PHYSICAL REVIEW E VOLUME 59, NUMBER 3 MARCH 1999

Properties of *AB* copolymers with a special adsorption-tuned primary structure

Ekaterina A. Zheligovskaya
Institute of Physical Chemistry, Russian Academy of Sciences, Moscow 117915, Russia

Pavel G. Khalatur
Department of Physical Chemistry, Tver State University, Tver 170002, Russia

Alexei R. Khokhlov
Physics Department, Moscow State University, Moscow 117234, Russia

(Received 20 August 1998)

As one of the realizations of our recently proposed idea of conformation-dependent sequence design (engineering) of *AB* copolymers, here we introduce the notion of adsorption-tuned copolymers (ATC's). Such copolymers are prepared by adsorbing a homopolymer chain onto a flat surface, after which its adsorbed segments are called type *A* ones, and its unadsorbed segments type *B* ones. Next, for the primary structure generated in this way, we study the adsorption of ATC's on (some other) surface for the case when only the type *A* segments are attracted to this surface. The properties of the ATC copolymers are compared with those of random copolymers (RC's) with the same content of type *A* segments, and random-block copolymers (RBC's) with the same content of type *A* segments and the same number of blocks. It is shown that the specific features of the ATC primary structure promote adsorption of the ATC chains, leading to an increase of their (negative) critical adsorption energy and the number of adsorbed segments as compared to that of RC and RBC chains. Some characteristics of the adsorbed single chains (statistics of trains, loops, and tails) are studied. The difference between the ATC, RBC, and RC ensembles in these characteristics is explained by the difference in the primary structures of the copolymers. The results obtained support the general idea of conformation-dependent sequence design of *AB* copolymers, and are probably connected with the theory of the early stages of macromolecular prebiological evolution. [S1063-651X(99)08602-X]

PACS number(s): 61.25.Hq, 36.20.Ey, 83.70.Hq, 87.15.By

I. INTRODUCTION

Recently [1–3], we proposed the idea of conformation-dependent sequence design (engineering) of *AB* copolymers consisting of monomeric units of two sorts, *A* and *B*. This approach includes the generation of a special primary sequence in *AB* copolymers from some specific conformation of a homopolymer chain (''parent'' conformation) via a ''coloring procedure'' of the instant image of such chain. In Refs. [1–3] we realized this idea for the case when the ''parent'' conformation corresponds to a dense globule of a homopolymer chain: the monomer units lying in the core of such globule were designated as *B* units and called hydrophobic, while the monomeric units belonging to the surface of the ''instant image'' of a globule were designated as *A* units and called hydrophilic. In this way we generated so-called ''proteinlike copolymers.'' Further studies of the properties of such copolymers have shown that their behavior is essentially different from that of the random copolymers with the same *A*/*B* composition, and random-block copolymers with the same *A*/*B* composition and the same ''degree of blockiness'' [1–3]. In particular, the coil-globule transition in proteinlike *AB* copolymers occurring under gradual increase of the attraction of *B* units takes place at higher temperatures, is more abrupt, and has faster kinetics, as compared to random and random-block copolymers with the same *A*/*B* composition and the same ''degree of blockiness.'' We concluded that the reason for this is that the proteinlike copolymer ''inherits'' (or ''memorizes'') some of the properties of the parent globule (in particular, the existence of well-formed single *B* core coupled with stabilizing *A* envelope) which are then manifested in other conditions [1–3], making the dense globule self-assembly easier to realize.

In the present paper we will consider another realization of the outlined idea. That is, we will assume that the parent conformation is that of adsorbed homopolymer chain. For some ''instant image'' of adsorbed chain, the adsorbed monomer units will be colored in *A* color, while the units in tails and loops will be assigned a *B* color. We will call the primary structure obtained in this way an ''adsorption-tuned *AB* copolymer.'' We initially assumed, and we will show that this is indeed the case, that in this way we obtain an *AB* sequence specially ''tuned'' for the adsorption on another flat surface in the case of attraction of *A* units to that surface.

More specifically, the primary structures of the adsorption-tuned copolymers are obtained by the following procedure. A homopolymer chain is allowed to adsorb onto a flat surface [see Fig. 1(a)]. Then its adsorbed segments (i.e., those lying on the surface) are called type *A* segments and all the other ones (i.e., those not lying on the surface) are called type *B* segments [Fig. 1(b)]. After the primary structure is fixed, two different potentials of interaction of type *A* and *B* segments with an adsorbing surface are introduced [Fig. 1(c)]. Type *B* segments interact with the surface only sterically, whereas type *A* segments are both sterically repelled from the surface and attracted to it with the energy ε.

Similar to proteinlike copolymers [1–3], adsorption-tuned

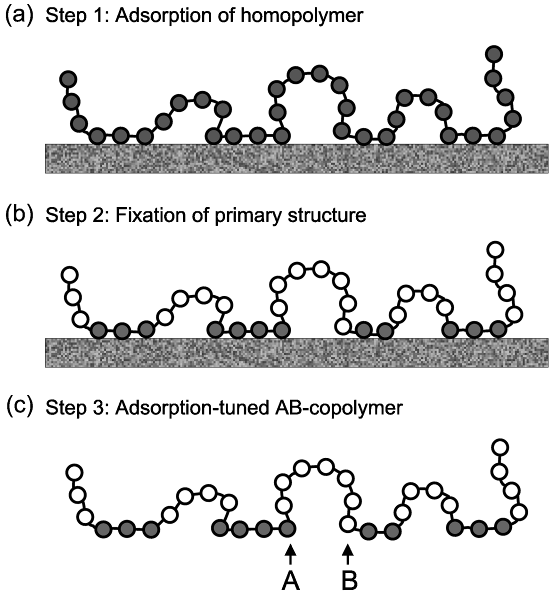

FIG. 1. Schematic representation of the sequence design procedure leading to an adsorption-tuned primary structure.

