6.18 Conformation-Dependent Design of Synthetic Functional Copolymers

J Genzer, North Carolina State University, Raleigh, NC, USA
PG Khalatur, Ulm University, Ulm, Germany; Russian Academy of Sciences, Moscow, Russia
AR Khokhlov, Moscow State University, Moscow, Russia

© 2012 Elsevier B.V. All rights reserved.

6.18.1 Introduction

Over the past few decades, we have witnessed increasing interest in developing new types of synthetic copolymers via design of their comonomer sequences. Most of the physical and mechanical properties of copolymers, which comprise two or more covalently bonded sequences of chemically distinct monomer species, depend not only on the chemical composition but also on the arrangement of the comonomers in the polymer chain. There may be significant differences, for example, between two polymer systems with the same chemical composition, but different comonomer distributions (i.e., one of which has the comonomers distributed randomly in the chain, while the other has long blocks of each monomer type). In many cases, just the copolymer sequence dictates the structure and properties of the copolymer. To illustrate this, we will mention only two familiar examples.

Synthetic block copolymers (BCPs) can self-assemble into highly ordered supramolecular structures, showing a rich morphological behavior. Depending on the architecture, chemical composition, chain flexibility, and mutual interaction of the chemically distinct segments, BCPs can be tailored in order spontaneously, via a thermally driven microphase separation process, form a number of equilibrium morphologies – periodic arrays of domains with tunable dimensions – such as body-centered-cubic (spherical), hexagonal (cylindrical), lamellar, and bicontinuous (gyroid) mesophases. These ‘weak crystalline’ structures, which are ‘coded’ in the monomer sequence distribution of single macromolecules, represent excellent candidates for preparing functional nanostructured materials with characteristic length scales from 1 to 10⁶ nm in a bottom-up procedure. In particular, those systems exhibiting long-range periodicity have shown feasibility for many applications including, but not limited to, a new generation of electronic and optoelectronic nanodevices, quantum dots, and catalyst arrays for nanotube growth. However, the realization of this potential depends to a large extent on our ability to control the lengths of blocks and their distribution along the chain.

Typically, proteins fold to organize a very specific globular conformation, known as the protein’s native state, which is, in general, reasonably stable and unique. It is this well-defined three-dimensional conformation of a polypeptide chain that determines the macroscopic properties and function of a protein. The folding mechanism and biological functionality are related directly to polypeptide sequence; a completely random amino acid sequence is unlikely to form a functional structure. In this view, polypeptide sequence forces a protein to be more than a collapsed heteropolymer, but rather to assume a highly specific three-dimensional structure. Hence, a fundamental issue is how functional protein sequences, which determine biologically active structures, differ from random sequences. Understanding the relationship between a protein’s sequence and its native structure is one of the hallmarks in modern science.

While some synthetic copolymers – alternating, di-, tri-, and multiblock copolymers – possess an ordered-sequence distribution of monomeric units along the chain backbone, random copolymers (RCPs) represent a special class of macromolecules, which possess a disordered monomer sequence distribution. A series of papers have been published which established that tuning the copolymer chemical composition and comonomer sequence distribution profoundly affects the characteristics of RCPs. Of those characteristics intrinsic to the polymeric materials, most work performed to date addressed the role of the chemical composition. By judiciously
choosing the polymerizing comonomers, RCPs have the potential to act, for example, as polymer blend compatibilizers\textsuperscript{13–15} or adhesion promoters,\textsuperscript{16–21} to mention a few. Bratko et al.\textsuperscript{22–24} investigated the bulk properties of RCPs and observed that the tunability of the polymer’s blockiness may endow RCPs with the ability to act as pattern recognition agents for similarly patterned surfaces.

Having good control over both the chemical composition and the comonomer sequence distribution in RCPs is essential for many applications. While tailoring the content of comonomers in RCPs is relatively straightforward, adjusting the distribution of monomers along the polymer backbone is not. Today, the majority of all polymeric materials is produced using the free-radical polymerization technique.\textsuperscript{25–31} For a limited range of comonomers, anionic and cationic polymerizations are also used.\textsuperscript{32,33} Unfortunately, however, in conventional free-radical copolymerization in solution, control of the incorporation of monomer species into a copolymer chain is practically impossible. Furthermore, in this process, the propagating macroradicals usually attack monomeric units in a random way, governed by the relative reactivities of polymerizing comonomers. This lack of control confines the versatility of the free-radical process, because the microscopic polymer properties, such as chemical composition distribution and tacticity are key parameters that determine the macroscopic behavior of the resultant product.

The absence of control of the incorporation of monomers into the polymeric chain implies that many macroscopic properties cannot be influenced to a large extent. Therefore, much effort has been directed toward the development of controlled radical polymerization (CRP) methods for the preparation of various copolymers (for a review, see Reference 31). CRPs offer the possibility of producing polymers with relatively well-defined properties, while at the same time maintaining the simplicity of radical processes.\textsuperscript{34} These methods are based on the idea of establishing equilibrium between the active and dormant species in solution phase. In particular, the methods include three major techniques called stable free-radical polymerization, degenerative chain transfer technique,\textsuperscript{35} and atom transfer radical polymerization, pioneered by Ando et al.\textsuperscript{35} and Matyjaszewski et al.\textsuperscript{36} Although such synthases pose significant technical problems, these difficulties have all been successively overcome in the past few years. Nevertheless, the procedure of preparation of the resultant copolymers with controlled monomer sequence distribution remains somewhat complicated.

Summarizing the above, presently there is no comprehensive understanding of the interrelation between chemical sequences in synthetic copolymers and the conditions of their synthesis. One has to glance merely at recent literature in synthetic polymer science to realize that the problem of sequence–property relationship is by no means studied comprehensively. In this contribution, we will discuss some new synthetic strategies, which have been developed recently to provide effective control of the chemical sequences. Our main focus will be on RCPs with tunable sequences.

Apart from the introductory section, the chapter is subdivided into two major parts: ‘theoretical approaches’ and ‘synthesis of designed copolymers’.

