Rigidity loss of protein macromolecule induced by force–Effective field theory

ARTICLE *in* PROTEINS STRUCTURE FUNCTION AND BIOINFORMATICS · JUNE 2014

Impact Factor: 2.63 · DOI: 10.1002/prot.24471 · Source: PubMed

2 AUTHORS, INCLUDING:

Evgeny Meilikhov
Kurchatov Institute
96 PUBLICATIONS **413** CITATIONS

SEE PROFILE

PROTEINS
STRUCTURE ■ FUNCTION ■ BIOINFORMATICS

Rigidity loss of protein macromolecule induced by force—Effective field theory

E.Z. Meilikhov[1,2]* and R.M. Farzetdinova[1]

[1] Kurchatov Institute, 123182 Moscow, Russia

[2] Department of General Physics, Moscow Institute of Physics and Technology (State University), 9, Institutsky lane, Dolgoprudny, 141707, Russia

ABSTRACT

In the framework of the effective field theory for the order parameter, which characterizes the degree of deviating the protein globule structure from its native state, the phase transition of the protein macromolecule from the elastic state into the plastic one under its mechanical stretching is considered. Elastic properties of a protein are studied as a function of the applied force, temperature, and the mean coordination number of the protein "network."

Proteins 2013; 00:000–000.

Key words: stretching a protein; mean-field theory; sharp drop of protein elasticity; percolation nature of protein rigidity.

INTRODUCTION

Recently the methods to investigate protein unfolding at the single-molecule level have been developed.[1] In particular, in experiments on the mechanical unfolding of proteins the external force F is applied between a pair of amino-acid residues, and the dependence of the distance x between them is measured as a function of the force. This method allows to study that region of the energy landscape which is not accessible in classical denaturation experiments but could be important physiologically. At this, it turns out, particularly, that β-proteins are more "strong" (i.e., less stretchable), than α- or α/β-proteins. In addition, the protein extension depends considerably on the stretching direction and the points of applying the force, as well.[1,2]

Dependencies $x(F)$ of the single protein molecule extension on the applied force demonstrate some "fine structure," associated with successive ruptures of individual bonds between residues.[2] However, broadly speaking, one could reveal that all those dependencies have some general feature—each of them demonstrates the saturation: when the force becomes strong enough the protein begins to "flow," that is to lengthen rapidly without further growth of the applied force.[3–5] It seems as though the sharp (phase?) transition occurs under which the protein "skeleton" loses the stability and transits from the elastic state into the plastic one. It is illustrated by Figure 1, where several typical dependencies $x(F)$ are presented for different proteins. Each of them demonstrates above-mentioned sharp transformation of protein mechanical properties.

Considering that question qualitatively as the rigidity problem for an assembly of linked inflexible rods, one could use the classic approximate analytic approach ascending to Maxwell,[6] who has shown that such a system loses its rigidity if the mean number n_0 of links, being accounted for a single site, lowers down to the value of $n_0 = 12/5 = 2.4$. The corresponding phase transition is studied effectively with the aid of computer calculations in the framework of the percolation model[7,8] which supplies us with multiple numerical data but is scant of analytic results. In addition, those works have not touched the question of the stiffness lowering and the rigidity loss of the system *under the action of the external force*, which the present study is largely devoted to.

The analysis of experiments on the mechanical unfolding of proteins is based on the idea of slanting the energy landscape by the applied force which first has been discussed in Ref. 9. The polypeptide chain modeled as a sequence of connected beads on a cubic lattice has been considered in Ref. 10. The response of the system to stretching has been calculated *numerically*. In Ref. 11 the equilibrium response of random heteropolymers to mechanical deformation has been analyzed in the

Grant sponsor: Russian Foundation of Basic Researches; Grant number: 12-02-00550-.
*Correspondence to: Evgeny Meilikhov, Kurchatov Institute, 123182 Moscow, Russia. E-mail: meilikhov@yandex.ru
Received 2 July 2013; Revised 24 October 2013; Accepted 4 November 2013
Published online 5 December 2013 in Wiley Online Library (wileyonlinelibrary.com). DOI: 10.1002/prot.24471