copolymers (ATC's) can exhibit some special properties. For example, they must differ in their primary structure from random copolymers (RC's) with the same A/B composition or random-block copolymers (RBC's) with the same A/B composition and average lengths of type A and B blocks. Because the ATC primary structures were prepared from adsorbed chains, in an ATC chain the type A segments occupy the sites most ''suitable'' for adsorption. Thus it is probable that adsorption of ATC chains would proceed more easily, which would result in a reduced absolute value of the critical adsorption energy or possibility of adsorption at higher temperatures. In addition to this, the difference in primary structure can result in some specific features of the adsorbed layer, as compared to the cases of random and random-block AB copolymers with the same fraction of type A segments (and, in the case of RBC's, with the same number of blocks). In the present work, we study these effects in the case of adsorption of single chains using the Monte Carlo simulation technique and a lattice model of a polymer chain.

When preparing chains of a given type (i.e., the ATC, RC, or RBC chains), we obtain a set of different primary structures, i.e., an ensemble of chains characterized by a certain distribution of primary structures. It is clear that any primary structure is in principle present in each of the three ensembles; however, in general, the weights of a certain primary structure in the ensembles are different. So, to compare the properties of ATC, RC, and RBC primary structures, for each of the ensembles we should either indicate a characteristic primary structure and compare the properties of these characteristic chains, or calculate the characteristics of interest for each chain and then perform averaging over the ensembles. In fact, we found it difficult to indicate a typical primary structure for each of the ensembles. So we compared the ensembles as a whole by performing averaging over a representative sample of primary structures for each of the ensembles.

In the next sections, we describe the lattice model used in this study and the procedure of preparation of the three ensembles, and then compare these ensembles from the point of view of the properties of the copolymer structure, the critical adsorption energy, and the characteristics of adsorbed chains.

II. MODEL

A. Lattice model of a copolymer chain with a given primary structure

We simulate an AB copolymer molecule using a modification of the bond-fluctuation model [4–8]. The polymer chain consisting of type A and B segments is placed in a three-dimensional (3D) cubic lattice, which is infinite in the x and y directions and limited in the z direction by the two planes $z=0$ and $z=z_{\max}$. The distance between planes is large, so that the chain cannot interact simultaneously with both planes. Both type A and B segments are not allowed to move to the layers $z=0$ and $z=z_{\max}$. In addition to this, there is a short-range potential of attraction of type A segments to the surface $z=0$: the energy ε is attributed to each type A segment lying in the layer $z=1$. This is the only respect, in which type A and B segments are different.

The excluded-volume condition is enforced by that each segment of a chain occupies one lattice site and other segments are not allowed to occupy both this site and its nearest environment (i.e., 26 adjacent sites in the 3D space). The length of a bond between two neighboring segments, b, can fluctuate between $b=2$ and $b=\sqrt{8}$ lattice spacings, so that a polymer chain is represented by a sequence of effective bond vectors, $\mathbf{b}$, belonging to the set

$$\{\mathbf{b}\}\in\{\mathbf{P}(2,0,0);\mathbf{P}(2,1,0);\mathbf{P}(2,1,1);\mathbf{P}(2,2,0)\}. \tag{1}$$

where $\mathbf{P}(i_x,j_y,k_z)$ stands for the set of all permutations and sign combinations of $\pm i_x,\pm j_y,\pm k_z$. In this case, the set of bond vectors $\{\mathbf{b}\}$ connecting two successive segments contains 66 allowed vectors $\mathbf{b}$ with four different choices of the bond length

$$b\in\{2,\sqrt{5},\sqrt{6},\sqrt{8}\} \tag{2}$$

Evolution of a chain is simulated by a common Monte Carlo procedure, the details of which are well known and described elsewhere [4–8].

B. Generation of the ensembles of adsorption-tuned, random, and random-block copolymers

Using the model described in Sec. II A, we generated different configurations of a single homopolymer chain containing $\mathcal{N}=32$ segments (all of type A) and interacting with a flat surface $z=0$ with the segment-surface attraction energy ε ($\varepsilon=-1.0$ in kT units). At this value of ε, 1 000 000 configurations were obtained and each 1000th of them was recorded by assigning unity to each adsorbed and zero to each unadsorbed segment. In subsequent studies, these records were treated as ATC primary structures, where unity indicated a type A segment and zero corresponded to a type B segment.

It is evident that some of the obtained primary structures consist of only type B segments. The frequency of appearance of these latter primary structures of $\mathcal{N}=\mathcal{N}_B$ type B segments depends on the parameter of the simulation box $z_{\max}$. So, in order for the ensemble averages to be independent of such an artificial parameter as the simulation box size, and to reflect the pure effect of the attraction energy ε at which the ATC ensemble was generated, we excluded these primary structures from subsequent consideration.

Estimation of the average fraction p_A of type A segments for the ATC ensemble generated at $\varepsilon=-1.0$ gave $p_A=0.205\pm0.05$. At this value of p_A, we generated a sample of 1000 primary structures of random copolymers. We note that by random copolymers we mean those for which the probability p_A of each monomer to be of type A is constant throughout the whole polymer. This means, for example, that the conditional probability $p_A(i+1,i)$ that the $(i+1)$th segment is of type A when the ith segment is also of type A is equal to p_A. The average fraction of type A segments in random copolymers is equal to the probability p_A.