### 6.18.2 Theoretical Approaches

#### 6.18.2.1 Two Paradigms in Sequence Design

Broadly speaking, sequence design may be defined as an approach aimed at finding an optimum sequence that provides desired properties of the resultant polymer. This requires a scoring function that may typically be based on physical principles, knowledge-based approaches, or a specifically designed function. Alas, insofar as the terms ‘sequence design’ and ‘sequence engineering’ imply, a rational, planned approach to the creation or modification of copolymer structure and function, both still remain beyond our capabilities in a general way.

There are two main paradigms in the sequence design problem. In protein science, the \textit{de novo} sequence design problem consists of finding a sequence of amino acids that fold into a target globular structure. This problem is sometimes called the inverse protein folding problem. Many current methods for \textit{de novo} protein sequence design consist of numerically mutating a sequence until a maximum stability is achieved for the target structure that is usually considered as a ground state. There are a number of excellent reviews that cover this subject.\textsuperscript{37–42} In synthetic polymer chemistry, emphasis is on the development of new methods of synthesis, on the control of (co)polymer stereochemistry and architecture, and on the design of high-performance polymeric materials tailored for specific uses and properties. The difference between the two sequence design concepts is related to several essential differences between natural and nonnatural copolymers. We mention here only a few of them.

The order of amino acids in a polypeptide chain produced by the synthetic apparatus of the living cell is always the same for a particular protein so that all the protein sequences of a given type are structurally identical copies in every cell in a living organism. We cannot distinguish one individual protein sequence from another. For most synthetic copolymers, produced industrially or synthesized in research laboratories, the occurrence of a certain degree of sequence disorder is inevitable. Therefore, if we are speaking about a synthetic copolymer sequence, we mean, explicitly or implicitly, that averaging over many different sequences has been carried out. For a protein to function, it must be in its highly specific native conformation that is stable only in a narrow temperature region. Secondary and tertiary structural motifs are the central issues to protein structure. In contrast, the properties and functions of synthetic copolymers are not so tightly related to their conformation. Moreover, we are mainly interested in the nonunique polymer conformations.

This list is far from exhaustive, but should rather be taken as simply a set of examples. In the present chapter, we will deal mainly with synthetic macromolecules and practically will not touch on biopolymers.

#### 6.18.2.2 How Can We Characterize Copolymer Sequence?

##### 6.18.2.2.1 HP model

There are many kinds of monomers used to make up copolymers. These differ in physical and chemical properties. One of the most important differences (essential features) is their solubility, that is, how much they like or dislike a given solvent (e.g., water). For example, while protein structure is generally determined by permutations of the 20 amino acids, the
primary distinction between amino acid side chains can be regarded from the very rough viewpoint to be depended only on their hydrophobicity. This proposition is supported by \textit{de novo} protein design experiments showing that the binary pattern of hydrophilic (more polar) and hydrophobic (less polar) amino acids plays a significant role in secondary\cite{44} and tertiary\cite{45} structure determination. Indeed, the most important difference between the monomeric units of globular proteins is that some amino acid residues are hydrophobic, while others are hydrophilic or charged.\cite{46-48} Following this line, simple binary monomer alphabets have been widely exploited in simulation studies of the thermodynamics and stability of folded lattice proteins\cite{42,51,52} and random heteropolymer globules.\cite{53,54} Hence, the chemical and atomistic details of different monomeric units may not be necessary to understand many basic properties of copolymers, including synthetic ‘two-letter’ copolymers.

In what follows, we will mainly use the so-called HP model\cite{51} that distinguishes only two kinds of monomeric units, namely hydrophobic (H) and hydrophilic (or polar, P). This coarse-grained model of a linear hydrophobic/hydrophilic macromolecule reflects the spirit of minimalist models, in that it is simple yet based on a physical principle.

Despite the apparent simplicity of the HP model, the information complexity of an $N$-unit sequence encoded only by H and P symbols can be arbitrarily high for large $N$ (strictly speaking, in the $N \rightarrow \infty$ limit). Indeed, using the two-letter ‘*--’ Morse alphabet, we can write the whole human history in a message of reasonable length.

\subsection{6.18.2.2.2 Long-range correlations versus short-range correlations}

Diblock and repeated-block copolymers are the simplest examples of binary copolymers made up of two different monomer species. More sophisticated distributions of chemically different groups along the chain are characteristic of random and random-block copolymers, including uncorrelated or ideal RCPs and so-called correlated RCPs. In the former class, the corresponding chemical sequences are uncorrelated that corresponds to Bernoulli or zeroth-order Markov processes.\cite{59} In the latter class, the correlation in the sequences of both types of segments is defined by means of a first-order Markov process. It is important to emphasize that in both cases the correlations characterizing the distribution of monomeric units along copolymer chains decay exponentially. The type of primary structure that emerges in these copolymers can be characterized as ‘random with short-range correlations (SRCs)’. Such correlations will always show up after the copolymerization process, if the probability of addition of chemically distinct units to the growing chain depends on the type of the unit, which was added on the previous polymerization step.\cite{51} There are, however, such copolymers for which this is not the case. In this review, we will consider just these copolymers. To clarify the point in question, we will undertake again a brief excursion into biology.

The key biological macromolecules – proteins, DNA, and RNA – are responsible for functions, which are incomparably more complex and diverse than the functions, which we are normally discussing for synthetic polymers. The molecular basis for this ability to perform sophisticated functions is associated with unique chemical sequences of these biopolymers. In particular, a protein sequence as a whole determines the globular conformation and hence biological function. If this sequence is cut into two pieces, those pieces normally neither correspond to a soluble globule nor have any biological function (\textbf{Figures 1(a) and 1(b)}). The same is true for statistically complex DNA sequences, which encode in a four-letter alphabet all genetic information and exhibit significant correlation on different scales.\cite{56}

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure1.png}
\caption{(a, b) If a protein sequence is cut into two pieces, the pieces neither correspond to a native globule nor have any biological function. (c, d) The sequence distribution for rather large sections of a synthetic RCP is practically identical to that of the whole polymer.}
\end{figure}
All these peculiarities are connected to long-range correlation (LRC) effects, and the corresponding sequence, which cannot be divided into shorter subsequences with similar statistical patterns and functional features, may be termed an ‘inseparable sequence’. Such sequence integrity and LRCs are not characteristic of the majority of synthetic linear copolymers, the primary structure of which is chemically homogeneous on a large scale. Indeed, the sequence distribution for rather large sections of synthetic copolymers is normally practically identical to that of the whole polymer (Figures 1(c) and 1(d)).

