Energy landscape, structure and rate effects on strength properties of alpha-helical proteins

2010 J. Phys.: Condens. Matter 22 035102

(http://iopscience.iop.org/0953-8984/22/3/035102)

IOP Publishing Journal of Physics: Condensed Matter

J. Phys.: Condens. Matter **22** (2010) 035102 (14pp) doi:10.1088/0953-8984/22/3/035102

Energy landscape, structure and rate effects on strength properties of alpha-helical proteins

Jérémie Bertaud[1], Joshua Hester[1], Daniel D Jimenez[1] and Markus J Buehler[1,2,3]

[1] Laboratory for Atomistic and Molecular Mechanics, Department of Civil and Environmental Engineering, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Room 1-235A&B, Cambridge, MA 02139, USA

[2] Center for Computational Engineering, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA 02139, USA

E-mail: mbuehler@MIT.EDU

Received 29 July 2009, in final form 18 November 2009
Published 21 December 2009
Online at stacks.iop.org/JPhysCM/22/035102

Abstract

The strength of protein domains is crucial to identify the mechanical role of protein domains in biological processes such as mechanotransduction, tissue mechanics and tissue remodeling. Whereas the concept of strength has been widely investigated for engineered materials, the strength of fundamental protein material building blocks and how it depends on structural parameters such as the chemical bonding, the protein filament length and the timescale of observation or deformation velocity remains poorly understood. Here we report a systematic analysis of the influence of key parameters that define the energy landscape of the strength properties of alpha-helical protein domains, including energy barriers, unfolding and refolding distances, the locations of folded and unfolded states, as well as variations of the length and pulling velocity of alpha-helical protein filaments. The analysis is facilitated by the development of a double-well mesoscale potential formulation, utilized here to carry out a systematic numerical analysis of the behavior of alpha-helices. We compare the results against widely used protein strength models based on the Bell model, one of the simplest models used to characterize the strength of protein filaments. We find that, whereas Bell-type models are a reasonable approximation to describe the rupture of alpha-helical protein domains for a certain range of pulling speeds and values of energy barriers, the model ceases to hold for very large energy barriers and for very small pulling speeds, in agreement with earlier findings. We conclude with an application of our mesoscale model to investigate the effect of the length of alpha-helices on their mechanical strength. We find a weakening effect as the length of alpha-helical proteins increases, followed by an asymptotic regime in which the strength remains constant. We compare strand lengths found in biological proteins with the scaling law of strength versus alpha-helix filament length. The mesoscale model reported here is generally applicable to other protein filaments that feature a serial array of domains that unfold under applied strain, where a similar length-dependent strength could be observed.

(Some figures in this article are in colour only in the electronic version)

1. Introduction

The strength of protein filaments is crucial to identify the mechanical role of protein constituents in key biological

[3] Author to whom any correspondence should be addressed.
http://web.mit.edu/mbuehler/www/

processes such as mechanotransduction, tissue mechanics and tissue remodeling [1–7]. In addition to medical and biological applications, a better understanding of alpha-helices and alpha-helix-based protein networks and their resulting mechanical properties could possibly lead to the design and synthesis of *de novo* synthetic alpha-helix-based peptide and protein materials [8–10]. Specifically, alpha-helices are found in many structural fibrous protein materials such as hair, hoof and intermediate filaments, and there they play an important biomechanical role. Thus, a thorough understanding of the mechanical properties is critical to improve biomechanical models of key biological materials and processes, and in particular an understanding of how the geometry of alpha-helices is linked to their mechanical properties. Whereas the concept of strength has been widely investigated for engineered materials, the strength of protein materials, and how it depends on the structure and properties of basic biological protein building blocks (e.g. the chemical bonding, protein filament length, etc) is not understood as well. Specifically, open questions remain with respect to the strength properties depending on the geometry (e.g. length) of protein filaments and the effects of the strength of key chemical bonds (such as H-bonds) that can vary as a function of the solvent environment of the protein.

This paper is focused on a multi-scale simulation analysis of strength properties of alpha-helical (also abbreviated as AH) protein domains. The geometry of an alpha-helical protein is shown in figure 1(a), showing its characteristic coiled structure. As stated above, the alpha-helix represents a fundamental building block of a broad class of protein materials (including hoof, hair, wool and cells), where it typically forms assemblies with other alpha-helices (such as coiled-coils) or other secondary structures (such as beta-sheets, random coils, etc). Here we focus on the analysis of single alpha-helix proteins, selected here as a model system, to elucidate the influence of key parameters that characterize the energy landscape of the unfolding process of alpha-helical protein filaments, including energy barriers, unfolding and refolding distances, the locations of folded and unfolded states, as well as the geometry (here: length) of the protein filament. The analysis is facilitated by the development of a simple double-well mesoscale potential formulation to model the behavior of alpha-helices, which is used to carry out a systematic parameter study based on molecular dynamics.

The plan of this paper is as follows. We first validate the double-well potential mesoscale model for single alpha-helical protein molecules against pulling experiments of a single alpha-helix that were carried out earlier with full atomistic simulations. After establishing this reference system, we measure the influence of energy landscape parameters on the rate (i.e. pulling velocity) dependence of the rupture force and compare it with the predictions of the Bell model for a range of pulling velocities. We identify the sensitivity of all parameters in our mesoscale model, and show the limitations of the applicability of the Bell model to describe the rupture strength. Finally, we apply our double-well potential mesoscale model to study the length effect of the strength of an individual alpha-helical protein, providing a structure–property link of the strength properties of alpha-helical proteins.

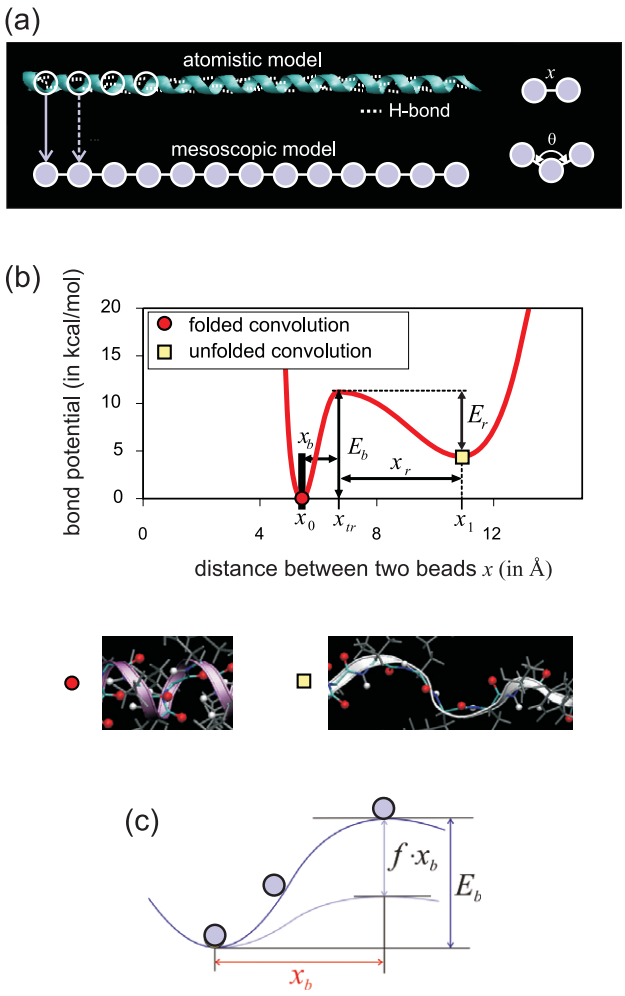

Figure 1. Mesoscale model set-up, geometry and parameters. Subplot (a) schematic of the coarse-graining procedure, in which we replace the full atomistic representation by a mesoscopic bead model. A pair of beads represents one turn in the alpha-helix (also called a convolution), and thus 3.6 residues (each bead also has the corresponding mass). In the atomistic representation, the folded states of the turns are stabilized by the presence of 3.6 H-bonds between turns. In the mesoscopic bead model, this is represented by using a double-well potential to describe the energy landscape under bond stretching. Subplot (b) double-well profile of the bond stretching potential in the bead model, representing the energy landscape associated with unfolding of one convolution. The numerical values of the equilibrium states (x_0 and x_1), energy barriers (E_b and E_r) and the transition state x_{tr} are obtained from geometric analysis of the alpha-helix geometry, as well as full atomistic simulations. The transition state (local energy peak of the potential) corresponds to the breaking of the 3.6 H-bonds between two convolutions of the alpha-helix. After failure of these weak bonds, the convolution unfolds to a second equilibrium state with a larger interbead distance. Under further loading, its covalent bonds begin to be stretched which leads to a second increase of the potential at large deformation. Subplot (c) evolution of the energy landscape of a bond subjected to a force according to Bell's theory. The rupture of the bond occurs via thermally assisted crossing of an activation barrier E_b which is reduced by $f x_b$ as the applied force f increases.

2. Theoretical and computational methods

This section contains a detailed discussion of the classical Bell model [11] and related methods used to describe the strength of protein domains, as well as details of the implementation of

our numerical mesoscale simulation approach. Several recent papers have introduced advanced models for the strength of proteins, including works by Szabo, Dudko, Hummer and co-workers [12–15], Klafter and co-workers [16] and Lin *et al* [17], as well as Friddle [18]. In our study, however, we focus solely on the simple Bell model [11] and related approaches. The Bell model, despite its limitations, is one of the most well-known, simplest and most widely used strength models of proteins and, as such, serves as a suitable starting point for the development of a coarse-grained model and subsequent comparison with a strength model. Studies that include a comparison with more advanced models is left to future work [12–18], where a similar systematic approach as used here could be pursued.

2.1. Theoretical strength models

The 'classical' Bell model is a simple phenomenological model that describes the frequency of failure of reversible bonds [11]. The concept of reversibility thereby means that an individual bond can break under no force if one waits a sufficiently long time, and that it can reform spontaneously. Such bonds may be associated with electrostatic, van der Waals (vdW) or H-bond interactions (as it applies to alpha-helices). The frequency of failure, also called dissociation rate or off rate, is defined as the inverse of the bond lifetime and is used as a concept to describe the dynamical behavior of such bonds.