The next step was to generate the corresponding RBC ensemble. By definition, the RBC chains are characterized by the Poisson distribution of the type A and B block lengths,

$$f(x)=e^{-\lambda}\lambda^x/x!, \tag{3}$$

where λ is the average length (λ_A for type A blocks and λ_B for type B blocks). Systematic consideration and comparison of the properties of the ensembles implies that we must determine $\langle\lambda_A\rangle$ and $\langle\lambda_B\rangle$ for the ATC ensemble generated previously, and create the RBC ensemble using these parameters in the corresponding Poisson distributions. In this case, the next step is the calculation and comparison of the ensemble-averaged critical adsorption energies and the structural characteristics of adsorbed chains for the three ensembles.

However, both ensembles contain a rather large fraction of chains with a small content of type A segments (see Fig. 2). For such chains, determination of the critical adsorption energy and some other characteristics is rather difficult. Therefore, we restricted our consideration by the ATC and RC subensembles of chains containing $\mathcal{N}_A=8$ type A segments. The average content of such chains is 0.07 ± 0.01 in the ATC ensemble and 0.13 ± 0.02 in the RC ensemble. Of course, the properties of subensembles differ slightly from those of the whole corresponding ensembles (see the discussion below on the data in Table I). Nevertheless, such separation of a certain fraction of chains from an ensemble still allows us to solve the problem of interest, i.e., to study the effect of the primary structure on the process of adsorption.

So, from the ATC and RC ensembles we separated the chains containing $\mathcal{N}_A=8$ type A segments. After that, we calculated $\langle\lambda_A\rangle$ and $\langle\lambda_B\rangle$ for the ATC subensemble of chains with $\mathcal{N}_A=8$, and used the obtained characteristics to generate a sample of 2000 chains of the RBC ensemble. From this sample, we also separated the fraction with $\mathcal{N}_A=8$. There were 259 such chains. Table I summarizes the average block lengths for the three ensembles as a whole, for the subensembles with $\mathcal{N}_A=8$, and for the samples of 50 primary structures with $\mathcal{N}_A=8$ used in subsequent calculations of the adsorption properties.

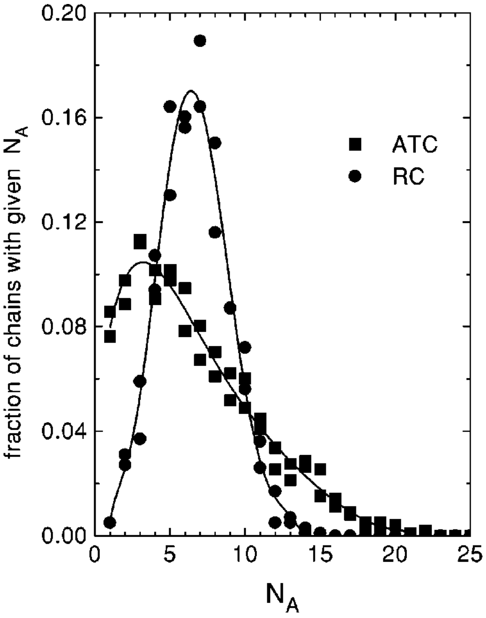

FIG. 2. The distribution of chains with respect to the number of type A segments for ATC and RC ensembles of chains of length $\mathcal{N}=32$. The ATC ensemble was obtained at $\varepsilon=-1.0$ (in units of kT), which corresponds to the average content of type A segments $p_A=0.205$. Two sets of ensemble-averaged data are shown to estimate the statistical error.

In addition for the average block lengths, we separately calculated the lengths of the internal and end blocks of both types. By an end block we mean the last or first block of a copolymer. By an internal block we mean a block that is not an end one. So a copolymer of $k\geqslant3$ blocks has two end blocks and $k-2$ internal blocks.

As can be seen from Table I, the RC ensemble as a whole and its subensemble with $\mathcal{N}_A=8$ are characterized by the shortest block lengths. For RC chains, the difference between the lengths of internal and end blocks of the same type is small, and can be explained by statistical errors. The ATC ensemble as a whole and its fraction with $\mathcal{N}_A=8$ are characterized by very long type B end blocks. In this case, the difference between end blocks and internal type B blocks cannot be explained by insufficient statistics. In fact, the reasons for this difference are apparent. As is known, in the weak adsorption regime, the average length of unadsorbed tails is larger than that of adsorbed fragments of a chain [9], and those were the conditions of the ATC ensemble preparation. As follows from the data of Ref. [8], where we studied the same lattice model for a homopolymer chain, the critical adsorption energy is around -1.0 (in kT units), and the ATC ensemble was obtained just at $\varepsilon=-1.0$. As for the generated sample of the RBC ensemble, it has nearly the same average block lengths as the ATC subensemble with

TABLE I. Average lengths of end and internal type *A* and *B* blocks for the ATC, RBC, and RC ensembles, and their fractions with $\mathcal{N}_A=8$.