Thus, there are two fundamentally different classes of sequences: those that exhibit LCR and those that do not possess this property.

It is instructive to discuss briefly the quantities used to measure LRC. To monitor the long-range statistical properties of sequences, the method developed by Peng et al. in their search for LRC in DNA sequences is usually employed. In this approach, each HP copolymer sequence is transformed into a sequence of +1 and −1 symbols, 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 +1 and −1 symbols inside the window is counted at each step. This number g(λ) is a new random variable that depends on the position k of the window along the sequence. The variable g(λ) has a certain distribution. Its average is determined by the overall sequence composition, and its variance is given by Dg = ⟨g2⟩ − ⟨g⟩2, where ⟨…⟩ represents the average over all windows of size λ and generated sequences.

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 Dg scales as 1/λα with a window of a sufficiently large λ. A power law Dg ∼ λ−α with a > 1/2 will then manifest the existence of LRC. At this point, we note that LRC is not equivalent to ‘random-blocky’ sequences. In the case of the latter, as we will see, the exponent α is always equal to 1/2 if the sequence length is large enough.

In some cases, because of large fluctuations, conventional scaling analysis cannot be applied reliably to the relatively short sequences characteristic of natural/synthetic polymers or generated in typical simulations. To avoid this problem, one can use the so-called detrended fluctuation analysis method, which was specifically adapted to handle problems associated with short nonstationary sequences. This approach leads to a special function Fg(λ) that characterizes the detrended local fluctuations within a window of length λ. Generally, Fg(λ) shows the same behavior as Dg. There are many other methods for monitoring LRCs in copolymers.

### 6.18.2.2.3 Information complexity

In a sense, the presence of LRC in a given copolymer sequence characterizes the ‘information complexity’ of the sequence. It is natural to expect that the content of information in the sequences of biopolymers (i.e., proteins, DNA, RNA) is relatively high in comparison with random sequences where it should be almost zero. If we then analyze the primary structures of the globular proteins and compare them with the simple primary structure of conventional synthetic copolymers, we should draw the perceptive conclusion that ‘protein-originated texts’ are much more informative and specific.

Intuitively, complexity lies somewhere between order and disorder or between regularity and randomness. Complexity has been measured by logical depth, metric entropy, information content (Shannon’s entropy and related characteristics), fluctuation complexity, and many other techniques. These measures are well suited to specific physical or chemical applications, but none describe the general features of complexity.

A simple measure capable of indicating how far a copolymer sequence differs from trivial random sequences is based on the so-called Jensen–Shannon (JS) divergence measure. Let us explain how it can be defined.

Let S = {s1, ..., sn} be a sequence of N symbols. For two subsequences S1 = {s1, ..., sk} and S0 = {s0,1, ..., s0} of lengths n and N − n, the difference between the corresponding discrete probability distributions f1(s1, ..., sk) and f0(s0,1, ..., s0) is quantified by the Jensen–Shannon divergence

$$\mathcal{JS}(S_1, S_2)/N = h(S) - \frac{n}{N} h(S_1) + \frac{N-n}{N} h(S_2)$$

where S = S1 ⊕ S2 (concatenation) and h(S) is Shannon’s entropy of the empirical probability distribution obtained from block frequencies in the corresponding subsequence. Of course, Shannon’s entropy depends on the definition of a set of words in the sequence. For two-letter HP copolymers, one can adopt the following set of words (uniform blocks): H, HH, HHH, ..., P, PP, PPP, ...; that is, word (block) is defined by its length t and type. In this case, Shannon’s entropy per monomer can be written as

$$h = \frac{N_w}{2N} \sum_{t=1}^{N_w} \left[ f_H(t) \log_2 f_H(t) + f_P(t) \log_2 f_P(t) \right]$$

where f_H(t) and f_P(t) are the frequencies of blocks of length t composed of letters H and P, respectively, and N_w is the total number of words.

The Jensen–Shannon divergence JS is zero for sequences with the same statistical characteristics; it takes higher values for increasing differences between the statistical patterns in the subsequences, and reaches its maximum value for a certain set of distributions. In particular, both purely random and any regular (multiblock) copolymers of infinite length show JS = 0. We normally expect that a completely random sequence or a sequence with long uniform blocks contains less information than a sequence containing many different blocks (words) of medium length. Of course, we cannot unambiguously distinguish only by sequence analysis between what might be called quantity and quality of information.

It is clear that information complexity and LRCs cannot emerge just as a result of random permutations of monomeric units in a copolymer chain. Some coupling of chemical modifications to other factors is necessary. It is most straightforward to introduce the interrelation of chemical modifications and conformational changes. Such a consideration leads naturally to a concept of conformation-dependent sequence design (CDS) in synthetic polymer chemistry.

### 6.18.2.3 Conformation-Dependent Sequence Design: Structure Dictates Sequence

In a series of publications, a concept of CDS of functional copolymers has been introduced (for reviews, see...
Reference 72–77). The essence of the approach is based on the assumption that a copolymer obtained under some preparation conditions is able to ‘remember’ features of its original (parent) conformation in which it was built up and to store the corresponding information in the resultant chemical sequence. In other words, this concept takes into account a strong coupling between the conformation and primary structure of copolymers during their synthesis.

Ideologically, the approach69–71 bears some similarities with that proposed earlier in the context of the problems of protein physics42,46–50 however, it aims at synthetic copolymers rather than biopolymers. The original idea for protein design consists of running through sequences of amino acids to determine which sequence (or sequences) has the lowest energy in a unique (target) conformation. We have stressed the differences between sequence design characteristic of natural and synthetic copolymers.

