Modeling of Radical Copolymerization near a Selectively Adsorbing Surface: Design of Gradient Copolymers with Long-Range Correlations

Nataliya Yu. Starovoitova,[†] Anatoly V. Berezkin,[†] Yurii A. Kriksin,[‡] Olga V. Gallyamova,[§] Pavel G. Khalatur,*[,†,⊥] and Alexei R. Khokhlov[§,⊥]

Department of Physical Chemistry, Tver State University, Tver 170002, Russia, Institute for Mathematical Modeling, Russian Academy of Sciences, Moscow 125047, Russia, Physics Department, Moscow State University, Moscow 119899, Russia, and Department of Polymer Science, University of Ulm, Ulm D-89069, Germany

Received June 28, 2004; Revised Manuscript Received November 12, 2004

ABSTRACT: A new approach to synthesis of copolymers with long-range correlations is proposed. Using Monte Carlo simulations and the lattice bond-fluctuation model, we perform the computer-aided sequence design of a two-letter (AB) copolymer with quenched primary structure near a chemically homogeneous impenetrable surface. We simulate an irreversible radical copolymerization of selectively adsorbed A and B monomers with different affinity to the surface, allowing for a strong short-range monomer− (A−) surface attraction. To describe the chain growth analytically, we introduce and investigate a simple theoretical model based on stochastic processes and probabilistic statistics. We find that this model provides a close approximation to the simulation data and explains a number of statistical properties of copolymer sequences. It is shown that, under certain conditions, the chain propagation near the adsorbing surface proceeds as a randomly alternating growth, leading to a copolymer with a specific quasi-gradient primary structure and power-law long-range correlations in distribution of different monomer units along the chain. The gradient extends along the entire chain for any chain length. We find that the statistical properties of the copolymer sequences correspond to those of a one-dimensional fractal object with scale-invariant correlations. Thus, just by radical copolymerization of two monomers with different affinity to a certain plane surface, it is possible to obtain copolymers with a gradient primary structure.

1. Introduction

Copolymers have been studied extensively for several decades, partly because of their biological and industrial importance, and partly because of their interesting and sometimes perplexing properties. Although recent years have witnessed an impressive confluence of experiments, simulations, and analytic theories, presently there is no comprehensive understanding what role do copolymer primary sequences play for the structural and functional properties of copolymer systems.

In a series of publications,[1−5] a concept of conformation-dependent sequence design of copolymers has been introduced. The essence of the proposed approach[1−5] is based on the assumption that a copolymer obtained under some bare ("parent") preparation conditions is able to "remember" features of its original conformation in which it was built up and can then manifest these features when the conditions are changed. In other words, this concept takes into account a strong coupling between the conformation and primary structure of copolymers during their synthesis. Ideologically, the approach[1−5] bears some similarities with that proposed earlier in the context of the problems of protein physics;[6−9] however, it aims at synthetic copolymers rather than biopolymers with a unique (target) conformation.

Initially, the general idea connected with conformation-dependent sequence design was realized in computer simulations[1−5] and developed theoretically[10,11] for the simplest model of a two-letter copolymer, i.e., a copolymer consisting of only two comonomers. In the model, an initial homopolymer in its original dense globular state corresponding to the bare conformation is modified into two different (A and B) units by assigning (insoluble) A type to the units in the core of the globule, and (soluble) B type to those on its surface. The heteropolymer obtained as a result of such simple one-step procedure can self-assemble into a segregated core−shell microstructure, thus resembling some of the basic properties of globular proteins, in particular, their solubility in water.

In ref 10, it was shown that the corresponding primary sequence is nonalternating and demonstrates the specific long-range correlations (LRC), which can be described by the statistics of the Lévy-flight type.[12] These correlations are shown to be directly related to the conformation-dependent sequence design scheme. Certainly, owing to finite size of the bare globule, the longest correlations that can be found in this case are also finite and should be understood as long-wave fluctuations of the chemical composition.[10,11]

The presence of LRC in designed sequences is due to the fact that assigning of the type of chain segment (A or B) at preparation conditions depends on the conformation of the bare globule as a whole, not on the conformation of small sections of the initial homopolymer chain. From this viewpoint one may say that such a sequence encodes the spatial (core−shell) structure of a copolymer globule. Obviously, this important functional feature can be realized if and only if a general statistical pattern is attributed to the sequence as a whole, and cannot be obtained by joining of independent statistical patterns of two or more subsequences of

* To whom correspondence should be addressed. E-mail: khalatur@germany.ru.
† Tver State University.
‡ Russian Academy of Sciences.
§ Moscow State University.
⊥ University of Ulm.

10.1021/ma0487094 CCC: $30.25
Published on Web 02/22/2005

smaller length.[5] In this respect, protein sequences are similar: sequence as a whole determines globular structure and hence biological function, while if this sequence is cut into two pieces, those pieces normally neither correspond to a soluble globule nor have any biological function.[13] This peculiarity is directly connected with LRC effects, and the corresponding sequence, which cannot be divided into shorter subsequences with similar statistical patterns and possible functional features, may be termed "inseparable sequence." Such kind of sequence integrity as well as LRC are not characteristic of the majority of synthetic linear copolymers whose primary structure is chemically homogeneous on a large scale.[14] Indeed, the sequence distribution for rather large sections of such copolymers should be practically identical to that of the whole polymer. On the other hand, one may anticipate that if scale-invariant correlations extend along the entire copolymers sequence, that is, the sequence is a "true" fractal object, its sufficiently large parts would have the same statistical pattern and large-scale compositional inhomogeneities. In the present paper, we will focus just on such sequences, showing both strong chemical inhomogeneity and LRC.

After the concept of conformation-dependent sequence design, leading to copolymers with LRC, was formulated,[1−3,15,16] several possible ways of experimental realization of this concept were proposed.[17−23] In the literature, there are the following two main approaches: (i) chemical modification ("coloring") of reactive groups in a certain homopolymer[17−19] and (ii) the formation of a designed conformation of AB copolymers directly in the course of the radical copolymerization of appropriate comonomers.[20−23] By "coloring" is implied, e.g., a polymer-analogous reaction, which converts the segments at the surface of the homopolymer globule (e.g., from a hydrophobic to a polar type: A → B). This approach was realized in the series of papers.[17−19] There are a few papers in which a direct chemical synthesis of functional copolymers via conformation-dependent design is described.[20−23]

The possibility to obtain the copolymers capable of forming core−shell microstructures via radical copolymerization of the monomers differing in hydrophilicity/hydrophobicity was confirmed by recent simulations[24,25] where the process of copolymerization was modeled in a selective (polar) solvent. The composition of an emerging copolymer chain was such that macromolecule adopted a globular conformation, and the preferential absorption of hydrophobic monomers in the core of the globule was taken into account. It was shown[24,25] that such copolymerization process, based on the so-called "bootstrap effect",[26] automatically leads to the formation of the core−shell microstructure in the resulting globule and to the well-pronounced LRC in the primary sequences. These results are in qualitative agreement with experimental data.[20−23]

The idea of conformation-dependent sequence design can be generalized. Indeed, the special primary sequence can be obtained not only from globular conformation; any specific polymer chain conformation can play the role of a parent one. The simplest example of this kind is connected with the conformation of a homopolymer partly adsorbed on a flat substrate. Let us assume that the chain segments being in direct contact with the surface in some typical instant conformation are chemically modified. This can take place when the surface catalyses some chemical transformation of the adsorbed segments. One can expect that after desorption, such a copolymer will have special functional properties: it will be "tuned to adsorption". Following this line, Zheligovskaya et al.[4] have compared the adsorption properties of copolymers with special "adsorption-tuned" primary structures (adsorption-tuned copolymers, ATC) with those of truly random copolymers and random-block copolymers. Monte Carlo simulations revealed that specific features of the ATC primary structure promote the adsorption of ATC chains, compared to their random and random-block counterparts under the same conditions. In other words, the resulting copolymer sequence "memorizes" the original state of adsorbed homopolymer chain.[4] It was shown that its statistical properties exhibit LRC of the Levy-flight type similar to those known for copolymers obtained via "coloring" of a homopolymer globule.[27] Recently, Velichko et al.[28] have suggested the model of so-called molecular dispenser that is a further development in the direction of conformation-dependent sequence design.

In the present paper, we will consider the computer-aided synthesis of copolymers with strongly correlated primary structure near a macroscopic planar substrate. This effort is intended as a step toward developing the approach discussed above further. Our study is aimed at computer-aided synthesis of copolymers from two types of selectively adsorbed monomers differently distributed in the reaction system. We will simulate the irreversible growth process in the dilute regime for polymerizing monomers, A and B, with different affinity to the substrate. The polymerization will be modeled as step-by-step chemical reaction of addition of A and B monomer units to the growing copolymer chain, assuming that depolymerization reaction is not allowed. Then, we will investigate the statistical properties of prepared sequences to compare with those known for other types of copolymers (random and random-block). Also, we introduce and investigate a simple probabilistic model describing the copolymerization process near an impenetrable surface. It should be noted that the polymerization procedure discussed in this paper can be considered as a prototype of template polymerization based on noncovalent binding of polymerizing monomers to the template. We will show that under certain conditions, surface-induced copolymer formation can result in copolymers with well-pronounced intramolecular chemical inhomogeneity and long-range correlations.

