Sequence Design of Biomimetic Copolymers: Modeling of Membrane Proteins and Globular Proteins with Active Enzymatic Center

A.V. Chertovich, V.A. Ivanov, A.A. Lazutin, A.R. Khokhlov *

Physics Department, Moscow State University, 119899 Moscow, Russia

E-mail: khokhlov@polly.phys.msu.su

SUMMARY: The biomimetic approach to the sequence design of synthetic AB-copolymers has been developed further by means of new series of Monte Carlo computer simulation. The approach is based on using of some particular conformation of a homopolymer chain for "coloring" of monomeric units into two "colors" (or types) A and B depending on the spatial position of particular monomeric unit. We present recent data of our Monte Carlo computer simulation studies of properties of designed AB-copolymers which mimic membrane proteins, and designed ABC-copolymers which mimic proteins with active enzymatic center. We have found further evidences for the fact that designed copolymer chain preserves the "memory" about its "parent" spatial conformation and shows the well-pronounced tendency to restore main features of the "parent" conformation.

Introduction

The so-called biomimetic approach has caused increasing interest in the last decade. This is connected with high promises of these ideas both for industrial applications (production of synthetic polymeric materials exhibiting complex functional properties) and for possible understanding of early stages of pre-biological molecular evolution. This approach consists in moving towards the construction (or design) of synthetic polymer systems having rather sophisticated and complex functions by implementing some ideas based on observation of functioning of real biopolymers in vivo.

Recently, a new method of the so-called conformation-dependent sequence design of AB-copolymers has been proposed in our group [1, 2)] which lies in the framework of this biomimetics paradigm. This sequence design scheme is based on the "coloring" of monomeric units of a homopolymer taken in some well-defined "parent" conformation (globular conformation [1, 2, 4)], conformation of adsorbed chain [2, 3)], etc. [4)]), depending on the spatial

 CCC 1022-1360/00/$ 17.50+.50/0

position of the unit in this "parent" conformation. It was shown [1-4] that copolymers with two-color 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.

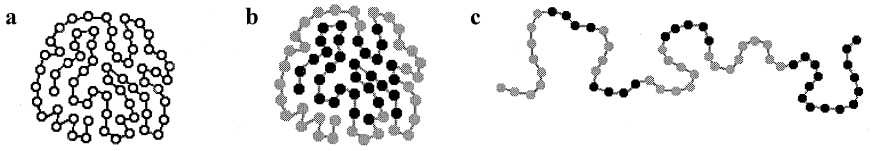

Fig.1. Main steps of the sequence design scheme for protein-like copolymers:
(a) homopolymer globule; (b) the same globule after "coloring" procedure;
(c) protein-like copolymer in the coil state.

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 the type A to the monomeric units belonging to the surface of the globule and the type B to the monomeric units inside inner core of the globule (we called the chain generated in this way 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 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).

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. However, we would like to emphasize here once more that we are not speaking about the description of the real biopolymers but rather about the most universal principles of evolution which could lead to formation of currently existing biomacromolecules.

In principle, any special macromolecular conformation can play the role of a parent one. For example, the conformation of a polymer chain adsorbed on a plane surface has been considered in this context [3]. The monomer units closest to the surface in some instant snapshot conformation were assigned to be A-units, others became B-units. We called thus obtained AB-chain an adsorption-tuned copolymer. It was shown that this copolymer adsorbs on another plane surface (to which only A-units are attracting) more efficiently than random and random-block copolymers with the same AB-composition and the same degree of blockiness.

In the present paper we developed further this method to the design of specific primary sequences for AB- and ABC-copolymers on the basis of the spatial position of a particular monomeric unit inside globular conformation of a homopolymer chain. We consider two new coloring criteria which are described below calling the resulting sequences correspondingly AB-copolymers mimicking membrane proteins and ABC-copolymers mimicking proteins with active enzymatic center.