Bell's model explains the force dependence of the off rate and thus shows the significant role of mechanical force in biological chemistry. For instance, this model can be applied to describe the forced unbinding of biological adhesive contacts such as adhesion of cells to cells [11]. Bell's model is an extension of the transition state theory for reactions in gases developed by Eyring and others [19]. Inspired also by Zhurkov's work on the kinetic theory of the strength of solids [20], Bell predicted for the first time that the off rate of a reversible bond, which is the inverse of the bond lifetime, increases when subjected to an external force f. Indeed, the rupture of bonds occurs via thermally assisted crossing of an activation barrier E_b which is reduced by fx_b as the applied force f increases, x_b being the distance between the bound state and the transition state (see figure 1(c)). Thus, the Bell off rate expression [21] is given by

$$k = \omega_0 \exp\left(-\frac{E_\mathrm{b} - fx_\mathrm{b}}{k_\mathrm{B}T}\right), \tag{1}$$

where ω_0 is the natural vibration frequency of the bond and $k_\mathrm{B}T$ is the thermal energy (see figure 1(c) for an overview of the energy landscape as considered in Bell's model; it contains only the energy barrier associated with unfolding).

The rupture force can be predicted based on different formulations. In this paper, we will utilize three different models and compare them with our mesoscale simulation results. Two of these methods are directly derived from the concept put forth in Bell's basic model (discussed in section 2.1). The rupture force

$$f_\mathrm{crit} = E_\mathrm{b}/x_\mathrm{b}, \tag{2}$$

corresponds to the force to completely diminish the energy barrier, defined as the critical force (see also the schematic in figure 1(c)), leading to instantaneous rupture of the bond. For forces smaller than f_crit ($f < f_\mathrm{crit}$), there still exists a finite probability that the bond breaks. The dynamics of bond rupture for these cases can be estimated based on Bell's model. The following expression is derived from Kramers' theory by Evans and Ritchie, to express the rupture force at a constant loading rate $r = Kv$ [22], where

$$f(v, E_\mathrm{b}, x_\mathrm{b}) = \frac{k_\mathrm{B}T}{x_\mathrm{b}}\ln(v) - \frac{k_\mathrm{B}T}{x_\mathrm{b}}\ln\left(\frac{k_\mathrm{B}T\,w_0}{Kx_\mathrm{b}}\right) + \frac{E_\mathrm{b}}{x_\mathrm{b}} = a\ln(v) + b. \tag{3}$$

In equation (3), K is the stiffness of the force transducer and $v = \Delta x/\Delta t$ is the constant pulling speed at which the protein structure is deformed. This relation predicts that the strength depends logarithmically on the pulling velocity. In the following, we refer to this equation as the 'Evans model'. Another model is given by

$$f(v, E_\mathrm{b}, x_\mathrm{b}) = \frac{k_\mathrm{B}T}{x_\mathrm{b}}\ln(v) - \frac{k_\mathrm{B}T}{x_\mathrm{b}}\ln(x_\mathrm{b}\omega_0) + \frac{E_\mathrm{b}}{x_\mathrm{b}} = a\ln(v) + b. \tag{4}$$

This expression is a direct rearrangement of the Bell off-rate expression (equation (1)). The assumption made to include the velocity in the expression given by equation (1) is to equal the pulling velocity to the distance to break one bond divided by the lifetime of the bond, that is, set $v = \Delta x/\Delta t = kx_\mathrm{b}$. For further details regarding this model and its derivation we refer the reader to [23]. In the following sections, we refer to the model presented in equation (4) as the 'Bell model'.

Both expressions (equations (3) and (4)) predict a logarithmic dependence of the rupture force with respect to the pulling speed, where the parameter a denotes the slope in the $f - \ln(v)$ domain and the parameter b the intercept. Thus the bond energy landscape characteristics x_b and E_b can be determined by fitting one of these equations with the $f - \ln(v)$ plot obtained from experiments or simulations. By fitting the slope a, one can obtain the x_b value. Then, by considering that $\omega_0 = 10^{13}$ s^{-1} [11], one can fit the intercept b and obtain the E_b value.

2.2. Computational methods

Here we describe the details of our atomistic-based multi-scale simulation approach used to develop a mesoscale description of alpha-helical protein domains. The coarse-grained model for alpha-helical protein domains is fitted to results from full atomistic simulations in explicit water that have been reported earlier by some of the authors of this paper [23]. All simulations and calculations are performed at a temperature of $T = 300$ K.

2.2.1. Coarse-graining approach for alpha-helical protein filaments. Multi-scale simulation models for protein structures have become increasingly popular in recent years, and have enabled us to seek a direct link between experiment

and theoretical bottom-up descriptions of materials. Single-bead models are the most direct approach taken for studying macromolecules. The term 'single bead' derives from the idea of using single beads, that is, point masses, for describing each amino acid in a protein structure. The elastic network model (ENM) [24], Gaussian network model (GNM) [25] and Go models [26–28] are well-known examples that are based on such bead model approximations. These models have been extremely successful in explaining thermal fluctuations of proteins [29] and have also been implemented to model the unfolding problem to elucidate atomic-level details of deformation and rupture that complement experimental results [30–33]. A more recent direction is coupling of ENM models with a finite element-type framework for mechanistic studies of protein structures and assemblies [34]. Even coarser-level multi-scale modeling methods have been reported more recently, applied to model biomolecular systems at larger time and length scales. These models typically employ superatom descriptions that treat clusters of amino acids as 'beads'. In such models, the elasticity of the polypeptide chain is captured by simple harmonic or anharmonic (nonlinear) bond and angle terms. These methods are computationally quite efficient and capture shape-dependent mechanical phenomena in large biomolecular structures [35], and can also be applied to collagen fibrils in connective tissue [36] as well as mineralized composites such as nascent bone [37]. In this paper, we apply such a coarse-level description of alpha-helical protein domains.

To achieve the coarse-grained description of alpha-helical protein filaments with a precise control over the associated unfolding energy landscape we propose an alternative mesoscale description. In this model, the entire sequence of amino acids that makes up the alpha-helical structure is replaced by a collection of beads, as shown in figure 1(a). The structure of an alpha-helix consists of a series of turns (also referred to as convolutions), whereby each turn features 3.6 amino acid residues. This protein secondary structure is stabilized through the presence of H-bonds between the O atom of residue n and the N atom of residue $n+4$, and hence there are 3.6 H-bonds between turns, on average.

In our coarse-grained model, we aim at capturing the main structural and energetic features of an alpha-helical protein domain, as it switches between the folded and unfolded state, by explicitly considering the discrete makeup of the alpha-helical protein. Therefore one bead represents one turn of the alpha-helix and has the same mass. Moreover, the beads interact according to a bond potential and an angle potential. We choose a double-well bond potential in order to capture the existence of two equilibrium states for a convolution, corresponding to the folded and unfolded configuration (see figure 1(b) for the energy landscape and snapshots of atomistic geometries of the folded and unfolded states). The model does not involve explicit solvent nor a friction coefficient; rather, the effect of solvent on the breaking dynamics of alpha-helical convolutions is captured in the effective double-well potential. Through this formulation, the bond potential can describe the microscopic details of the rupture mechanism of the 3.6 H-bonds between each convolution under force, and the transition from the folded states to the unfolded states of convolutions through an energy barrier that separates the two states. Yet, the description is sufficiently 'coarse' so that it enables a significant computational speed-up and efficiency compared with the full atomistic description, suitable for the systematic analysis carried out here that requires a very large number of simulations.

The mathematical expression for the total energy of the system is given by

$$E = E_{\mathrm{T}} + E_{\mathrm{B}}, \tag{5}$$

where E_{T} is the total tensile energy and E_{B} is the total bending energy.

The total bending energy is given by the sum over all triples of beads:

$$E_{\mathrm{B}} = \sum_{\text{triplets}} \phi_{\mathrm{B}}(x). \tag{6}$$

The angle potential is given by

$$\phi_{\mathrm{B}}(\theta) = \tfrac{1}{2} K_{\mathrm{B}}(\theta - \theta_0)^2, \tag{7}$$

where K_{B} relates to the bending stiffness of the molecule EI, θ as the interbead angle (in triplets of atoms) and θ_0 as the equilibrium angle. The bending stiffness parameter K_{B} is given by

$$K_{\mathrm{B}} = \frac{3EI}{x_0}, \tag{8}$$

with x_0 denoting the equilibrium bead distance which corresponds to the equilibrium distance of one folded convolution and EI as the bending stiffness of the alpha-helix. In order to distinguish the bending stiffnesses of a folded alpha-helix and an unfolded alpha-helix (which entails a severe structural change of the protein), we use a stiffness parameter K_{B} that depends on the distance between neighboring beads x as

$$K_{\mathrm{B}} = K_{\mathrm{B,fold}}\left[(1-\alpha)\frac{1}{\pi}\left(-\arctan(100(x-x_{\mathrm{tr}})) + \frac{\pi}{2}\right) + \alpha\right] \tag{9}$$

with

$$\alpha = \frac{K_{\mathrm{B,unfold}}}{K_{\mathrm{B,fold}}} \tag{10}$$

defined as the ratio between the bending stiffness parameters of the unfolded state $K_{\mathrm{B,unfold}}$ and folded state $K_{\mathrm{B,fold}}$. The two bending stiffness parameters are given by

$$K_{\mathrm{B,fold(unfold)}} = \frac{3EI_{\mathrm{fold(unfold)}}}{x_{0(1)}} \tag{11}$$

with $x_{0(1)}$ denoting the equilibrium bead distance which corresponds to the equilibrium distance of one folded (unfolded) convolution and $EI_{\mathrm{fold(unfold)}}$ is the bending stiffness of a folded (unfolded) alpha-helix. This formulation based on a reaction coordinate (here the parameter x) is similar in spirit to reactive force field formulations based on chemical bond orders (such as Tersoff, ReaxFF and others, as reviewed, for example, in [38]).

