature reports suggest, because hydrogen bond breaking is both extremely frequent and does not *per se* induce loss of helical structure. Instead, we found that the opposite occurs: loss of helical structure determines the permanent breakage of backbone hydrogen bonds. We also found that hydrogen bonds overwhelmingly break individually rather than in clusters. The unraveling of individual residues occurs *via* multiple pathways, and may include transitions to turn or to metastable 3_{10}-helix configurations. These transitions are stick-slip events; their presence is consistent with the appearance of peaks in individual force–strain curves. Entropy clearly softens the mechanical strength of the helix and the cooperative nature of the unfolding process suggests that the contribution of entropy arises at amino acid level. For uncooperative unfolding processes this will not be the case, and adequately modeling force–extension curves under those conditions must account for chain entropy. At present, experimental data on the mechanical response of SAHs to tension is scarce under quasi-static conditions, and does not exist for non-equilibrium pull speeds. Our results using the sticky chain model[28] to fit the force–extension curves obtained using the simulations point to the importance of obtaining non-equilibrium data. If that data exists, the sticky chain model can be used to fit the experimental data and should yield unprecedented insight into the free energy landscape and mechanism associated with unfolding of SAHs. Such insight is critical to understand the role of SAHs in biology and their potential applications in materials science.

Conflicts of interest

There are no conflicts of interest to declare.

Acknowledgements

Notes and references

1 B. J. Spink, S. Sivaramakrishnan, J. Lipfert, S. Doniach and J. A. Spudich, *Nat. Struct. Mol. Biol.*, 2008, **15**, 591–597.

2 D. Süveges, Z. Gáspári, G. Tóth and L. Nyitray, *Proteins*, 2009, **74**, 905–916.

3 M. Peckham and P. J. Knight, *Soft Matter*, 2009, **5**, 2493–2503.

4 Z. Gáspári, D. Süveges, A. Perczel, L. Nyitray and G. Tóth, *Biochim. Biophys. Acta, Proteins Proteomics*, 2012, **1824**, 637–646.

5 D. Simm, K. Hatje and M. Kollmar, *PLoS One*, 2017, **12**, e0174639.

6 C. A. Barnes, Y. Shen, J. F. Ying, Y. Takagi, D. A. Torchia, J. R. Sellers and A. Bax, *J. Am. Chem. Soc.*, 2019, **141**, 9004–9017.

7 P. Lyu, P. J. Gans and N. R. Kallenbach, *J. Mol. Biol.*, 1992, **223**, 343–350.

8 C. A. Olson, E. J. Spek, Z. Shi, A. Vologodskii and N. R. Kallenbach, *Proteins*, 2001, **44**, 123–132.

9 A. Idiris, M. T. Alam and A. Ikai, *Protein Eng., Des. Sel.*, 2000, **13**, 763–770.

10 M. A. Lantz, S. P. Jarvis, H. Tokumoto, T. Martynski, T. Kusumi, C. Nakamura and J. Miyake, *Chem. Phys. Lett.*, 1999, **315**, 61–68.

11 M. Kageshima, M. A. Lantz, S. P. Jarvis, H. Tokumoto, S. Takeda, A. Ptak, C. Nakamura and J. Miyake, *Chem. Phys. Lett.*, 2001, **343**, 77–82.

12 M. Wolny, M. Batchelor, P. J. Knight, E. Paci, L. Dougan and M. Peckham, *J. Biol. Chem.*, 2014, **289**, 27825–27835.

13 F. Berkemeier, M. Bertz, S. Xiao, N. Pinotsis, M. Wilmanns, F. Gräter and M. Rief, *Proc. Natl. Acad. Sci. U. S. A.*, 2011, **108**, 14139–14144.

14 I. Schwaiger, C. Sattler, D. R. Hostetter and M. Rief, *Nat. Mater.*, 2002, **1**, 232.

15 G. Ozer, S. Quirk and R. Hernandez, *J. Chem. Theory Comput.*, 2012, **8**, 4837–4844.

16 N. Bavi, O. Bavi, M. Vossoughi, R. Naghdabadi, A. P. Hill, B. Martinac and Y. Jamali, *Channels*, 2017, **11**, 209–223.