Presently, there are three known CDSD techniques: (1) the chemical modification (polymer-analogous transformation (PAT)) of homopolymers; (2) the step-growth copolymerization of monomers with different properties under special conditions; (3) the evolutionary and theoretically assisted sequence design approaches. We will address all these techniques.

6.18.2.3.1 Polymer-analogous transformations

Controlling the chemical composition and architecture of polymeric materials is important in many technological applications. While in most instances the resultant polymer can be prepared conveniently using existing synthetic methods, in other situations generating functional polymers can be challenging due to various reasons. These include the difficulty of finding a suitable polymerization media, the inability of existing synthetic methods to polymerize certain monomers, and complications arising from uncontrolled chemical side-reactions on a given monomer. These barriers can be overcome to some extent by implementing postpolymerization protocols (PPPs), also referred to as PATs, where a homopolymer is modified chemically with a selective reagent. Modification of polymers by PAT has been known for a long time (see, e.g., the Nobel Prize lecture of Hermann Staudinger78) and presently is well investigated. This synthetic strategy can be exemplified, for instance, by the bromination of polystyrene (PS).79 Polymer-analogous reactions represent deliberate changes of functional groups that are carried in macromolecular chains with the general objective of maintaining the polymerization degree of the original macromolecules.

Recent years have witnessed increased interest in utilizing PAT to create specialty polymers. In fact, some industrial processes, such as modification of polyolefins,80 employ PATs routinely. Progress in this field has recently been documented in a comprehensive review by Gauthier et al.81 While PAT can be used to convert a polymer made of a given type of segments, say A type, to a macromolecule made of chemically different units, say B type, by converting A into B, there exists a vast class of ‘in between’ materials that involve A1−xBx copolymers, where x is a fraction of the A units that have been converted chemically into B moieties. Thus, by adjusting the degree of chemical modification of the bare A-homopolymer, one can prepare heteropolymers consisting of various degrees of chemical modification along the polymer backbone. In most instances, the distribution of the newly added B units in resulting chains is random, thereby leading to A1−xBx RCPs. However, by fine-tuning the condition of the PAT chemical reaction, one can also alter the degree of randomness of the A and B species. In particular, this can be done using a synthetic strategy based on the CDSD approach.

6.18.2.3.1(i) Protein-like copolymers

It is known that the hydrophilic P units in globular proteins are mainly covering the surface of the globule, giving rise to their stability against intermolecular aggregation, while hydrophobic H units are forming primarily the core of the globule.82 Such requirement is rather restrictive, that is, it is satisfied only for a very small fraction of all possible primary structures. Moreover, since the HP correlations defined in such a way depend on the conformation of the globule as a whole (i.e., on the protein ternary structure), they should be characterized as long-range ones. From a practical viewpoint, the question is whether such primary structures can be obtained for binary copolymers, not obligatorily of biological origin.

In the late 1990s, Khokhlov and Khalatov69,70 devised a clever PAT-based methodology for adjusting the spatial distribution of comonomer species in two-letter RCPs. It is easy to do this by computer simulation, and much more difficult in real experiments. In both cases, however, the corresponding procedure should involve the four stages that are schematically depicted in Figure 2.

Stage 1: We take a bare homopolymer chain with excluded volume interactions in a good solvent (see Figure 2(a)).

Stage 2: Strong attraction between all monomeric units is switched on and a homopolymer globule (‘parent’ conformation) is formed (Figure 2(b)). Of course, when we are speaking about real experiments, by switching on the attraction we should understand the jump of temperature, addition of poor solvent, and so on.

Stage 3: This step is much more easily realizable in computer experiments. We should simply consider the ‘instant snapshot’ of the globule and ‘color’ the units on the surface and in the core in different ‘colors’, that is, we assign the index P to those units that are on the surface of the globule and call these units hydrophilic and assign the index H to the units in the core of the globule and call these units hydrophobic. By doing so, we fix this primary structure (Figure 2(c)). In real experiments, the ‘coloring’ of the surface can typically be done with a chemical reagent Z entering in the reaction with monomeric units and converting them from hydrophobic to hydrophilic, H + Z → P. In chemical language, it is called PAT. If the amount of the reagent is small enough, only surface monomeric units will be contaminated, the core remaining hydrophobic. Another important feature is to have fast enough ‘coloring reaction’ and slow enough intermolecular aggregation (which will always take place under the conditions when globules are formed). To slow down the aggregation, the neutral thickeners of the aqueous solutions may be used.
Stage 4: This last step is necessary for computer realization. The uniform strong attraction of units should be switched off, and different interaction potentials should be introduced for H and P units (Figure 2(d)). As a result, we obtain a specific copolymer called ‘protein-like copolymer’ (PLC).\textsuperscript{69} Initially, the protein-like HP sequences were generated in References \textsuperscript{69} and \textsuperscript{70} for a lattice chain, using a Monte Carlo method and the bond-fluctuation model.\textsuperscript{83} When the chain is a random heteropolymer, an average over many different sequences distribution must be carried out explicitly to produce the final properties. Therefore, the sequence design scheme\textsuperscript{69,70} was repeated many times, and the results were averaged over different initial configurations.

It turned out that many statistical properties of protein-like and RCWs with the same HP composition are very different. In order to distinguish whether this difference is due to the special sequence design described above, or just due to the different degree of blockiness, one can introduce for comparison also a ‘random-blocks’ primary sequence. We assume that random-blocky HP copolymers have the same chemical composition and the same average length \(L\) of uninterrupted H or P sequence as PLWs, but in other respects, the HP sequence is random. In Reference \textsuperscript{69}, the distribution of block length \(\ell\) was taken in the Poisson form: \(f(\ell) = e^{-\lambda\ell}/\ell!\).

In Figure 3, we present the typical distributions of H and P monomers along the chain for regular (multiblock) copolymer as well as for truly random copolymers, random-blocky copolymers, and PLWs. From the comparison of the primary structures of the random and PLWs, we can see that the average lengths of H and P blocks in the PLCs are notably longer. In contrast, the copolymer with the random-blocky architecture, having the same average length of H and P sections, \(L_{H}\) and \(L_{P}\), as for the PLC, exhibits a different distribution of these blocks along the chain. The main feature of the protein-like sequences is the presence of rather long uniform H and P sections.

