EPJ Web of Conferences **75**, 08001 (2014)
DOI: 10.1051/epjconf/20147508001
© Owned by the authors, published by EDP Sciences, 2014

Collaps of protein macromolecule induced by a force as an analog of remagnetization

E.Z. Meilikhov[a], R.M. Farzetdinova

Kurchatov Institute, 123182 Moscow, 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".

1 Introduction

Recently the methods to investigate protein unfolding at the single-molecule level have been developed [1]. 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.

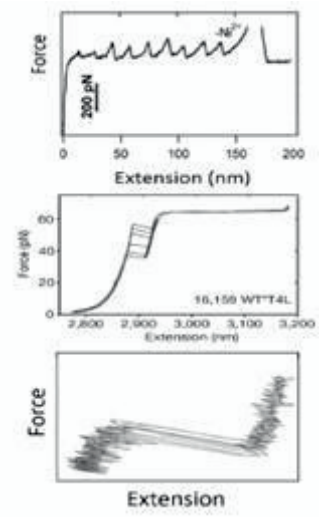

Fig. 1. Dependencies of the protein elongation on the applied force. Above – unfolding of GB1 bi-His mutant [5], middle – force-extension curves of 16,159 WT*T4L protein [3], below – P5ab RNA hairpin from different pulling experiments.

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, one could reveal that all those dependencies have some general feature – each of them demonstrates the saturation: when the force becomes to be strong enough the protein begins to "flow", that is to lengthen rapidly with-out 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 Fig. 1, where typical dependencies $x(F)$ are presented for some proteins. Each of them demonstrates above-mentioned sharp transformation of protein strength.

The corresponding phase transition is studied effectively with the aid of computer calculations in the framework of the percolation model [6, 7] 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 work is largely devoted to.

To consider qualitatively the process of protein stretching we introduce the order parame-ter – thermodynamic value, characterizing the long-range order in a system. 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

[a] Corresponding author: meilikhov@yandex.ru

This is an Open Access article distributed under the terms of the Creative Commons Attribution License 2.0, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

with the magnetic ordering in the system of Ising magnetic moments having two possible orientations only – parallel or anti-parallel to some appropriate direction ("up" or "down").

In the case considered , the dimensionality of the order parameter is tremendously high. Therefore, to catch though some total, integral features of those processes one should simplify the prob-lem. As a first approximation one could take the Ising model with a scalar order parameter.

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 thermal fluctuations and forced stretching, the rest N^- residues are in states that far-off from equilibrium ones. The degree of that "farness" is defined not by absolute values of space and/or angle deviations of residues from the equilibrium state but by the condition that in the perturbed state energies of weak residues' interaction are significantly lower than equilibrium energies. Then the order parameter ε is defined on the analogy of the said above:

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

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 [8,9].

Though the interaction energies of various pairs of residues are different [10], all of them are practically of the same order. 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: $W_{nm} = -Js_n s_m \quad (n, m = 1, 2, ..., N;\ n \neq m)$.

where $s_{n,m}$=1 and $s_{n,m}$= −1 for residues being, correspond-dingly, 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

$$W_n = \sum_{neighbors} W_{nm} = -Js_n \sum_{neighbors} s_m. \qquad (1)$$

where the summation is produced over all residues 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 (1) is suggested to be equal to its average value, 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 para-meter ε as a function of temperature T, reads [11]

$$\varepsilon = \tanh\left(zJ\varepsilon / k_B T\right) \qquad (2)$$

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

2 Generalized mean-field theory

Below, we follow our paper [12]. 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 identic on 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 equation (2) is replaced, for instance, by such a generalized analog [18]

$$\varepsilon = \int_0^\infty \tanh\left(\frac{W}{k_B T}\right) \Phi(\varepsilon; W) dW. \qquad (3)$$

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 for a lot of proteins. Those functions are more or less universal for various residues and proteins, and their envelopes are reasonably good described by the Gauss formulae

$$f(n) = \exp\left[-\frac{(n-n_0)^2}{2\sigma^2}\right] \Big/ \sum_{n=1}^{\infty} \exp\left[-\frac{(n-n_0)^2}{2\sigma^2}\right] \quad (n = 1, 2, ...),$$

where the mostly probable bond number n_0 depends on the residue type [10]. Average bond number varies in the range n_0=4–8, 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.

To simplify the further consideration, we shall neglect the difference of n_0 for various residues and accept n_0 =7, σ=2 [13]. Notice, however, that the average bond number for residues situated in the globule core is much higher than that in the near-surface layer [13], which should result in lesser stability of this layer with regard to thermal fluctuations. Taking account of these details will later allow to refine the model considered.