In the literature, there are some computer models that are aimed at simulating polymerization processes at a microscopic level. In particular, Rouault and Milchev[29] have applied the highly efficient lattice bond-fluctuation model for Monte Carlo simulations of living polymerization, where the reacting chain ends remain activated throughout the process of reversible equilibrium polymerization (see also refs 30 and 31). The formation of polymer networks (gels) was investigated by the Monte Carlo method using a model of free-radical cross-linked polymerization.[32]

The rest of the paper is organized as follows: We begin with defining the model and assumptions used in the simulation in section 2.1, and formulate the theoretical approach in the next subsection, section 2.2. Then, we describe the results in section 3 where the numerical calculations are compared to the analytical prediction. Our main aim here is to explore in detail

the statistical properties of generated sequences. We summarize and conclude in section 4.

2. Methodology

2.1. Simulation Technique and Model Assumptions. For the sake of computational efficiency, we do not take into account chemical details of simulated polymer chains, addressing the generic features of surface-induced sequence formation only. Thus, all simulations are performed with the standard lattice bond-fluctuation model, which is described in detail in refs 33 and 34. The motion of particles (polymerizing monomers and chain segments) is generated by the Monte Carlo (MC) technique: A particle and a lattice direction are chosen at random, and a move is attempted in the proposed direction. The move is successful, if the targeted lattice sites are empty (excluded volume interaction), if it allows to be compatible with chain connectivity, and if it lowers the potential energy. On the other hand, if the energy difference, Δu, between the final and the initial state is positive, the move is only accepted with probability $\exp[-\Delta u/T]$ according to the standard Metropolis algorithm,[35] where T denotes the temperature (Boltzmann's constant $k_B = 1$).

The results are expressed in internal units: lengths are measured in units of the lattice spacing, σ, and time, τ, is in units of MC steps (MCS) per particle (on the average each particle attempts to jump once within a Monte Carlo step, successfully or unsuccessfully). Temperature was set to $T = 1$.

A total of N monomers A and B were placed in a self-closed slab (with periodic boundary conditions in the x and y directions) of size $b_x = b_y = 400$ and $b_z = 120$, in units of σ. The planes at $z = 0$ and $z = b_z$ are considered as reflecting boundaries. The $z = 1$ plane was chosen as an adsorbing surface. To model adsorption, we used the following slowly decaying adsorption potential: $u(z) = \epsilon/z^3$ (where ϵ is the characteristic adsorption energy at $z = 1$) that acts only on the polymerizing A monomers and the corresponding chain segments. Therefore, we simulated selective adsorption with a fixed adsorption energy ($\epsilon_A < 0$), while the B monomers had no attraction to the surface, $\epsilon_B = 0$. In this study, the value of ϵ_A was set to $\epsilon_A = -2.08$. With this choice, the density of A monomers far from the attractive surface was negligibly small, and thus the precise location of the repulsive walls had a negligible influence on the results.

The simulation was performed in two stages. First, to calculate equilibrium adsorption profiles, we considered free (nonreacting) monomers inside the reaction box. Among the monomers, N_A were the monomers of type A and N_B were the monomers of type B. We placed into the simulation box a fixed number of A monomers, while the number of B monomers was a variable parameter. The values of N_A and ϵ_A were chosen such that about 95% of all the polymerizing A monomers were in strongly adsorbed state for each N_B. The bulk number density of unadsorbed B monomers, ρ_B, was always larger than the volume fraction of A monomers in the system, $\rho_B > \rho_A$. Because of the strong adsorption, however, the equilibrium volume fraction of A monomers within a narrow-width adsorption layer near the surface was always much larger than ρ_B. The ρ_B value was varied from 0.01 to 0.25; the volume fraction ρ_A was fixed at 0.008. After equilibration, the simulations were carried out during 3×10^6 MC steps. After each 10^2 MC steps the monomer density profiles $\rho(z)$ normal to the $z = 0$ plane were calculated. To this end, we divided the distance b_z into 120 bins and sampled how many monomers fall into each bin. Then, the average density profiles for A and B monomers, $\rho_A(z)$ and $\rho_B(z)$, were found. Thus, choosing the volume fraction of B monomers in the bulk, ρ_B, we can obtain various density profiles $\rho_B(z)$. Note that $\rho_A(z)$ depends on ρ_B very weakly.

During the second phase of the simulation, we studied the irreversible polymerization process, which was considered as a step-by-step chemical reaction of addition of monomer units to the growing polymer chain. In our polymerization model, the presence of low-molecular-weight species (both solvent molecules and free monomers) was taken into account only implicitly through the density profiles of A and B monomers obtained at the first stage. In the implicit-monomer model, these profiles define the probabilities of addition of monomer units to the growing macroradical. It is assumed that the reactants are ideally mixed and that diffusion is not important. In principle, such conditions can be realized if typical diffusive time, τ_d, and the characteristic time of adsorption equilibrium are essentially less than the reaction time τ_r. It is clear that this assumption corresponds to the kinetically controlled regime.

The usual radical polymerization process consists of the following three parts: initiation, chain propagation, and termination. In this study, the initiator of polymerization process was modeled as a fixed point of type A with coordinates $x = b_x/2$, $y = b_y/2$, and $z = 1$ and was fixed on the attractive surface. Since only one radical initiator was put into reaction mixture, spontaneous chain terminating reactions, such as radical combination and disproportionation, were excluded. In the present study, the growing chain was automatically terminated at the number of monomer units $N = 1024$.

The implicit-monomer model involves two characteristic time scales. First, the reaction time scale is set by the time between two reaction trials, τ_r. Second, the equilibrium motion time, τ_{rel}, includes conformation relaxation of growing chain. Internal motion of the chain mainly depends on the current chain length, N_τ; that is, the length of growing chain corresponding to a given time τ. It should be noted here that the Rouse relaxation time $\tau_{rel} \sim N^2$ for an ideal chain. For simplicity, we take the following fixed relaxation time $\tau_{rel} = 1024N_\tau$ in the present study. Any bias introduced by the growing method toward a certain distribution of configurations should become unimportant as the chain has enough time to relax from the nonequilibrium state between two successive reactions.

In the copolymerization mechanism, the rate of addition of a monomer to a growing free radical depends only on the nature of the end group on the chain. We model an "ideally polymerizing" system in which the rate constants are equal to each other. In this case, the end group on the growing chain has no influence on the rate of monomer addition, and the two types of monomer units are arranged along the chain in relative amounts determined by the local concentrations of A and B monomers near the end radical of the chain. Thus, the time of addition of monomer to the chain and probabilities of A or B monomer additions are determined by combination of the equilibrium density profiles, $\rho_A(z)$ and $\rho_B(z)$, and by the growing chain end location.

During the simulation, we found the current position of free ("active") end of the chain along the z direction,

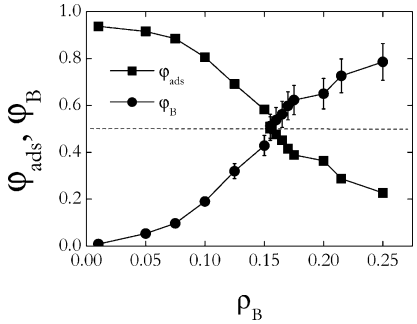

Figure 1. (■) Average fraction of adsorbed segments, φ_{ads}, and (●) average fraction of B segments in the resulting copolymer, φ_B, as a function of the volume fraction of monomers B in the bulk, ρ_B. Chain length $N = 1024$. Dashed line corresponds to $\varphi_{ads} = \varphi_B = 1/2$.

$z_p(\tau)$, and added to it the new monomer, according to relative probabilities to occur the monomer A or monomer B in the z layer: $p_A \equiv \tilde{\rho}_A(z_p)$ and $p_B \equiv \tilde{\rho}_B(z_p)$, where $\tilde{\rho}_A(z_p)$ and $\tilde{\rho}_B(z_p)$ are the normalized density profiles and $p_A + p_B = 1$. Thus, in the model used here, we have only one variable parameter of the copolymerization process, viz., the volume fraction of monomers B in the bulk, ρ_B.

The algorithm used for modeling of chain growth was as follows. A random number uniformly distributed between 0 and 1 was generated and then compared to the probability p_A of having an addition of the monomer unit to the growing chain. If the random number occurred to be less than the value of p_A, monomer A was linked with the "active" end of the N_r-unit macroradical; otherwise, monomer B was taken. After each such step, we performed chain relaxation during the time τ_{rel}. After termination of the growing chains at $N = 1024$, their statistical analysis was carried out. To gain better statistics, 10^3 fixed-length copolymers for each density ρ_B were generated and then required average characteristics were found.

