

ELASTIC BEHAVIOR OF PROTEIN-LIKE SINGLE CHAIN^{*}

Wei-qi Yi and Lin-xi Zhang^{**}

Department of Physics, Wenzhou Normal College, Wenzhou 325027, China

Abstract The conformational properties and elastic behaviors of protein-like single chains in the process of tensile elongation were investigated by means of Monte Carlo method. The sequences of protein-like single chains contain two types of residues: hydrophobic (H) and hydrophilic (P). The average conformations and thermodynamics statistical properties of protein-like single chains with various elongation ratio λ were calculated. It was found that the mean-square end-to-end distance $\langle R^2 \rangle_r$ increases with elongation ratio λ . The tensor eigenvalues ratio of $\langle L_2^2 \rangle : \langle L_1^2 \rangle$ decreases with elongation ratio λ for short (HP)_n protein-like polymers, however, the ratio of $\langle L_2^2 \rangle : \langle L_1^2 \rangle$ increases with elongation ratio λ , especially for long (H)_n sequence. Average energy per bond $\langle U \rangle$ increases with elongation ratio λ , especially for (H)_n protein-like single chains. Helmholtz free energy per bond also increases with elongation ratio λ . Elastic force (f), energy contribution to force (f_U) and entropy contribution to force (f_S) for different protein-like single chains were also calculated. These investigations may provide some insights into elastic behaviors of proteins.

Keywords: Elastic behavior; Protein-like single chains; Conformation.

INTRODUCTION

Numerous functions in biological organisms are attributed to the behaviors of proteins. Elastic behavior of protein molecule is of great significance in understanding the mechanical characteristic of proteins, such as the capability of silk, even the outstanding mechanical properties of spider silk. Recently a series of single molecules experiments, in which proteins are stretched by mechanical forces, reveal lots of information about the mechanical properties of structural proteins and provide detailed insight into the intermolecular and intramolecular force^[1–4]. These experiments can help us directly to observe the mechanical behavior of proteins.

Proteins are compact polymers^[5], therefore some methods of studying the elastic behaviors of polymers can be adopted in investigating the mechanical properties of proteins. In fact, much works have been done on the elasticity of polymers. Much effort was put into network theories based on the non-Gaussian models^[6]. Later, Abe and Flory put forward the rotational-isomeric state (RIS) theory based on the effect of elongation of a polymer chain on the apportionment of its bonds and bond sequences among various RIS^[7]. Curro and Mark provided a non-Gaussian theory of rubber-like elasticity based on rotational isomeric state simulations of network chain configurations and investigated the rubber-like elasticity from the distribution functions for the end-to-end separation r of the chains using the Monte Carlo method^[8, 9]. Recently, Makarov *et al.* used Monte Carlo simulations of unfolded polypeptide chains to examine the average distance between the ends of a polypeptide chain as a function of the force that pulls these ends apart, and compared with the results of cantilever-type unfolding experiment^[10]. However, the interactions between residues in protein molecules are ignored in their calculation. In fact, the interactions between residues in protein molecules are very important. Perhaps, some special behaviors of elastic forces in protein molecules depend mainly on the residue interactions. In our previous

^{*} The research was financially supported by the National Natural Science Foundation of China (Nos. 20174036, 20274040), and the Natural Science Foundation of Zhejiang Province (No. R404047).

^{**} Corresponding author: Lin-xi Zhang (章林溪), E-mail: lxzhang@hzcnc.com

Received April 27, 2004; Revised July 6, 2004; Accepted July 9, 2005

work, we have discussed the elastic behavior of short compact polymers using the enumeration calculation method and the HP model^[11]. Nevertheless the actual protein chains are long, and always consist of 100-400 residues. So in this work, we will discuss the elastic behavior of long protein-like single chains by considering the long-range contact interactions. In fact, single proteins always exist in biological organisms. Our purpose here is to investigate the conformation properties and elastic behaviors of protein-like single chains in the process of tensile elongation in more detail.