In Figure 3, we also show the variance \(D_{\lambda}\) as a function of \(\lambda\) for protein-like sequences with a 1:1 HP composition. For comparison, the data for two other types of sequences are also presented. One of them is a purely random 1:1 sequence; it demonstrates \(D_{\lambda} \propto \lambda^{1/2}\) scaling, as expected. Comparing this curve with simulation results,\textsuperscript{84} we see immediately that the protein-like sequence is not random; rather well-pronounced correlations exist in it. Thus, it is interesting to compare the simulation data with the Poisson distribution adjusted to achieve the same 1:1 composition and the same degree of blockiness as for the corresponding PLC. This model sequence exhibits a somewhat more rapid variation of \(D_{\lambda}\) at small \(\lambda\), but ultimately the law \(D_{\lambda} \propto \lambda^{1/2}\) is obeyed for large values of \(\lambda\). Nevertheless, this random-blocky model is also seen to be unsatisfactory for the statistical behavior of a protein-like sequence throughout the interval of \(\lambda\) examined, \(2 < \lambda < 512\). Although the simulation data do not fit accurately to any power law \(D_{\lambda} \propto \lambda\), the slope of the observed \(D_{\lambda}\) dependence corresponds to a significantly larger value than 1/2, up to about \(\alpha = 0.85\), thus indicating pronounced LRCs in a protein-like sequence. Thus, the primary structure emerging in the case of PLWs can be characterized as ‘quasi-random with long-range correlations’. Here, we stress again that there is the fundamental difference between the sequences with LRC and random-blocky sequences.

The findings discussed above are surprisingly similar to those known for DNA sequences, which appeared as a mosaic of coding and noncoding patches\textsuperscript{56–58}. Indeed, like DNA chains, the PLC sequence 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 vice versa), and the DNA chain is fully...
correlated throughout its whole length. As a result, the $D_n^2$ (or $D_n^2$) curve does not contain the linear portion $D_n^2 \propto n$. In principle, the same behavior is observed for protein-like sequences.\(^{77}\)

Govorun et al.\(^{84}\) have shown, both by exact analytical theory and by computer simulations, that the corresponding chemical sequence is nonalternating and demonstrates the specific LRC, which can be described by the statistics of the Lévy-flight type.\(^{85}\) For such probabilistic processes, an observable stochastic variable $x$ exhibits large jumps (flights), called ‘Lévy flights’ (LFs), characterized by power-law (rather than exponential) probability distribution function $f(x)$. The possibility for exact analytical description of sequences resulting from surface ‘coloring’ (Figure 2(c)), comes from the fact that the statistics of polymer chains inside dense globule is Gaussian, that is, it is described by the ordinary diffusion equation. One has only to worry about correct boundary conditions, and this problem was resolved in Reference 84. The probability distribution of block lengths $\ell$ in PLC can be written as\(^{84}\)

$$f_H(\ell) = \frac{\pi b^2}{3R^2} \sum_{n=1}^{\infty} \frac{n^{-1} \sin \left( \frac{n\pi b}{R} \right) \exp \left[ -\frac{b^2}{6} \left( \frac{n\pi}{R} \right)^2 \ell \right]}{\ell} \quad [3]$$

for H units and

$$f_P(\ell) = \frac{1}{3(2^{1/3} - 1)} \sum_{n=1}^{\infty} \frac{\zeta_n^2 \sin \left( \zeta_n b/\zeta_n R^{-1} \right)}{\zeta_n \cos \zeta_n}$$

$$\exp \left[ -\frac{b^2}{6} \left( \frac{\zeta_n}{R^{(2^{1/3} - 1)}} \right)^2 \ell \right] \quad [4]$$

for P units. Here, $b$ is the bond length, $R = b(3N/8)\^{1/3}$ defines the characteristic size of the bare homopolymer globule, and $\zeta_n$ satisfies $\zeta_n = (1 - 2^{-1/3})\tan \zeta_n$.

Analytical theory\(^{84}\) suggests that the LF-type statistics, albeit with a broader crossover region, is expected even if parental (globular) conformation used to generate protein-like molecules is not maximally compact, but rather a globule somewhat closer to the 0-point. Note that, strictly speaking, this distribution is not a ‘true’ LF distribution, which falls to zero as $1/|x|^{\alpha + 1}$ (where $0 < \alpha < 2$) and, therefore, has an infinite variance, but represents the so-called truncated Lévy flight (TLF) with a finite variance for finite chain lengths $N$.\(^{84,85}\)

As has been shown,\(^{84}\) the copolymers generated with the use of eqns [3] and [4] reproduce the coloring procedure described in References 69–71: monomers located in the core of globule are set to be H-type, while monomers belonging to a globular surface are assigned to be of P-type. An analogous result was obtained in a later work\(^{86}\) for the model that takes into account more chemical details.

The presence of LRCs in designed sequences is a very important feature. It is easy to understand that these correlations are due to the fact that assigning of the type of chain segments (H or P) under the preparation conditions described above depends on the conformation of the parent 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 in two-letter alphabet the spatial (core–shell) structure of a copolymer globule. Obviously, this 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 smaller lengths.

Because of the finite size of the bare globule subjected to CDS, the longest correlations that can be found in this case are also finite. However, the range of these correlations is much larger than for usual synthetic copolymers obtained, for example, via radical copolymerization under homogeneous conditions.\(^{31}\)

The question of whether natural proteins originate from random sequences of amino acids was addressed in many

\[\text{Figure 3.} \text{ Left panel: Typical distributions of H and P monomeric units along the chain for regular (multiblock) sequence having fixed block length of 8 units, as well as for purely random sequence with average block length} L = 2, \text{ random-blocky sequence with} L = 6.4, \text{ and protein-like sequence with} L = 6.4. \text{ The H units are denoted as } *1 \text{ and the P units as } -1. \text{ Right panel: Variance of the number of different monomeric units in the fragment of sequence of size } n \text{ for protein-like, random, and random-blocky copolymers. The sequence length } N = 1024.\]
publications. It was demonstrated that protein sequences are not completely random sequences. In particular, the statistical distribution of hydrophobic residues along chains of functional proteins is nonrandom. Furthermore, protein sequences derived from corresponding complete genomes display a distinct multifractal behavior characterized by the so-called generalized Rényi dimensions (instead of a single fractal dimension as in the case of self-similar processes).