Depending on ρ_B, the following characteristics were calculated for the ensemble of the generated copolymers: the average fraction of adsorbed segments, φ_{ads}; the average fractions of A and B segments in the chain, φ_A and φ_B ($\varphi_A + \varphi_B = 1$); the average numbers of chemically uniform A and B sections (blocks), n_A and n_B, as well as the average lengths of these sections, L_A and L_B; and the distributions of segments along the chain. We note that, first, an average of each of these quantities is found for a chain with a certain chemical sequence and then, the quantities are averaged over the sample of the sufficiently large number of independently generated sequences.

Intuitively, one may expect that the average fraction of B units φ_B in the resulting copolymer should increase with their bulk concentration ρ_B, thereby leading to a decrease in the number of strongly adsorbed segments. As seen from Figure 1, where we plot the values of φ_{ads} and φ_B vs ρ_B, this is indeed the case. At some point $\rho_B^* \approx 0.155$ the curves φ_{ads} and φ_B cross each other. We find that in this case, about half of chain segments is in adsorbed state ($\varphi_{ads} \approx 1/2$), and the AB composition of synthesized copolymer is close to equimolar ($\varphi_A = 0.512$ and $\varphi_B = 0.488$). At $\rho_B < \rho_B^*$, the growing chain is enriched with A monomers; when $\rho_B > \rho_B^*$ the growing chain end deeply penetrates into the bulk, and the B segments prevail in the resulting copolymer. Thus, the change in the solution concentration of unadsorbed (or weakly adsorbed) reactive monomer allows varying the chemical composition in a wide range. Because of a one-to-one correspondence between φ_A (or φ_B) and ρ_B (Figure 1), our model can operate only with one parameter governing the chain properties at a fixed degree of polymerization N, specifically with average chemical composition φ_A.

For the sake of simplicity, the discussion presented in this paper will be focused mainly on the systems where the chemical AB composition of resulting copolymers is close to 1:1. As a matter of fact, the equimolar composition is of special interest.

Throughout the simulation, snapshots of the system were collected. To give a visual impression of the simulated system, Figure 2 presents a typical snapshot picture of an 1024-unit copolymer chain synthesized at $\rho_B = 0.155$.

2.2. The Probabilistic Model of Copolymerization. One of the ways to get a correct insight into the nature of correlations in a complex polymerizing system consists of an ability of constructing a simple mathematical (probabilistic) object (e.g., a correlated sequence of symbols) possessing the same statistical properties as the initial system. To this end, we consider in this subsection the model of an ideal, self-intersecting linear polymer chain consisting of two types of segments, A and B. In the model, there is no excluded-volume interaction of chain segments between themselves. The ideal structure of our polymer chain ensures the presence of spatial correlation only between the neighboring segments.

To describe radical copolymerization near a surface, we introduce two discrete stochastic processes, $\xi^{(A)}$ and $\xi^{(B)}$. They form a compound process ξ, which corresponds to a path associated with the primary sequence of an AB copolymer chain. Let us characterize the sequence of steps (monomer units) by $\sigma_i \in \{A,B\}$ with $i = 0, 1, 2, \ldots$, labeling the steps along the sequence so that the random variable ξ_i is the realization of the process ξ at the ith step. We take the origin of the path as the origin of coordinates. There is a starting point $i = 0$ which is associated with the initiator of polymerization (monomer A); that is, $\xi_0 = 0$ or $\sigma_0 = A$ in all realizations of the process ξ. This will not affect the asymptotic scaling behavior.

In the model describing a polymer chain growth by reaction of monomers A and B with a reactive end group on the growing chain, the discrete-state process $\xi^{(A)}$ characterizes attachment of strongly adsorbed monomers A. This process has two outcomes: success, with probability p_A, and failure, with probability $1 - p_A$. In other words, with probability p_A, the polymer chain is lengthened by addition of the next monomer unit of the type A to a terminal free-radical reactive site. If the process $\xi^{(A)}$ is terminated, then the process $\xi^{(B)}$, associated with attachment of monomers B in the bulk solution, starts. In our probabilistic polymer growth model, it is convenient to treat the process $\xi^{(B)}$ as an asymmetric random walk on the 1D half-line with an absorbing boundary at the origin. The random walk is defined as follows. At each ith step, the path goes further from the origin ($\Delta_i^{(B)} = \xi_i^{(B)} - \xi_{i-1}^{(B)} = +1$) with probability p_B, and goes closer to the origin ($\Delta_i^{(B)} = -1$) with probability $1-p_B$. When the random walk reaches the origin ($\sum_i \Delta_i^{(B)} = 0$ or $\xi_0^{(B)} = -1$), it is terminated (absorbed). After that the process $\xi^{(A)}$ starts again. Thus, the processes $\xi^{(A)}$ and $\xi^{(B)}$ change each other and

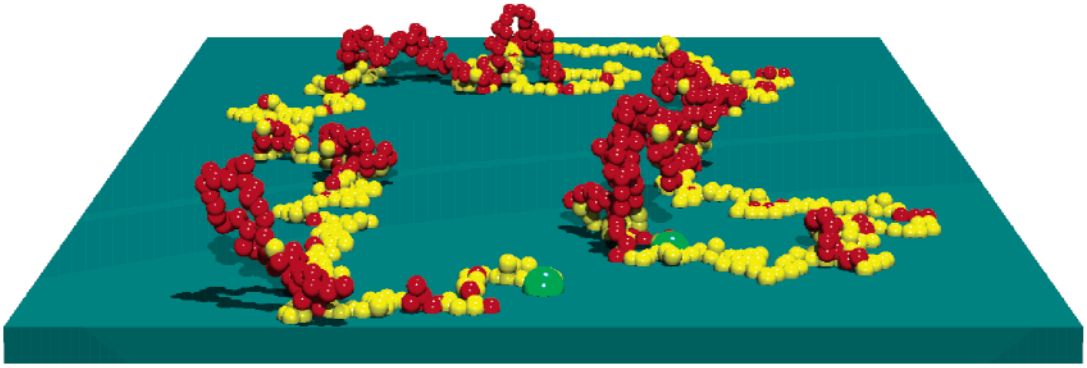

Figure 2. Snapshot of the copolymer chain with $N = 1024$ obtained from copolymerizing monomers at $\rho_B = 0.155$. The yellow and red spheres show adsorbed (A) and unadsorbed (B) monomer units, respectively. The place of connection of the chain with the surface ("initiator") and the chain end are depicted as larger green spheres.

continue until the path reaches a desired total length; that is, we deal with a randomly alternating growth. Hence, the random variable ξ_i, obeying specific probability distribution function, takes value $\xi_i = 0$ if $\sigma_i =$ A and $\xi_i = k$ if $\sigma_i =$ B, where $k = 0, 1, ...$

If the step i in the infinite path ξ is of the type A, then the conditional probabilities p_{AA} and p_{AB} for the next step $i + 1$ to be of the type A or B are defined as

$$p_{AA} = p_A + (1 - p_A)(1 - p_B)p_A + (1 - p_A)(1 - p_B)(1 - p_A)(1 - p_B)p_A + ... = \frac{p_A}{1 - (1 - p_A)(1 - p_B)} \quad (1)$$

$$p_{AB} = (1 - p_A)p_B + (1 - p_A)(1 - p_B)(1 - p_A)p_B + ... = \frac{(1 - p_A)p_B}{1 - (1 - p_A)(1 - p_B)} \quad (2)$$

with $p_{AA} + p_{AB} = 1$. Below, we restrict ourselves to the simplest (symmetric) case: $p_B = {}^1/_2$ when each first step of the process B is absorbed by the boundary with probability ${}^1/_2$.

The probability to find an uninterrupted succession (block) of some length l of A steps in the infinite path ξ is given by the well-known geometric distribution $P_A{}^0(l) = (1 - p_{AA})p_{AA}{}^l$. However, taking into account that the very first step in the path should always be of the type A ($\sigma_0 =$ A), we have the following conditional probability distribution

$$P_A(l) \equiv P_A{}^0(l|i > 0) = P_A{}^0(l)/P_A{}^0(i > 0) = [(1 - p_{AA})p_{AA}{}^l]/p_{AA} = (1 - p_{AA})p_{AA}{}^{l-1} \quad (3)$$

where $l = 1, 2, ...,$ and p_{AA} is defined by eq 1. Therefore, the average length of blocks A for the infinite total path length is given by

$$L_A = \sum_{l=1}^{\infty} l(1 - p_{AA})p_{AA}{}^{l-1} = (1 - p_{AA})^{-1} \quad (4)$$

One may say that the conditional probability p_{AA} defines the strength of persistent correlations in the sequence. If the persistent correlations are extremely strong ($p_{AA} \to 1$), then $L_A \to \infty$. In the polymerization process, such a situation is realized when the bulk concentration of B monomers approaches zero. If $p_{AA} = p_{AB} = {}^1/_2$, one arrives at the known trivial result for the Bernoullian statistics without correlations that corresponds to a random copolymer. In this case, the probability p_A of each segment to be of type A is constant throughout the whole sequence. This means, for example, that the conditional probability p_{AA} 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 φ_A in a purely random sequence is equal to the probability p_A. In particular, for a purely random sequence with $p_A = {}^1/_2$ we have $L_A = L_B = 2$. Such a copolymer is obtained in the course of solution copolymerization of A and B monomers having identical solubility and reactivity. Finally, at $p_{AA} = 0$ ($p_{AB} = 1$) we deal with a regular sequence with alternating distribution of A and B segments in the sequence for which $L_A = L_B = 1$ (eq 4).