METHOD OF CALCULATION

The hybrid (310) lattice model of Kolinsk *et al.* for protein-like single chain is adopted here^[12, 13]. According to this model, the residues in the protein-like single chains are represented by C^α only^[12, 13]. The positions of residues are limited to a quasi crystalline lattice. A very flexible hybrid (310) lattice with a high coordination number $z = 90$ is chosen^[12, 13]. This lattice is constructed from the vectors belonging to one of the several sets — namely (3,1,1), (3,1,0), (3,0,0), (2,1,1) and (2,2,0) — with all coordinates of the position and a sign. Here all short-range interactions have been neglected, and the long-range interactions are counted for $|j-i| > 3$. The potential of the long-range interactions has the form of a square well function given below^[12, 13]:

$$E_{ij} = \begin{cases} E_{rep} & \text{for } r_{ij} < 4 \\ \epsilon_{ij} & \text{for } 4 \leq r_{ij} \leq 6 \\ 0 & \text{for } r_{ij} > 6 \end{cases} \quad (1)$$

where $E_{rep} = 5k_B T$ strong repulsive potential, and r_{ij} is the distance between the i^{th} and j^{th} residues. The values of ϵ_{ij} depends on the types of i^{th} and j^{th} residues, and it can take the values of $\epsilon_{HH} = -2k_B T$, $\epsilon_{PP} = -1k_B T$ and $\epsilon_{HP} = 0$ ^[12, 13]. All energy parameters are defined in dimensionless $k_B T$ units. In all cases, the total conformational energy is the sum of the pairwise contributions for protein-like polymers.

The classical Monte Carlo sampling algorithm is employed here. The model chains undergo a series of local micro-modifications. The motion is based on vectors belonging to one of the several types: (100) and (110) with a coordination number $z = 18$, with which the new dynamics of the chain still belong to the hybrid (310) lattice model. The two-bond kink motion and chain-end motion are used here. During the Monte Carlo simulation, the system is gradually to be in the balanceable states because a set of elementary motions are efficient and reproduce the new dynamics of the chain which is adopted following the Metropolis sampling algorithm.

When a force acts on a protein-like single chain, the atoms in the chain move a distance r along the direction of the force. Then a lot of conformations vanish and the number of the conformations decreases. If the force $f = 0$, this is an isolated chain, and we assume that the minimum of end-to-end distance is $r_{\min}$. If the force $f \neq 0$, the minimum of end-to-end distance in the elongated direction becomes $r_{\min} + r$, and those conformations whose end-to-end distance are less than $r_{\min} + r$ will vanish.

If the force acts on the chain in the direction of x -axis, the partition function of the system is

$$Z(r) = \sum_i \exp(-E_i / k_B T) \quad (2)$$

where $\sum_i$ is the sum of conformations whose x -axis component of end-to-end distance is greater than $r_{\min} + r$. As we know that the partition function is very important because all thermodynamic properties can be expressed through $\ln Z(r)$ and its partial differential. We define the elongation ratio λ as

$$\lambda = \frac{\sqrt{\langle R^2 \rangle_r}}{\sqrt{\langle R^2 \rangle_0}} \quad (3)$$

where $\langle R^2 \rangle_r$ is the mean-square end-to-end distance of protein-like chains with moving r along the direction of force and $\langle R^2 \rangle_0$ is the mean-square end-to-end distance of protein-like chains without elongation. In fact, the mean-square of end-to-end distance $\langle R^2 \rangle_r$ is obtained from

$$\langle R^2 \rangle_r = \frac{\sum_i R_i^2 e^{-E_i / k_B T}}{\sum_i e^{-E_i / k_B T}} = \frac{\sum_j R_j^2 e^{-E_j / k_B T}}{M} \quad (4)$$

where $\sum_i$ is the sum of conformations whose x -axis component of end-to-end distance is greater than $r_{\min} + r$ in all conformations, and $\sum_j$ is the sum of conformations whose x -axis component of end-to-end distance is greater than $r_{\min} + r$ in our Monte Carlo simulation samples. Here M is the number of samples.