The total tensile energy is given by the sum over all pairwise interactions:

$$E_{\mathrm{T}} = \sum_{\text{pairs}} \phi_{\mathrm{T}}(x). \tag{12}$$

The double-well bond potential $\phi_T(x)$ is given by (see also figure 1(b) for a schematic and an illustration of all parameters):

$$\phi_T(x) = \begin{cases} \dfrac{E_b}{x_b^4}(x - x_{tr})^2(x - x_{tr} - \sqrt{2}x_b) \\ \times (x - x_{tr} + \sqrt{2}x_b) + E_b, & x < x_{tr} \\ \dfrac{E_r}{x_r^4}(x - x_{tr})^2(x - x_{tr} - \sqrt{2}x_r) \\ \times (x - x_{tr} + \sqrt{2}x_r) + E_b, & x_{tr} \leqslant x. \end{cases} \quad (13)$$

The first equilibrium with reaction coordinate x_0 (first potential minimum) corresponds to the folded state of one turn of an alpha-helix under no force. The transition state (energy barrier E_b), with position x_{tr} (peak of the potential between two wells), corresponds to the breaking of the 3.6 H-bonds between two turns of the alpha-helix. After failure of these weak bonds, the turn unfolds to a second equilibrium state. This corresponds to the second potential minimum with a larger interbead distance, x_1. Under further loading, its backbone bonds begin to be stretched which leads to a second increase of the potential. The parameters x_b and E_b represent the distance and energy barrier required to unfold one convolution. Similarly, x_r and E_r correspond to the refolding process. It is noted that the development of a model that enables us to tune the key energy landscape parameters would not have been possible with other, existing mesoscale models (e.g. united atom method and others as discussed above).

2.2.2. Parameter identification: linking atomistic and mesoscale for the reference model. The parameters are determined through a fitting procedure against geometric properties of alpha-helices as well as full atomistic MD simulation results in an explicit solvent. We fit the energy barrier measured from MD simulation to the energy barrier in the mesoscale model formulation. The mass M of each bead corresponds to the approximate average mass of one turn or 3.6 residues, leading to 400 amu. The two parameters of the angle potential are introduced in equation (7). The value of the equilibrium angle θ_0 is 180°, based on the geometry of the alpha-helical structure. The bending stiffness parameter $K_{B,fold}$ is linked to molecular parameters as described by equation (8) and therefore can be determined from full atomistic simulations of bending studies of alpha-helical protein domains (we use the results reported in [39]). We find $K_{B,fold} = 21.589$ kcal mol^{-1} rad^{-2}, which corresponds to a persistence length of approximately 6.5 nm. For the unfolded configuration, we use the typical persistence length of a free polypeptide chain, which is about 0.4 nm [40]. Thus we find $K_{B,unfold} = 0.665$ kcal mol^{-1} rad^{-2}.

The parameters of the tensile double potential are introduced in equation (10) (see also the schematic shown in figure 1(b)). We find $x_0 = 5.4$ Å for the equilibrium bead distance of the folded state, which corresponds to the length of one folded convolution. The distance x_b between the folded state equilibrium and the transition state corresponds to the distance to break 3.6 H-bonds which leads to the unfolding of the convolution, and the parameter E_b is the corresponding energy barrier. It is noted that earlier full atomistic MD simulations [23] have confirmed this mechanism for stretching alpha-helical domains, that is, H-bonds in turns indeed break in clusters of 3–4. These two parameters are determined from fitting against full atomistic simulations of tensile loading experiments of alpha-helical domains of the vimentin filament protein family (with Protein Data Bank identification code 1gk6), for a range of pulling velocities below 0.3 m s^{-1} [23] (with further experimental validation [41]). We find $x_b = 1.2$ Å and $E_b = 11.1$ kcal mol^{-1} from these full atomistic studies by using the Bell model. The equilibrium bead distance of the unfolded state, x_1, is determined by fitting the mesoscopic force–strain curve against the atomistic simulation results in large deformation. In particular, this parameter describes at what strain levels an alpha-helix convolution is completely unfolded, and when further strain leads to significant stiffening due to stretching of the protein backbone. Thus we adjust x_1 so that the angular point between the plateau regime and the backbone stretching regime occurs at the same strain as in the atomistic simulation curve. We find $x_1 = 10.8$ Å, which corresponds to twice the length of the folded state. Then, since we know the values of the parameters x_0, x_b and x_1, we can determine the distance between the unfolded state and the transition state x_r. We obtain $x_r = 3.5$ and $x_b = 4.2$ Å. Finally, the energy barrier E_r to refold a 'broken' convolution must be smaller than E_b, since the folded state is the most favorable state for a convolution in equilibrium. Based on a suggestion put forth in [42], we determine $E_r = 0.6E_b$ (it is noted, however, that the resulting mechanical properties of the alpha-helix is insensitive to variations of choices of E_r, as long as $E_r < E_b$, as discussed below in section 3.3).

The complete set of parameters of the mesoscopic model and their physical meaning is summarized in table 1. In the following we refer to this set of parameters as the 'reference model' (which specifically represents the alpha-helical protein domain found in vimentin intermediate filaments from which the mesoscopic model parameters are derived [23]).

2.2.3. Mechanical characterization. We carry out tensile deformation of alpha-helical proteins. Loading is applied using the steered molecular dynamics (SMD) approach with a transducer spring constant $K = 10$ kcal mol^{-1} A^{-2} used for all simulations unless noted otherwise. The force versus extension data is recorded for subsequent analysis of mechanical properties.

3. Results and discussion

3.1. Validation of the double-well potential mesoscale model

The first step of the study is the validation of the single-molecule mesoscale results against full atomistic simulation results. Figure 2(a) presents the validation of the mesoscale model by direct comparison of the strength of alpha-helical protein domains with full atomistic results of the rupture mechanics of an alpha-helix protein domain [23], for varying

Table 1. Summary of all parameters of the mesoscale model. All parameters are derived from geometric analyses and atomistic simulations, corresponding to equations (5)–(10), as well as the discussion presented throughout section 2.2.2. This set of parameter values defines the 'reference model'.

Parameters	Numerical values
Equilibrium bead distance of the folded state x_0 (Å)	5.4
Distance between folded state and transition state x_b (Å)	1.2
Distance between unfolded state and transition state x_r (Å)	4.2
Energy barrier between folded state and transition state E_b (kcal mol^{-1})	11.1
Energy barrier between unfolded state and transition state E_r (kcal mol^{-1})	6.7
Equilibrium angle θ_0 (deg)	180
Bending stiffness parameter $K_{B,fold}$ (kcal mol^{-1} rad^{-2})	21.6
Bending stiffness parameter $K_{B,unfold}$ (kcal mol^{-1} rad^{-2})	0.665
Mass of each mesoscale bead M (amu)	400

pulling velocities (in this figure the rupture force is plotted over the log of the pulling velocity). The straight line in this plot corresponds to the predictions by the Bell model (equation (4)), discussed in section 2.2.

Figure 2(b) depicts the entire force–strain curve for a stretching experiment on the 14-bead mesoscopic model of an alpha-helix with a length of 70.2 Å at a temperature of 300 K and a pulling velocity of 0.1 m s^{-1}. The curve shows the three typical regimes observed in full atomistic simulations: an elasticity regime at low strain, a regime in which the convolutions of the alpha-helix unfold (referred to as the energy dissipation regime, with 13 peaks on the curve), and the subsequent regime of stretching of the backbone bonds. The mean alpha-helix unfolding force, which corresponds to the mean force of the energy dissipation regime, fits atomistic results closely. Both full atomistic and mesoscale model simulations predict a rupture force of approximately 350 pN. Here and throughout this paper, we refer to 'unfolding force' (also referred to as strength or rupture force) as the average unfolding force during the entire unfolding regime (mean unfolding force, MUF), unless noted otherwise. In the studies of the length dependence of the strength of alpha-helices, we refer to the 'unfolding force' as the force at the first peak (FFP), in order to accurately capture the size dependence of the strength. This is because we cannot use an averaged unfolding force in these studies as the effective length of the alpha-helix decreases continuously and thus does not provide an accurate measure of the actual strength dependence for a specific length. For most studies, we prefer the MUF method against measuring only the first peak, because it averages among all bond rupture events and avoids strong sensitivity to the resolution of the data. As a consequence we suggest using an energy barrier slightly higher than the one from the reference model (13 kcal mol^{-1} instead of 11.1 kcal mol^{-1}) in order to compensate the decrease of the measured unfolding force due to force relaxations after the peaks, if the mesoscale is used to fit closely to full atomistic simulation results.