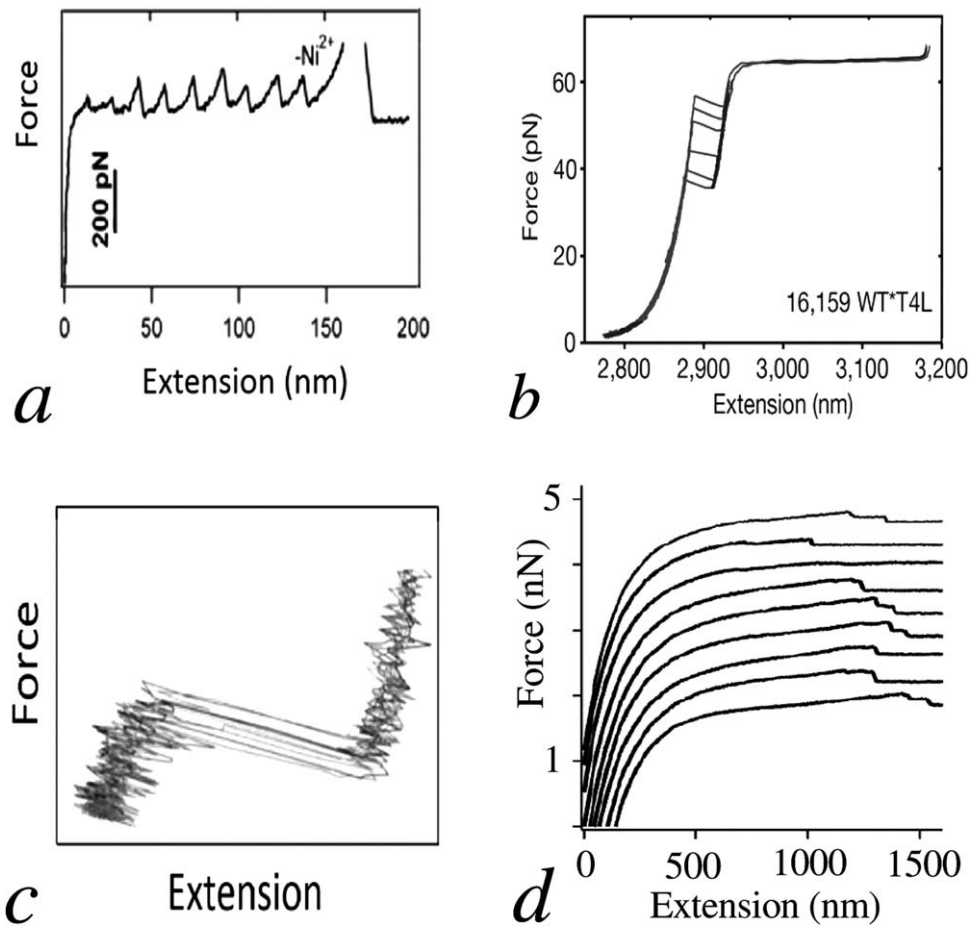

Figure 1

Dependencies of the protein elongation on the applied force. (**a**) Unfolding of GB1 bi-His mutant enhanced by the binding of Ni^{2+}.[5] (**b**) Force-extension curves of 16,159 WT*T4L protein.[3] (**c**) Measured forces as a function of extension for P5ab RNA hairpin from different nonequilibrium pulling experiments.[4] (**d**) Membrane proteins pulled away from the red blood cell surface.[38] Jumps in extension, associated with a slight drop in force, indicate unfolding of the molecule.

framework of a microscopic model, in which monomers are represented as beads on an elastic string. The phase diagram has been constructed, but temperature and force dependencies of some order parameter (for instance, the fraction of native contacts) have not been studied.

Our aim is to describe *analytically* the process of protein mechanical unfolding. To this end, we investigate the hypothetic protein phase transition in terms of graph (network) theory which, putting aside the nature of interacting elements composing the network, describes general features of such systems, allows to classify them by certain characteristics, to estimate qualitatively their stability, etc.[12–14] In the framework of that approach, a protein in the native globular state is the three-dimensional network structure. Further, we use the Ising model which, as known, has no analytical solution in the three-dimensional case. Thus, we simplify the problem even more, and find the analytical solution in the framework of the properly generalized mean-field theory.

MATERIALS AND METHODS

Order parameter

To consider qualitatively the process of protein stretching we introduce some order parameter, such as thermodynamic value, characterizing the long-range order in a system.[15] The steady-state order parameter equals zero in the disordered phase, and nonzero in the ordered one. To define such a parameter for a protein we could use the analogy with the magnetic ordering in the system of Ising magnetic moments having two possible orientations only, such as parallel or anti-parallel to some appropriate direction ("up" or "down"). Similarly, one could use the analogy with melting molecular liquid crystals which are described through the use of the parameter of the orientational order.[16]

In fact, the order parameter could be not only scalar one (as in the classic Ising model) but also of the vector

type (for instance, the magnetization), of complex number type (e.g., the wave function in superconductivity) and so on. Up to now, nobody succeeded to solve analytically the problem of protein folding/melting (denaturation/renaturation) of such a high dimensionality. Therefore, to catch though some total, integral features of those processes one should simplify the problem.

As a first approximation one could take the Ising model with a scalar order parameter. That model is one of the so called ensemble-based methods which are more efficient than the molecular dynamics because they retain only essential properties of the protein. It focuses on the most important driving forces through simplified representations of how elements within the protein interact (see the recent instructive review[17]). Similar approach was used in Ref. 18, however, it was limited to two-dimensional systems only and did not account of inter-residue links number' variety (cf., below). Originally created to study ferromagnetism, the Ising model is based on discrete variables ("spins") that can be in one of two states ("up" or "down"). The spins are arranged on a lattice, and they interact directly with an external force and neighboring "spins" through coupling interactions. The translation of the relevant definitions for the system of amino-acid protein residues is straightforward. Specifically, let N be the total number of residues in the protein chain, and only N^+ of them are in the "equilibrium" states corresponding to the native conformation while, due to the thermal fluctuations and forced stretching, the rest N^- residues are in non-native states. Energies of those non-native states differ significantly from equilibrium native energies. Then the order parameter ε is defined on the analogy of the said above:

$$\varepsilon \equiv \frac{N^+ - N^-}{N}, \quad \frac{N^\pm}{N} = \frac{1 \pm \varepsilon}{2}. \tag{1}$$

Our choice for this parameter is associated unambiguously with the fraction of the right (native) links between residues. This simple approach is similar to that considered in seminal papers[19,20] whose authors declared: "The non-native interactions of adjoining residues will have a distribution of energies with mean $-J''$. The native interactions must satisfy the principle of minimal frustration, and the simplest way to ensure this is to set all native tertiary interaction energies equal to $-J'$ where $J' > J''$." Obviously, that approach is equivalent to ours.