The radius of gyration tensor S is also considered here and it is used to measure the shape of protein-like polymer chains

$$S = \frac{1}{n+1} \sum_{i=0}^n S_i S_i^T = \begin{pmatrix} S_{xx} & S_{xy} & S_{xz} \\ S_{yx} & S_{yy} & S_{yz} \\ S_{zx} & S_{zy} & S_{zz} \end{pmatrix} \quad (5)$$

where $S_i = \text{col}(x_i, y_i, z_i)$ is the position of bead i in a frame of reference with its origin at the center of mass, and the tensor S can be diagonalized to form a diagonal matrix with three eigenvalues L_1^2 , L_2^2 and L_3^2 ($L_1^2 \leq L_2^2 \leq L_3^2$).

The energy of protein-like polymer chain is important, and the average energy is define as

$$\langle U \rangle = \frac{\sum_i E_i e^{-E_i / k_B T}}{\sum_i e^{-E_i / k_B T}} = \frac{\sum_j E_j e^{-E_j / k_B T}}{M} \quad (6)$$

where $\sum_i$ and $\sum_j$ are the same as Eq. (4).

From the partition function, we can obtain Helmholtz free energy $\langle A \rangle$ of protein-like chains

$$\langle A \rangle = -k_B T \ln Z(r). \quad (7)$$

Here $Z(r)$ is defined by Eq. (2). The number of conformations is very large for long protein-like polymer, therefore it is difficult to count all conformations according to Eq. (2). In fact, $Z(r)$ is also calculated through^[14-17]

$$Z(r) = Z_0 \times \frac{M_r}{M} \quad (8)$$

where Z_0 is the partition function of the system without elongation ($\lambda = 1.0$), M is total number of samples, and M_r is the number of samples whose x -axis component of end-to-end distance is greater than $r_{\min} + r$. Therefore,

$$\langle A \rangle = -k_B T \ln Z_0 - k_B T \ln \frac{M_r}{M} \quad (9)$$

For simplicity, we define $\langle A' \rangle$ as

$$\langle A' \rangle \equiv \langle A \rangle + k_B T \ln Z_0 = -k_B T \ln \frac{M_r}{M}. \quad (10)$$

In fact, Z_0 is independent of elongation ratio, and only depends on chain length and sequence of protein-like single chain. In the meantime, the force value f can be obtained from the dependence of Helmholtz free energy $\langle A \rangle$ on the elongated distance along the force direction^[14-17]

$$f = \frac{\partial \langle A \rangle}{\partial r} \equiv \frac{\partial \langle A' \rangle}{\partial r}. \quad (11)$$

According to Newton's third law, force f is the elastic force stored in the polymer chains. The energy contribution and the entropy contribution to the elastic force f_U and f_S are also defined as

$$f_U = \frac{\partial \langle U \rangle}{\partial r} \quad (12)$$

$$f_S = \frac{\partial (-TS)}{\partial r} = f - f_U. \quad (13)$$

RESULTS AND DISCUSSION

Average Conformations

The properties of protein-like single chains are analyzed during the final period of simulation of about 10^7 MCS for each run, and there are 100 runs in our simulation. Here the number of samples M is 100000 for each sequence. In order to find out the interior conformation changes in the process of tensile elongation, we first calculate the mean-square end-to-end distance $\langle R^2 \rangle_r$, and the radius of gyration tensor S for 40-bond, 60-bond, and 80-bond chains with $(H)_x$ and $(HP)_x$ sequences, and the results are given in Figs. (1)–(3).

The mean-square end-to-end distance $\langle R^2 \rangle_r$ as a function of elongation ratio λ for 40-bond, 60-bond, and 80-bond protein-like single chains with $(H)_x$ and $(HP)_x$ sequences are shown in Fig. 1. In Fig. 1 $\langle R^2 \rangle_r$ increases with elongation ratio λ in either $(H)_x$ or $(HP)_x$ sequences. In the process of tensile elongation, the force acting on protein-like single chains does work, therefore the average distances between atoms become large, leading to the increase of mean-square end-to-end distance $\langle R^2 \rangle_r$. For 40-bond protein-like single chains, $\langle R^2 \rangle_r$ increases slowly with elongation ratio λ in either $(H)_x$ or $(HP)_x$ sequences. For 80-bond protein-like single chains $\langle R^2 \rangle_r$ increases fast with elongation ratio λ in either $(H)_x$ or $(HP)_x$ sequences. We also notice that the mean-square end-to-end distance $\langle R^2 \rangle_r$ of $(H)_x$ protein-like single chain is larger than that of $(HP)_x$ chains at the same elongation ratio.