It was also found that the coil-to-globule transition for PLs, induced by the strong attraction of H segments, occurs at higher temperatures and leads to the formation of denser globules and has faster kinetics than for random and random-block counterparts. The reason for this is illustrated in Figure 4 where typical snapshots of globules formed by protein-like and RCPs are shown.

One can see that the HP heteropolymer obtained as a result of the simple one-step ‘coloring procedure’ can self-assemble into a segregated core–shell microstructure, thus resembling some of the basic properties of globular proteins. The core of protein-like globule is much more compact and better formed as compared to that observed for RCPs; it is surrounded by the loops of hydrophilic segments, which stabilize the core.

Apparently, this is due to some ‘memory effect’: the core, which existed in the parent globular conformation (Figure 4(a)), was simply reproduced upon refolding caused by the attraction between H units. One may say that the features of parent conformation are inherited by the PLC. Looking at the conformations of Figure 4, it is natural to argue that PLC globules could be soluble in water and thus are open to further chemical modification, while RCP globules will most probably precipitate. It is clear that the main reason for the deviation of the properties of PLCs from those for RCPs is the special sequence design scheme, not just difference in the degree of blockiness.

6.18.2.3.1(ii) Statistically controlled monomer sequence distributions

A drawback of the simplest sequence design scheme described above is that either a copolymer sequence is templated or it is random, with little control in between. In their initial work, Khokhlov and Khalatur assumed that the polymer-analogous reaction was instantaneous, that is, the parent homopolymer before chemical modification was ‘frozen’. It is easy to understand that instantaneously modification of the parent homopolymer is almost impossible to realize experimentally since any reaction between the homopolymer units and the reactant species is kinetically limited and will take place on a much longer time scale than the monomer and coil relaxation times.

Semler and Genzer extended the original method of Khokhlov and Khalatur by allowing the copolymer to reconfigure during the chemical modification process, based on the reasonable expectation that the copolymer’s conformation would evolve during this process. To this end, they used Monte Carlo simulations to model the formation of PLs with tunable monomer sequence distributions.

The simulations were carried out in a two-step computational scheme. First, using a Monte Carlo method and a lattice chain model with an implicit solvent, the collapse transition was mapped out via varying the reduced monomer–monomer interaction potential \( k_B T \). This procedure provided the initial homopolymer configurations for the subsequent ‘coloring’ simulation. The degree, to which the homopolymer coil is collapsed, was achieved by adjusting the inter- and intramolecular potentials acting among the chain monomers. In the second process, another Monte Carlo simulation was conducted, which involved the PAT performed directly on the homopolymer chain of a given state of collapse. This simulation scheme used an explicit low-molecular-weight reactant that had a predetermined reaction probability with the monomeric units along the homopolymer. A threefold excess of reactant was added at random to the simulation box. During this simulation, a monomer unit along the length of the chain had a certain probability to either react with the reactant or make a conformational move. Reactions were allowed to occur with a predetermined probability (set to either 5% or 25%) if the distance between the reactant and monomer along the chain was less than a predetermined reaction radius. The simulation progressed until a specified percentage of the chain was colored (two degrees of coloring was considered: 40% and 60%). Both the chain collapse and coloring simulations were performed many times in order to access the statistics of the process.

The simulation results revealed that regardless of the length of the parent homopolymer, all monomers along the fully expanded chains possessed an equal probability of coloring. In contrast, the probability of coloring monomer units in collapsed chains was distributed widely. Increasing the chain length resulted in monomer sequences with broader distributions of colored units. Furthermore, expanded coils exhibited an upsurge in the frequency of occurrence of short sequence lengths of the colored units, while collapsed coils contained larger fractions of long sequential monomer lengths. A higher frequency of long-colored blocks was also detected upon increasing the chain length of the parent homopolymer. A simple statistical analysis revealed that while the sequence distribution of the colored species in polymers colored in an expanded state was random, chains colored under collapsed conformations transformed into heteropolymers, which possessed statistically self-similar sequence distributions. The latter sequences can be identified as protein-like ones.

Figure 4 Typical snapshots of the globular conformations of (a) protein-like, (b) random, and (c) random-block copolymers of the same length \((N=512)\). Hydrophobic segments are shown in light-red color and hydrophilic segments in green color.
The fact that distinct coloring sequences can be achieved by coloring parent homopolymers over time (rather than performing instantaneous coloring as performed in the original sequence design scheme of Khokhlov and Khalatur[26-27]) has important implications for preparation of RCPs with distinct and adjustable monomer sequences.

**6.18.2.3.1(iii) Hydrophobic modification of hydrophilic polymers**

The segregated core–shell microstructures, consisting of a hydrophobic core surrounded by a hydrophilic shell, are of great practical interest as their mechanical properties are mainly influenced by the core polymer and the chemical properties and solubility mainly by the shell monomer units. These materials have attracted increased attention because they not only maintain the functions of both the core and shell components but also exhibit additional excellent optical, electrical, and magnetic properties. Polymeric core–shell microstructures can be obtained through phase separation, which is realized by inductive polymerization, solvent extraction and evaporation, self-assembly of amphiphilic block copolymers, or sequential precipitation, which is intrinsically a self-assembly and phase-separation process. However, these methodologies are usually complicated and tedious.

A simple CDSD-based approach to synthesis of copolymers with a specific sequence capable of forming core–shell microstructures has been suggested. The idea is rather similar to that discussed in the previous sections but bears an important difference: instead of the hydrophilization of globular surface, it is suggested to perform a sequential hydrophobization of a hydrophilic polymer chain (poly-P) dissolved in a good solvent, using a low-molecular-weight hydrophobic modifier (H) poorly soluble in this solvent. The approach is illustrated schematically in Figure 5.