The probability distribution over the length of blocks B is related to the well-known ballot problem or to a classical probabilistic "first return" problem.[36] Here we do not have a underlying analytical process like the geometrical one. In the present context, the ballot problem can be solved by using a conditional probability argument leading to a recurrence relation.

In our formulation, $\xi_i^{(B)}$ is the result of the ith step on the half-line: $\xi_i^{(B)} = \xi_{i-1}^{(B)} + 1$ for a step that jumps further from the origin and $\xi_i^{(B)} = \xi_{i-1}^{(B)} - 1$ for a step that jumps closer to the origin. One may then use standard techniques for one-dimensional random walks. It is clear that each block of the type B should contain the even number of steps. Therefore, we consider the one-dimensional walks that start at origin and take $2m$ steps ($m = 1,2,...$). The total number of random walks with the positive first step is thus 2^{2m-1} and we consider that all these walks to be equally probable (since $p_B = {}^1/_2$). Let us collect the fraction of random walks that visit origin for the last time at the step $2n$ and have the nonnegative first term. The fraction of such walks of length $2m$ can be written as

$$\phi_m = 2^{1-2m} \sum_{n=1}^{m} \frac{n}{m} \binom{2m}{n+m}, \; m = 1,2,.. \quad (5)$$

with $\phi_0 = 1$. The function ϕ_m, given above, is related to the probability distribution over the length of process $\xi^{(B)}$. This distribution is similar to the so-called discrete arc-sine distribution of order m.[36] It can be calculated using the following relation

$$P_{\xi^{(B)}}(k) \equiv P_{\xi^{(B)}}(k = n|k > 0) = \begin{cases} 0, & n = 2m - 1 \\ \phi_{m-1} - \phi_m, & n = 2m \end{cases} m = 1,2,... \quad (6)$$

with

$$\phi_m = \sum_{j=1}^{m} b_{j,m}, \; m = 1,2,... \quad (7)$$

where

$$b_{j,m} = \frac{j}{j-1}\frac{m-j+1}{m+j} b_{j-1,m}, j = 2,...,m \quad (8)$$

$$b_{1,m} = \frac{a_m}{m+1} \quad (9)$$

$$a_m = \left(1 - \frac{1}{2m}\right) a_{m-1}, \; m = 2,3,... \quad (10)$$

$$a_1 = 1 \quad (11)$$

The distribution over the length of blocks B, $P_B(l)$, is more complex than $P_{\xi^{(B)}}(k = n|k > 0)$. This is due to the fact that B blocks of some length can be obtained by "gluing together" two or more shorter B sequences, which are neighbors in the total path. This happens when after terminating the process $\xi^{(B)}$, the process $\xi^{(A)}$ is terminated immediately at the first step, and the current B sequence has a chance to be continued. In other words, we take into account the presence of unadsorbed B monomers on the surface.

Let us call the B block of length l "irreducible block", if it is originated from the process $\xi^{(B)}$ that is terminated at the step $2m$. It is clear that any blocks of the type B, including those located at the end of a finite path, can be composed from the irreducible blocks by "gluing" them together. Then the distribution of the B blocks can be written as

$$P_B(l = 2m) = \frac{p_{AA}}{1 - p_{AA}} \sum_{n_j} \prod_{j=1}^{m} R_{n_j}, \sum_{j=1}^{m} n_j = m \quad (12a)$$

$$P_B(l = 2m + 1) = 0, \; m = 1,2,... \quad (12b)$$

where

$$R_n = (1 - p_{AA}) P_{\xi^{(B)}}(k = 2n|k > 0), \; n = 1,2,...,m \quad (13a)$$

$$R_0 = 1 \quad (13b)$$

Here the function $P_{\xi^{(B)}}(k = 2n|k > 0)$ is given by eq 6. To calculate $P_B(l = 2m)$, it is more convenient to use the recurrence relations

$$P_B(l = 2m) = \sum_{j=1}^{m} q_j^{(m)}, \quad m = 1,2,... \quad (14)$$

$$q_j^{(m)} = \sum_{n=j-1}^{m-1} R_{m-n} q_{j-1}^{(n)}, \quad j = 2,3,...,m \quad (15)$$

$$q_1^{(n)} = (1 - p_{AA})^{-1} p_{AA} R_n, \; n = 1,2,...,m \quad (16)$$

One can show that the block length distribution $P_B(l)$ is characterized for asymptotically large $P_B(l)$ by the power-law decay of its density $P_B(l) \propto l^{-\alpha}$ with the exponent $\alpha = 3/2$.

Formally, the average length of blocks B for the infinite total path length is defined as

$$L_B = \sum_{m=1}^{\infty} 2m P_B(l = 2m) \quad (17)$$

Note that the second moment of $P_B(l)$ diverges for $l \to \infty$, but in a simulation or experiment it should be always finite.

Now we are interested in so-called truncated probability distributions describing a finite path of a given length. A distribution is truncated if observed values must fall within a restricted range, instead of the expected range over all possible values. In our case, the longest path (sequence) can have a length not more than N. If a probability distribution function $P(l)$ is known for an infinite path, then the corresponding truncated distribution $P(l \le N)$ is given by

$$P(l \le N) = \begin{cases} P(l), & l = 1,2,..., N-1, \\ Q(l), & l = N \end{cases} \quad (18)$$

where

$$Q(l) = \sum_{n=l}^{\infty} P(n) = 1 - \sum_{n=1}^{l-1} P(n) \quad (19)$$

and $P(l)$ is defined from (3) and (12) for blocks A and B, respectively. The average length of the blocks is

$$L_A^{(N)} = \sum_{l=1}^{N} l P_A(l) \quad (20a)$$

$$L_B^{(N)} = \sum_{l=1}^{N-1} l P_B(l) \quad (20b)$$

Note that in eq 20b we take into account that $\sigma_0 = A$, and contribution from the very first step A always takes place. It is easily seen from eqs 19 and 3 that for blocks A, the $Q_A(l)$ function reads

$$Q_A(l) = p_{AA}^{\;l} \quad (21)$$

Therefore, we have

$$L_A^{(N)} = \frac{1 - p_{AA}^{\;N}}{1 - p_{AA}} \quad (22)$$

In the $N \to \infty$ limit, eq 22 coincides with eq 4a. For $L_B^{(N)}$ there is no simple analytical expression.

As has been noted above, of special interest is a copolymer with a fixed (in particular, equimolar) chemical composition. For the model considered, the corresponding condition can be written as

$$1 + \sum_{i=1}^{N-1} p_A^{(i)} = N\varphi_A \quad (23)$$

Here $p_A^{(i)}$ is the probability that the variable σ_i is in the state A at the ith step of the finite path ξ of a given length N and $\varphi_A = N_A/N$ is the fraction of the A-type states. By definition, $p_A^{(0)} = 1$. Using the relation for total probability, we arrive at

$$p_A^{(i)} = p(\sigma_i = A|\xi_{i-1} = 0)p(\xi_{i-1} = 0) + p(\sigma_i = A|\xi_{i-1} > 0)p(\xi_{i-1} > 0) \quad (24)$$

Taking into account that $p(\sigma_i = A|\xi_{i-1} = 0) = p_{AA}$ and $p(\sigma_i = A|\xi_{i-1} > 0) = 0$, we rewrite eq 24 for $N \geq 4$

$$p_A^{(i)} = p_{AA}p(\xi_{i-1} = 0) \quad (25)$$

where in general case

$$p(\xi_i = 0) = p_{AA}p(\xi_{i-1} = 0) + (1 - p_B)p(\xi_{i-1} = 1)$$

$$p(\xi_i = 1) = (1 - p_{AA})p(\xi_{i-1} = 0) + (1 - p_B)p(\xi_{i-1} = 2)$$

$$p(\xi_i = k) = p_B p(\xi_{i-1} = k - 1) + (1 - p_B)p(\xi_{i-1} = k + 1) \text{ for } k > 1 \quad (26)$$

with $p(\xi_0 = 0) = 1$ and $p(\xi_0 = k) = 0$ for $k > 0$. Because of the composition constraint (23), the conditional probability p_{AA} (and therefore p_A, see eq 1) turns out to be a function of N. In this case, it can be shown that the probability p_A converges to unity as $p_A \propto 1 - N^{-1/2}$ for sufficiently large N. Thus, the probability distributions $P(l)$ and the corresponding moments, L_A and L_B, become complex functions of two parameters N and φ_A.

In the next section, our analysis based on the probabilistic model (PM) presented above will be semianalytical/semi-numerical. Although we will focus our attention on the case $\varphi_A = 1/2$, nevertheless, the major part of the results is valid for the whole region $0 < \varphi_A < 1$.