Although the interaction energies of various pairs of residues are different,[21] all of them are practically of the same order (by the interaction energy is meant here the energy renormed by the interaction with a solvent). Thus, remaining within the framework of a semi-qualitative model, we shall further reckon all those energies being the same and equal to J, so that the interaction energy W_{nm} of the two nearest residues (both, in the equilibrium and the perturbed states) could be written in the form characteristic to the Ising system:

$$W_{nm} = -Js_n s_m \quad (n, m = 1, 2, \ldots, N;\ n \neq m), \tag{2}$$

where $s_{n,m} = 1$ and $s_{n,m} = -1$ for residues being, correspondingly, in the native and the perturbed states (these parameters are analogous to spins of the Ising model). The total interaction energy W_n of the n-th residue with all its nearest neighbors (excepting, naturally, its strong covalent bond in the peptide chain) equals

$$W_n = \sum_{\text{neighbors}} W_{nm} = -Js_n \sum_{\text{neighbors}} s_m, \tag{3}$$

where the summation is produced over all residues in the protein globule which are nearest neighbors of a given one.

Within the standard mean-field theory, the number z of nearest neighbors is the same for all network sites, and the last sum in Eq. (3) is suggested to be equal to its average value $\langle \sum s_m \rangle = z\varepsilon$, so that the average interaction energy (per a residue) turns to be equal to $\langle W_n \rangle = -Jz\varepsilon$. The corresponding standard mean-field equation, determining the order parameter ε as a function of temperature T, reads[22]

$$\varepsilon = \tanh\left(\frac{zJ}{k_B T}\varepsilon\right) \tag{4}$$

(where k_B is the Boltzmann constant). In particular, that equation defines the phase transition temperature $T_c = zJ/k_B$.

Generalized mean-field theory

Below, we follow our article.[23] In the mean-field theory, by the "field" is meant some effective field h, influencing a given network site (in our case: amino-acid residue) on the part of its neighbors. Quantitatively, that field is proportional to the order parameter ε and could be expressed through the average interaction energy: $h = -\langle W_n \rangle = \lambda\varepsilon$, where $\lambda = zJ$ is the so called mean-field constant. In the traditional theory, the equivalence of all network sites is suggested. This results in the fact that the effective field (or, what is the same,—the interaction energy W_n) is identical in all sites, although, in fact, it changes from site to site randomly and that randomness should be (though, approximately) taken into account. To do that, the standard Eq. (4) is replaced, for instance, by such a generalized analog[23]

$$\varepsilon = \int_0^\infty \tanh\left(\frac{W}{k_B T}\right)\Phi(\varepsilon; W)\,dW, \tag{5}$$

where $\Phi(\varepsilon; W)$ is the distribution function of residues' interaction energies with their neighbors in the system

with the order parameter ε. Thus, the problem reduces to calculating the distribution function over *bond energies* for the system considered.

For this purpose, one should use distribution functions $f(n)$ of *bond numbers* known (proceeding from the information on the spatial protein structure which contains in special banks, see for instance Ref. 24) for a lot of proteins.[21,25–28] Those functions are more or less universal for various residues and proteins, and their envelopes are reasonably good described by the Gauss formulae

$$f(n)=\frac{\exp[-(n-n_0)^2/2\sigma^2]}{\sum_{n=1}^{\infty}\exp[-(n-n_0)^2/2\sigma^2]} \quad (n=1,2,\ldots), \qquad (6)$$

where the mostly probable bond number n_0 depends on the residue type (n_0 is minimal for Gln, Ser, Asn, Asp, Glu, Lys and maximal for Leu, Phe, Ile, Met, Val), and the scatter σ is practically the same for all residues.[21] Average bond number varies in the range $n_0=4-8$,[21,28] so the protein network is similar to that one which corresponds to the simple cubic lattice. It is natural to assume that proteins consisting of large number of residues with small numbers of bonds have a tendency to be partially or fully disordered. That is confirmed by available data.[29]

To simplify the further consideration, we shall neglect the difference of n_0 for various residues and accept $n_0=7$, $\sigma=2$ (numbers which are known from experiments).[28] Notice, however, that the average bond number for residues situated in the globule core is much higher than that in the near-surface layer,[28] that should result in lesser stability of this layer with regard to thermal fluctuations. It is also of interest that bonds between residues, being distant from each other along the peptide chain (more than 12 intermediate chain links), constitute although a small but substantial part of all bonds (~10%[26] that is typical for *small world* networks.[30] Taking account of these details will later allow to refine the model considered. Specifically, one could generalize the model by (i) dividing a protein into surface and core regions with different interaction energies and numbers of links per residue or (ii) by using the two-sublattice approach with residues being divided into two classes: those which strongly interact with each other (e.g., via disulfide links), and those with weak interaction (e.g., hydrogen bonds). The interaction between residues of different types is a third parameter. After that one could investigate a variety of scenarios.

Let us now find which could be the set of possible interaction energies of a given residue with its neighbors in protein network, if it has n bonds and each of them provides an equal interaction energy $\pm J$ [the sign depends on the mutual orientation of interacting residues' "spins," see Eq. (2)]. To be definite, let the "spin" of a given residue equals $s_n=+1$. The probability $p_n(k)$, that out of n neighboring residues k ones have "spins" of the opposite sign, is

$$p_n(k)=C_n^k\left(\frac{N^+}{N}\right)^{n-k}\left(\frac{N^-}{N}\right)^{k}=C_n^k\left(\frac{1+\varepsilon}{2}\right)^{n-k}\left(\frac{1-\varepsilon}{2}\right)^{k}$$
$$=\frac{1}{2^n}C_n^k(1+\varepsilon)^{n-k}(1-\varepsilon)^{k}, \qquad (7)$$

where C_n^k is the binomial coefficient giving the number of combinations of n elements taken k at a time. For that configuration of bonds, the interaction energy of a given residue with its neighbors equals $W_n(k)=(n-2k)J$. Taking into account, that the probability to have just n bonds for that residue equals $f(n)$, one finds the distribution function of interaction energies for residues possessing n bonds:

$$\Phi_n(\varepsilon;W)=\frac{1}{2^n}f(n)\sum_{k=0}^{n}C_n^k(1+\varepsilon)^{n-k}(1-\varepsilon)^{k}\delta[W-(n-2k)J], \qquad (8)$$

where $\delta[W-(n-2k)]$ is the delta-function.