In a dilute solution, when the polymer is in a coiled state (Figure 5(a)), the diffusion of hydrophobic particles into the coil is normally faster than the chemical reaction. In this case, the local concentration of particles H inside the swollen coil is practically the same as in the bulk. Therefore, we expect that at the initial stage, the reaction will lead to a RCP: some of the P monomeric units will attach H reagent and thereby they will acquire amphiphilic (A) properties: P + H → A (Figure 5(b)). As long as the number of the modified A units is not too large, the chain remains in a swollen coil-like conformation (Figure 5(b)). However, when this number becomes sufficiently large, the hydrophobically modified polymer segments tend to form intrachain micelle-like aggregates (Figure 5(c)). This is due to the loss in translational entropy of covalently bonded H species after their grafting to the hydrophilic backbone. Structurally, the intramolecular micelles are similar to reverse micelles formed by free low-molecular-weight surfactants, and like ordinary micelles they should solubilize hydrophobic species. The presence of the intrachain aggregates can dramatically alter the reaction conditions. Because of preferential adsorption, the poorly soluble modifier will diffuse inside the A-rich regions, thus leading to spatially inhomogeneous distribution of the concentrations (Figure 5(c)). One can expect that further modification of the hydrophilic sections of the polymer chain will occur predominantly within these microglobular regions. In the course of chemical reaction, they would progressively increase in size and then coalesce. As a result, a nonrandom microblock distribution of the chemically modified chain segments emerges. Finally, we should have hydrophobically modified segments inside rather compact conformation formed by the resultant hydrophilic–amphiphilic (PA) copolymer. In that way, one can expect the formation of core–shell morphologies with inner (poorly soluble) core and outer (well-soluble) hydrophilic cover (Figure 5(d)). In principle, the polymer-analogous reaction can be terminated at any desirable time by quenching (e.g., by lowering the temperature) or by stopping the supply of reactive compounds.

The algorithm used in Reference 98 simulated a reaction in a three-dimensional cube, utilizing periodic boundary conditions. Initially, an N-unit hydrophilic homopolymer (poly-P) in a coil state and N freely diffusing hydrophobic (H) monomers were placed in the cube. For the conditions considered, despite the attraction between H monomers, they were soluble due to their high translational entropy. To treat the hydrophobic interactions, an approach based on the solvation free energy (SASA) model was used. Upon contact between a monofunctional reactant H particle and a P bead on the chain, a bond could be formed between the two. This led to the formation of an amphiphilic’ monomer unit (PH amphiphilic dumbbell): P + H → A. The reaction was considered as a sequence of the alternating steps: grafting of a new H monomer to the chain and the subsequent long relaxation of the chain. The process was terminated when the required number of the hydrophilic monomer units was transformed into the A type. The composition of the resultant copolymer was constrained so that there were 75% hydrophilic and 25% amphiphilic monomer units. As a result, a hydrophilic–amphiphilic PA copolymer with certain distribution of P and A units along the hydrophilic polymer backbone was obtained. To gain better statistics, ~10^3 hydrophobically modified copolymer chains for each set of the parameters were independently generated and then the required average characteristics were found.

To give a visual impression of the simulated system, Figure 6 presents a typical snapshot of an amphiphilic copolymer having 256-unit hydrophilic backbone with 64 attached hydrophobic side groups. It is seen that, using the synthetic
strategy described above, one can indeed end up with a copolymer having a dense hydrophobic core surrounded by a hydrophilic shell.

From the picture presented in Figure 6, one can expect that the sequential hydrophobization of a polymer coil should lead to a copolymer with a nonrandom sequence distribution. This is indeed the case. As an example, let us consider the average number fractions of blocks consisting of $\ell$ neighboring amphiphilic monomers, $f_{\ell}(t)$, occurring in a copolymer chain. Some results are shown in Figure 6 on a semilogarithmic scale.

We expect that for a random distribution of monomers A incorporated into the polymer chain, the function $f_{\ell}(t)$ should decay exponentially with increasing $\ell$. In fact, such a behavior is observed for the copolymer modified in a solvent, which is good for the low-molecular-weight modifier, that is, at $\gamma = 0$. When the solvation parameter $\gamma$ is increased and the solvent becomes selectively poorer for the modifying agent, the values of $f_{\ell}(t)$ are skewed toward A sections of greater length, which implies a copolymer with blocky tendencies. A further worsening of the solvent quality leads to a strong deviation from the exponential decay of $f_{\ell}(t)$, thus indicating a nonrandom copolymer sequence distribution.

Thus, copolymers of the same composition can have qualitatively different sequence distributions depending on the solvent in which the chemical transformation is performed. In a solvent selectively poor for modifying agent, hydrophobically modified copolymers were found to have the sequence distribution with LRCs, whereas in a nonselective (good) solvent, the reaction always leads to the formation of random (Bernoullian) copolymers. In the former case, the chemical microstructure cannot be described by any Markov process, contrary to the majority of conventional synthetic copolymers.

In general, the microsegregated structures observed for hydrophobically modified polymers (Figure 6) are similar to core–shell globules obtained via ‘coloring’ procedure (Figure 2). However, it should be kept in mind that the experimental realization of the hydrophilic modification of the surface of a hydrophobic globule was shown to be rather unreliable, because of the difficulty to stabilize dense globules in the solution for the time sufficient to implement a PAT. On the other hand, the simulations show that the method based on the hydrophobic modification of soluble polymers should be quite universal and robust.