3. Results and Discussion

In this section we investigate the factors controlling the formation of copolymers near a selectively adsorbing surface and describe some sequence characteristics. We begin our discussion with an analysis of average block lengths observed for the generated copolymer sequences.

a. Block Lengths. The average lengths of A and B blocks, L_A and L_B, are shown in Figure 3 as a function of copolymer composition φ_A. As expected, when φ_A is increased the A blocks become longer while the B blocks shorten. Our probabilistic model predicts qualitatively similar behavior. At $\varphi_A = 0.57$ the values L_A and L_B become approximately equal to each other, $L_A \approx L_B \approx 8.5$ (in this case, $n_A \approx n_B \approx 60$ so that $2nL \approx N$). Surprisingly, the equality $L_A = L_B$ is observed in PM for almost the same composition ($\varphi_A = 0.60$), although the length of A and B blocks predicted for 1024-unit sequence is considerably larger as compared to that from MC simulations. For $\varphi_A = 0.512$ (the value which is closest to equimolar composition in the simulations performed in this study) we find $L_A = 8.4$, $L_B = 9.3$, and $L = (L_A + L_B)/2 \approx 9$.

b. Detrended Fluctuation Analysis. Let us now turn to the question of long-range correlations in the copolymer sequences obtained in the present work. For statistical analysis of copolymer sequences, different mathematical techniques have been used. We will use the method developed by Stanley and co-workers[37−39] in their search for LRC in DNA sequences. In this approach, each AB copolymer sequence is transformed into a sequence of symbols +1 and −1 which are considered as steps of a one-dimensional random walk. Shifting the sliding window of length λ along this sequence step by step, the number of symbols +1 and −1 inside the window is counted at each step. This number

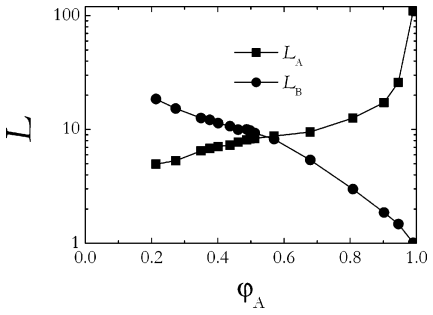

Figure 3. Average lengths of (■) A and (●) B blocks, L_A and L_B, as a function of copolymer composition φ_A. Chain length $N = 1024$.

$$\gamma_k(\lambda) = \sum_{i=k}^{k+\lambda-1} \eta_i$$

is a new random variable, depending on the position k of the window along the sequence; here η_i is associated with every symbol i, such that $\eta_i = +1$ if chain segment i is A and $\eta_i = -1$ if it is B. This random variable has certain distribution. Its average is determined by the overall sequence composition, and its dispersion is given by $D_\lambda^2 = \langle\gamma^2(\lambda)\rangle - \langle\gamma(\lambda)\rangle^2$, where $\langle ...\rangle$ stands for the average over all windows of size λ.

If the sequence is uncorrelated (normal random walk) or there are only local correlations extending up to a characteristic range (Markov chain) then the value of D_λ scales as $\lambda^{1/2}$ with the window of sufficiently large λ. A power law $D_\lambda \propto \lambda^\alpha$ with $\alpha > 1/2$ would then manifest the existence of LRC. However, due to large fluctuations, conventional scaling analysis cannot be applied reliably to the relatively short sequences generated in typical simulations. To avoid this problem, we employ the so-called detrended fluctuation analysis (DFA), the method specifically adapted to handle problems associated with nonstationary sequences.[38,39] In brief, DFA involves the following steps. The entire N-symbol sequence is divided into N/λ subsequences, each consisting of λ symbols. The best linear trend for each subsequence is then defined. The difference between the original random walk $\gamma(\lambda)$ and the corresponding local trend is found. Finally, the variance about the detrended walk for each subsequence is calculated. The average of these variances over all the subsequences of size λ, denoted $F_D^2(\lambda)$, characterizes the detrended local fluctuations within the window of length λ. Generally, the $F_D(\lambda)$ function shows the same behavior as D_λ.

Figure 4a presents the results of the statistical analysis performed as described above for 10^3 generated sequences. On average, these sequences have approximately 1:1 AB composition ($\varphi_A = 0.512$). For comparison, we show in the same figure the data obtained for a purely random 1:1 sequence and a random-block 1:1 sequence, both with $N = 1024$. By definition, a random-block sequence is characterized by the Poisson distribution of the type A and B block lengths, $f(x) = e^{-L}L^x/x!$, where L is the average block length. Systematic consideration and comparison of the properties of the ensembles implies that we must determine L_A and L_B for the ensemble of simulated sequences and then create the random-block ensemble using these parameters in the corresponding Poisson distribution. Thus, we gener-

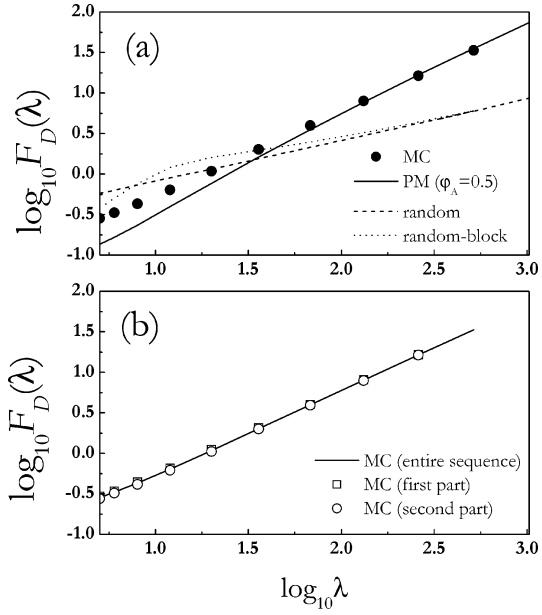

Figure 4. (a) Detrended fluctuation functions for (●) designed, (-- --) random, and (·····) random-block sequences of length $N = 1024$. Solid line shows the PM results at $\varphi_A = 0.5$. (b) Detrended fluctuation functions for the generated copolymers: (−) the entire 1024-unit sequence, (□) the first 512-unit half, and (○) the second 512-unit half.

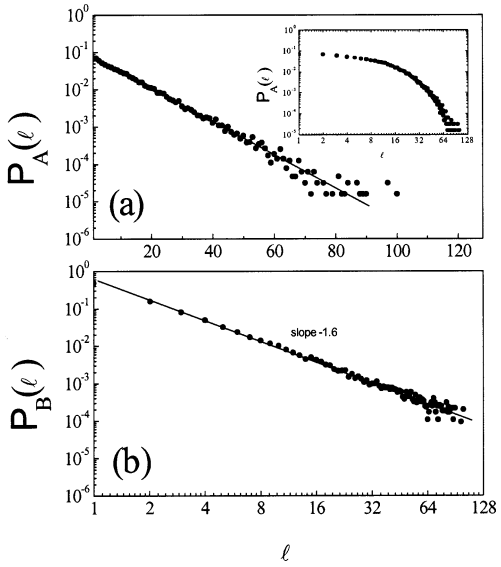

Figure 5. Block length distribution functions obtained from simulations: (a) $P_A(l)$ plotted on a semilog scale and on a log−log scale (the insert) and (b) $P_B(l)$ plotted on a log−log scale, at $\varphi_A = 0.512$. Solid lines show the best fits on the simulation data.

ated the Poisson distribution adjusted to achieve 1:1 composition and the same "degree of blockiness" as for a simulated copolymer. Also, we present in Figure 4a the $F_D(\lambda)$ function predicted by our probabilistic model for $N = 1024$ and $\varphi_A = 1/2$.

It is seen from Figure 4a that the data obtained for a purely random sequence demonstrates $F_D(\lambda) \propto \lambda^{1/2}$ scaling throughout the interval of λ examined, as expected. For a random-block sequence having nearly the same composition and average block length as simulated sequences, there are well-pronounced correlations in the region corresponding to small λ's, which are close to the average block length. However, as λ increases, we observe dumping of the correlations, and for sufficiently large λ, the behavior demonstrated by the random and random-block sequences is almost identical. Comparing these curves with the simulation results we see immediately that designed sequences do not correspond to random and random-block statistics, and strong correlations do exist in these sequences. For sufficiently large λ, very good agreement between the MC simulation and the analytical PM result is observed. In both cases, the long-range correlations persist up to the windows with length close to N. Moreover, the correlations turn out to be more pronounced as λ is increased: the dependence of $\log[F_D(\lambda)]$ on $\log \lambda$ becomes a nearly linear function whose slope approaches unity with increasing λ. These sequences do not show dumping of the correlations and a wide crossover region with varying slope, the behavior that was found for "protein-like" copolymers obtained via "coloring" of a homopolymer globule.[10] On the other hand, our findings are surprisingly similar to those known for DNA sequences, which appeared as a mosaic of coding and noncoding patches.[37−39] Indeed, similar to DNA chains containing coding and noncoding regions, the copolymer under consideration also contains two types of alternating sections forming a certain pattern. It is known that the noncoding regions in DNA do not interrupt the correlation between the coding regions (and visa versa), and the DNA chain is fully correlated throughout its whole length. As a result, the D_λ^2 curve does not contain the linear portion $D_\lambda^2 \propto \lambda$. In a sense, such a behavior reflects the statistical properties of nonstationary fragments of a fixed length. In principle, the same behavior is observed in the present study.