At last, the total distribution function of residue interaction energies in a system, where they have different number of bonds with their neighbors ($n=1, 2, 3,\ldots$), is obtained by summing partial (for a given n value) distribution functions Eq. (8):

$$\Phi(\varepsilon;W)=\sum_{n=1}^{\infty}\Phi_n(\varepsilon;W)$$
$$=\sum_{n=1}^{\infty}\frac{1}{2^n}f(n)\sum_{k=0}^{n}C_n^k(1+\varepsilon)^{n-k}(1-\varepsilon)^{k}\delta[W-(n-2k)J] \qquad (9)$$

The lower limit of summation over n has been chosen to be $n = 1$, since at $n = 0$ the interaction energy equals to zero and such a residue without bonds does not influence the order parameter. As for the non-physical choice of the upper limit ($n \to \infty$) it is the convenient mathematical procedure only, based on the Gaussian form of the bond number distribution (with $\langle n\rangle \approx 7$). The contribution of all terms in Eq. (9) with $n \geq 12$ is negligible (<1%–2%) due to the rapid drop of the Gaussian.

Now we include in our scheme the force F applied to the protein. If the latter is single-domain, we choose the points of the force exerting so that this force would stretch a whole protein but not a single part of it. (For a protein consisting of several domains the case is possible when one of them is effectively stretched only.)

On a first approximation, one could assume that for a given residue the stretching affects one of n_0 bonds only and, namely, just that of them whose direction is mostly close to the direction of the stretching force. That estimate is obtained as follows. Of the total number Nn_0 of

bonds between amino-acid residues in a globule (N is the total number of residues in the protein), those of them "resist" to the protein extension only, whose directions are close to the direction of the applied force. There are about $n_0/4$ of such actual bonds per a single residue. For $n_0 \approx 6-8$ (residue "lattice" is similar to the simple cubic one) this leads to $n_0/4=1-2$, that just validates the assumption made. Then the force f, stretching every actual bond, equals $f=F/N$. It is self-evident, this conclusion is valid for the force being applied to the opposite sites of a protein globule. For more local stretching (with the force applied to some nearby residues) this output should be reconsidered because the number of bonds, resisting to the force F, is much lower than N in that case and, hence, the force f, stretching each of them, is much higher than F/N.

Influence of the applied force on the system energy depends on the states of residues at the ends of that actual bond. If both of those residues are in "stable" states ($s_n, s_m=1$), then the stretching force should result in increasing energy of that bond, so that the total energy of coupling that residue with its network neighbors equals

$$W \equiv W_1 = -(n-2k)J + f\delta x,$$

where δx is the bond extension under the action of the force f. The typical value of that extension in the "elastic" regime of the protein unfolding is much less than the inter-residue distance a: $\delta x << a$. In the opposite case ($s_n=1, s_m=-1$), the coupling energy of a given residue does not change under stretching:

$$W \equiv W_2 = -(n-2k)J.$$

Taking into account that probabilities of both above-mentioned cases are equal, respectively, to $p_1=(n-k)/n$ and $p_2=k/n$, one could rewrite the distribution function Eq. (8) in the form

$$\Phi(\varepsilon; W) = \sum_{n=1}^{\infty} \frac{1}{2^n} f(n) \sum_{k=0}^{n} C_n^k (1+\varepsilon)^{n-k} (1-\varepsilon)^k \{p_1 \delta[W-W_1] + p_2 \delta[W-W_2]\}. \quad (10)$$

In fact, this function determines all features of our model. It depends on (i) statistics of weak links between residues [through the parameters n_0 and σ of the Gaussian Eq. (6)], (ii) interaction energy J, and (iii) the force f (through the coupling energy W_1). All these parameters should be individually suited to various proteins to describe their properties quantitatively. However, in the present article we are interested mainly in universal behavior of stretched proteins irrespective of specific values of the mentioned parameters.

After substituting the distribution function of Eq. (10) in Eq. (5), one obtains the generalized mean-field equation which determines the order parameter in the system considered:

$$\varepsilon = \sum_{n=1}^{\infty} \frac{1}{2^n} f(n) \times \sum_{k=0}^{n} C_n^k \left\{ \frac{n-k}{n} \tanh\left[\frac{(n-2k)-\mathcal{F}}{\tau}\right] + \frac{k}{n} \tanh\left[\frac{n-2k}{\tau}\right] \right\} (1+\varepsilon)^{n-k} (1-\varepsilon)^k, \quad (11)$$

where $\tau = k_B T/J$ and $\mathcal{F} = F\delta x/JN$ are the reduced force and temperature, respectively.

RESULTS AND DISCUSSION

Equation (11) yields, in the mean-field approximation, the dependence $\varepsilon(\mathcal{F})$ of the order parameter on the applied force. That dependence could give a qualitative idea of integral elastic properties which determine the process of the protein stretching under the force action. The considered approach describes a quasi-static process corresponding to the series of steady states under the force that changes infinitely slow. To associate the length variation ΔL of the unfolding protein with the order parameter ε, notice that every broken bond enhances the length by the value on the order of δx,[2] so that $\Delta L \sim N_- \cdot \delta x = (1-\varepsilon)N\delta x/2$. As it could be seen from the results of calculating the dependence $\varepsilon(\mathcal{F})$ (see Fig. 2), the protein elasticity diminishes with the force increasing and drops sharply at some threshold value $\mathcal{F}=\mathcal{F}_c$ of that force, to which critical value ε_c of the order