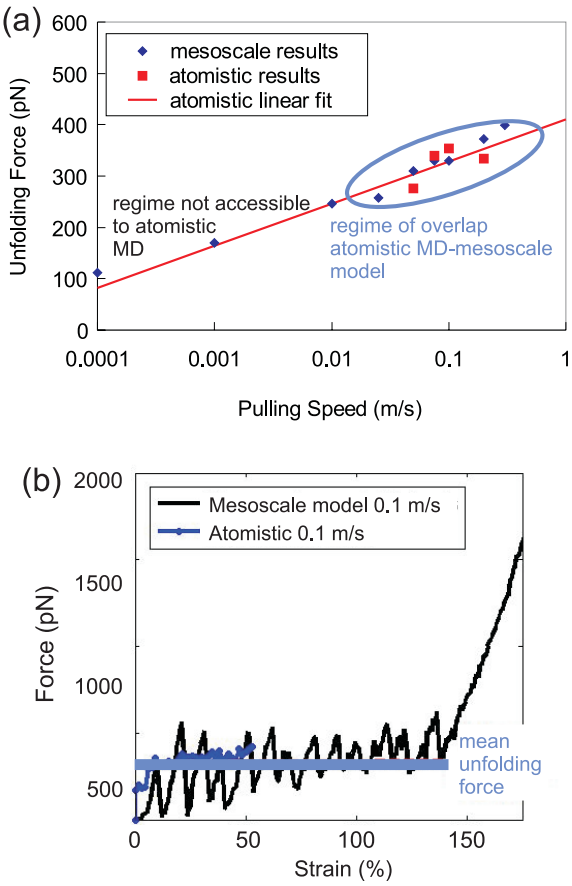

Figure 2. Subplot (a) validation of the double-well potential mesoscopic model of single molecules by comparison with full atomistic results of alpha-helical protein domains (full atomistic explicit water MD results taken from [23]). The plot shows the pulling velocity dependence of the unfolding force for both the mesoscopic and atomistic models. The mesoscale model is in very good agreement with the full atomistic simulations, validating the fitting of the mesoscopic bond potential. The figure further illustrates that the mesoscale model is capable of reaching much slower pulling velocities than those accessible to full atomistic simulation studies, here shown for the slowest pulling speed of 0.0001 m s^{-1}. Subplot (b) entire force–strain curve for a stretching experiment on the 14-bead mesoscopic model of an alpha-helix (with a length of 70.2 Å at a temperature of 300 K and a pulling velocity of 0.1 m s^{-1}). The mesoscale curve shows the three typical regimes observed in full atomistic simulations: an elasticity regime (before the first peak occurs), an energy dissipation regime which corresponds to the unfolding of the 13 bonds (corresponding to the 13 peaks on the curve) and the regime of stretching of the backbone bonds. The peaks in the mesoscale model, corresponding to the rupture forces, agree very well with the rupture force measured from full atomistic simulation. The dotted line approximates the increase of rupture force with increasing strain. In agreement with atomistic simulations reported earlier [39], the third regime sets in at approximately 135% strain.

It is noted that the results shown in figure 2(a) illustrate a key advantage of the coarse-grained model in reaching much longer timescales than what could be achieved in full atomistic simulations (the current limit in MD simulations is approximately 0.01 m s^{-1}, whereas we have easily reached a more than 100-fold increase in accessible timescales by using our mesoscale model). The model is capable of

reaching timescales of several microseconds and longer with a quantitative accuracy comparable with full atomistic MD simulations. Such relatively long simulations can be carried out within several days of computational time (on a single Intel Xeon CPU). In comparison, MD simulations of the dynamical behavior at fractions of microseconds can take weeks and months of computational time (even on a large parallelized simulation set-up). This reflects a considerable speed-up due to the coarse-graining approach, while the model is still capable of describing the small- and large-deformation force–strain response characteristics (e.g. softening at ≈10% strain and stiffening at ≈135% strain) as well as strength values quite accurately. Experimental results of stretching and breaking single alpha-helix domains [43, 44] (with a alpha-helix length of less than 100 Å) report forces between 140 and 240 pN during unfolding, close to the force level predictions at slow pulling speeds shown in figure 2(a). We emphasize that, despite the improved capacity of the coarse-grained model to reach longer timescales, the experimental studies and simulations are still performed at different pulling speeds and, as such, prevent us from carrying out a direct comparison. Achieving the direct comparison remains computationally intractable, despite the speed-up provided by the mesoscale model.

We now proceed with a systematic variation of the parameters in the mesoscale model, including: E_b, x_b, E_r and x_r. We vary these parameters individually, while fixing all others at their standard values as defined in the reference model (see table 1 and section 2.2.2).

3.2. Influence of E_b and x_b on the velocity dependence of the rupture force

Figure 3 shows plots of the unfolding force versus pulling speed obtained with the double-well potential mesoscale model for different values of x_b and E_b. The results illustrate the influence of x_b and E_b on the velocity dependence of the rupture force as predicted by equations (3) and (4). The range of pulling speeds is taken from 0.01 to 0.3 m s^{-1}, which correspond to the atomistic simulation regime where 3–4 H-bonds (that is, one whole convolution) break simultaneously [23].

Figure 3(a) shows the plots of the unfolding force versus pulling speed obtained for different x_b values, varying from 0.5 to 2 Å. The other potential parameters remain at their standard values as in the reference model. According to equations (3) and (4), changes in the parameter x_b lead to a change in slope and a change in the intercept with the y axis. This behavior is observed in figure 3(a). Figure 3(b) shows the plots of the unfolding force versus pulling speed obtained for different E_b values, varying from 5 to 20 kcal mol^{-1}. As before, the other potential parameters remain at their standard values as in the reference model. According to equation (2), changes in the parameter E_b lead to a change in the intersection value, while the slope is maintained. This behavior is observed in figure 3(b), albeit this prediction fails for relatively small values of the energy barrier ($\leqslant$7 kcal mol^{-1}). For the sake of clarity, we do not show the plots for high energy barriers

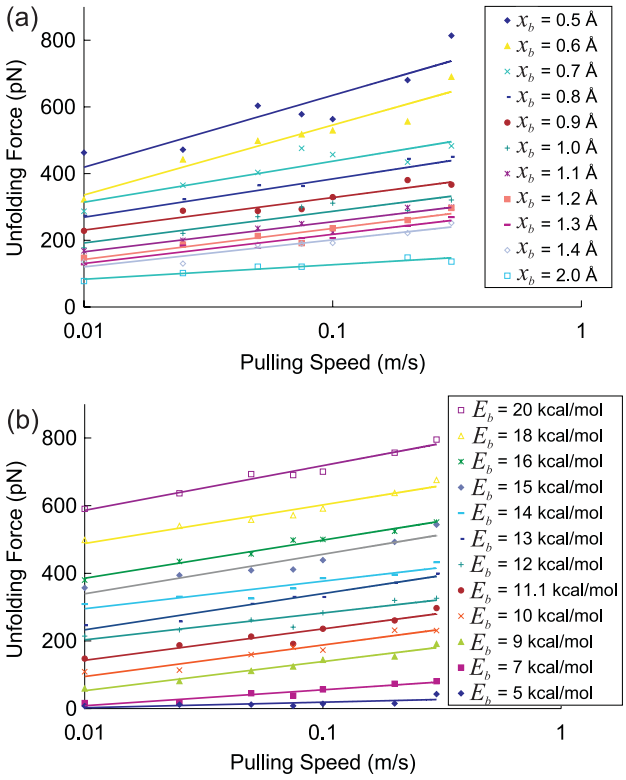

Figure 3. Plots of the unfolding force versus pulling speed obtained with the double-well potential mesoscale model for different values of x_b and E_b. The range of pulling speeds is taken from 0.01 to 0.3 m s^{-1} which corresponds to the atomistic simulation regime where 3.6 H-bonds (one whole convolution) break simultaneously [23]. Subplot (a) shows the plots of the unfolding force versus pulling speed obtained for different x_b values varying from 0.5 to 2 Å, while the other potential parameters remain at their standard values. Subplot (b) shows the plots of the unfolding force versus pulling speed obtained for different E_b values varying from 5 to 20 kcal mol^{-1} (that is, between 8 k_BT and 34 k_BT), while the other potential parameters remain at their standard values. The linear fits of these two plots show the well-known logarithmic velocity dependence of the unfolding force. The variations of slopes and intercepts between these linear fits show that both x_b and E_b have an influence on the logarithmic pulling velocity dependence.

(E_b values between 20 and 100 kcal mol^{-1}), which correspond to much larger rupture forces. We note that, because of the scale we use for the plot in figure 3(b), it appears as if the slopes for the E_b values of 12 and 14 kcal mol^{-1} are slightly different from the others. However (as will be discussed shortly in more detail), an in-depth study of this dependence is done in figure 4(d), which clearly shows that the slope remains relatively constant within the range 7–20 kcal mol^{-1} compared to the lower and higher ranges. The linear fits of these two plots show the well-known logarithmic velocity dependence of the unfolding force. Overall, the variations of slopes and intercepts between these linear fits show that E_b and x_b have an influence on the logarithmic velocity dependence. We will now measure their influence carefully and quantitatively compare them with the predictions of the Bell model.

Figure 4 depicts the plots of the slope a and the intercept b of the logarithmic velocity dependence of the unfolding force,

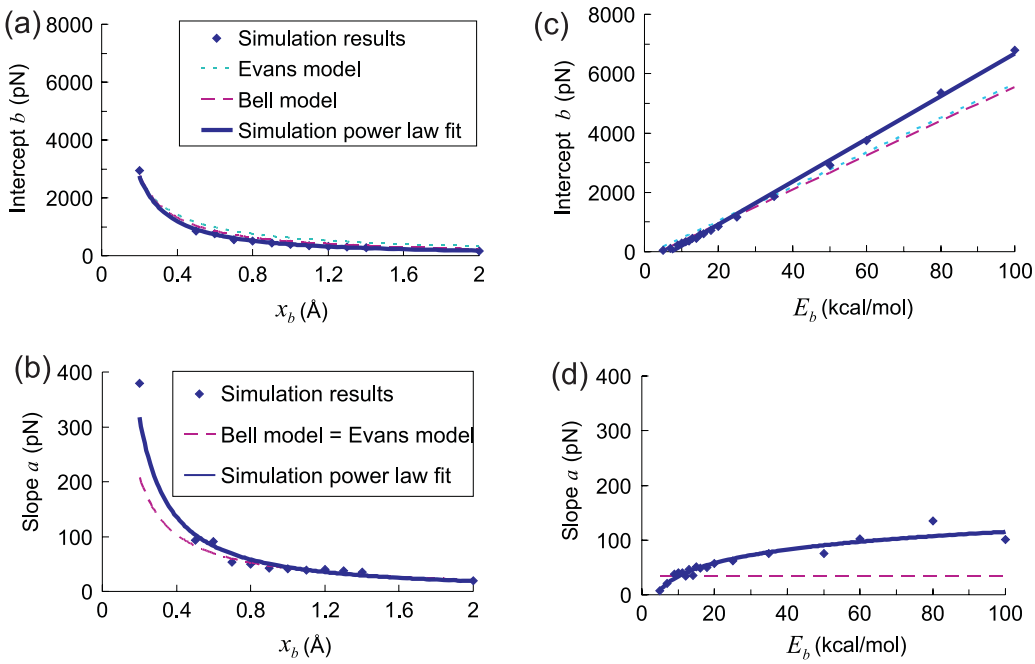

Figure 4. Plots of the slope a and the intercept b of the logarithmic pulling velocity dependence of the unfolding force versus x_b and E_b, obtained with the mesoscale model, the Bell model and the Evans model. Subplots (a) and (b) show the dependence of the intercept b and slope a on x_b, respectively. Subplots (c) and (d) show the dependence of the intercept b and slope a on E_b, respectively.

