Computer Simulation of Globules with Microstructure.

Victor A. Ivanov[1], Alexander V. Chertovich[1], Alexei A. Lazutin[1],
Nadezhda P. Shusharina[1], Pavel G. Khalatur[2], Alexei R. Khokhlov[1*]

[1]Physics Department, Moscow State University, 117234 Moscow, Russia

[2]Department of Physical Chemistry, Tver State University, 170002 Tver, Russia

SUMMARY: We present recent data of our Monte Carlo computer simulation study of properties of AB-copolymer globules which depend strongly on the primary sequence of A and B monomeric units. Different primary sequences have been studied: random, random-block, regular and designed ones by using some particular spatial conformation of a homopolymer chain (we have compared here three models: protein-like copolymers, AB-copolymers modeling membrane proteins and ABC-copolymers modeling proteins with active enzymatic center). We have found several evidences for the fact that an AB-copolymer chain with a primary sequence prepared on the basis of a particular conformation of a homopolymer chain by some "coloring" procedure preserves the "memory" about its "parent" spatial conformation. Analyzing the power spectra of AB-sequences, we find the existence of long-range power-law correlations for the copolymers with specially designed primary sequences.

Introduction

In this letter, we develope futher the recently proposed new approach[1,2] to the design of specific primary sequences for the copolymers consisting of monomeric units of two types (A and B). This approach is based on the "coloring" in two "colors" monomeric units of a homopolymer taken in some well-defined conformation (globular conformation[1,2,4], conformation of adsorbed chain[2,3], etc.[4]), depending on the spatial position of the unit in this "parent" conformation. It was shown in[1-4] that copolymers with AB-sequences generated in this way acquire a number of special functional properties which distinguish them from the AB-copolymers with random or block primary structures. In a sense, we can say that some features of the "parent" conformation are "memorized" (or "inherited") and then manifested in other conditions.

This general idea has been first illustrated taking as an example a globular conformation of a polymer chain [1,2]. To prepare the primary sequences we have taken some particular conformation of a usual homopolymer globule and assigned two different types A and B to monomeric units which differ by their spatial positions inside that globule due to some of criteria. We have used in the present paper three different criteria to split monomeric units into two or even three types. First, as it was also done in [1,2,4], we assigned the type A to

 CCC 1022-1360/99/$ 17.50+.50/0

the surface monomeric units and the type B to the monomeric units inside inner core of the "parent" globule (we called such chain a protein-like AB-copolymer). Computer realization of this procedure is illustrated in Figure 1. At first, we prepare a dense globular conformation of a homopolymer chain by switching on strong attraction between all monomeric units (Figure

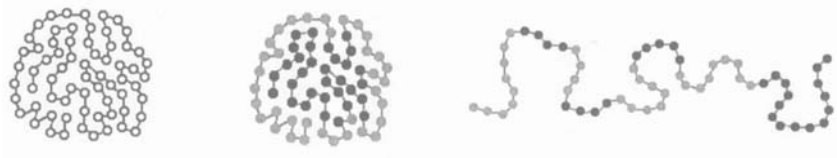

Fig.1. Main steps of the sequence design scheme for protein-like copolymers:
(a) homopolymer globule; (b) the same globule after "coloring" procedure;
(c) protein-like copolymer in the coil state.

1a). Then, we take the "instant snapshot" of the globule and assign the index A to those units that are on the surface of the globule and call these units hydrophilic, and assign the index B to the units in the core of the globule and call these units hydrophobic. Then we fix this primary structure (Figure 1b; the structure shown in this figure will be further referred to as "parent" globule). Finally, the last step was to remove a uniform strong attraction of monomeric units and to acquire different interaction potentials for A- and B-units (Figure 1c).

However, we consider in the present paper the two other new coloring criteria which are described below in two last sections calling them correspondingly AB-copolymers modeling membrane proteins and ABC-copolymers modeling proteins with active enzymatic center.

It should be mentioned here that real globular proteins are of course much more complicated objects than AB-copolymers. However, the most essential distinction between different monomeric unit of proteins is that some of these units are hydrophobic, while others are hydrophilic or charged[5], thus in a very rough approximation it is possible to represent a globular protein as a kind of AB-copolymer. The spatial (ternary) structure of such a copolymer in the native state would then normally correspond to the structure in which hydrophilic units (A-type) cover the globular surface and prevent different globules from aggregation, while hydrophobic units (B-type) constitute the globular core. The reader can find the detailed discussion of relation between the problem of construction of synthetic functional AB-copolymers and the problem of protein folding in paper[2] (and references therein).