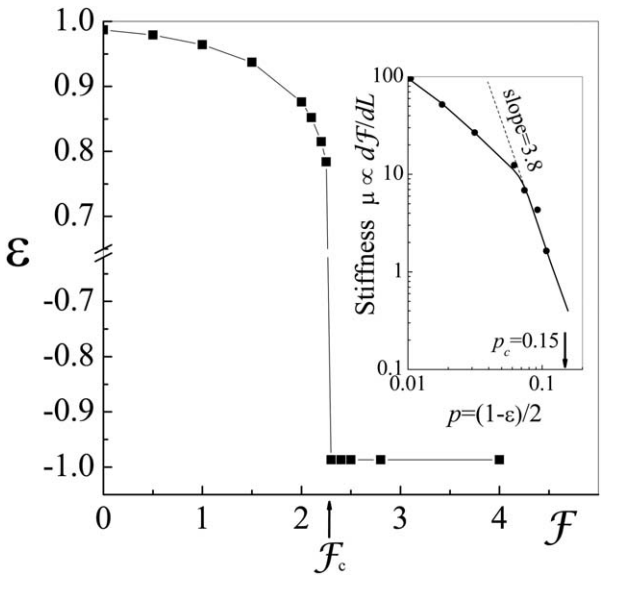

Figure 2

Calculated on the base of Eq. (11) dependence of the order parameter on the applied force. The set of parameters $\tau=2$, $n_0=7$, $\sigma=2$ have been accepted. In the inset, dependence of the protein strength on the fraction $p=(1-\varepsilon)/2$ of broken bonds.

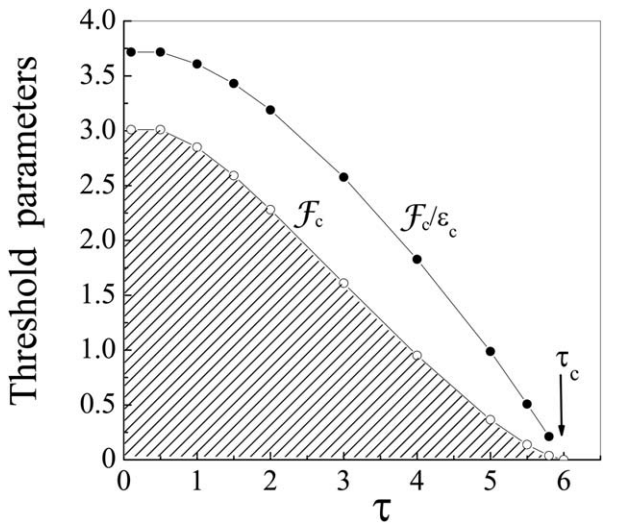

Figure 3

Temperature dependence of the critical force and the integral relative protein elasticity $\mathcal{F}_c/\varepsilon_c$. The *dashed area* corresponds to the elastic state. The set of parameters $n_0=7$, $\sigma=2$ have been accepted.

parameter corresponds. Values $\varepsilon < 0$ have no physical meaning: they correspond to the regime with infinitesimally low elasticity, when the protein construction "flows" freely. Notice, that at $\mathcal{F}=0$ but $T \neq 0$, the order parameter is different from unit and tends to it at $T \to 0$. This means that even with no force there is a finite extension $\Delta L_0(T)$ of the molecule relative to that at $T = 0$ due to the thermal fluctuations.

Near the transition in the plastic state mechanical properties of a medium, which is strongly non-uniform relative to the elastic modulus, are described by the percolation theory. According to that theory, with increasing the fraction p of the "soft" component (in our case, the fraction of broken bonds) the medium rigidity drops, by a threshold fashion, according to the law[31]

$$\frac{\mu}{\mu_0} = (p_c - p)^f, \qquad (12)$$

where μ, μ_0 are the effective modulus of elasticity for the composite and the modulus of elasticity for the original "matrix," respectively; $f \approx 3.8$ is the critical index; and $p_c \approx 0.15$ is the critical concentration of soft inclusions.[32]

The parameter $(1-\varepsilon)/2$ is a measure of the protein extension (see above), and the same parameter defines the fraction p of broken bonds: $p=N_-/N=(1-\varepsilon)/2$, so that the critical value of the order parameter corresponding to the rigidity loss equals $\varepsilon_c=1-2p_c \approx 0.7$. As a measure of rigidity in our model, one could take the rate $d\mathcal{F}/dL$ of varying the force with the extension. In Figure 2 (inset), the dependence of that derivative on the fraction p of broken bonds is presented. As seen, near the threshold it is quite adequately described by the percolation law Eq. (12).

Naturally, the critical force decreases with temperature, as with heating a protein becomes less and less strong. The respective dependence is shown in Figure 3 which, in fact, defines the phase diagram of a protein: in the dashed region it is compacted and outside it is "disrupted." In the same figure, the temperature dependence of the protein integral relative elasticity $\mathcal{F}_c/\varepsilon_c$ is depicted. The phase diagram boundary has a "tail" of the length $\Delta\tau \approx 2$. Such a tail is the effect of broadening distribution function $f(n)$ and results from the critical temperature τ_c spread. It is standard for mean-field theory that $\tau_c \sim n$, and this leads to $\Delta\tau_c \sim \sigma=2$, in accordance with Figure 3.

The same tail has been demonstrated in Ref. 33 where this diagram has been obtained for some proteins obtained by numerical simulation. That evidences the validity of our analytical model.

It may be of interest to compare this feature of the phase diagram with temperature dependencies of Young's modulus for some proteins (lysozyme, albumin, and myoglobin).[34] They have the high-temperature tails of the length $\Delta T \approx 100\,\mathrm{K}$ with denaturation temperature $T_c \approx 300\,\mathrm{K}$. That corresponds to $\Delta T/T_c \approx 0.3$ and agrees with our above-mentioned estimation $\Delta\tau/\tau_c \approx 0.3$.