$\langle\lambda_{A,\mathrm{int}}\rangle$	$\langle\lambda_{A,\mathrm{end}}\rangle$	$\langle\lambda_A\rangle$	$\langle\lambda_{B,\mathrm{int}}\rangle$	$\langle\lambda_{B,\mathrm{end}}\rangle$	$\langle\lambda_B\rangle$
Averaging over 1000 ATC chains with arbitrary $\mathcal{N}_A$ (two realizations)					
2.49	2.47	2.49	3.59	12.74	7.97
2.46	2.44	2.45	4.08	12.30	8.03
Averaging over ATC chains with $\mathcal{N}_A=8$ (69 of 1000 primary structures)					
2.59	2.30	2.54	4.22	10.21	6.65
Averaging over a sample of 50 ATC chains with $\mathcal{N}_A=8$					
2.47	2.14	2.41	4.57	9.44	6.42
Averaging over 1000 RC chains with arbitrary $\mathcal{N}_A$ (two realizations)					
1.25	1.20	1.25	4.16	4.66	4.29
1.25	1.30	1.26	4.22	4.81	4.38
Averaging over RC chains with $\mathcal{N}_A=8$ (157 of 1000 primary structures)					
1.28	1.16	1.27	3.58	3.38	3.54
Averaging over a sample of 50 RC chains with $\mathcal{N}_A=8$					
1.35	1.19	1.33	3.84	3.26	3.70
Averaging over 2000 RBC chains with arbitrary $\mathcal{N}_A$					
2.67	2.63	2.66	6.73	5.86	6.42
Averaging over RBC chains with $\mathcal{N}_A=8$ (259 of 2000 primary structures)					
2.52	2.16	2.45	7.19	5.85	6.70
Averaging over a sample of 50 RBC chains with $\mathcal{N}_A=8$					
2.57	2.29	2.50	7.60	5.87	6.98

$\mathcal{N}_A=8$, and the difference between the end and internal block lengths is not large. However, in this case the end blocks are noticeably shorter than the internal ones. This is also quite understandable, because the length of the last segment in a chain is always not larger (more often, it is smaller) than the corresponding generated Poisson random number (if the generated Poisson random number is smaller than the length of the remaining part of the chain, then this segment is not last). In the next sections, we estimate the subensemble-averaged critical adsorption energies and some characteristics of the adsorbed chains, and compare these properties for the three ensembles.

III. RESULTS AND DISCUSSION

A. Calculation and comparison of the ensemble-averaged critical adsorption energies

To estimate the critical adsorption energy of a single polymer chain, we determine the average number of type *A* chain segments lying on the adsorbing surface, $N_{A,\mathrm{ads}}$, as a function of the energy ε of attraction of a type *A* segment and the adsorbing surface. Interception of the *x* axis with the straight line tangential to a sigmoidal approximation of the obtained dependence $N_{A,\mathrm{ads}}(\varepsilon)$ at the inflection point gives an estimate of the critical adsorption energy of the chain.

In Ref. [8], we had an experience of calculating the dependences of N_{ads} on ε for a similar model of a polymer chain. We have seen that an accurate determination of such dependences for a single copolymer chain requires averaging over the trajectory of about 10^6 Monte Carlo (MC) passes. (We recall that on average, each particle makes one, successful or unsuccessful, trial move per MC pass.)

In the case of this paper, the problem is more difficult, because we have to perform averaging with respect to a sufficiently large number of different primary structures, and for each of these primary structures the accuracy in determination of $N_{A,\mathrm{ads}}(\varepsilon)$ must be sufficiently high. So, for the required computer time to be acceptable, we performed averaging over 50 conformations chosen randomly from the ATC, RC, and RBC subensembles with $\mathcal{N}_A=8$. For each of the chains, the averaging was performed over a MC trajectory of 150 000 MC passes. This provided sufficient accuracy in determination of the $N_{A,\mathrm{ads}}$ against ε curves.

Because the fraction with $\mathcal{N}_A=8$ of the generated sample of 1000 ATC chains consists of only 69 primary structures, our limit of consideration by only 50 primary structures seems reasonable. For each sample of 50 primary structures, for which the subsequent calculations were performed, we also calculated the average block lengths $\langle\lambda_{A,\mathrm{int}}\rangle$, $\langle\lambda_{A,\mathrm{end}}\rangle$, $\langle\lambda_A\rangle$, $\langle\lambda_{B,\mathrm{int}}\rangle$, $\langle\lambda_{B,\mathrm{end}}\rangle$, and $\langle\lambda_B\rangle$, and summarized these values in the table. As can be seen, the difference in these parameters for the whole subensemble and its portion of 50 primary structures is never more than 10%. So, from this point of view, use of a sample of 50 primary structures to represent the subensemble with $\mathcal{N}_A=8$ for each ensemble seems reasonable.

Figure 3 presents the dependence of the subensemble-average number $N_{A,\mathrm{ads}}$ of type *A* segments lying on the adsorbing surface as a function of the energy ε of attraction between a type *A* segment and the adsorbing surface for the three subensembles. At each ε studied, the inequality $N_{A,\mathrm{ads\,ATC}}>N_{A,\mathrm{ads\,RBC}}>N_{A,\mathrm{ads\,RC}}$ is fulfilled. The estimates of the critical adsorption energy found as the points of interception of the *x* axis with the corresponding tangential straight lines satisfy similar inequality $\varepsilon_{\mathrm{cr\,RC}}<\varepsilon_{\mathrm{cr\,RBC}}<\varepsilon_{\mathrm{cr\,ATC}}$.