AB-copolymers mimicking membrane proteins

We introduce here the following 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% of the whole number) of the aminoacid units (mainly the hydrophobic ones) are located inside the bilipid layer of the membrane, while other aminoacid units (mainly hydrophilic) are located outside the area of bilipid layer. In our rough model the main steps of the sequence generation that we perform to mimic such polymers are the same as in Figure 1 except we assigned the type B ("hydrophobic units") 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 2a the snapshot of an original conformation of just prepared AB-copolymer globule (this Figure 2a

replaces the Figure 1b in the preparation scheme). In Figure 2a we marked 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 refolding of such a globule. We have studied the conformations of such AB-copolymer chain for the following values of attraction interaction energies $\varepsilon_{AA} : \varepsilon_{AB} : \varepsilon_{BB} = 1 : 1 : 2$, i.e. there is an attraction between all types of monomeric units, but the attraction between B-units is twice stronger. We have performed the procedure described

Fig.2. Originally colored membrane-protein-like globule (a) and the globule finally obtained after refolding (b).

above for the chain of N=256 units using for simulations the Monte Carlo (MC) method and the bond fluctuation model [6-8]. 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 refolding of that chain is shown in Figure 2b). The spherical B-core is formed instead of original disk-like B-core; that is, of course, natural due to isotropy of interaction potential. But one can see definitely that the grey A-units have much more contacts with each other than with the black A-units and vice versa, i.e. the grey and black A-units are segregated from each other although the interaction potential is the same for contacts between all A-units and between B-units and all A-units.

To answer the question what is the reason for this microsegregation we have divided all blocks of B-units in the primary sequence in the two groups. In the first group we have included those B-blocks which have on their both ends the A-units from the same hemisphere in the "parent" conformation, i.e. the A-units of the same "subcolor" (A-units can be either grey or black). We called such B-blocks the "loop" B-blocks: the first and the last B-units of this block contact with the same "parent" A-hemisphere. The second group corresponds to the

B-blocks which are connecting different "parent" A-hemispheres, we call them "bridging" blocks. We have analyzed the distribution functions of the length of "loop" and "bridging" B-blocks. The histograms normalized by the total number of "loop" and "bridging" B-blocks over which the averaging has been performed are presented in Figure 3 for three different A/B compositions. One can see that the "loop" B-blocks of short length are the most probable ones (Figure 3a). Contrary, the histogram of "bridging" B-blocks (Figure 3b) shows a

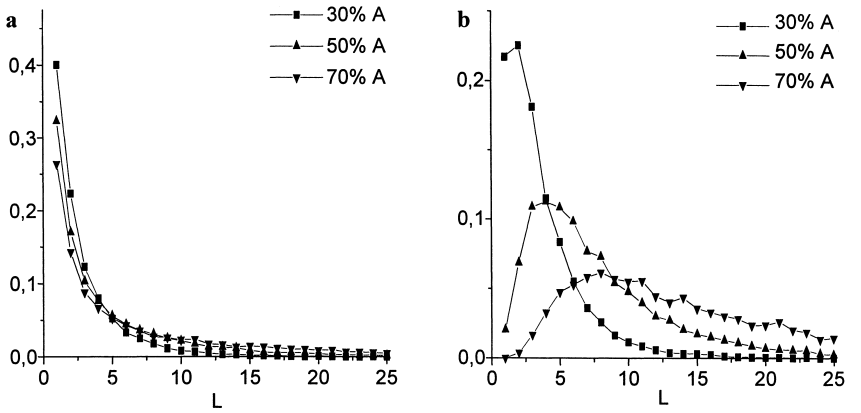

Fig. 3. Histograms of the length of "loop" (a) and "bridging " (b) B-blocks for a membrane protein-like copolymer.

maximum at some length representing the fact that the segment should have a certain length to be able to bridge two different hemispheres. Upon increasing the fraction of B-units in the sequence from 30% to 70% the histogram of "loop" segments does not show any significant changes (Figure 3a) while the position of the maximum on the histogram of "bridging" blocks is shifted to larger length (Figure 3b) that can be easily explained by the increase of the width of B-sheet in the "parent" conformation.

The results of Figure 3 can be used to explain the effect of re-emerging of two hemispheres upon refolding (Figure 2b). Indeed, "loop" B-segments are very small, therefore after refolding they automatically find themselves on one of the A/B interfaces, therefore they remain "loop" segments. "Bridging" B-blocks are much longer, they have a good chance to remain "bridges" after refolding, and they normally use this chance in order not to violate the overall topology of parent-like spatial arrangement where different parts of the globule are fitted to each other.

ABC-copolymers mimicking proteins with active enzymatic center

To check the possibility to "memorize" the more complicated internal microstructure we have studied the designed 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 ("hydrophilic") monomeric units, the type B to the inner ("hydrophobic") monomeric units (as it has been previously done for protein-like copolymers), and the type C to those inner (i.e. also "hydrophobic") 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 4). Our idea was to check whether such a "parent" conformation can be recovered upon refolding. In particular, we intended to check 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.