In Figure 4, presented are temperature dependencies (i) of the order parameter critical value $\varepsilon_c(\tau)$, (ii) of the protein order parameter $\varepsilon(\mathcal{F}=0)$ at zero stretching force, and (iii) of the difference $\varepsilon-\varepsilon_c$ between them, characterizing the maximum possible elongation of a protein before its transition to the flow regime. All those parameters are, naturally, turn to zero at the protein

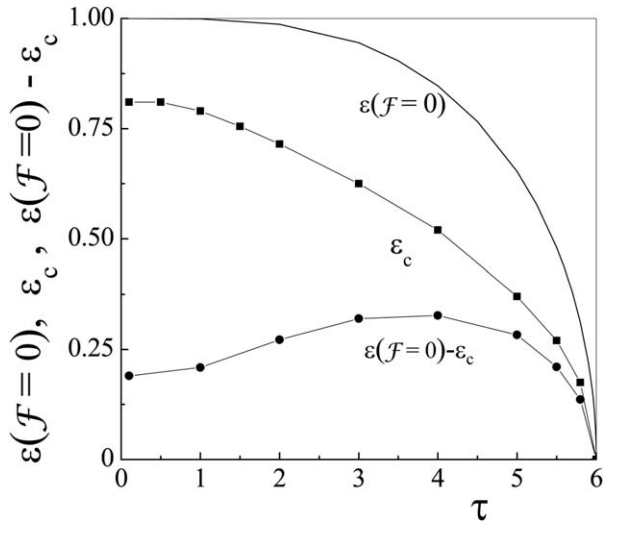

Figure 4

Temperature dependencies of the order parameter critical value $\varepsilon_c(\tau)$, order parameter $\varepsilon(\mathcal{F}=0)$ with no force, and their difference $\varepsilon-\varepsilon_c$, characterizing the maximum possible protein elongation before jumping into the flow regime. The set of parameters $n_0=7$, $\sigma=2$ have been accepted.

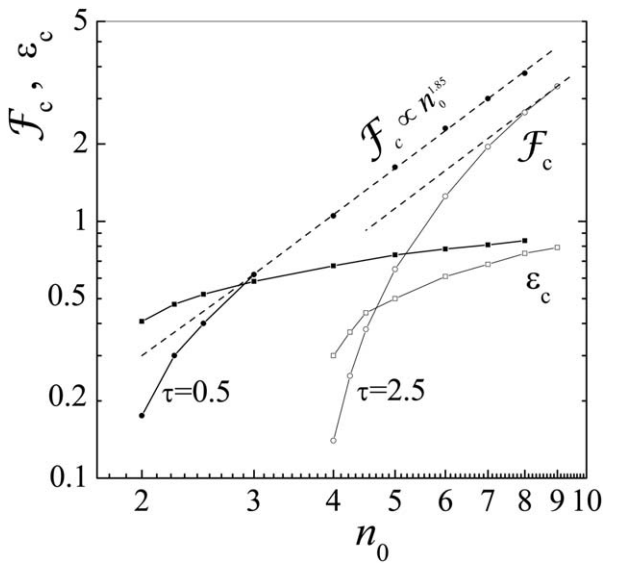

Figure 5

Dependencies of the critical force $\mathcal{F}_c$ and the order parameter critical value ε_c on the average number n_0 of residues' bonds for $\tau=0.5$, 2.5 at $\sigma=2$. *Dashed line* is the power-law dependence $\mathcal{F}_c \propto n_0^{1.85}$.

melting temperature $\tau=\tau_c$. Notice the non-trivial temperature dependence of the possible order parameter drop (the protein "solidity factor") which is defined by the difference $\varepsilon(\mathcal{F}=0)-\varepsilon_c$ and determines the possible protein extension: it is maximum at some intermediate temperature $\tau \approx 0.7\tau_c$. That means: for optimizing their functionality some proteins require the temperature which is by approximately 30% lower than the denaturated one. To observe that phenomenon, one needs to measure the maximum protein elongation as a function of temperature. Unfortunately, we do not know relevant experiments, so this result could be considered as a prediction. We hope, it could be inspected experimentally.

Quantitatively, critical parameters (critical value of the order parameter, critical force) depend significantly on the average number n_0 of bonds for amino-acid protein residues; the higher that number, the higher the force required for the phase transition to the plastic state. Corresponding dependencies of those parameters are presented in Figure 5 which, particularly, shows that at $n_0 >> 1$ the critical force $\mathcal{F}_c \propto n_0^{1.85}$, and the critical value of the order parameter $\varepsilon_c \to 0.9$ irrespective of temperature. In addition, notice more rapid drop of the critical force with decreasing n_0 in the vicinity of $n_0 \approx 2$ at low temperature ($\tau=0.5$). In fact, it is near the result predicted by Maxwell.[6] At higher temperature ($\tau=2.5$), the rigidity loss occurs at significantly higher $n_0 \approx 4$, that is the consequence of thermal fluctuations.

One could also see that protein networks with the average "coordination number" n_0, smaller than some critical value $n_0^{(c)}$, could not exist in the globular state—any, arbitrarily small, force destroys that state. The threshold value $n_0^{(c)}$ raises with temperature: $n_0^{(c)} \approx 1.5-2$ at $\tau=0.5$, $n_0^{(c)} \approx 3.5-4$ at $\tau=2.5$, and $n_0^{(c)} \approx 7$ at $\tau=6$ (see Fig. 3). Generally, $n_0^{(c)} \approx 1+\tau$ at $\tau >> 1$, so that the maximum reduced temperature at which a protein globule could exist (for the close packing of amino-acid residues, where $n_0=12$) equals $\tau_{\max} \approx 11$. That corresponds to the absolute temperature $T_{\max}=(J/k_B)$ $\tau_{\max} \approx 350\,\mathrm{K}$ for $J=0.5\times10^{-21}$ J.