So, the RC subensemble, which is characterized by the shortest average lengths of both type *A* and *B* blocks, has a

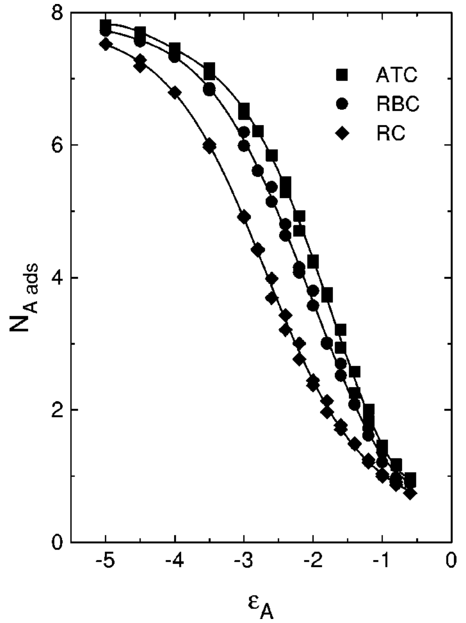

FIG. 3. The average number of adsorbed type *A* segments against the attraction energy ε_A (in units of kT) for ATC, RBC, and RC chains of length $\mathcal{N}=32$ with the content of type *A* segments $\mathcal{N}_A=8$. Two sets of ensemble-averaged data are shown to estimate the statistical error.

considerably lower critical adsorption energy (with the largest absolute value), as compared to other ensembles. This is consistent with the analytical result obtained previously for regular *AB* copolymers [10], according to which the absolute value of the critical adsorption energy decreases with the increasing block length at the same fraction of type *A* segments.

Comparison of ε_{cr} for the ATC and RBC subensembles is of interest, because these two subensembles have the same average block lengths, and the difference in their critical energies can be due only to the difference in primary structures. In fact, the studied sample of 50 RBC chains has even slightly longer type *A* and *B* block lengths than the studied sample of 50 ATC chains. Nevertheless, the estimate of the ATC critical adsorption energy is larger than the corresponding RBC value, and this difference cannot be explained by statistical errors (see Fig. 3). The same conclusion can be formulated for the number of adsorbed segments at a given value of ε_A. We relate this difference to specific features of the ATC primary structures. For example, the ATC chains show a considerable difference in the internal and end type *B* block lengths. This indicates that in ATC chains type *A* segments form more compact regions (as compared to RBC chains), thus favoring adsorption.

The difference in the primary structures influences not only the critical adsorption energy but the characteristics of adsorbed chains as well (especially in the strong adsorption regime). We study this effect in Sec. III B.

B. Comparison of properties of adsorbed ATC, RBC, and RC chains

In Sec. III A, we studied the ensemble-averaged numbers of type *A* segments lying on the surface in the case of adsorption of a single chain for the ATC, RBC, and RC sub-

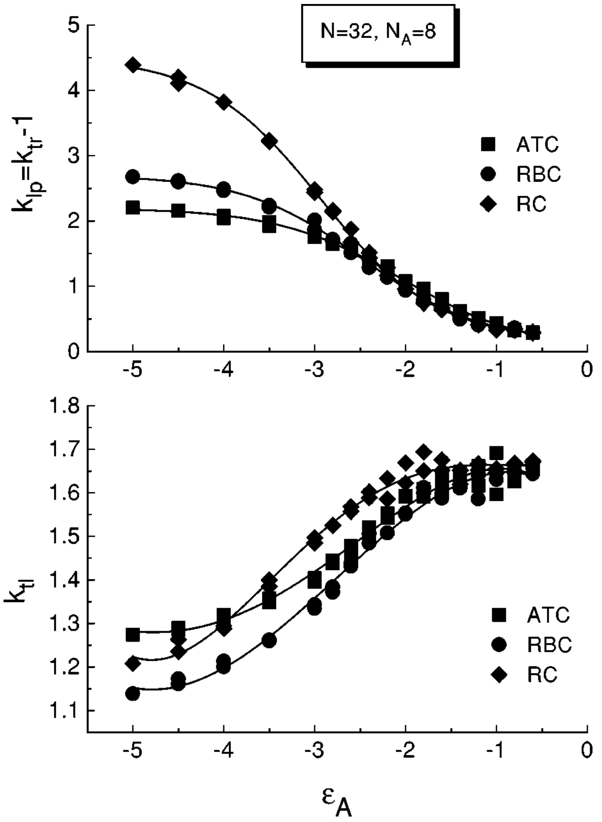

FIG. 4. The average numbers of train, loop, and tail sections k_{tr}, k_{lp}, and k_{tl}, respectively, in ATC, RBC, and RC chains of length $\mathcal{N}=32$ with the content of type *A* segments $\mathcal{N}_A=8$ as functions of the attraction energy ε_A (in units of kT). Two sets of ensemble-averaged data are shown to estimate the statistical error.

ensembles. In this section, we compare some other average characteristics of adsorbed single chains of the RC, RBC, and ATC subensembles with the same content of type *A* segments ($\mathcal{N}_A=8$). We calculate the following parameters: the average numbers of segments belonging to train, loop, and tail sections of a chain, $n_{\text{tr}}(\varepsilon)$, $n_{\text{lp}}(\varepsilon)$, and $n_{\text{tl}}(\varepsilon)$, respectively; and the average numbers of train, loop, and tail sections in a chain, $k_{\text{tr}}(\varepsilon)$, $k_{\text{lp}}(\varepsilon)=k_{\text{tr}}(\varepsilon)-1$, and $k_{\text{tl}}(\varepsilon)$, respectively. (By a train section of a chain we mean its adsorbed section; a loop is an unadsorbed section between two trains, and a tail is the unadsorbed end section of a chain (cf. Ref. [9]). So, we always have $n_{\text{tr}}+n_{\text{lp}}+n_{\text{tl}}=\mathcal{N}$.) We recall that, first, a thermodynamic average of each of these values is found for a chain with a certain primary structure, and, then, the obtained thermodynamic average is averaged over the sample of 50 primary structures of the corresponding subensemble. After that, we find the average lengths of train, loop, and tail sections $l_{\text{tr}}(\varepsilon)=n_{\text{tr}}(\varepsilon)/k_{\text{tr}}(\varepsilon)$, $l_{\text{lp}}(\varepsilon)=n_{\text{lp}}(\varepsilon)/k_{\text{lp}}(\varepsilon)$, and $l_{\text{tl}}(\varepsilon)=n_{\text{tl}}(\varepsilon)/k_{\text{tl}}(\varepsilon)$, respectively.