From the facts presented above, it is evident that the sequences discussed here are correlated throughout their whole length. Also, this implies that if the averaged sequence is cut into a number of short fragments and then randomly shuffled, the $F_D(\lambda)$ dependence obtained in this way would not correspond to that observed in Figure 4a. On the other hand, it was found that any sufficiently large part of the averaged sequence has practically the same correlation properties as the entire sequence. Figure 4b demonstrates the fluctuation functions $F_D(\lambda)$ calculated separately for the first and second 512-unit parts of the 1024-unit sequences. One sees that the $F_D(\lambda)$ functions for these sections and for the entire sequence are nearly the same. This means that the generated sequences show scale invariance, a feature typical of fractal structures. Thus, the question is: what kind of statistics corresponds to the generated copolymer?

c. Distribution Functions. To gain some insight into the discussed problem, we present in Figure 5 the block length distribution functions $P_A(l)$ (plotted on a semilog scale) and $P_B(l)$ (plotted on a log−log scale) obtained from simulations at $\varphi_A = 0.512$. It is seen that the $P_A(l)$ function decays exponentially with increasing block length l while the $P_B(l)$ function exhibits a power-law decay $P_B(l) \propto l^{-\alpha}$. As has been noted, the same behavior follows from the probabilistic model. Moreover, the exponent α estimated form the simulation data ($\alpha \approx 1.6$) is, up to the "experimental" uncertainty, quite close to that predicted by our analytical model ($\alpha = 3/2$). Such power-law decay is a characteristic of Lévy probabilistic processes.[12] For these processes an observable stochastic variable x exhibits large jumps ("flights"), termed "Lévy flights", characterized by power-law (rather

than exponential) probability distribution function, $P(x) \propto x^{-\mu}$ ($1 < \mu < 3$). In this case, the variance D_l^2 is a nonlinear function of l, in contrast to Markovian processes for which D_l^2 is linearly dependent upon l for large l . In this context, one may mention that statistical properties of many two-letter copolymers can adequately be described by the ergodic n-step Markov process.[14] Markovian statistics have often been utilized to fit polymer sequence distribution data. For Markovian binary copolymers there is a characteristic length ("memory length") of the order of n, which defines the scale of correlations between segments in a copolymer sequence. For $l \gg n$, however, such correlations are washed out, and it is expected that $D_l^2 \propto l$ (or $F_D(\lambda) \propto \lambda^{1/2}$) and $\ln P(l) \propto -l$. As seen from Figures 4 and 5b, this is definitely not the case for the copolymers generated in the present study. Therefore, one cannot define the finite value of "memory length" which is analogous to n, and the statistics of these copolymers cannot be described by a Markov process. Of course, it is possible to interpret any sequence of length N in terms of an (N-1) Markov process, but the definition of this maximum order model is nothing but a rephrasing of the original sequence and one has no guarantee that the model will work for longer sequences.

d. Intrachain Correlations. Another useful quantity that can be defined is the normalized correlation function

$$C_\eta^{(i)} = \frac{(N-i)\sum_{j=1}^{N-i}\eta_j\eta_{j+1} - \sum_{j=1}^{N-i}\eta_j\sum_{j=1}^{N-i}\eta_{j+1}}{\{(N-i)\sum_{j=1}^{N-i}\eta_j^2 - [\sum_{j=1}^{N-i}\eta_j]^2\}^{1/2}\{(N-i)\sum_{j=1}^{N-i}\eta_{j+1}^2 - [\sum_{j=1}^{N-i}\eta_{j+1}]^2\}^{1/2}} \quad (27)$$

It quantifies how the stochastic variable η at one position i_1 is correlated to its value at some other position i_2, both separated by a gap of $i = |i_1 - i_2|$ segments along the chain. By definition, $C_\eta^{(1)} = 1$. When all η_i are statistically independent (random copolymer), $C_\eta^{(i)} = 0$ for $i > 1$. If the $C_\eta^{(i)}$ function decreases fast with "chemical distance" i, then faraway positions are relatively uncorrelated and the sequence is dominated by its short-range structure. On the other hand, a slow decrease of $C_\eta^{(i)}$ would imply that faraway positions along the chain have a large degree of correlation or influence on each other. In Figure 6a, we present the $C_\eta^{(i)}$ function averaged over all the sequences generated by MC method for $N = 1024$ and $\varphi_A = 0.512$ as well as the analogous PM results for a few values of N. It is evident from the PM data that $C_\eta^{(i)}$, when plotted against scaling variable i/N, is a universal function of "chemical distance" for any sufficiently long sequence. The fact that no effect of sequence length is seen here reflects the strict fractal structure of the sequence. There is good agreement between the MC and PM data; moreover, at $i/N > 0.25$ they practically coincide. In both cases, no simple exponential decay of $C_\eta^{(i)}$ is observed. One sees that there are two regimes. On short scales, there is a fast decay. On longer scales, however, there is a long negative tail, which is determined by the overall structure of the whole sequence. The presence of negative correlations in $C_\eta^{(i)}$ bears some similarities with the behavior of a regularly alternating multiblock copolymer for which $C_\eta^{(i)}$ has an oscillatory structure with persistent period corresponding to block length. However, the sequences under consideration have rather randomly alternating statistical pattern than regularly alternating one.

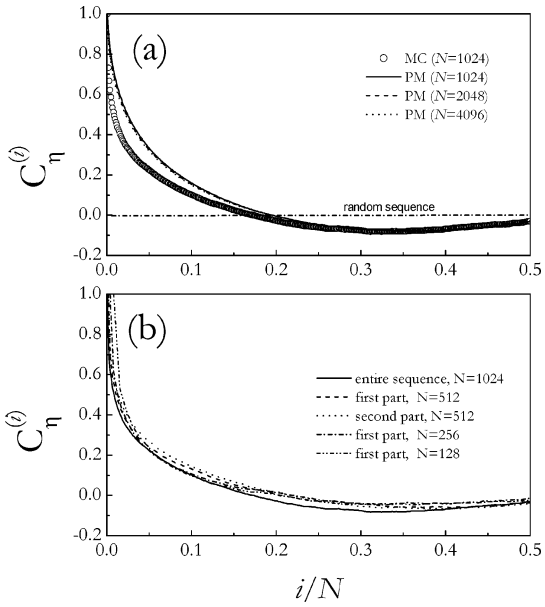

Figure 6. (a) Correlation function (eq 27) plotted vs i/N. (○) Simulation data ($\varphi_A = 0.512$ and $N = 1024$); solid, dashed, and dotted lines correspond to the PM results for different N, as marked in the figure. (b) MC results for subsequences of different length.

In addition, we show in Figure 6b the correlation functions (27) calculated separately for different subsequences of the generated 1024-unit sequences. It is seen that the statistical properties of the subsequences are close to each other and to those of the entire sequence. Again, such a behavior mirrors fractal properties of this two-letter sequence and is the signature of scale invariance.

e. Compositional Inhomogeneity. We calculated the intramolecular composition profile $\varphi_A^{(i)} = 2p_A^{(i)} - 1$ which characterizes intramolecular chemical inhomogeneity along the chain and is related to the probabilities $p_A^{(i)}$ to find a monomer unit A at the ith position from the beginning of a growing macromolecule. The present definition of $\varphi_A^{(i)}$ assumes that the A-type segments are coded by symbol +1 while symbol −1 is prescribed to the B-type segments. For an ideal random copolymer in which chemically different segments follow each other in statistically random fashion, the $p_A^{(i)}$ function should coincide with the average fraction φ_A of A segments for any i. For a random-block copolymer, the fraction of one component averaged over many chains should also be uniform along the chain. Figure 7 demonstrates a typical composition profile calculated as a function of i/N for $\varphi_A = 0.512$. First, we observe rather good agreement between MC and PM results. Second, both profiles show a monotonic decrease with the segment position. Therefore, we deal with a specific copolymer whose chemical sequence is similar to that known for so-called "*tapered*" or *gradient* copolymers exhibiting strong composition inhomogeneity along their chain.[40−42] Such copolymers can be considered as a special type of block copolymers in which composition of one component varies along the chain. With a decreasing difference in the monomer reactivity rations, the formation of gradient statistical copolymers rather

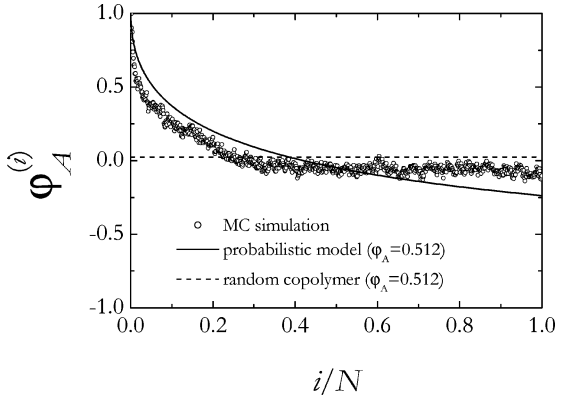

Figure 7. Composition profiles calculated as a function of i/N for $\varphi_A = 0.512$ and $N = 1024$: (○) simulation data, (−) probabilistic model, (- -) random copolymer.