Our results which are based on the *analytic* model, leading to Eq. (11), could be compared with results of *numerical* calculations,[35] performed in the framework of the (slightly modified) lattice model known as "Partially Directed Self-Avoiding Walk" (PDSAW).[36] (Analogous dependence follows from numerical calculations relating to the one-strand DNA stretching,[37] as well.) In Figure 6 one could see corresponding dependencies of the stretching force on the induced extension of a macromolecule at $\tau << \tau_C$. As long as the relative elongation does not exceed approximately 50% that dependence is logarithmic one in both models:

$$\mathcal{F}=\alpha\ln[\Delta L/\Delta L_0(T)], \qquad (13)$$

where $\Delta L_0(T)$ is the finite extension of the molecule with no force ($\mathcal{F}=0$), and α is a constant.

The same logarithmic law describes some *experiments*. For example, in Figure 7 the upper graph of Figure 1 is represented again in the form which reveals the law of Eq. (13). Similar logarithmic force-extension

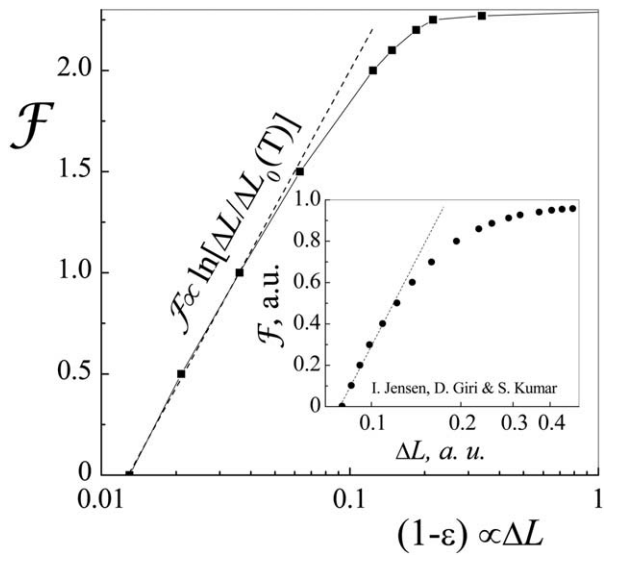

Figure 6

Dependencies of the stretching force on the protein elongation resulting from our model and the PDSAW-model[35] (inset; data has been obtained by digitizing the initial section of the graph for $T=0.2$ in Figure 13a from Ref. 35. *Dashed lines* are logarithmic dependencies $\mathcal{F}=\alpha\ln[\Delta L/\Delta L_0(T)]$.

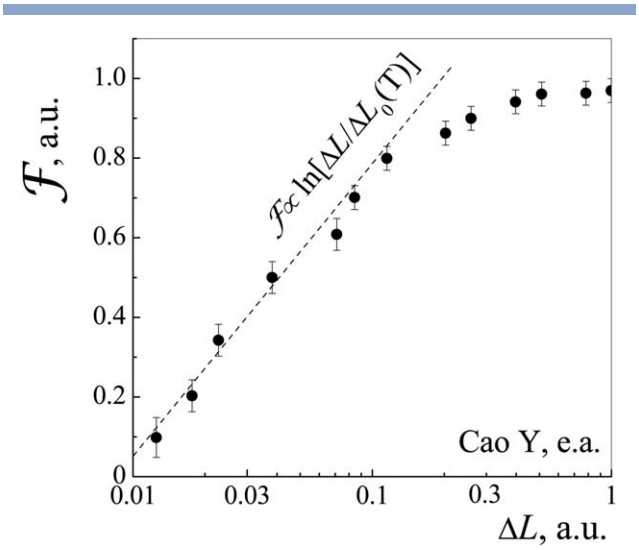

Figure 7

Experimental evidence[5] of the logarithmic law Eq. (13) by the example of stretching the protein mutant GB1 bi-His enhanced by the binding of Ni^{2+} (see the upper graph in Fig. 1).

dependencies have been also observed in Ref. 38 for some membrane proteins pulled away from the red blood cell surface.

This logarithmic dependence could be obtained on the base of the natural suggestion that the differential protein "compliance" (elasticity), characterized by the derivative $dL/d\mathcal{F}$, is proportional to the fraction p of broken bonds. Since the protein elongation ΔL is also proportional to p (see above), that statement could be written in the form

$$\frac{d\Delta L}{d\mathcal{F}} = \alpha^{-1}\Delta L, \qquad (14)$$

and herefrom the relation of Eq. (13) follows.

It has been shown that long-range links (between residues, widely spaced along the peptide chain but closely situated in the native globular conformation) play the noticeable role in the thermal stability of protein macromolecules.[39] In some proteins (as a rule, thermophilic ones) the fraction of those links reaches several percents that could result in elevating their denaturation temperature by 10–20K. One should expect that such proteins possess higher mechanical strength, as well.

CONCLUSION

In the present article, we have studied (in the framework of the effective field theory) the phase transition of a protein macromolecule from the elastic state to the plastic one, which arises from mechanical stretching. It is shown that with increasing the stretching force the sharp drop of the protein elasticity occurs, it losses the rigidity, and the protein construction begins to "flow" freely. Near that threshold, the rigidity falls according to the percolation theory. The critical force depends significantly on the average number of bonds for amino-acid residues in a protein—the higher that number, the stronger the force inducing the phase transition in the plastic state. Protein networks with average coordination number n_0 smaller than some critical value could not exist in the globular state—arbitrarily small force destroys that state. With lowering temperature the threshold value of n_0 decreases. On the whole, the stated results agree with experiments. This gives promise that the effective field theory could adequately describe integral mechanical properties of protein macromolecules.