for variations of E_b and x_b. The graphs show results obtained with the mesoscale model (diamonds for the data and solid line for the fit which is either linear, power law or logarithmic), the Bell model (medium-dashed line) and the Evans model (short-dashed line). For the slope plots, the Evans model is not represented because it is identical to the Bell model. Figure 4(a) plots the intercept b as a function of x_b. The results reveal a close agreement between simulations, Bell and Evans models to describe the influence of x_b on the intercept b (note that, as before for this data, $E_b = 11.1$ kcal mol^{-1}). In the given range of x_b, for a fixed value of $E_b = 11.1$ kcal mol^{-1}, the maximum error of the Bell model to predict the measured x_b value from the intercept b is below 20% and is found at larger values of x_b. For the Evans model the maximum error is below 40%. Figure 4(b) plots the slope a as a function of x_b and shows a close agreement between simulations and the Bell model (note that, for this data, $E_b = 11.1$ kcal mol^{-1}). In the given range of x_b, the maximum error of the predicted value (according to the Bell model) against the measured x_b value (from the slope) is below 25% and tends to maximize as x_b reaches rather small values (corresponding to very shallow energy barriers). Figure 4(c) plots the intercept b as a function of E_b (note that, as before for this data, $x_b = 1.2$ Å fixed). The comparison reveals a close agreement between our coarse-grained simulations and both the Bell and Evans models. Within the E_b value ranges corresponding to weak bonds, the highest relative errors for the Bell and the Evans models are below 15% and 30%, respectively. For larger E_b values, the highest error is below 15% for both the Bell and Evans models.

Figure 4(d) plots the slope a as a function of E_b (note that for this data $x_b = 1.2$ Å fixed). The results show that the Bell model fails to describe the influence of E_b on the slope as observed in the simulations. Most importantly, the Bell model predicts no dependence of the slope a on E_b, whereas the simulation data clearly reveals a dependence of a on E_b. This dependence could be fitted empirically by the following logarithmic relation:

$$a = (35.48 \ln(E_b/(\text{kcal mol}^{-1})) - 48.31)\ \text{pN}. \qquad (14)$$

The disagreement between the Bell model prediction and the measurement from simulation is largest for very small values of E_b and for very large values of E_b. Within the range of energy barriers between 9 and 18 kcal mol^{-1} (that is, for a range of 15–30 k_BT), which corresponds to the range of weak bonds (e.g. H-bonds and clusters of H-bonds), the slope a does not vary significantly since the variation is smaller than 45%. If one uses the Bell model to determine x_b from the measurement of the slope a, then, given that x_b is inversely proportional to the slope a, the error of x_b is equal to the error on the slope a, which leads to an error on x_b smaller than 45% as well (thus comparable to the error observed in the previous plots). This result suggests that, in the range of weak bond energies, the Bell model is a suitable model to describe the energy landscape. We note that many applications of the Bell model in the literature are indeed within the range of weak bond energies [42, 45, 46]. For E_b values lower than 9 kcal mol^{-1}, we find a strong decrease of the slope when E_b decreases. A possible reason is that the critical velocity below which the Bell model predicts a negative force value enters the studied range of velocities for values below $E_b = 7$ kcal mol^{-1} ($E_b = 5$ kcal mol^{-1} for the Evans model). Thus, in this range of velocities and for low E_b values, we would expect a deviation from the Bell model prediction (for instance, through

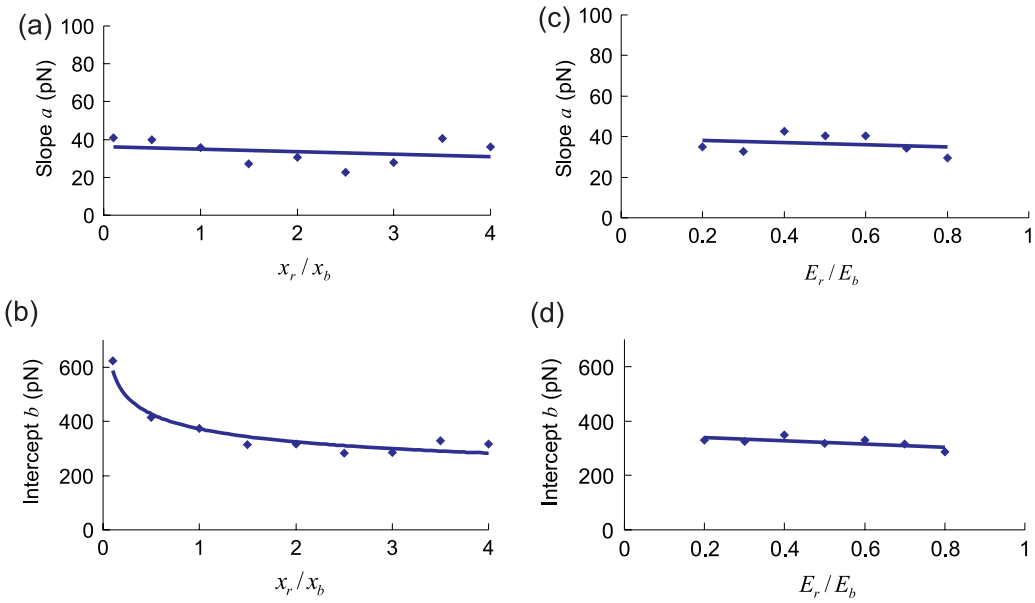

Figure 5. Plots of the slope a and the intercept b of the logarithmic pulling velocity dependence of the unfolding force versus x_r and E_r (properties associated with the second local unfolded equilibrium), obtained with the mesoscale model (diamonds for the data and solid line for the fit which is either linear, power law or logarithmic). Subplots (a) and (b) show the dependence of the slope a and intercept b on x_r, respectively. Subplots (c) and (d) show the dependence of the slope a and intercept b on E_r, respectively. The plots show that the parameters of the second potential well do not significantly influence the pulling velocity dependence of the unfolding force except for x_r values below x_b.

a smooth transition between the Bell linear fit to a very low rupture force in an asymptotic regime [41]).

In light of Evans' earlier derivation, this analysis may also suggest that the assumption made in Kramers' theory that the energy barrier is much greater than the thermal energy breaks down for E_b below 9 kcal mol^{-1} in this studied range of velocities. The disagreement at very large values of E_b could be explained by the fact that, for values larger than 20 kcal mol^{-1}, the Bell model for 'weak bonds' is no longer applicable and other protein strength models suitable for strong bonds must be used.

3.3. Influence of E_r and x_r on the velocity dependence of the rupture force

Figure 5 shows the plots of the slope a and the intercept b for variations of potential parameters that relate to the second local unfolded equilibrium (see figure 1(b) for a schematic and corresponding atomistic geometry as well as a description of parameters used to characterize this part of the energy landscape). These are parameters x_r and E_r (corresponding to the distance and energy barrier for refolding of a convolution). It is noted that these two parameters are not included in Bell's or Evans' models, respectively. The parameter x_r varies between $0.1x_b$ and $4x_b$. The parameter E_r varies within a range between 0 and E_b, so that the first equilibrium remains the most probable state (that is, a global minimum).

Figure 5(a) plots the influence of variations of x_r on the slope a and shows that x_r does not have a significant influence on the slope. This is evident from the fact that the linear fit to the simulation data is almost horizontal. Figure 5(b) plots the influence of variations of x_r on the intercept b. The results suggest that x_r does not have a significant influence on the intercept, except for values lower than x_b. For instance, the intercept is approximately 100% higher for $x_r = 0.1x_b$ than for $x_r = 2x_b$ (the plot shows a fit of a power law to the data). The results suggest that the Bell model may not be relevant to describe an energy landscape where the width of the second potential equilibrium is shorter than the width of the first potential equilibrium. Figures 5(c) and (d) show that E_r does not have any significant influence on either the slope a or the intercept b. In conclusion, the parameters of the second unfolded equilibrium do not affect the logarithmic velocity dependence of the unfolding force, provided that x_r is higher than x_b. Conversely, if x_r is significantly lower than x_b, the Bell model is not relevant.

3.4. Influence of E_r and x_r on the rupture force

We proceed with an analysis of the influence of x_r and E_r on the rupture force at a given pulling speed of 0.1 m s^{-1} (as before, since these parameters are not included in the Bell model there is no influence of these parameters predicted from the Bell model). We focus on two aspects, first the value of the rupture force at the first peak in the force–strain plot (see figure 2), and second on the average unfolding force during the entire unfolding regime (as the strain increases until stiffening occurs due to stretching of the protein backbone). Figures 6(a) and (b) show the influence of x_r and E_r on the first peak in the force–strain plot. The data clearly shows that these parameters do not influence the first force peak. Figure 7 depicts the influences of x_r and E_r on the average unfolding force. Figure 7(a) shows that E_r does not significantly influence the unfolding force. In contrast, figure 7(b) shows that x_r influences the unfolding force for x_r values below x_b.