According to [12], the distribution function of interaction energies for residues possessing n bonds reads

$$\Phi(\varepsilon;W)=\sum_{n=1}^{\infty}\Phi_n(\varepsilon;W)= \quad (4)$$
$$\sum_{n=1}^{\infty}\frac{1}{2^n}f(n)\sum_{k=0}^{n}C_n^k(1+\varepsilon)^{n-k}(1-\varepsilon)^k\delta\left[W-(n-2k)J\right]$$

with δ being the delta-function.

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.) In 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 $f=F/N$.

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 (4) in the form

$$\Phi(\varepsilon;W)=\sum_{n=1}^{\infty}\frac{1}{2^n}f(n)\sum_{k=0}^{n}C_n^k(1+\varepsilon)^{n-k}(1-\varepsilon)^k\times \quad (5)$$
$$\left\{p_1\delta\left[W-W_1\right]+p_2\delta\left[W-W_2\right]\right\}.$$

After substituting the distribution function (5) in Eq. (3), one obtains the generalized mean-field equation which determines the order parameter in the system considered:

$$\varepsilon=\sum_{n=1}^{\infty}\frac{1}{2^n}f(n)\sum_{k=0}^{n}C_n^k\times \quad (6)$$
$$\left\{\frac{n-k}{n}\tanh\left[\frac{(n-2k)-\mathrm{F}}{\tau}\right]+\frac{k}{n}\tanh\left[\frac{n-2k}{\tau}\right]\right\}(1+\varepsilon)^{n-k}(1-\varepsilon)^k,$$

where $\tau=k_BT/J$ and $=F\delta x/JN$ are the reduced force and temperature, respectively.

3 Results and discussion

Equation (6) yields, in the mean-field approximation, the dependence ε(Φ) 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 ε(Φ) (see Fig. 2), the protein elasticity diminishes with the force increasing and drops sharply at some threshold value $\Phi=\Phi_c$ of that force, to which critical value ε_c of the order parameter corresponds. Values ε<0 have no physical meaning: they correspond to the regime with infinitesimally low elasticity, when the protein construction "flows" freely. Notice, that at Φ=0 but $T\neq0$, the order parameter is different from unit and tends to it at $T\to0$. 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 thermal fluctuations.

The parameter (1–ε)/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 $\varepsilon_c=1-2p_c\cong0.7$. As a measure of rigidity in our model, one could take the rate $d\Phi/dL$ of varying the force with the extension. In Fig. 2 (inset), the dependence of that derivative on the fraction p of broken bonds is presented.

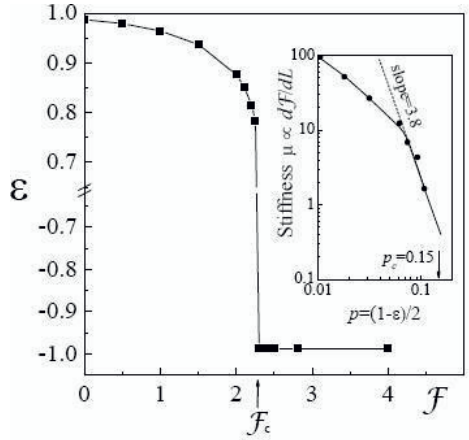

Fig. 2. Dependence of the order parameter on the applied force calculated on the base of Eq. (6). The set of parameters τ=2, n_0=7, σ=2 have been accepted. In the inset – dependence of the protein strength on the fraction p=(1–ε)/2 of broken bonds.

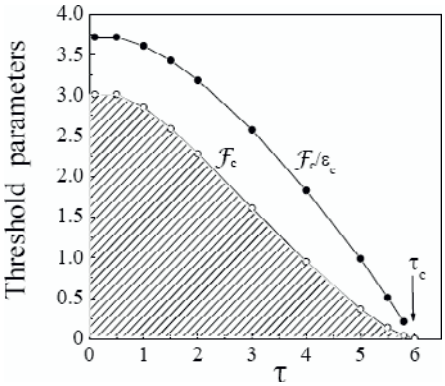

Fig. 3. Temperature dependence of the critical force and the integral relative protein elasticity F_c/ε_c. The dashed area corresponds to the elastic state. The set of parameters n_0=7, σ=2 have been accepted.

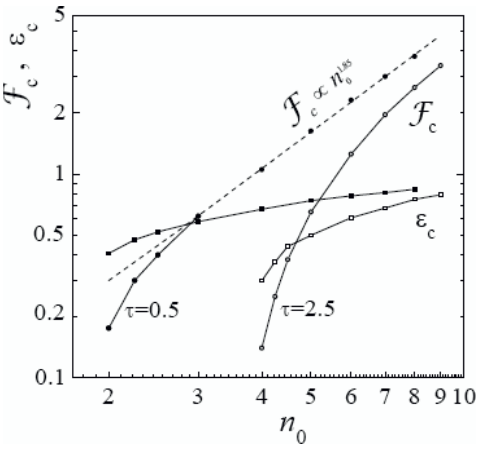

Fig.4. Dependencies of the critical force F_c and the order parameter critical value ε_c on the average number n_0 of residues' bonds for τ=0.5, 2.5 at σ=2. Dashed line is the power-low dependence $F_c \propto n_0^{1.85}$.