REFERENCES

1. Brockwell DJ. Force denaturation of proteins - an unfolding story. Curr Nanosci 2007;3:3–15.
2. Borgia A, Williams PM, Clarke J. Single-molecule studies of protein folding. Annu Rev Biochem 2008;77:101–125.
3. Shank EA, Cecconi C, Dill JW, Marqusee S, Bustamante C. The folding cooperativity of a protein is controlled by its chain topology. Nature 2010;465:637–640.
4. Hummer G, Szabo A. Free energy profiles from single-molecule pulling experiment. PNAS 2010;107:21441–21446.
5. Cao Y, Yoo T, Li H. Single molecule force spectroscopy reveals engineered metal chelation is a general approach to enhance mechanical stability of proteins. PNAS 2008;105;11152–11157.
6. Maxwell JC. On the calculation of the equilibrium and stiffness of frames. Phil Mag 1864;27:294–299.
7. Thorpe MF, Jacobs DJ, Chubynsky NV, Rader AJ. Generic rigidity of network glasses. In: Thorpe MF, Duxbury PM, editors. Rigidity theory and applications. New York: Plenum Publishing; 1999.
8. Rader AJ, Hespenheide BM, Kuhn LA, Thorpe MF. Protein unfolding: Rigidity lost. PNAS 2002;99;3540–3545.
9. Bell GI. Models for the specific adhesion of cells to cells. Science 1978;200:618–627.
10. Klimov DK, Thirumalai D. Stretching single-domain proteins: phase diagram and kinetics of force-induced unfolding. PNAS 1999;96; 6166–6170.
11. Geissler PL, Shakhnovich EI. Reversible stretching of random heteropolymers. Phys Rev E 2002;65:056110.
12. Barabási AL, Oltvai ZN. Network biology: understanding the cell's functional organization. Nat Rev Genet 2004;5:101–113.
13. Newman MEJ. The structure and function of complex networks. SIAM Rev 2003;45:167–256.
14. Atilgan C, Okan OB, Atilgan AR. Network-based models as tools hinting at non-evident protein functionality. Ann Rev Biophys 2012; 41:205–225.
15. Landau LD, Lifshitz EM. Statistical physics. Oxford: Butterworth-Heinemann; 1980.
16. Chandrasekhar S. Liquid crystals. Cambridge: Cambridge University Press; 1994.
17. Jacobs DJ. Ensemble-based methods for describing protein dynamics. Curr Opin Pharmacol 2010;10:760–769.
18. Imparato A, Pelizzola A, Zamparo M. Ising-like model for protein mechanical unfolding. Phys Rev Lett 2007;98:148102.
19. Bryngelson JD, Wolynes PG. Spin glasses and the statistical mechanics of protein folding. Proc Natl Acad Sci USA 1987;84:7524–7528.
20. Bryngelson JD, Wolynes PG. Intermediates and barrier crossing in a random energy-model (with applications to protein folding). J Phys Chem 1989;93:6902–6915.

21. Miyazawa S, Jernigan RL. Residue-residue potentials with a favorable contact pair term and an unfavorable high packing density term, for simulation and threading. J Mol Biol 1996;256:623–644.
22. Kittel C. Introduction to solid state physics. New York: Wiley; 2004.
23. Meilikhov EZ, Farzetdinova RM. Network model of protein globule. J Biol Phys 2013;39:673–685.
24. Grosberg AY, Khokhlov AR. Statistical physics of macromolecules. Berlin: Springer; 1994.
25. Berman HM, Henrick K, Nakamura H. Announcing the worldwide Protein Data Bank. Nat Struct Biol 2003;10:980.
26. Bagler G, Sinha S. Assortative mixing in Protein Contact Networks and protein folding kinetics. Bioinformatics 2007;23:1760–1767.
27. Huang J, Kawashima SI, Kanehisa M. New amino acid indices based on residue network topology. Genome Inform 2007;18:152–161.
28. Atilgan AR, Akan P, Baysal C. Small-world communication of residues and significance for protein dynamics. Biophys J 2004;86:85–91.
29. Romero P, Obradovic Z, Kissinger C, Villafrance JE, Dunker AK. In: Proceedings of the international conference on neural networks; 1997. pp 91–95.
30. Watts DJ, Strogatz SH. Collective dynamics of 'small-world' networks. Nature (Lond) 1998;393:440–442.
31. Bergman DJ. Exact relations between elastic and electrical response of d-dimensional percolating networks with angle bending forces. J Stat Phys 2003;111:171–199.
32. Shklovskii BI, Efros AL. Electronic properties of doped semiconductors. Berlin-Heidelberg-New York-Tokyo: Springer-Verlag; 1984.
33. Hur JS, Darve E. Annual research briefs. 425 (2003), Center for Turbulence Research, Stanford University (http://ctr.stanford.edu/ResBriefs03/joehur.pdf).
34. Morozov VN, Gevorkian SG. Low-temperature glass transition in proteins. Biopolymers 1985;24:1785–1799.
35. Jensen I, Giri D, Kumar S. Role of pulling direction in understanding the anisotropy of the resistance of proteins to force-induced mechanical unfolding. Mod Phys Lett B 2010;24:379.
36. Vanderzande C. Lattice models of polymers. Cambridge: Cambridge University Press; 1998.
37. Kumar S, Mishra G. Force-induced stretched state: effects of temperature. Phys Rev E 2008;78:011907.
38. Ikai A, Afrin R, Sekiguchi H. Pulling and pushing protein molecules by AFM. Curr Nanosci 2007;3:17–29.
39. Srivastava A, Granek R. Cooperativity in thermal and force-induced protein unfolding: integration of crack propagation and network elasticity models. Phys Rev Lett 2013;110:138101.