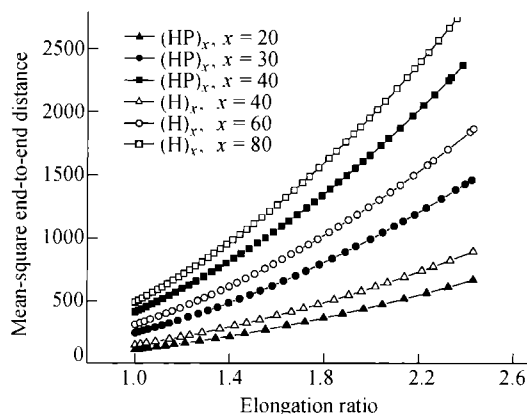


Fig. 1 Mean-square end-to-end distance $\langle R^2 \rangle_r$ as a function of elongation ratio λ for different protein-like single chains

To investigate the shape of protein-like single chains in more detail, here we also calculate the radius of gyration tensor S . The tensor S can be diagonalized to form a diagonal matrix with three eigenvalues L_1^2 , L_2^2 and L_3^2 ($L_1^2 \leq L_2^2 \leq L_3^2$). Solc and Stockmayer first used the ratio of $\langle L_1^2 \rangle : \langle L_2^2 \rangle : \langle L_3^2 \rangle$ to describe the shape of polymer chains, and they estimated the ratio to be 1:2.7:11.7 based on a random walk of polymer chain^[18]. We also calculate the ratio of $\langle L_1^2 \rangle : \langle L_2^2 \rangle : \langle L_3^2 \rangle$ of protein-like polymer chains, and the results are shown in Fig. 2. We find that the ratio of $\langle L_1^2 \rangle : \langle L_2^2 \rangle : \langle L_3^2 \rangle$ depends mainly on the sequence of protein-like single chains. In

Fig. 2(a), the ratio of $\langle L_2^2 \rangle / \langle L_1^2 \rangle$ decreases with elongation ratio λ for short $(HP)_x$ protein-like single chains, and increases with elongation ratio λ for long protein-like chains, especially for $(H)_x$ sequence. In the meantime, the ratio of $\langle L_2^2 \rangle / \langle L_1^2 \rangle$ increases with chain length of protein-like chains, especially for $(H)_x$ sequence. For example, the ratio of $\langle L_2^2 \rangle / \langle L_1^2 \rangle$ without elongation ($\lambda = 1.0$) is 1.52, 1.61, and 1.88 for 40-bond, 60-bond, and 80-bond $(HP)_x$ protein-like chains, and 1.61, 2.02, and 2.51 for 40-bond, 60-bond, and 80-bond $(H)_x$ protein-like chains, respectively. However, the ratio of $\langle L_2^2 \rangle / \langle L_1^2 \rangle$ at $\lambda = 1.80$ is 1.47, 1.55, and 2.00 for 40-bond, 60-bond, and 80-bond $(HP)_x$ protein-like chains, and 1.47, 2.30, and 3.06 for 40-bond, 60-bond, and 80-bond $(H)_x$ protein-like chains, respectively. Therefore, the shape of $(H)_x$ protein-like single chain is different from $(HP)_x$ one. In fact, $(H)_x$ protein-like polymer is homopolymeric chain, and $(HP)_x$ protein-like single chain is heteropolymeric with type pattern $-HP-^{[12, 13]}$. In Fig. 2(b), the ratio of $\langle L_3^2 \rangle / \langle L_1^2 \rangle$ increases with elongation ratio λ , especially for $(H)_x$ chains. This means that it is easy to form a flat structure in the force direction for $(H)_x$ protein-like single chains.