Figure 4 shows variation of k_{tl} and $k_{\text{lp}}=k_{\text{tr}}-1$ with ε. As can be seen, the average number of tails per chain is nearly the same for the three ensembles. In all three cases, the number of tails decreases slightly with ε, indicating that in the weak adsorption regime more than half of chains have two unadsorbed tail sections, whereas in the strong adsorption regime most of the chains have only one unadsorbed tail section.

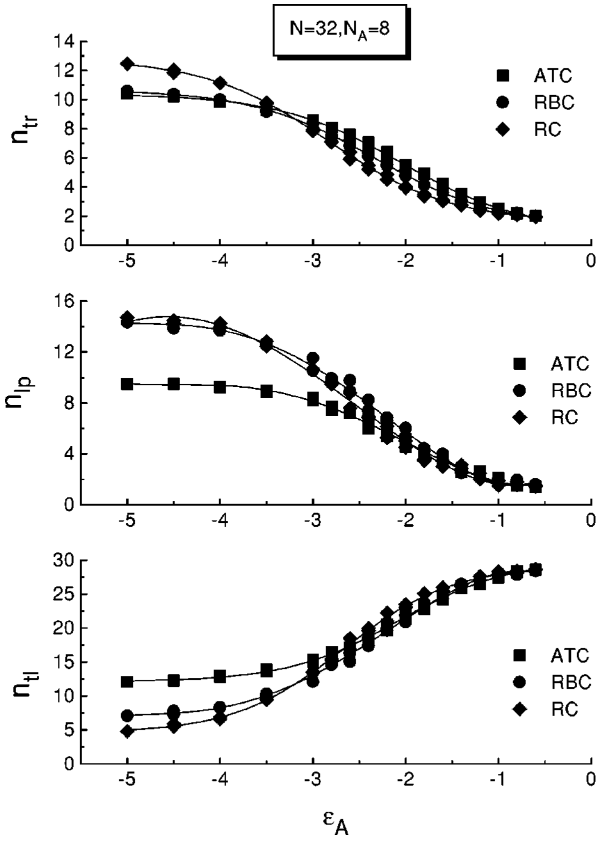

FIG. 5. The average numbers of segments in train, loop, and tail sections n_{tr}, n_{lp}, and n_{tl}, respectively, for ATC, RBC, and RC chains of length $\mathcal{N}=32$ with the content of type A segments $\mathcal{N}_A=8$ as functions of the attraction energy ε_A (in units of kT). Two sets of ensemble-averaged data are shown to estimate the statistical error.

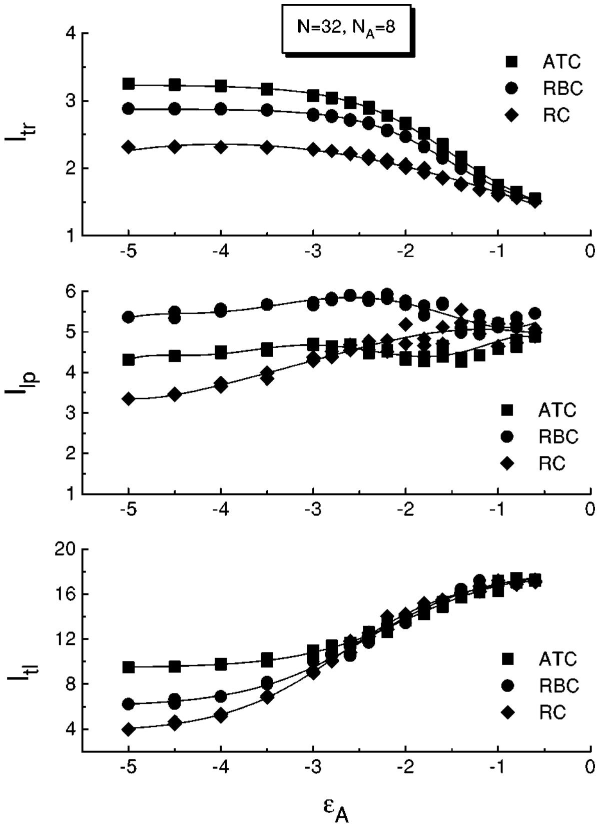

FIG. 6. The average lengths of train, loop, and tail sections l_{tr}, l_{lp}, and l_{tl}, respectively, for ATC, RBC, and RC chains of length $\mathcal{N}=32$ with the content of type A segments $\mathcal{N}_A=8$ as functions of the attraction energy ε_A (in units of kT). Two sets of ensemble-averaged data are shown to estimate the statistical error.

In the absence of adsorption, the average numbers of train and loop sections are nearly the same for the three ensembles, whereas in the strong adsorption regime these characteristics are considerably different. In the strong adsorption regime, nearly all type A segments and some portion of type B segments lie on the surface. So loops can be formed only from type B segments. As the RC chains have the largest number of type B blocks, they exhibit the largest number of loops in the adsorbed state. Because in the ATC chains type A segments are placed more compactly than in the RBC chains (i.e., the average lengths of the internal type B blocks are shorter for ATC chains), they exhibit the lowest number of loops in the strong adsorption regime.