Long-range power-law correlations in primary sequence

An important question is whether the "memory effect" is manifested in some long-range power-law correlations in primary sequences of protein-like AB-copolymers. To answer this question, we have applied two scaling methods to study these primary sequences. The first method is a standard power spectrum analysis (PSA)[6], but with power spectrum estimation using the maximum entropy (all poles) procedure[7]. The second method is the detrended fluctuation analysis (DFA)[8,9], which was designed to treat sequences with statistical heterogeneity. If a sequence has long-range correlations, then the power spectrum S(f) behaves as $S(f) \propto f^{-\beta}$, and the corresponding log-log plot of S(f) versus f (=1/N) is a straight line with slope $-\beta$. On the other hand, for sequences with power-law long-range correlations, the detrended fluctuation, $F_D(L)$, can be approximated by power law[8,9] $F_D(L) \propto L^{\alpha}$. For ideal power-law correlations, the two exponents α and β are related by $\alpha = (1 + \beta)/2 =$ constant[8,9]. In Figure 2, we show a log-log plot of the power spectrum S(f) for protein-like copolymer with N=1024. We note in this figure the presence of three spectral regions corresponding to relatively low-frequency, mid-frequency, and high-frequency regions. For the mid-frequency region, one has $\beta = 0.80 \pm 0.05$. In contrast, random AB-sequences have no correlations (S(f) = 2 for any f), i.e., the spectral exponent β is equal to zero. The same value of β is obtained for protein-like copolymers in the low-frequency region. Using the

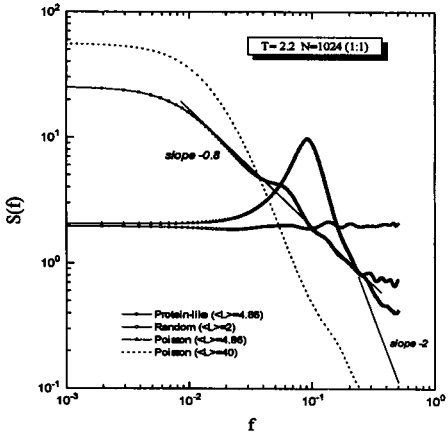

Fig.2.: The averaged power spectrum for AB-sequence of 1024-unit protein-like copolymer chain.

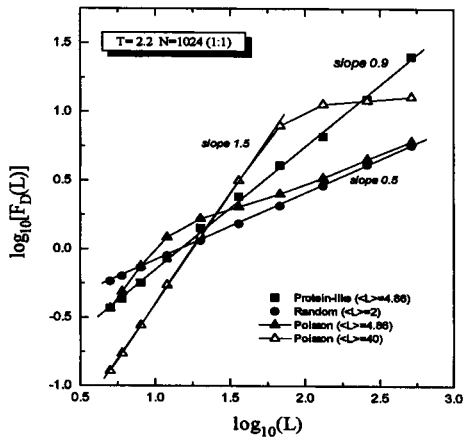

Fig.3.: Detrended fluctuation analysis.

DFA method (Figure 3), we find $\alpha = 0.90 \pm 0.05$ for protein-like copolymers and $\alpha = 1/2$ for random and random-block (Poisson)[1-4] copolymers. Moreover, for random-block copolymer there is a crossover between two different regimes at some characteristic length L^*. It should be stressed that there is no characteristic length scale for protein-like copolymers where power law with exponent larger than 0.5 holds on all length scales. Thus, we observe distinct long-range power-law correlations which are presented in protein-like AB-copolymers.

AB-copolymers modeling membrane proteins

As the second criteria for coloring monomeric units inside the dense homopolymer globule, we have introduced the model for AB-copolymers which mimic some properties of real membrane proteins. It is well known that the real membrane proteins are located inside the membrane in such a way that some part (about 30% from the whole number) of the aminoacid units (mainly the hydrophobic and uncharged ones) are located inside the bilipid layer of the membrane when the other aminoacid units are located in water environment

Fig.4. Originally colored membrane-protein-like globule and finally obtained after decollapse and further collapse procedure.

inside and outside the cell. In our rough model we assigned the type B to monomeric units which lie inside a cross-section of a parent globule by a narrow flat layer. So, the B-part of a parent conformation has a form of a narrow disk. We have taken 30% of all links to be of B-type. We present in the left part of Figure 4 the snapshot of an original conformation of just prepared AB-copolymer globule. We marked the both hemispheres of outer A-links (70% from the whole amount) of original globule into two different colors (black and grey) to see whether the "parent" microsegregated structure can be reestablished after the equilibration procedure. We have studied the conformations of such AB-copolymer chain for two different sets of interaction potential $(\varepsilon_{AA},\varepsilon_{AB},\varepsilon_{BB}) = (-1,0,-1)$ and $(-1,-1,-2)$. We will call these both sets below correspondingly set 1 and set 2. We have performed the procedure described above for the chains of N=256 units using for simulations the Monte Carlo (MC) method and the

bond fluctuation model[10]. This sequence design scheme was repeated many times, and the results were averaged over ~10^6 MC steps and different initial configurations.

We have indeed found that such a chain shows the effect of stability of a "parent" microsegregated structure (a typical conformation obtained after the procedure of decollapse and following collapse of that chain is shown on the right part of Figure 4). Depending on the interaction potential the B-core of the final conformation can be located inside the A-fringe (if A and B links like each other, set 2) or two dense cores of A and B link can be separated from each other (in the opposite case, set 1). But in both cases the spherical B-core is formed instead of original disk-like B-core what is, of course, natural due to isotropy of interaction potential. We have found also that for set 2 of interaction potential the effect of stability of "parent" microsegregated structure is more pronounced than for set 1. In other words, we can say that the protein-like copolymer "inherited" (or "memorized") some important structural features of the "parent" globule which were then reproduced in the other conditions.