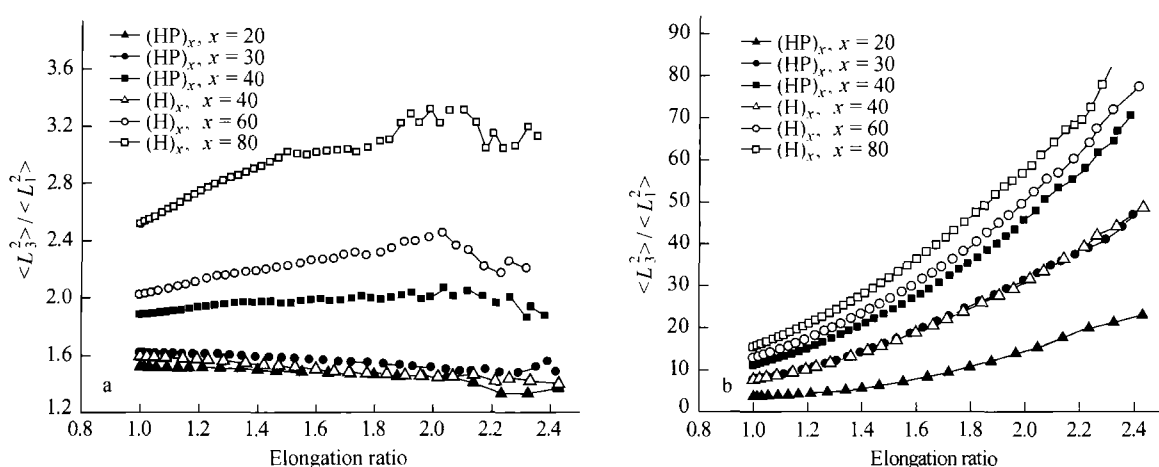


Fig. 2 Ratio of $\langle L_2^2 \rangle / \langle L_1^2 \rangle$ (a) and $\langle L_3^2 \rangle / \langle L_1^2 \rangle$ (b) as a function of elongation ratio λ for different protein-like single chains

Thermodynamics Properties

Energy changes always exist in the process of tensile elongation. In fact, change of conformations leads to the change of thermodynamics property of protein-like polymer chains. Figure 3 presents the average energy per bond $\langle U \rangle$ as a function of elongation ratio λ for different protein-like single chains. First, the average energy per bond of $(HP)_x$ protein-like single chain is higher than that of $(H)_x$ chain. The reason is that there are $\epsilon_{HH} = -2k_B T$, $\epsilon_{PP} = -1k_B T$ and $\epsilon_{HP} = 0$, respectively, and there is more attraction between HH pair. Meanwhile, average energy per bond $\langle U \rangle$ increases with elongation ratio λ , especially for $(H)_x$ chains, and it decreases with chain length. The difference of average energy per bond with chain length for $(HP)_x$ protein-like single chain is less than that for $(H)_x$ one. It suggests that $(H)_x$ protein-like single chains may be combined firmer than $(HP)_x$ ones. According to Eq. (1), if $r_{ij} < 4$, the long-range contact potential will be repulsive ($\epsilon_{rep} = +5k_B T$). Therefore $(HP)_x$ protein-like single chain is more compact and its energy is higher. In addition, we can see that long protein-like single chain has a low energy.

Average value of $\langle A' \rangle$ per bond as a function of elongation ratio is also shown in Fig. 4. As Z_0 is independent of elongation ratio, and only depends on chain length and sequence of protein-like chain, average value of $\langle A' \rangle$ per bond also represents the change of Helmholtz free energy per bond in the process of tensile elongation. Therefore, average value of $\langle A' \rangle$ per bond is also important. In Fig. 4, $\langle A' \rangle$ increases with elongation ratio λ for different protein-like single chains. In the process of tensile elongation, the number of vanishing conformation increases little by little. The curve of 40-bond chain is greater than that of 60-bond chain and the curve of 60-bond is greater than that of 80-bond at the same elongation ratio for either $(H)_x$ or $(HP)_x$

protein-like single chains. However, with the increasing of elongation ratio λ , the gap of the curves gets bigger. Obviously, the value of $\langle A' \rangle$ for 40-bond $(HP)_x$ protein-like single chain is greater than that of 40-bond $(H)_x$ one. However, there is no difference for 60-, 80-bond $(HP)_x$ and 60-, 80-bond $(H)_x$ protein-like single chain.