Figure 5 presents a variation of the average numbers of segments in train, loop, and tail segments with ε. In the absence of adsorption, the difference in these characteristics is not large for the three subensembles. In the strong adsorption regime, the average number of segments in train sections (i.e., the average number of chain segments lying on the adsorbing surface) is nearly the same for RBC and ATC chains, whereas for RC chains it is noticeably higher. This is quite understandable, because in the strong adsorption regime all type A segments and some portion of type B segments lie on the adsorbing surface. For all types of chains, we have $\mathcal{N}_A=8$. The probability that a type B segment adjacent to a type A block is adsorbed is higher than the same probability for a type B segment not adjacent to a type A block. For the RBC and ATC chains, the number of blocks is close, whereas for RC chains it is considerably higher. So, for RC chains, the number of adsorbed type B segments (and, therefore, n_{tr}) is considerably higher than for RBC and ATC chains, for which the difference in n_{tr} is rather small.

In the strong adsorption regime, the average number of segments in loop sections is nearly the same for the RBC and RC subensembles. At the same time, the corresponding ATC value is considerably smaller. This is due to the fact that for ATC chains the end type B blocks are considerably longer than the internal ones, and the average number of type B blocks is the same for ATC and RBC chains. The relations between the numbers of segments in tail sections for the three ensembles can be explained similarly.

Figure 6 shows variation of the average lengths of train, loop, and tail sections for the three subensembles. We recall that in the strong adsorption regime trains are formed preferably from type A blocks with some contribution of type B segments; loops are formed from internal type B blocks, and tails are formed from end type B blocks. The data of Fig. 6 exhibit a correlation with the average lengths of the corresponding blocks (see Table I). As expected, in all cases the average train lengths are slightly greater than the corresponding type A blocks (for the studied samples of 50 primary structures) and the average loop and tail length are slightly shorter than the lengths of internal and end type B blocks, respectively. So, we conclude that the difference in the primary structure of copolymers manifests itself in the characteristics of adsorption of the copolymers.

IV. CONCLUSIONS

The ATC ensemble was compared with a RC ensemble of chains with the same *AB* composition and a RBC ensemble with the same *AB* composition and average lengths of type *A* and *B* blocks. The ATC and RBC chains consist of smaller number of different blocks, as compared to the RC chains. In addition to this, the ATC chains have longer end type *B* blocks and shorter internal type *B* blocks, as compared to the RBC chains. This difference in the primary structure of chains of the three ensembles leads to the difference in the critical adsorption energy and the characteristics of adsorbed chains. The fact that the RC ensemble is characterized by a lower (the largest by the absolute value) critical adsorption energy than the RBC and ATC ensembles can be explained by the fact that the type *A* and *B* blocks of the RC chains are considerably shorter than those of the RBC and ATC chains. At the same time, the difference in the critical adsorption energy for the RBC and ATC ensembles (which are characterized by the same average block lengths) can be explained only by the details of the ATC primary structure. Because for the ATC chains the end type *B* blocks are longer than the internal type *B* blocks, the type *A* segments are placed more compactly in a chain. This specific feature of the ATC primary structure promotes adsorption of ATC chains. The studied characteristics of the adsorbed chains are different for the three ensembles. This difference is consistent with the details of the copolymer primary structures.

Thus the difference in the adsorption behavior of ATC, RC, and RBC chains can be rationally explained by taking into account the specific features of the primary structure of these chains. On the other hand, the obtained results support the general idea of a conformation-dependent sequence design of *AB* copolymers proposed in Refs. [1–3]. That is, the generated ATC sequence ''memorized'' some features of the specific ''parent'' conformation of the adsorbed homopolymer chain. In particular, the position of *A* segments turned out to be ''tuned'' in the best way for subsequent adsorption on another surface. It is not surprising, therefore, that this memorized hidden information became apparent as soon as we considered the adsorption of *AB* copolymers on another surface. Among the three types of *AB* copolymers which we studied, ATC chains adsorb better at a given adsorption energy (see Fig. 3). In other words, the *AB* chain ''learns to be adsorbed'' in the parent conformation, and this ''experience'' is used in the subsequent ''life'' of this copolymer. There are obvious implications of this result for a theoretical description of the early stages of macromolecular prebiological evolution.

It should be noted that although the problems discussed in this paper are different from those encountered in the protein folding area, these are obvious parallels in approaches with this rapidly developing field. Indeed, it is generally accepted that nature proteins are the result of an evolutionary sequence selection to optimize the stability and kinetic accessibility of the native conformation [11–28]. The subject of the present paper is also connected with a well-known problem of the successful design of amino acid sequences that control the native protein structure at the coarse-grained level [29–40]. In this case, the simplest model of proteins, the so-called two letter $\mathcal{HP}$ model, consists of sequences made up of just two kinds of amino acids ($\mathcal{H}$ and $\mathcal{P}$ representing hydrophobic and polar residues) configured as self-avoiding chains on a three-dimensional lattice, and described by a model contact Hamiltonian [32–40]. In a sense, our sequence design procedure has some parallels in the general ideology to that proposed several years ago by Shakhnovich and co-workers for protein design problems [29]. There is also a very recent application [39] of design procedure which considers an off-lattice model similar to the one used in the present study (see also Refs. [1–3]), as well as the idea of an initial choice of a homopolymer conformation and subsequent design with a two letter code.

[1] A. R. Khokhlov and P. G. Khalatur, Physica A **249**, 253 (1998).

[2] P. G. Khalatur, V. A. Ivanov, N. P. Shusharina, and A. R. Khokhlov, Russ. Chem. Bull **47**, 855 (1998).