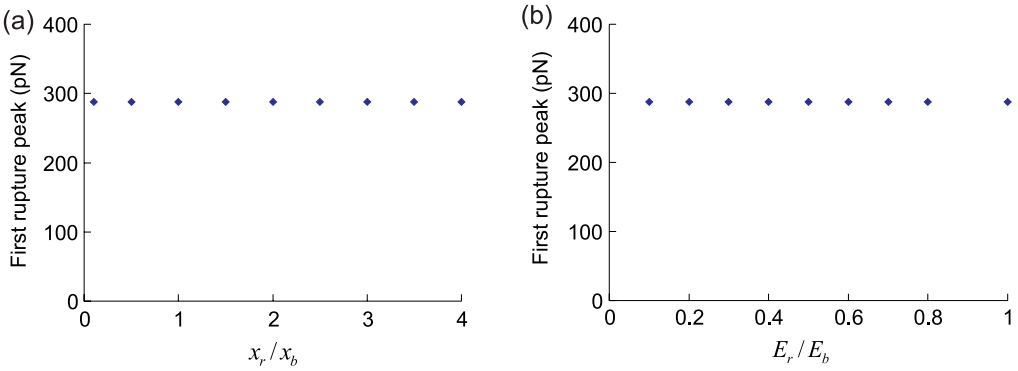

Figure 6. Influence of x_r and E_r on the first rupture peak. Subplot (a) and (b) show that the parameters of the second equilibrium do not influence the first bond rupture event (first force peak).

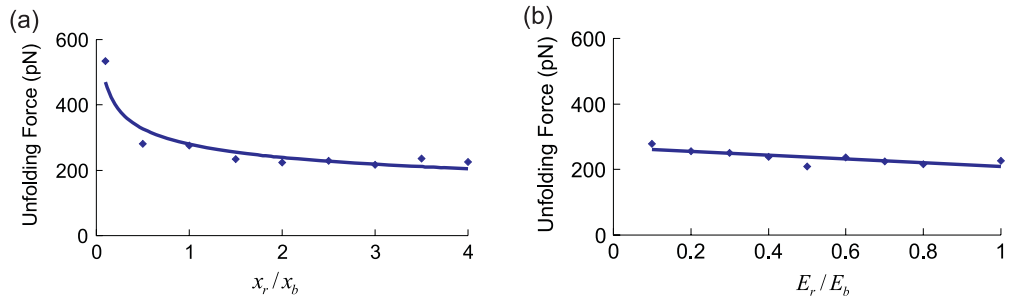

Figure 7. Influence of x_r and E_r on the average unfolding force. Subplot (a) shows that E_r does not significantly influence the unfolding force. Subplot (b) shows that x_r influences the unfolding force more strongly for x_r values below x_b.

3.5. Influence of E_b and x_b on the rupture force

We proceed with an analysis of the influence of E_b and x_b on the rupture force at a pulling speed of 0.1 m s^{-1}. Figures 8(a) and (b) show the dependence of the unfolding force on E_b and x_b. As predicted by both Bell's and Evans' models, we observe a linear force dependence on E_b and a relative linear force dependence on $1/x_b$. The critical force expression f_{crit} given by equation (2) qualitatively agrees with the scaling behavior found in the simulations. We note that the critical force typically yields an overestimation of the rupture force compared with simulations and the other models, except for very large energy barriers. This is due to the fact that this expression does not take into account that the probability of the bond to break for a force below the critical force is not zero.

3.6. Strength dependence on length of alpha-helical protein filament

We report an application of the mesoscale model to a study of the effect of the length of an individual alpha-helix protein on its strength properties, with protein lengths ranging from $L = 10.8$ to 4001.4 Å (that is, from 7 residues to 2668 residues) at a pulling velocity of 0.1 m s^{-1}. Figure 9 shows the strength obtained from our model for varying the length of alpha-helices. Here the strength of alpha-helices is measured as the force value of the first rupture peak (force at first peak, FFP; see figure 2(b)). We observe two distinct regimes. For short alpha-helices, the rupture strength decreases as the length increases. For long alpha-helices, the rupture strength reaches an asymptotic value of approximately 180 pN (indicated with a dashed line in figure 9). The transition between these two strength regimes occurs between 200 and 550 Å. We note that the simulation of the system with such long lengths would not have been possible with a full atomistic simulation (the full atomistic simulation of a 70.2 Å protein structure took several weeks of computational time). For short alpha-helices, the first force peak of the two convolution system (the smallest one considered) is approximately 600 pN and is thus almost four times as high than the 61-turn structure (that is, 329.4 Å, so around the transition of regimes), where the strength approaches 150 pN. We have used the short alpha-helix simulation results to fit an empirical equation of the form

$$f(L) = a \ln(L/L_0) + b, \tag{15}$$

where we find $a = -117.97$ pN, $b = 864.6$ pN and $L_0 = 1$ Å. In summary, our model predicts that the strength of alpha-helices decreases as the length increases and then reaches an asymptotic regime for lengths larger than 200–550 Å. The decreasing behavior can be explained based on the fact that longer alpha-helices contain more serial bonds, each of which can break with the same probability. Since failure of one turn is sufficient to initiate failure of the entire system, we expect that longer molecules are weaker, as observed in our simulations and in [47]. This finding is also supported by the plot of the prevalence of the length of alpha-helices that shows that short alpha-helices are more prevalent in naturally found biological

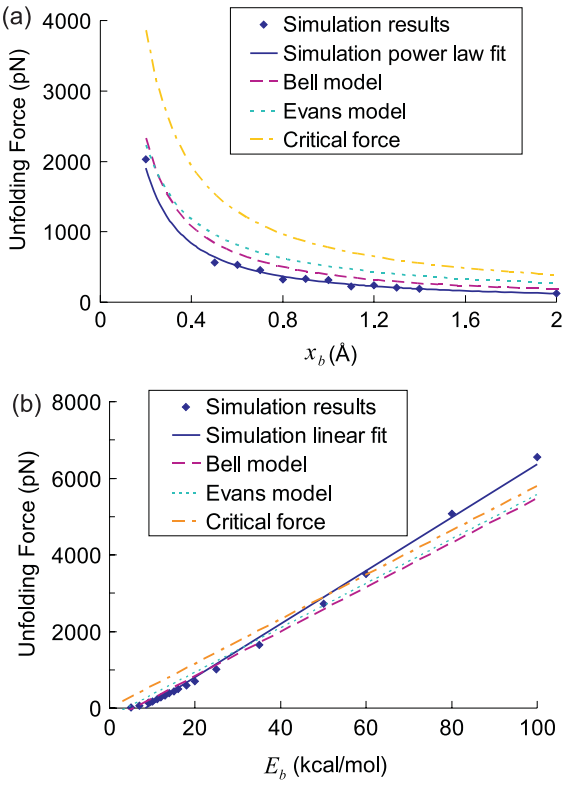

Figure 8. Dependence of the unfolding force on x_b and E_b obtained with the mesoscale model, the Bell model, the Evans model and the critical force expression at a pulling speed of 0.1 m s^{-1}. As predicted by Bell's and Evans' models as well as the simple critical force expression, we observe a linear force dependence on E_b and a relative linear force dependence on $1/x_b$.

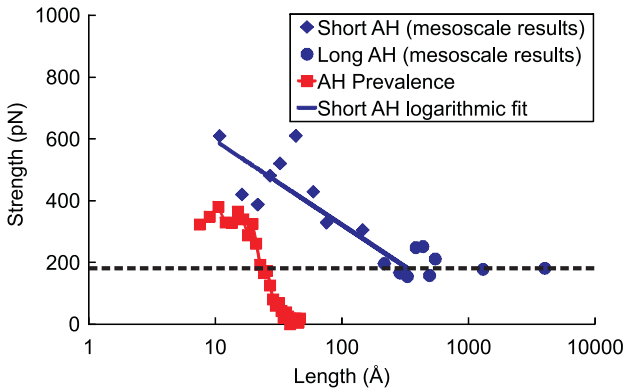

Figure 9. Size effects, strength and prevalence of length of alpha-helices. This plot shows the strength properties of alpha-helices with different lengths, ranging from $L = 10$ to 4000 Å (from 1 to 400 nm), at 0.1 m s^{-1} pulling velocity. The results illustrate two different regimes. For short alpha-helices, the strength decreases as the length of the alpha-helix increases. The continuous line shows a logarithmic fit to the data obtained from mesoscale simulations (see equation (12) for the equation and numerical fitting parameters). For longer alpha-helices, the strength reaches an asymptotic regime of about 180 pN. The transition between these two regimes is between 200 and 550 Å. The plot of the prevalence over the alpha-helix length (distribution of alpha-helix single-strand lengths found in biological proteins) illustrates that shorter alpha-helices are more prevalent [48]. This plot shows a correlation between the mesoscale results and the prevalence that is the most prevalent alpha-helices correspond to the strongest ones.

proteins [48]. However, this analysis does not explain why we observe an asymptotic strength for long lengths and why the transition occurs between 200 and 550 Å. An explanation could be that the deformation of longer systems is smoother than that of shorter systems because they have much more interactions, and thus the peak of failure is smoothed out more. As a consequence, the more bonds the system has, the closer to the average (the asymptote) the force peak is. This hypothesis seems to be supported by the observation that the more bonds the system has (or, equivalently, a larger initial length), the less significant the fluctuations are during the unfolding regime (that is, the flatter is the force–strain curve during unfolding).

3.7. Multi-timescale analysis of strength

Our coarse-grained model enables us to carry out pulling simulations over very large ranges of timescales. Figure 10 depicts the rupture force versus pulling velocity of the mesoscale model of an individual alpha-helix over seven orders of magnitude of pulling speeds, for a fixed value of $E_b = 7.04$ kcal mol^{-1} and an SMD spring constant of $K = 0.1$ kcal mol^{-1} A^{-2} (parameters are chosen to shift the critical velocity to larger values in order to enable the investigation of the breakdown of the Bell model within the range of accessible pulling speeds). We utilize two definitions of the rupture force for analysis of the data: first, the force at first peak (FFP) and, second, the mean unfolding force (MUF), and compare the results with the Bell model predictions.