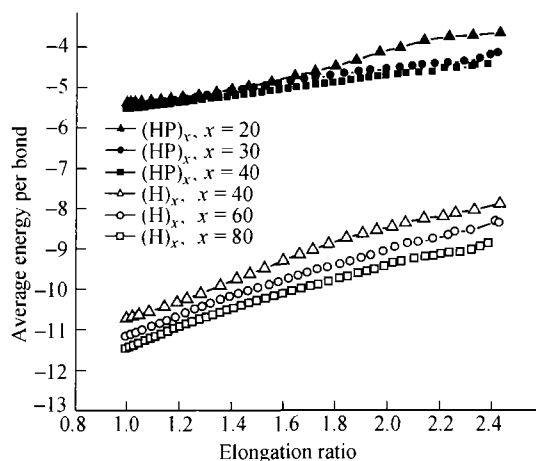


Fig. 3 Average energy per bond $\langle U \rangle$ as a function of elongation ratio λ for different protein-like single chains

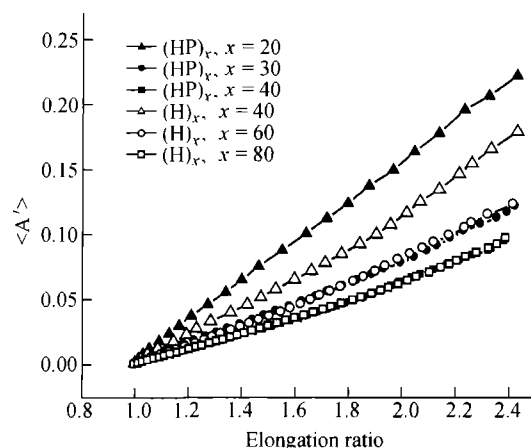


Fig. 4 Average value of $\langle A' \rangle$ per bond as a function of elongation ratio λ for different protein-like single chains

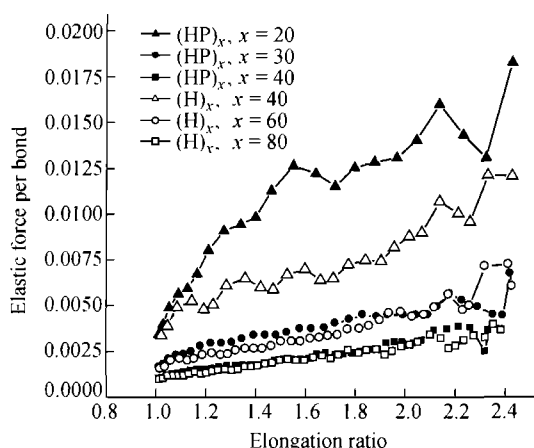


Fig. 5 Elastic force per bond f as a function of elongation ratio λ for different protein-like single chains

The elastic force stored in protein-like polymer chain is calculated according to Eqs. (11–13), and the results are given in Figs. 5 and 6. In Fig. 5 we can see that the elastic force per bond f increases slowly with elongation ratio λ . The elastic force per bond for 40-bond $(HP)_x$ protein-like single chain is greater than that for 40-bond $(H)_x$ one and the elastic force per bond for 60-bond and 80-bond $(HP)_x$ and $(H)_x$ protein-like single chains are very close.

It is more important for the energy contribution to force f_U in the process of tensile elongation. In Fig. 5(a), the curves increase first, then they increase slowly to some extent, at last, they drop gradually. The curve of 40-bond chain is greater than that of 60-bond chain and the curve of 60-bond chain is greater than that of 80-bond chain, especially for short protein-like single chain. This result is similar to short compact chains^[11], and is related with the protein structure. Comparing with Figs 5 and 6(a), we can know that the value of the energy contribution to force f_U is greater than that of the elastic force per bond f . Therefore, it is necessary to consider the interactions between residues in studying the elastic behaviors of protein molecules.