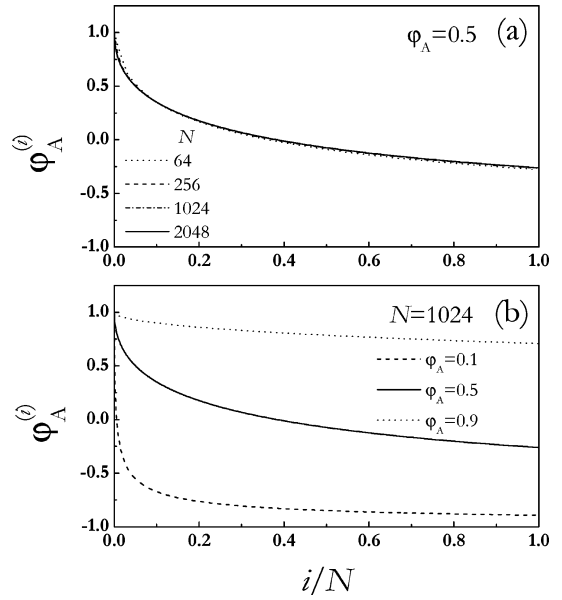

Figure 8. Intramolecular composition profiles calculated using PM (a) for chains of different length (at $\varphi_A = 0.5$) and (b) for several compositions (at $N = 1024$).

than tapered block copolymers occurs.[14,42] In our model polymerization process, however, all the polymerizing species have the same reactivities, and the monomer concentrations remain unchanged during synthesis. Therefore, the changes in the probabilities of addition of components to the growing macroradical are due to the evolution of its chemical composition.

Additionally we present in Figure 8 the intramolecular composition profiles calculated using PM for chains of different length (at $\varphi_A = 0.5$, Figure 8a) and for several compositions (at $N = 1024$, Figure 8b). First observation from the data shown in Figure 8a is that at a fixed φ_A and sufficiently large N, the profiles $\varphi_A^{(i)}$ are universal functions of scaling variable i/N. Indeed, the curves coincide reasonably well establishing the validity of the scaling ansatz. The results obtained for $N = 64$ deviate, though not considerably, from the scaling curve. Also, it is interesting to note the following fact: it seems that the value $p_A^{(N)}$ corresponding to the end-segment converges to e^{-1} in the $N \to \infty$ limit when $\varphi_A = 0.5$. As seen from Figure 8b, the gradient structure of the copolymer becomes less pronounced as the average composition φ_A deviates from equimolar. However, for nonzero φ_A the gradient extends along the entire chain for any N. Therefore, we observe compositional scale invariance.

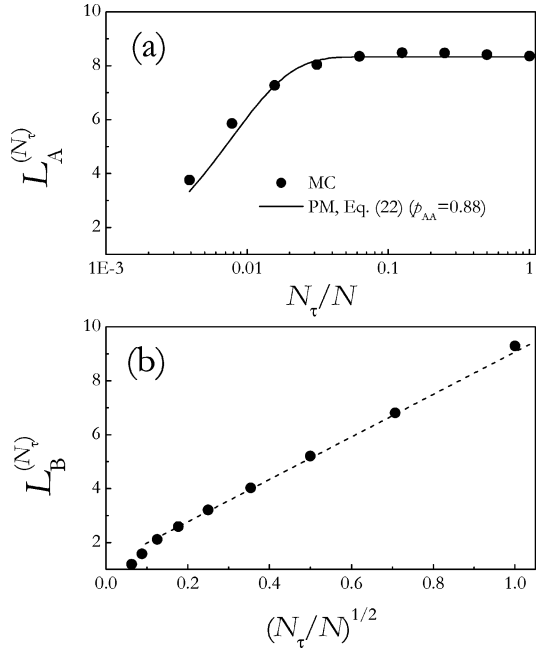

Figure 9. Current average block lengths, (a) $L_A^{(N\tau)}$ and (b) $L_B^{(N\tau)}$, vs the length of growing copolymer fragment, for 1024-unit chains: (●) simulation data, (−) probabilistic model (eq 22, $p_{AA} = 0.88$), (- -) the best fit on the simulation data.

f. Conversion Dependencies. To further explore the origin of the gradient-like primary structure observed in our simulation and PM calculations, we have monitored the variation of AB composition during copolymerization. To this end, we define the "degree of conversion" N_τ/N which varies from N^{-1} to 1. Analysis of the MC data shows that at the very beginning of copolymerization ($N_\tau/N < \sim 0.05$), the growing chain radical gathers mainly monomers A strongly adsorbed at the surface. Then, as the chain becomes longer and the number of connected B units increases, the probability to find growing end-radical in B environment rises (reactions with the participation of unadsorbed B monomers dominate in the bulk). This causes the changes both in the current AB composition and in the average block lengths in growing chains of the current length N_τ. Figure 9 illustrates these trends. Here, we present the current average block lengths, $L_A^{(N_\tau)}$ and $L_B^{(N_\tau)}$, vs N_τ/N. Note that the data in Figure 9a are given on a semilog scale while Figure 9b shows the dependence of $L_B^{(N_\tau)}$ on $(N_\tau/N)^{1/2}$. In Figure 9a the data of MC simulations are compared with the theoretical model (see eq 22 where we used adjustable parameter $p_{AA} = 0.88$). In Figure 9b, the best fit on the simulation data is presented.

First of all, we emphasize again qualitative agreement between MC and PM results. In particular, as follows from eqs 18−20 for truncated probability distribution, the $L_B^{(N)}$ value scales with N as $L_B^{(N)} \propto N^{1/2}$. It is evident from Figure 9a that the $L_A^{(N\tau)}$ value reaches its limiting level L_A rather quickly (at $N_\tau > \sim 10^2$). On the other hand, the current average length of blocks B increases with N_τ up to $N_\tau = N$ (Figure 9b). As a result, for our copolymer which was adjusted to achieve approximately 1:1 composition and required length of $N = 1024$, we obtain the primary structure of gradient-like type. In this copolymer, the content of B monomer units and the

corresponding block length gradually increase from the beginning of chain to its end, while the same quantities observed for A monomer units remain practically unchanged along the chain, except its very beginning.

It is clear that the behavior discussed above can take place only if both the total chain length and AB composition are constrained simultaneously. Without these constraints we would have: $L_B \rightarrow \infty$ and $L_A \rightarrow$ constant for the $N \rightarrow \infty$ limit. However, it should be stressed that if average copolymer composition is fixed, then both L_A and L_B scale with N as $N^{1/2}$, as directly follows from the PM prediction. This is due to the fact that the probability p_{AA} entering eqs 12 and 18 itself depends upon N via the constraint eq 22 so that increasingly large values of p_{AA} must be used to achieve desired AB composition for each given N. Such a composition constraint is precisely the major reason behind the specific chain growth and affects the global statistical nature of sequences on large scales.

4. Concluding Remarks

In this work, we have used a Monte Carlo simulation technique to study two-letter (AB) quasirandom copolymers with quenched primary structure, which were generated via surface-induced computer-aided sequence design. Our approach represents an irreversible radical copolymerization (chain growth) of selectively adsorbed A and B monomers with different affinity to a surface. In the model, one of the types of monomers and chain segments were preferentially distributed near the adsorbing surface while other monomers and chain segments were preferentially located in the solution. As a result, the chemical sequence and conformation of growing macroradical become mutually dependent. The growth of polymer chains during the addition copolymerization produced a monodispersed system. Also, for describing the same chemical reaction in the presence of adsorbing surface, we have developed an analytical model based on stochastic processes and probabilistic statistics. Despite the relatively simple formulation of this model, its predictions are in good qualitative (in some cases, semiquantitative) agreement with the simulation. The differences between the analytical and simulation results may be mainly due to the fact that our probabilistic model does not include intrachain excluded volume effects.

It was shown that copolymerization near the adsorbing surface can lead to a copolymer with a specific quasigradient chemical sequence and power-law long-range correlations in distribution of different monomer units along the chain. The gradient extends along the entire chain for any chain length. Moreover, we have found that the sequence corresponding to the ensemble of the copolymers generated in this way can be viewed as a one-dimensional fractal object with scale-invariant properties. This means that for such a sequence there is no characteristic length, and the correlations should become infinite for infinitely long chain. Inhomogeneous spatial distribution of polymerizing monomers and composition constraints in the resulting copolymer are major reasons behind the specific chain growth and the global statistical nature of sequences on large scales. These numerical results are all consistent with our theoretical predictions.