[3] A. R. Khokhlov and P. G. Khalatur, Phys. Rev. Lett. (to be published).

[4] I. Karmesin and K. Kremer, Macromolecules **21**, 2819 (1988).

[5] H. P. Deutsch and K. Binder, J. Chem. Phys. **91**, 2294 (1991).

[6] P.-Y. Lai, Macromol. Theory Simul. **5**, 255 (1996).

[7] P. G. Khalatur, A. R. Khokhlov, I. A. Nyrkova, and A. N. Semenov, Macromol. Theory Simul. **5**, 713 (1996).

[8] E. A. Zheligovskaya, P. G. Khalatur, and A. R. Khokhlov, J. Chem. Phys. **106**, 8598 (1997).

[9] J. M. H. M. Scheutjens and G. J. Fleer, J. Phys. Chem. **84**, 178 (1980).

[10] E. B. Zhulina, A. M. Skvortsov, and T. M. Birshtein, Vysokomol. Soedin., Ser. A **23**, 304 (1981).

[11] C. Anfinsen, E. Haber, M. Sela, and F. H. White, Proc. Natl. Acad. Sci. USA **47**, 1309 (1961).

[12] C. Levinthal, J. Chem. Phys. **65**, 44 (1968).

[13] I. M. Lifshitz, Zh. Éksp. Teor. Fiz. **55**, 2408 (1968) [Sov. Phys. JETP **28**, 1280 (1969)].

[14] N. Go, Annu. Rev. Biophys. Bioeng. **12**, 183 (1969).

[15] K. F. Lau and K. Dill, Macromolecules **22**, 3986 (1989).

[16] J. D. Bryngleson and P. G. Wolynes, Biopolymers **30**, 177 (1990).

[17] E. Shakhnovich, G. Farztdinov, A. Gutin, and M. Karplus, Phys. Rev. Lett. **67**, 1665 (1991).

[18] P. E. Leopold, M. Montal, and J. N. Onuchic, Proc. Natl. Acad. Sci. USA **89**, 8721 (1992).

[19] C. Camacho and D. Thirumalai, Proc. Natl. Acad. Sci. USA **90**, 6369 (1993).

[20] V. S. Pande, A. Y. Grosberg, and T. Tanaka, Proc. Natl. Acad. Sci. USA **91**, 12 972 (1994).

[21] A. Sali, E. Shakhnovich, and M. Karplus, J. Mol. Biol. **235**, 1614 (1994); Nature (London) **369**, 248 (1994).

[22] M. H. Hao and H. A. Scheraga, J. Chem. Phys. **98**, 4940 (1994).

[23] V. S. Pande, A. Y. Grosberg, and T. Tanaka, Macromolecules **28**, 2218 (1995).

[24] J D. Bryngelson, J. N. Onuchic, N. D. Socci, and P. G. Wolynes, Proteins: Struct., Funct., Genet. **21**, 167 (1995).

[25] E. M. Boczko and C. L. Brooks, Science **269**, 393 (1995).

[26] P. G. Wolynes, Z. Luthey-Schulten, and J. N. Onuchic, Chem. Biol. **3**, 425 (1996).

[27] P. G. Wolynes, Proc. Natl. Acad. Sci. USA **93**, 14 249 (1996).

[28] J. N. Onuchic, Z. Luthey-Schulten, and P. G. Wolynes, Annu. Rev. Phys. Chem. **48**, 539 (1997).

[29] E. Shakhnovich and A. M. Gutin, Proc. Natl. Acad. Sci. USA **90**, 7195 (1993); Protein Eng. **6**, 793 (1993); E. Shakhnovich, Phys. Rev. Lett. **72**, 3907 (1994); I. Abkevich, A. M. Gutin, and E. I. Shakhnovich, J. Mol. Biol. **252**, 460 (1995); K. Yue, K. M. Fiebig, P. D. Thomas, H. S. Chan, E. I. Shackhnovich, and K. A. Dill, Proc. Natl. Acad. Sci. USA **92**, 325 (1995).

[30] S. Kamtekar, J. M. Schiffer, H. Xiong, J. M. Babik, and M. H. Hecht, Science **262**, 1680 (1993).

[31] M. H. J. Cordes, A. R. Davidson, and R. T. Sauer, Curr. Opin. Struct. Biol. **6**, 310 (1996).

[32] K. F. Lau and K. A. Dill, Macromolecules **22**, 3986 (1989).

[33] K. F. Lau and K. A. Dill, Proc. Natl. Acad. Sci. USA **87**, 6388 (1990).

[34] H. S. Chan and K. A. Dill, J. Chem. Phys. **95**, 3775 (1991).

[35] K. A. Dill, S. Bromberg, S. Yue, K. Fiebig, K. M. Yee, P. D. Thomas, and H. S. Chan, Protein Sci. **4**, 561 (1995).

[36] J. N. Onuchic, P. G. Wolynes, and N. D. Socci, Proc. Natl. Acad. Sci. USA **92**, 3626 (1995).

[37] F. Seno, M. Vendruscolo, A. Maritan, and Jayanth R. Banavar, Phys. Rev. Lett. **77**, 1901 (1996).

[38] C. Micheletti, F. Seno, A. Maritan, and J. R. Banavar, Phys. Rev. Lett. **80**, 2237 (1998).

[39] C. Clementi, A. Maritan, and J. R. Banavar, Phys. Rev. Lett. **81**, 3287 (1998).

[40] F. Seno, A. Maritan, and J. R. Banavar, Proteins: Struct., Funct., Genet. **30**, 244 (1998).