ABC-copolymers modeling proteins with active enzymatic center

As the third criteria of preparation of primary structure of copolymer chain, we have studied the ABC-copolymer prepared in the course of a "triple coloring" of some particular homopolymer globule in the following way: we assigned the type A to the surface monomeric units, the type B - to the inner monomeric units (as it has been previously done for protein-like copolymers), and the type C to those inner monomeric units which lie inside a small sphere which center does not coincide, however, with the center of mass of the parent homopolymer globule (see Figure 5). Our idea was to prove whether such a "parent"

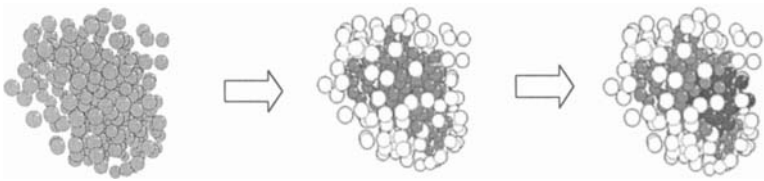

Fig.5. Coloring procedure for ABC-copolymer with active C-center.

conformation can be recombined in coarse of equilibration procedure for different set of interaction parameters including the restoration of the originally given distance between the centers of B-core and C-core. We intended to prove the idea whether the position of C-links

inside the primary sequence together with specially chosen interaction potential can lead to stable reconstruction of spatial conformation of the whole chain.

We performed computer simulations for the chain of N=256 monomeric units using the following two sets of interaction potentials ($\varepsilon_{BB},\varepsilon_{BC},\varepsilon_{CC}$) = (-1,0,-2) and (-1,-1,-2) (sets 1 and 2 correspondingly). We have found in our computer experiment that such ABC-copolymers normally restore their original structure with B- and C-cores although we have not succeeded up to now in determination of interaction potential which would allow us to get the center of C-core at the same distance from the center of B-core as in original conformation. For set 1 we found the distance between C- and B-core to be larger than in "parent" conformation (C-core is pushed outside the B-core) while for set 2 we found the decrease of this distance (C-core is involved deeper inside the B-core). Nevertheless, we have definitely found the effect of restoration of active center after the following procedure: we switch off the attraction between C-links and let them "dissolve" inside the dense B-core; and after the switching of attraction between C-links we observe the restoration of C-core again.

Conclusion

The properties of AB-copolymer chains with conformation-dependent primary AB-sequences have been studied by means of a Monte Carlo computer simulation. Different primary sequences have been studied: random, random-block, regular and designed ones by using some particular spatial conformation of a homopolymer chain. We have found several evidences for the fact that an AB-copolymer chain with a primary sequence prepared on the basis of a particular conformation of a homopolymer chain by some "coloring" procedure preserves the "memory" about its "parent" spatial conformation. We have found the existence of long-range power-law correlations for the copolymers with specially designed primary sequences. We have introduced two new models for AB-copolymers modeling membrane proteins and for ABC-copolymers modeling proteins with active enzymatic center. We have preformed computer simulation for these models and compared the effects of "memory" about the parent structure. It is important also to emphasize the effect of "memorizing" of some features of the "parent" conformation by the AB-copolymers generated according to our sequence design scheme. These features are then manifested in other conditions. Such an interrelation can be regarded as one of the possible mechanisms of molecular evolution: polymer acquires some special primary sequence in the "parent" conditions and then (in other conditions) uses the fact that primary structure is "tuned to perform certain functions".

References

1. A.R.Khokhlov, P.G.Khalatur, *Physica A* **249**, 253 (1998)
2. A.R.Khokhlov, P.G.Khalatur, *Phys.Rev.Lett.* **82**, 3456 (1999)
3. E.A.Zheligovskaya, P.G.Khalatur, A.R.Khokhlov, *Phys.Rev.E* **59**, 3071 (1999)
4. A.R.Khokhlov, V.A.Ivanov, N.P.Shusharina, P.G.Khalatur, in: *The Physics of Complex Liquids,* F.Yonezawa et al (Eds.), World Scientific, Singapore 1998, p.155
5. R.E.Dickerson and I.Geis, *The Structure and Action of Proteins*, Harper & Row, New York 1969
6. R.Voss, *Phys. Rev. Lett.* **68**, 3805 (1992)
7. W.H.Press, B.P.Flannery, S.A.Teukolsky, and W.T.Vetterling, *Numerical Recipes. The Art of Scientific Computing*, Cambridge University Press, Cambridge 1986, p. 430
8. C.-K.Peng, S.V.Buldyrev, S.Havlin, M.Simons, H.E.Stanley, and A.L.Goldberger, *Phys. Rev.* E **49**, 1685 (1994)
9. S.V.Buldyrev, A.L.Goldberger, S.Havlin, R.N.Mantegna, M.E.Matsa, C.-K.Peng, M.Simons, and H.E.Stanley, *Phys. Rev.* E **51**, 5084 (1995)
10. I.Carmesin and K.Kremer, *Macromolecules* **21**, 2819 (1988)