Actually, the polymerization procedure discussed in this paper can be considered as a prototype of template polymerization (also connected with molecular imprinting) that is one of the most productive and promising approaches to the synthesis of various polymeric materials. The main idea of this technique is to form the desired (inhomogeneous) spatial distribution of monomers, e.g., via binding them to a template, and then to fix this distribution by polymerization. In the present study we had marked concentration inhomogeneities in the reaction bicomponent system. Therefore, by varying the average copolymer composition, defined by the effective monomer-substrate interaction and surface coverage, the copolymers with different chemical sequences can be designed and synthesized in a controlled fashion. Moreover, just by radical copolymerization of two monomers with different affinity to a certain plane surface it is possible to obtain copolymers with a gradient chemical sequence.

Acknowledgment. We would like to thank S. I. Kuchanov for stimulating discussions. The financial support from Alexander-von-Humboldt Foundation, Program for Investment in the Future (ZIP), INTAS (Project No. 01-607), SFB 569, and Russian Foundation for Basic Research is highly appreciated. A.V.B. acknowledges the "Deutscher Akademischer Austauschdienst" (DAAD) for financial support.

References and Notes

(1) Khokhlov, A. R.; Khalatur, P. G. *Physica A* **1998**, *249*, 253.
(2) Khalatur, P. G.; Ivanov, V. I.; Shusharina, N. P.; Khokhlov, A. R. *Russ. Chem. Bull.* **1998**, *47*, 855.
(3) Khokhlov, A. R.; Khalatur, P. G. *Phys. Rev. Lett.* **1999**, *82*, 3456.
(4) Zheligovskaya, E. A.; Khalatur, P. G.; Khokhlov, A. R. *Phys. Rev. E* **1999**, *59*, 3071.
(5) Khalatur, P. G.; Novikov, V. V.; Khokhlov, A. R. *Phys. Rev. E* **2003**, *67*, 051901.
(6) Shakhnovich, E. I.; Gutin, A. M. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 7195. Shakhnovich, E. I.; Gutin, A. M. *Protein Eng.* **1993**, *6*, 793. Abkevich, V. I.; Gutin, A. M.; Shakhnovich, E. I. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 839.
(7) Pande, V. S.; Grosberg, A. Yu.; Tanaka, T. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 12972. Pande, V. S.; Grosberg, A. Yu.; Tanaka, T. *Rev. Modern Phys.* **2000**, *72*, 259.
(8) Irbäck, A.; Peterson, C.; Potthast, F.; Sandelin, E. *Phys. Rev. E* **1998**, *58*, R5249. Irbäck, A.; Peterson, C.; Potthast, F.; Sandelin, E. *Structure* **1999**, *7*, 347.
(9) Gupta, P.; Hall, C. K.; Voegler, A. C. *Protein Sci.* **1998**, *7*, 2642. Giugliarelli, G.; Micheletti, C.; Banavar, J. R.; Maritan, A. *J. Chem. Phys.* **2000**, *113*, 5072.
(10) Govorun, E. N.; Ivanov, V. A.; Khokhlov, A. R.; Khalatur, P. G.; Borovinsky, A. L.; Grosberg, A. Yu. *Phys. Rev. E* **2001**, *64*, 040903.
(11) Kuchanov, S. I.; Khokhlov, A. R. *J. Chem. Phys.* **2003**, *118*, 4672.
(12) *Levy Flights and Related Topics in Physics*; Shlesinger, M. F., Zaslavskii, G. M., Frisch, U., Ed.; *Lecture Notes in Physics,* Springer-Verlag: Berlin, 1996.
(13) Grosberg, A. Yu.; Khokhlov, A. R. *Giant Molecules: Here and There and Everywhere...*; Academic Press: New York, 1997.
(14) Kuchanov, S. I. *Adv. Polym. Sci.* **2000**, *152*, 157.
(15) Van den Oever, J. M. P.; Leermakers, F. A. M.; Fleer, G. J.; Ivanov, V. A.; Shusharina, N. P.; Khokhlov, A. R.; Khalatur, P. G. *Phys. Rev. E* **2002**, *65*, 041708.
(16) Kriksin, Yu. A.; Khalatur, P. G.; Khokhlov, A. R. *Macromol. Theory Simul.* **2002**, *11*, 213.
(17) Virtanen, J.; Baron, C.; Tenhu, H. *Macromolecules* **2000**, *33*, 336.
(18) Virtanen, J.; Tenhu, H. *Macromolecules* **2000**, *33*, 5970.
(19) Virtanen, J. Self-assembling of Thermally Responsive Block and Graft Copolymers in Aqueous Solutions. Ph.D. Thesis, Department of Chemistry, University of Helsinki, Helsinki, 2002; 42 pp.
(20) Lozinsky, V. I.; Simenel, I. A.; Kurskaya, E. A.; Kulakova, V. K.; Grinberg, V. Y.; Dubovik, A. S.; Galaev, I. Y.; Mattiasson, B.; Khokhlov, A. R. *Dokl. Chem.* **2000**, *375*, 273.

(21) Wahlund, P.-O.; Galaev, I. Y.; Kazakov, S. A.; Lozinsky, V. I.; Mattiasson, B. *Macromol. Biosci.* **2002**, *2*, 33.
(22) Siu, M. H.; Zhang, G.; Wu, C. *Macromolecules* **2002**, *35*, 2723.
(23) Lozinsky, V. I.; Simenel, I. A.; Kulakova, V. K.; Kurskaya, E. A.; Babushkina, T. A.; Klimova, T. P.; Burova, T. V.; Dubovik, A. S.; Grinberg, V. Ya.; Galaev, I. Yu.; Mattiasson, B.; Khokhlov, A. R. *Macromolecules* **2003**, *36*, 7308.
(24) Berezkin, A. V.; Khalatur, P. G.; Khokhlov, A. R. *J. Chem. Phys.* **2003**, *118*, 8049.
(25) Berezkin, A. V.; Khalatur, P. G.; Khokhlov, A. R.; Reineker, P. *New J. Phys.* **2004**, *6*, 44.
(26) Harwood: H. J. *Macromol. Symp.* **1987**, *10/11*, 331.
(27) Starovoitova, N. Yu.; Khalatur, P. G.; Khokhlov, A. R. *Dokl. Chem.* **2003**, *392*, 242. Starovoitova, N. Yu.; Khalatur, P. G.; Khokhlov, A. R. In *Forces, Growth and Form in Soft Condensed Matter: At the Interface between Physics and Biology*; Skjeltorp, A. T., Belushkin, A. V., Eds.; NATO Science Series II: Mathematics, Physics and Chemistry; Kluwer: Dordrecht, The Netherlands, 2004; Vol. 160, 253 pp.
(28) Velichko, Yu. S.; Khalatur, P. G.; Khokhlov, A. R. *Macromolecules* **2003**, *36*, 5047.
(29) Rouault, Y.; Milchev, A. *Phys. Rev. E* **1995**, *51*, 5905.
(30) Milchev, A.; Landau, D. P. *Phys. Rev. E.* **1995**, *52*, 6431.
(31) Milchev, A.; Landau, D. P. *J. Chem. Phys.* **1996**, *104*, 9161.
(32) Nosaka, N.; Takasu, M. *J. Chem. Phys.* **2001**, *115*, 11333.
(33) Carmesin, I.; Kremer, K. *Macromolecules* **1988**, *21*, 2819. Deutsch, H. P.; Binder, K. *J. Chem. Phys.* **1991**, *94*, 2294. Paul, W.; Binder, K.; Heermann, D. W.; Kremer, K. *J. Phys. II (Fr.)* **1991**, *1*, 37.
(34) Binder, K. in *Monte Carlo Methods in Statistical Physics*; Binder, K., Ed.; Topics in Current Physics 7; Springer: Berlin, 1977.
(35) Allen, M. P.; Tildesley, D. J. *Computer Simulation of Liquids*; Claredon Press: Oxford, England, 1990.
(36) Feller, W. *An Introduction to Probability Theory and its Applications*, 3rd ed.; John Wiley & Sons: New York, 1970.
(37) Peng, C.-K.; Buldyrev, S. V.; Goldberger, A. L.; Havlin, S.; Sciortino, F.; Simon, M.; Stanley, H. E. *Nature* (*London*) **1992**, *356*, 168.
(38) Peng, C.-K.; Buldyrev, S. V.; Havlin, S.; Simons, M.; Stanley, H. E.; Goldberger, A. L. *Phys. Rev. E* **1994**, *49*, 1685. Hu, K.; Ivanov, P. Ch.; Chen, Z.; Carpana, P.; Stanley, H. E. *Phys. Rev. E* **2001**, *64*, 011114.
(39) Dokholyan, N. V.; Buldyrev, S. V.; Havlin, S.; Stanley, H. E. *Phys. Rev. Lett.* **1997**, *79*, 5182.
(40) Fay, R.; Jerome, R.; Tessie, Ph. *J. Polym. Sci., Polym. Phys. Ed.* **1982**, *20*, 2209.
(41) Hashimito, T.; Tsukahara, Y.; Tachi, K.; Kawai, H. *Macromolecules* **1982**, *16*, 648.
(42) Davis, K. A.; Matyjaszewski, K. *Statistical, Gradient, Block and Graft Copolymers by Controlled/Living Radical Polymerizations*, Springer: Berlin and New York, 2002.

MA0487094