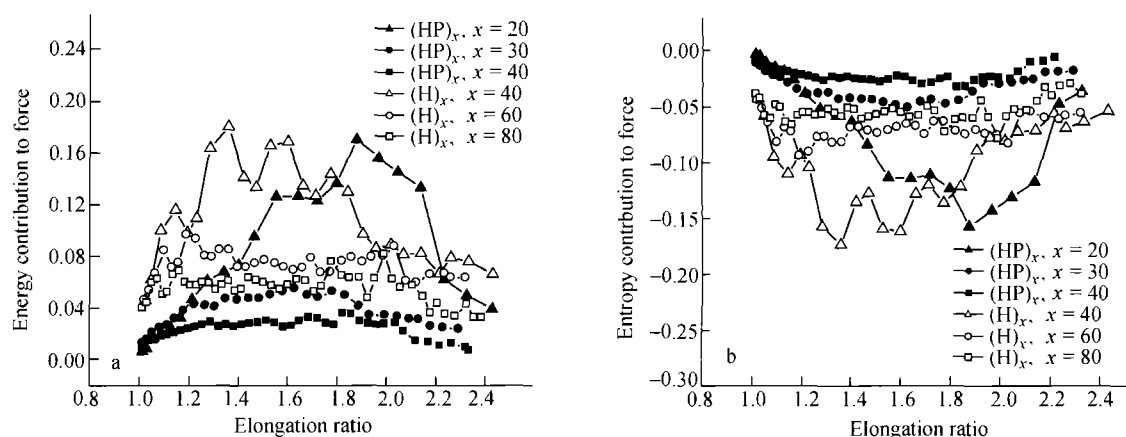


Fig. 6 Energy contribution to force f_U (a) and entropy contribution to force f_S (b) as a function of elongation ratio λ for different protein-like single chains

We also calculate the entropy contribution to force f_S according to Eq. (13), and the results are shown in Fig. 6(b). The curves drop first, then increase slowly. The entropy contribution to force f_S for 80-bond protein-like single chain is greater than that for 60-bond (or 40-bond) one. Although protein structure is very complex and it is difficult to understand proteins clearly, our investigation may be helpful to comprehend the microscopic mechanism of proteins.

REFERENCES

- 1 Pennisi, M.E., Science, 1999, 283: 168
- 2 Marsalek, P.E., Lu, H., Li, H., Vazquez, M.C., Oberhauser, A.F., Schulten, K. and Fernandez, J., Nature, 1999, 402: 100
- 3 Rief, M., Fernandez, J.M. and Gaub, H.E., Phys. Rev. Lett., 1998, 81: 4764
- 4 Rief, M., Gautel, M., Oesterhelt, F., Fernandez, J.M. and Gaub, H.E., Science, 1997, 276, 1109
- 5 Chan, H.S. and Dill, K.A., Macromolecules, 1998, 22, 4559-75
- 6 Treloar, L.R.G., Trans. Faraday Soc., 1946, 42, 77
- 7 Abe, Y. and Flory, P.J., J. Chem. Phys., 1970, 52, 2814
- 8 Curro, J.G. and Mark, J.E., J. Chem. Phys., 1984, 80, 4521
- 9 Mark, J.E., J. Phys. Chem. B 2003, 107, 903
- 10 Makarov, D.E., Wang, Z., Thompson, J.B. and Hansma, H.G., J. Chem. Phys., 2002, 116, 17
- 11 Sun, T.T., Zhang, L.X., Chen, J. and Shen, Y. J. Chem. Phys., 2004, 120, 5469
- 12 Kolinski, A. and Madziar, P., Biopolymers, 1997, 42, 537
- 13 Romiszowski, P. and Silorski, A., Biopolymers, 2000, 54, 262
- 14 Yang, X.Z. and Li, X.F., Chinese J Polym Sci., 1998, 16, 279
- 15 Zhang, L.X., Huang, Y.X. and Zhao, D.L. Macromol. Theory Simul., 2001, 10, 479
- 16 Zhang, L.X., Xia, A.G., Jiang, Z.T. and Zhao, D.L. Macromol. Theory Simul., 2001, 10, 651
- 17 Zhang, L.X., Jiang, Z.T. and Zhao, D.L. J. Polym. Sci. Polym. Phys., 2002, 40, 105
- 18 Solc, K. and Stockmayer, W. H. J. Chem. Phys., 1971, 54, 2756