17 T. Ackbarow and M. J. Buehler, *J. Mater. Sci.*, 2007, **42**, 8771–8787.

18 T. Ackbarow, X. Chen, S. Keten and M. J. Buehler, *Proc. Natl. Acad. Sci. U. S. A.*, 2007, **104**, 16410–16415.

19 M. Rief, J. M. Fernandez and H. E. Gaub, *Phys. Rev. Lett.*, 1998, **81**, 4764–4767.

20 J. Bertaud, J. Hester, D. D. Jimenez and M. J. Buehler, *J. Phys.: Condens. Matter*, 2010, **22**, 035102.

21 A. V. Savin, I. P. Kikot, M. A. Mazo and A. V. Onufriev, *Proc. Natl. Acad. Sci. U. S. A.*, 2013, **110**, 2816–2821.

22 M. L. Hughes and L. Dougan, *Rep. Prog. Phys.*, 2016, **79**, 076601.

23 R. W. Friddle, A. Noy and J. J. De Yoreo, *Proc. Natl. Acad. Sci. U. S. A.*, 2012, **109**, 13573–13578.

24 O. K. Dudko, G. Hummer and A. Szabo, *Phys. Rev. Lett.*, 2006, **96**, 108101.

25 B. Chakrabarti and A. J. Levine, *Phys. Rev. E: Stat., Nonlinear, Soft Matter Phys.*, 2005, **71**, 031905.

26 B. Chakrabarti and A. J. Levine, *Phys. Rev. E: Stat., Nonlinear, Soft Matter Phys.*, 2006, **74**, 031903.

27 Z. Qin and M. J. Buehler, *Phys. Rev. Lett.*, 2010, **104**, 198304.

28 I. L. Jäger, *Biophys. J.*, 2001, **81**, 1897–1906.

29 M. Andersson, E. Fällman, B. E. Uhlin and O. Axner, *Biophys. J.*, 2006, **91**, 2717–2725.

30 M. Wolny, M. Batchelor, G. J. Bartlett, E. G. Baker, M. Kurzawa, P. J. Knight, L. Dougan, D. N. Woolfson, E. Paci and M. Peckham, *Sci. Rep.*, 2017, **7**, 44341.

31 K. Lindorff-Larsen, S. Piana, K. Palmo, P. Maragakis, J. L. Klepeis, R. O. Dror and D. E. Shaw, *Proteins: Struct., Funct., Bioinf.*, 2010, **78**, 1950–1958.

32 E. F. Pettersen, T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng and T. E. Ferrin, *J. Comput. Chem.*, 2004, **25**, 1605–1612.

33 B. H. Zimm and J. K. Bragg, *J. Chem. Phys.*, 1959, **31**, 526–535.

34 M. Cieplak, T. X. Hoang and M. O. Robbins, *Proteins: Struct., Funct., Genet.*, 2002, **49**, 104–113.

35 F. C. Zegarra, G. N. Peralta, A. M. Coronado and Y. Q. Gao, *Phys. Chem. Chem. Phys.*, 2009, **11**, 4019–4024.

36 R. Rohs, C. Etchebest and R. Lavery, *Biophys. J.*, 1999, **76**, 2760–2768.

37 R. Afrin, I. Takahashi, K. Shiga and A. Ikai, *Biophys. J.*, 2009, **96**, 1105–1114.

38 J. P. Renn, S. Bhattacharyya, H. Bai, C. He, H. Li, A. F. Oberhauser, J. F. Marko, D. E. Makarov and A. Matouschek, *Sci. Rep.*, 2019, **9**, 11101.

39 D. D. Root, V. K. Yadavalli, J. G. Forbes and K. Wang, *Biophys. J.*, 2006, **90**, 2852–2866.

40 S. B. Smith, Y. Cui and C. Bustamante, *Science*, 1996, **271**, 795–799.

41 P. Cluzel, A. Lebrun, C. Heller, R. Lavery, J.-L. Viovy, D. Chatenay and F. Caron, *Science*, 1996, **271**, 792–794.

42 S. A. Harris, Z. A. Sands and C. A. Laughton, *Biophys. J.*, 2005, **88**, 1684–1691.

43 M. Goktas, C. Luo, R. M. A. Sullan, A. E. Bergues-Pupo, R. Lipowsky, A. Vila Verde and K. G. Blank, *Chem. Sci.*, 2018, **9**, 4610–4621.

44 A. E. Bergues-Pupo, K. G. Blank, R. Lipowsky and A. Vila Verde, *Phys. Chem. Chem. Phys.*, 2018, **20**, 29105–29115.

45 C. Valotteau, F. Sumbul and F. Rico, *Biophys. Rev.*, 2019, **11**, 689–699.

46 S. Park, F. Khalili-Araghi, E. Tajkhorshid and K. Schulten, *J. Chem. Phys.*, 2003, **119**, 3559–3566.

47 A. Torres-Sánchez, J. M. Vanegas, P. K. Purohit and M. Arroyo, *Soft Matter*, 2019, **15**, 4961–4975.

48 D. Frishman and P. Argos, *Proteins: Struct., Funct., Bioinf.*, 1995, **23**, 566–579.

49 M. Post, S. Wolf and G. Stock, *J. Chem. Phys.*, 2019, **150**, 204110.

50 P. E. Mason, P. Jungwirth and E. Duboue-Dijon, *J. Phys. Chem. Lett.*, 2019, **10**, 3254–3259.

51 H. Adalsteinsson, A. H. Maulitz and T. C. Bruice, *J. Am. Chem. Soc.*, 1996, **118**, 7689–7693.

52 A. Ikai, *Biopolymers: Lignin, Proteins, Bioactive Nanocomposites*, 2010, vol. 232, pp. 65–96.

53 J. M. Scholtz, H. Qian, E. J. York, J. M. Stewart and R. L. Baldwin, *Biopolymers*, 1991, **31**, 1463–1470.

54 R. E. Thomson and W. J. Emery, in *Data Analysis Methods in Physical Oceanography*, ed. R. E. Thomson and W. J. Emery, Elsevier, Boston, 3rd edn, 2014, pp. 425–591.

55 G. Bell, *Science*, 1978, **200**, 618–627.

56 E. Evans and K. Ritchie, *Biophys. J.*, 1997, **72**, 1541–1555.