Naturally, the critical force decreases with temperature – with heating a protein becomes less and less strong. The respective dependence is shown in Fig. 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 Φ/ε_c is depicted.

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 Fig. 4 which, particularly, shows that at n_0>>1 the critical force $F \propto n_0^{1.85}$, and the critical value of the order parameter $\varepsilon_c \to 0.9$.

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: $n_0^{(c)} \cong 1.5$–2 at τ=0.5, $n_0^{(c)} \cong 3.5$-4 at τ=2.5 and $n_0^{(c)} \cong 7$ at τ=6. Generally, $n_0^{(c)} \cong 1+\tau$ at τ>>1, so the maximum reduced temperature at which a protein globule could exist (for the close packing of amino-acid residues, where $n_0^{(c)}$=12) equals $\tau_{max} \cong 11$. That corresponds to the absolute temperature $T_{max}=(J/k_B)\tau_{max} \cong 350$K for $J=0.5 \cdot 10^{-21}$ J.

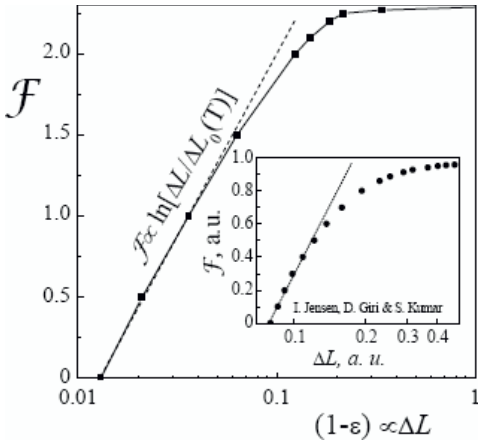

Fig.5. Dependencies of the stretching force on the protein elongation resulting from our model and the PDSAW-model [15] (inset; data has been obtained by digitizing the initial section of the graph for τ=0.2 in Fig. 13a from [15]). Dashed lines are logarithmic dependencies $F=\alpha\ln[L/L_0(T)]$.

Our results, which are based on the *analytic* model, could be compared with results of *numerical* calculations [14], performed in the framework of the (slightly modified) lattice model known as "Partially Directed Self-Avoiding Walk" (PDSAW) [15]. In Fig. 5 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 ~50% that dependence is logarithmic one in both models:

$$F = \alpha \ln\left[\Delta L / \Delta L_0(T)\right], \qquad (7)$$

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

This logarithmic dependence could be obtained on the base of the natural suggestion that the differential protein "compliance" (elasticity), characterized by the derivative $dL/d\Phi$, 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 $d\Delta L/d\Phi=\alpha^{-1}\Delta L$, and here from the Eq. (7) follows.

In conclusion, 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.

The work has been supported by the Grant # 12-02-00550-a of the Russian Foundation of Basic Researches.

References

1. D.J. Brockwell, Current Nanosci. **3**, 3 (2007).
2. A. Borgia, P.M. Williams, and J. Clarke, Annu. Rev. Biochem. **77**, 101 (2008).
3. E.A. Shank, et. al, Nature **465**, 637 (2010).
4. G. Hummer, and A. Szabo, PNAS **107**, 21441 (2010).
5. Y. Cao , T. Yoo , and H. Li, PNAS **105**, 11152 (2008).
6. M. F. Thorpe, et. al, in: Rigidity Theory and Applications, Eds: M.F. Thorpe and P.M. Duxbury, Plenum Publishing, New York (1999).
7. A. J. Rader, et. al, PNAS **99**, 3540 (2002).
8. J.D. Bryngelson, and P.G. Wolynes, PNAS. **84,** 7524 (1987).
9. J.D. Bryngelson, and P.G. Wolynes, J. Phys. Chem. **93**, 6902 (1989).
10. S. Miyazawa, and R. L. Jernigan , J. Mol. Biol. **256**, 623 (1996).
11. Ch. Kittel, Introduction to Solid State Physics, J. Wiley (2004).
12. E.Z. Meilikhov, and R.M. Farzetdinova, J. Biol. Phys. (2013, in print).
13. A.R. Atilgan, P. Akan, and C. Baysal, Biophys. J. **86**, 85 (2004).
14. I. Jensen, D. Giri, and S. Kumar, Mod. Phys. Lett. B **24**, 379 (2010).
15. C. Vanderzande, Lattice Models of Polymers Cambridge University Press, Cambridge (1998).