We observe two regimes for the force analysis based on FFP and three regimes for the force analysis based on MUF. The Bell model prediction line is consistent with both the FFP results and the MUF for pulling velocities above the critical velocity (that is, pulling velocities below which the Bell model rupture force is negative). This applies to velocities below 0.01 m s^{-1} for the parameters selected here. However, for very high pulling velocities (above 100 m s^{-1}), the Bell model does not fit well any longer since the MUF curve increases exponentially. For pulling velocities below the critical velocity, both MUF and FFP curves converge to an asymptotic positive strength value between 12 and 30 pN (the estimate based on MUF is ≈30 pN and the estimate based on FFP is ≈12 pN). This result agrees qualitatively with the theoretical multi-timescale model for alpha-helix protein domains developed in [41] and suggests that indeed there exists an asymptotic positive strength value and that the Bell model breaks down at small pulling velocities. Due to the choice of a relatively small energy barrier, the strength values found in this regime are smaller than those typically observed in experimental studies (i.e. the energy landscape parameters used here are not an accurate representation of the properties of alpha-helices).

3.8. Influence of transducer stiffness

Finally, we study the rupture force of alpha-helices as a function of the SMD spring constant, representing a variation of the transducer stiffness in a corresponding AFM experiment.

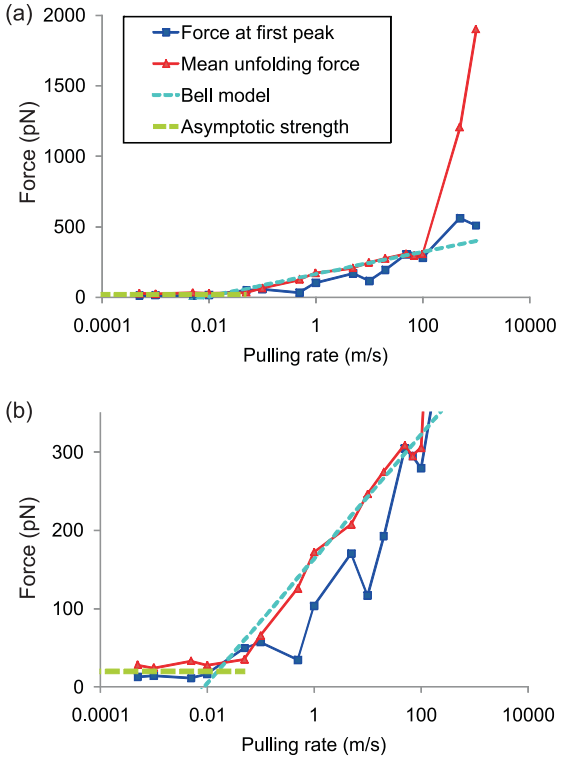

Figure 10. Rupture force versus pulling velocity of an individual alpha-helix over seven orders of magnitude. We plot the two rupture force definitions (force at first peak and mean unfolding force) and compare the curves with the Bell model predictions. Panel (a) is a zoom of panel (b) in the regime of small pulling velocities. We observe two regimes for the force at first peak (FFP) and three regimes for the mean unfolding force (MUF). The Bell model prediction line is consistent with the FFP results and the MUF for pulling velocities above the critical velocity (that is, the pulling velocity above which the Bell model rupture force is positive), that is, for velocities above 0.01 m s^{-1}. For very high pulling velocities (above 100 m s^{-1}), the Bell model does not fit any more and the MUF curve increases exponentially. For pulling velocities below the critical velocity, both MUF and FFP curves converge to a respective asymptotic positive strength (the estimate based on MUF is ≈30 pN and the estimate based on FFP is ≈12 pN). Our results agree qualitatively with the theoretical multi-timescale model for alpha-helices reported in [41].

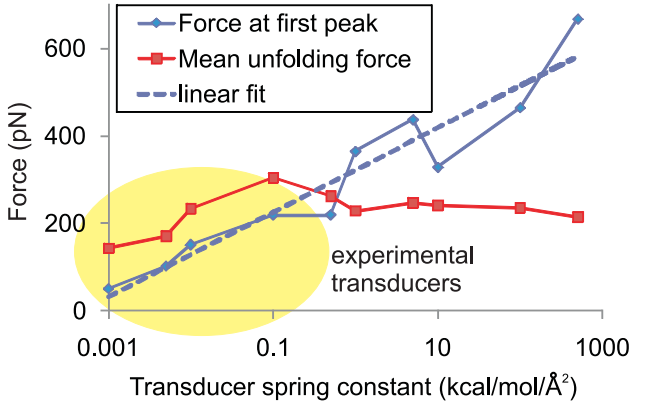

Figure 11. Rupture force versus transducer stiffness of the mesoscale model of an individual alpha-helix over six orders of magnitude, pulling at a velocity of 0.1 m s^{-1}. We plot the two rupture force definitions (force at first peak, FFP, and mean unfolding force, MFU). According to the FFP fit, the rupture force increases logarithmically with the transducer spring constant whereas the MFU curve shows a very low spring constant dependence of the rupture force. The circled area indicates the range of transducer spring constants typically used in experiments.

Figure 11 depicts the rupture force versus transducer stiffness of the mesoscale model of an alpha-helix over six orders of magnitudes, at a pulling velocity of 0.1 m s^{-1}. We plot the two rupture force definitions (force at first peak, FFP, and mean unfolding force, MFU) as defined earlier. According to the FFP fit, the rupture force increases logarithmically with the transducer spring constant whereas the MFU curve shows a very low spring constant dependence of the rupture force. The circled area underlines the range of transducer spring constants accessible in typical AFM experiments.

4. Conclusions

We have carried out a systematic analysis of the effects of energy landscape parameters on the strength properties of alpha-helix protein domains and compared the results with commonly used simple protein strength models. The most important contributions of this paper are:

- We have developed and calibrated, by fitting against full atomistic MD results, a mesoscale model of alpha-helical protein domains, representing one convolution as a pair of mesoscale bead particles (see the structure shown in figure 1). This represents a reference system, which agrees well with the predictions from Bell's or Evans' models (figure 2).
- This reference system provided us with the starting point for a systematic variation of all relevant parameters, to identify how strength properties depend on the details of the energy landscape and under which conditions the Bell model fails. By systematically varying the energy landscape parameters, we have shown that the Bell model is a reasonable approximation to predict the mechanical strength properties as long as the model parameters are in a certain range: E_b between 9 and 18 kcal mol^{-1} (weak bond energy barrier) and an x_r value above the x_b value (see, e.g., figure 4). Conversely, we find that the Bell model breaks down when the model parameters are not in this range. We have also analyzed the dependence of the unfolding force on E_b and x_b (figure 8) and confirmed the predictions from the Bell model.
- We have shown that the properties of the second, unfolded state of a alpha-helical turn do not strongly influence the strength properties, provided that x_r, the distance between the unfolded state and the transition state, is higher than x_b, the distance between the folded state and the transition state (see, e.g., figure 5). We have also shown that the energetic properties of the second unfolded state do not influence the rupture strength.
- We have applied our mesoscale model to predict the strength of individual alpha-helices with different lengths, showing two different regimes, to elucidate fundamental

scaling laws of the strength of alpha-helices. For short alpha-helices, we have shown a weakening effect as the molecule becomes longer. This could be explained by the larger number of serial coupled bonds, where failure of one bond is sufficient to initiate failure of the entire system (figure 9). For long alpha-helices, we have shown the existence of a transition from the strength decrease regime to a strength asymptotic regime. An explanation could be that the deformation of longer systems is smoother than shorter systems because they have much more interactions and thus the peak of failure is averaged. This might be one of the reasons why short single alpha-helices are more prevalent in nature than longer ones, as they provide an increased strength against failure.

- We have shown that, under extremely small pulling velocities, the rupture strength of alpha-helical proteins approach an asymptotic value (figure 10), where the strength does not depend on the pulling speed any more. This observation could be an important issue in reaching a direct link between experiment and simulation, since the results suggest that there will be no or only a marginal dependence of the unfolding force on the pulling speed below a critical pulling speed. Therefore, even though the molecular simulations are still carried out at larger pulling speeds than experiments, the predictions derived from them (such as those shown in figure 2(a)) could perhaps be utilized for a meaningful comparison of simulation and experiment (as discussed in section 3.1).
- We have also reported an analysis of the effect of varying transducer stiffnesses on the strength properties. Our results suggest that stiffer transducers generally lead to increased strength properties (figure 11).

Potential applications of the coarse-grained alpha-helix protein domain model presented here could be further studies of length scale effects on alpha-helix strength, elasticity and effects of hierarchical arrangements of alpha-helical-based protein domains. Further studies could focus on larger variations of timescales (e.g. to extend to cover very slow, experimental pulling speeds) and the development of similar formulations for coiled-coil proteins, larger-scale protein folds with tertiary structures or larger-level protein materials in general.

The results reported here confirm previous predictions that the Bell model is relevant for low pulling speeds, at deformation speeds just above the critical pulling speed [12–18]. It is also noted that our findings directly confirm suggestions that, at very low pulling speeds, rebinding processes are observed, a prediction already made in [14]. Future studies could be focused on an extensive comparison of the simulation results based on our model with various advanced single-molecule strength models reported in recent years [12–18]. The systematic molecular dynamics based approach as utilized here could provide a useful tool to pursue such studies.

Our analysis also provides a thorough study of the influence of key energy landscape parameters on the strength properties. Energy landscape parameters (such as the energy barrier E_b), in turn, are directly influenced by the solvent and other environmental conditions. As such, our study is important in understanding how different solvents could change the mechanical properties of alpha-helices, a problem of great relevance for biological applications and mechanomutability. Related to the previous point, changes in E_b could also be achieved by engineering the amino acid sequence of proteins. An application with great relevance could be *de novo* protein materials that are designed and created based on genetic engineering approaches. In such applications, our work could be useful to better understand which sequence would provide the greatest mechanical stability and what force levels such a system could withstand, or which sequences could be used to create hierarchical patterns of alpha-helical domains (e.g. with alternating strong–weak domains to create nanocomposite structures).