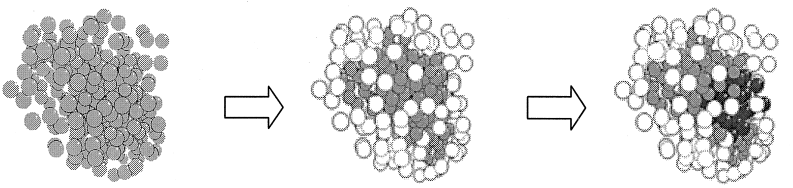

Fig. 4. Coloring procedure for ABC-copolymer with active C-center.

We performed computer simulations for the chain of N=256 monomeric units using the following values of attraction interaction energies $\varepsilon_{BB} : \varepsilon_{BC} : \varepsilon_{CC} = 1 : 1 : 2$, i.e. the contacts between two C-units were more favorable than the B-B and B-C contacts. We have found in our computer simulation that such ABC-copolymers normally restore their original structure with B- and C-cores. Moreover, we have definitely found the effect of restoration of active center after the following procedure: first, we switch off the attraction between C-links and let them "dissolve" inside the dense B-core; then, after the switching on attraction between C-links, we observe the restoration of C-core again. From our simulations it is more or less obvious that the "parent" conformation (including the distance between the centers of B- and C-core) can be reassembled quite exactly by using appropriate interaction potential.

We have compared the properties of ABC-copolymers with designed (as described above) primary sequence with the properties of quasi-random ABC-copolymers. The quasi-random sequence was prepared in the following way. At the first stage we prepared the protein-like

AB-copolymer according to the original scheme described in the Introduction. Then we have "recolored" some of the internal B-units into C-units randomly, i.e. the BC-sequence is a purely statistical one. The attraction between C-units in this case was also taken to be stronger than that between B-units. We have found that ABC-copolymers with our designed primary sequence forms larger clusters of active units than this quasi-random ABC-copolymer.

To investigate the influence of the primary sequence on the formation of the selected core in the globule we have considered the case of equally favourable attraction for B-B, B-C and C-C contacts. In this case there is no difference between B- and C-links from energetical point of view. We have found the pronounced difference between the histograms of C-cluster size for designed (Figure 5b) and quasi-random ABC-sequences (Figure 5a) in globular state. One can see from Figure 5 that in the same spatial conformation (at the same temperature) there

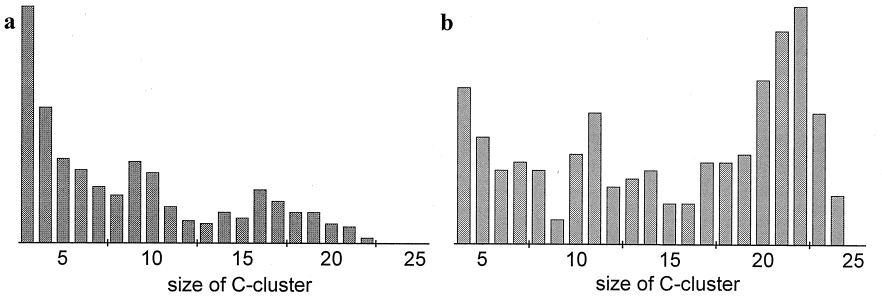

Fig. 5. Histograms (non-normalized) of the C-units cluster size for quasi-random (a) and designed (b) primary sequences of B- and C-units. (B- and C-units are equivalent from the viewpoint of energies of attraction).

are rather small C-clusters which dominate for the case of quasi-random sequence while for the designed sequence the large C-clusters which include almost all C-links have much larger probability (there were totally 25 C-units in the chain).

Conclusion

The properties of AB-copolymer chains with specially designed AB-sequences have been studied by means of Monte Carlo computer simulation for AB-copolymers which mimic some properties of membrane proteins and for ABC-copolymers which mimic some properties of proteins with active enzymatic center. The effect of "memorizing" of some features of the "parent" conformation by the copolymers generated according to our sequence design scheme has been observed. 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".

Acknowledgements

We acknowledge the financial support from RFBR (grant 98-03-33337a), NATO (grant PST/CLG.974956) and from INTAS (grant 97-0678).

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. I. Carmesin and K. Kremer, *Macromolecules* **21**, 2819 (1988)
7. H.P. Deutsch, K. Binder, *J. Chem. Phys.* **94**, 2294 (1991)
8. W. Paul, K. Binder, D.W. Heermann, and K. Kremer, *J. Phys. II* **1**, 37 (1991); *J. Chem. Phys.* **95**, 7726 (1991)