The coarse-graining approach used here works well for one-dimensional filamentous protein structures and could be extended easily to model parallel arrays of alpha-helices or larger-scale protein filament assemblies. However, for more complex protein structures other coarse-graining techniques such as elastic network models or shape-based coarse-graining could be more appropriate techniques [49–52] together with the techniques discussed in section 2.2.1 [24–28]. Moreover, individual, isolated alpha-helices are rarely found in biology. Therefore, further studies could be focused on applying the double-well potential to other structural proteins and filaments, or assemblies of alpha-helices into larger-scale assemblies. The conclusions put forth here should be generically valid for a broader class of structural proteins. This is because the energy landscape structure as presented in figures 1(b) and (c) is not unique to alpha-helices. Rather, other protein filaments featuring serially arranged domains that unfold under strain can be described based on the modeling framework proposed here, with appropriate parameterization of the model parameters (as listed in table 1). In light of this, the choice of alpha-helices in the studies reported can be considered as a model system that helped us to elucidate the generic behavior of a general class of protein filaments using a computationally efficient mesoscopic description. Specifically, the insight into the length dependence of strength properties (figure 9), the effects of pulling speed and the asymptotic behavior at vanishing rates (figure 10) and transducer stiffness (figure 11) might be useful for the interpretation of experimental results and the comparison with theoretical and numerical models for a variety of other protein filaments.

Acknowledgments

This research was supported by the Air Force Office of Scientific Research (AFOSR). Additional support from the National Science Foundation (NSF) and the Office of Naval Research (ONR) is acknowledged. JB acknowledges partial support from the Schoettler Fellowship Program at Massachusetts Institute of Technology. JH and DDJ acknowledge support from the Massachusetts Institute of Technology Undergraduate Research Opportunity Program (UROP).

Author contributions and conflict of interest. All authors of this paper have contributed equally to this work. The authors declare no conflict of interest of any sort.

References

[1] Alberts B *et al* 2002 *Molecular Biology of the Cell* (New York: Taylor and Francis)
[2] Gruber M and Lupas A N 2003 Historical review: another 50th anniversary—new periodicities in coiled coils *Trends Biochem. Sci.* **28** 679–85
[3] Moir R D and Spann T P 2001 The structure and function of nuclear lamins: implications for disease *Cell. Mol. Life Sci.* **58** 1748–57
[4] Wilson K L, Zastrow M S and Lee K K 2001 Lamins and disease: insights into nuclear infrastructure *Cell* **104** 647–50
[5] Buehler M J and Yung Y C 2009 Deformation and failure of protein materials in physiologically extreme conditions and disease *Nat. Mater.* **8** 175–88
[6] Makarov D E, Hansma P K and Metiu H 2001 Kinetic Monte Carlo simulation of titin unfolding *J. Chem. Phys.* **114** 9663–73
[7] Rathore N, Yan Q L and de Pablo J J 2004 Molecular simulation of the reversible mechanical unfolding of proteins *J. Chem. Phys.* **120** 5781–8
[8] Bryson J W *et al* 1995 Protein design—a hierarchical approach *Science* **270** 935–41
[9] Kirshenbaum K, Zuckermann R N and Dill K A 1999 Designing polymers that mimic biomolecules *Curr. Opin. Struct. Biol.* **9** 530–5
[10] Ackbarow T *et al* 2009 Alpha-helical protein networks are self protective and flaw tolerant *PLoS ONE* **4** e6015
[11] Bell G I 1978 Models for specific adhesion of cells to cells *Science* **200** 618–27
[12] Dudko O K, Hummer G and Szabo A 2006 Intrinsic rates and activation free energies from single-molecule pulling experiments *Phys. Rev. Lett.* **96** 108101
[13] Dudko O K 2009 Single-molecule mechanics: new insights from the escape-over-a-barrier problem *Proc. Natl Acad. Sci. USA* **106** 8795–6
[14] Dudko O K *et al* 2003 Beyond the conventional description of dynamic force spectroscopy of adhesion bonds *Proc. Natl Acad. Sci. USA* **100** 11378–81
[15] Hummer G and Szabo A 2003 Kinetics from nonequilibrium single-molecule pulling experiments *Biophys. J.* **85** 5–15
[16] Dudko O K *et al* 2002 Dynamic force spectroscopy: a Fokker–Planck approach *Chem. Phys. Lett.* **352** 499–504
[17] Lin H J *et al* 2007 Bell's expression and the generalized garg form for forced dissociation of a biomolecular complex *Phys. Rev. Lett.* **98** 088304
[18] Friddle R W 2008 Unified model of dynamic forced barrier crossing in single molecules *Phys. Rev. Lett.* **100** 138302
[19] Hanggi P, Talkner P and Borkovec M 1990 Reaction-rate theory: fifty years after Kramers *Rev. Mod. Phys.* **62** 251–341
[20] Zhurkov S N 1965 Kinetic concept of the strength of solids *Int. J. Fract. Mech.* **1** 311–23
[21] Bell G I 1978 Models for the specific adhesion of cells to cells *Science* **200** 618–27
[22] Evans E and Ritchie K 1997 Dynamic strength of molecular adhesion bonds *Biophys. J.* **72** 1541–55
[23] Ackbarow T *et al* 2007 Hierarchies, multiple energy barriers and robustness govern the fracture mechanics of alpha-helical and beta-sheet protein domains *Proc. Natl Acad. Sci. USA* **104** 16410–5
[24] Tirion M 1996 Large amplitude elastic motions in proteins from a single-parameter, atomic analysis *Phys. Rev. Lett.* **77** 1905–8
[25] Haliloglu T, Bahar I and Erman B 1997 Gaussian dynamics of folded proteins *Phys. Rev. Lett.* **79** 3090–3
[26] Hayward S and Go N 1995 Collective variable description of native protein dynamics *Ann. Rev. Phys. Chem.* **46** 223–50
[27] Sulkowska J I and Cieplak M 2008 Selection of optimal variants of $\tilde{G}$-like models of proteins through studies of stretching *Biophys. J.* **95** 3174–91
[28] Cieplak M, Pastore A and Hoang T X 2005 Mechanical properties of the domains of titin in a Go-like model *J. Chem. Phys.* **122** 54906
[29] Tozzini V 2005 Coarse-grained models for proteins *Curr. Opin. Struct. Biol.* **15** 144–50
[30] West D K *et al* 2006 Mechanical resistance of proteins explained using simple molecular models *Biophys. J.* **90** 287–97
[31] Dietz H and Rief M 2008 Elastic bond network model for protein unfolding mechanics *Phys. Rev. Lett.* **100** 098101
[32] Sulkowska J I and Cieplak M 2007 Mechanical stretching of proteins—a theoretical survey of the Protein Data Bank *J. Phys.: Condens. Matter* **19** 283201
[33] Noid W G *et al* 2008 The multiscale coarse-graining method. I. A rigorous bridge between atomistic and coarse-grained models *J. Chem. Phys.* **128** 244114
[34] Bathe M 2008 A finite element framework for computation of protein normal modes and mechanical response *Proteins-Struct. Function Bioinform.* **70** 1595–609
[35] Arkhipov A, Freddolino P L, Imada K, Namba K and Schulten K 2006 Coarse-grained molecular dynamics simulations of a rotating bacterial flagellum *Biophys. J.* **91** 4589–97
[36] Buehler M J 2006 Nature designs tough collagen: explaining the nanostructure of collagen fibrils *Proc. Natl Acad. Sci. USA* **103** 12285–90
[37] Buehler M 2007 Molecular nanomechanics of nascent bone: fibrillar toughening by mineralization *Nanotechnology* **18** 295102
[38] Buehler M J 2007 Hierarchical chemo-nanomechanics of stretching protein molecules: entropic elasticity, protein unfolding and molecular fracture *J. Mech. Mater. Struct.* **2** 1019–57
[39] Ackbarow T and Buehler M J 2007 Superelasticit, energy dissipation and strain hardening of vimentin coiled-coil intermediate filaments: Atomistic and continuum studies *J. Mater. Sci.* **42** 8771–87
[40] Keten S and Buehler M J 2008 Asymptotic strength limit of hydrogen bond assemblies in proteins at vanishing pulling rates *Phys. Rev. Lett.* **100** 198301
[41] Ackbarow T, Keten S and Buehler M J 2009 Multi-time scale strength model of alpha-helical protein domains *J. Phys.: Condens. Matter* **21** 035111
[42] Karcher H *et al* 2006 A coarse-grained model for force-induced protein deformation and kinetics *Biophys. J.* **90** 2686–97
[43] Lantz M A *et al* 1999 Stretching the alpha-helix: a direct measure of the hydrogen-bond energy of a single-peptide molecule *Chem. Phys. Lett.* **315** 61–8
[44] Kageshima M *et al* 2001 Insight into conformational changes of a single alpha-helix peptide molecule through stiffness measurements *Chem. Phys. Lett.* **343** 77–82
[45] Husson J and Pincet F 2008 Analyzing single-bond experiments: influence of the shape of the energy landscape and universal law between the width, depth, and force spectrum of the bond *Phys. Rev.* E **77** 026108
[46] Tshiprut Z, Klafter Y and Urbakh M 2008 Single-molecule pulling experiments: when the stiffness of the pulling device matters *Biophys. J.* **95** L42–4
[47] Qi H J, Ortiz C and Boyce M C 2006 Mechanics of biomacromolecular networks containing folded domains *J. Eng. Mater. Technol.-Trans. ASME* **128** 509–18
[48] Penel S, Morrison R G, Mortishire-Smith R J and Doig A J 1999 Periodicity in alpha-helix lengths and C-capping preferences *J. Mol. Biol.* **293** 1211–9
[49] Ayton G S, Noid W G and Voth G A 2007 Multiscale modeling of biomolecular systems: in serial and in parallel *Curr. Opin. Struct. Biol.* **17** 192–8
[50] Zacharias M 2008 Combining elastic network analysis and molecular dynamics simulations by Hamiltonian replica exchange *J. Chem. Theory Comput.* **4** 477–87
[51] Dietz H and Rief M 2008 Elastic bond network model for protein unfolding mechanics *Phys. Rev. Lett.* **100** 4
[52] Brown A E X *et al* 2007 Forced unfolding of coiled-coils in fibrinogen by single-molecule AFM *Biophys. J.* **92** L39–41