6.18.2.3.1(iv) Design of PLCs with tunable randomness

Discussion of the further development of the sequence design scheme based on PAT involving macromolecular substances modified by a low-molecular-weight reactant, it is of interest to consider how differences in the solubility of a homopolymer and modified monomeric units affect the sequence of the resulting copolymer. One would expect that as the modification reaction evolves, solubility of the newly created segments along the designed copolymer affects the chain conformation and ultimately the distribution of the chemically different units inside the designed copolymer. It is believed that while coloring the A homopolymer with B species that are more soluble in the solvent should result in opening up the chain, B monomers that are less soluble than the A units would move toward the center of the collapsed polymer. The situation when the modified monomer unit exhibits lower solubility in the solvent than the original unmodified monomer appears to have been the case in experimental work by Semler et al. who carried out a polymer-analogous reaction by brominating parent PS coils in selective solvents.

Recently, Strickland et al. have used computer simulations to examine situations in which the modified monomeric unit exhibited lower, identical, or greater solubility when compared with the original unmodified monomer. They presented the results of discontinuous molecular dynamics (also referred to as event-driven molecular dynamics) simulations of polymer-analogous reactions with special focus on the case in which the chemically modified monomers had different solubility than the parent homopolymer. A system of polymer
molecules, modeled as chains containing 100–300 square-well monomers of A type, was allowed to equilibrate at a selected temperature. After equilibration, athermal reactant particles were placed in the simulation box. When the reactant particles came in contact with A-type monomers, the monomer was modified to become a B-type monomer. This reaction was continued until a desired number of the parent monomers was modified. Since the solvent in the simulations was implicit, the solubility of A- and B-type monomers was adjusted by varying the A–A and B–B interaction strengths. The ratio of B–B interaction strength to A–A interaction strength, \( R_{BA} = |\epsilon_{BA}|/|\epsilon_{AA}| \), was varied from 0.5 to 10, thus allowing to vary the solubilities of A and B species. Interactions between non-bonded A and B units along the chain were modeled as athermal, that is, \( \epsilon_{AB} = 0 \).

The simulations showed that the sequence distribution of A and B species in the resultant copolymer is affected by the interaction strengths acting between the A–A and B–B monomers, that is, their solubilities. In general, when the solubility of the B monomers is not the same as the A monomers (\( \epsilon_{AA} \neq \epsilon_{BA} \)), the chain conformation changes during the coloring reaction by either expanding (when B is more soluble than A) or contracting (when B is less soluble than A). It was found that there exists a minimum degree of modification before substantial chain reconfiguration can take place; typically, reconfiguration requires a minimum of 20\% conversion of A to B. For high values of \( R_{BA} \) (i.e., very low B solubility), the B monomers tend to move toward the center of the globule, causing the globule to invert. Strong A–A interaction and good B solubility led to the formation of tight globules and hence produced copolymers with increased tendency toward random-blocky comonomer distribution. In contrast, weak A–A interaction and good B solubility led to extended homopolymer conformations, which produced copolymers with random comonomer distribution.

More interesting and somewhat unexpected behavior occurred when the solubility of B was poor compared to that of A. One might expect that in cases of poor B solubility, the copolymer would contract in size during the coloring process. Instead, as can be seen by examining the data in Figure 7 (left panel), under poor solvent condition, the globule size increases with increased degree of modification, an effect which, at first glance, is counterintuitive. This initial swelling of the copolymer is explained by the fact that B monomers move into the globule away from the solvent. Thus, as the poor-solvent homopolymer is modified, it begins to expand as if it were in good solvent due to the solubility of B. In essence, the good solubility of B reduces the effective theta temperature of the system as the chain is being modified. One can also observe the opposite trend when the solubility of B species is poor. In this case, the initially swollen chain transforms from coil-to-globule as the chain is modified. In other words, the poor solubility of B species increases the effective theta temperature of the system as the chain is being modified. This effect is similar to that discussed for a hydrophobically modified copolymer in the previous section. Some typical conformations of the resultant copolymers are shown in Figure 7 (right panel).

The results from the simulations are summarized in Table 1, which shows that the changes in chain conformation during the PAT can be divided into four major categories, depending on the initial conformation of the parent homopolymer (globule or coil) and the relative solubility of B (more or less soluble).

Statistical analysis of the generated copolymer sequences revealed that one can effectively control the degree of randomness by simply varying the value of \( k_BT/|\epsilon_{AA}| \).

6.18.2.3.1(v) **Chemical modification of surface-tethered homopolymers**

Very recently, Strickland et al.\(^{102} \) developed a sequence design methodology based on chemical modification of homopolymer macromolecules grafted to flat impenetrable surfaces. As discussed above, previous studies employed PAT by placing the parent homopolymers in a selective solvent and subsequently letting the chemical reagent modify the units on the
homopolymer that were sterically accessible. Since the parent homopolymer adopted either a coil or globule conformation and, most importantly, was always isolated, the chemical modification occurred along the entire backbone of the macromolecule. Selective shielding of specific portions of the chain was not possible. This limitation can be overcome by tethering the parent chains to a surface prior to coloring. Not only does the surface restrict the approach of the reactant to one end of the chain, but neighboring chains would also laterally shield each other depending upon the chain grafting density. While chemical modification of surface-tethered homopolymers has been performed experimentally by several groups, no quantitative information exists on the distribution of the modified species in the resultant copolymers.

In their simulations, Strickland et al. employed discontinuous molecular dynamics to study the coloring process in macromolecular tethers made of various lengths. Specifically, they explored the effect of the grafting density of the parent polymer and the interplay between the solubilities of the unmodified and modified segments along the polymeric grafts on the comonomer distribution in designed copolymers.

The computer simulation proceeded in the following way. First, M homopolymer chains comprising various numbers N of square well monomers (50-, 100-, and 300-mers) were attached to a surface (the z = 0 plane) at various surface densities, \( \rho = M/S \), where S is the area of the tethering surface. Then the homopolymer grafts were equilibrated at a select temperature and surface density. After equilibration, reactant particles \( (Z) \) were placed randomly in the simulation box with lateral \( (x,y) \) periodic boundary conditions. Upon contact between an A-monomer and an Z-particle, the A-monomer was ‘colored’ and converted from A into B species \( (A \rightarrow B) \). The Z-particle was removed and then randomly placed back into the box to maintain a constant reactant concentration. All A–Z interactions were considered effective and any A–B conversion was accepted. This polymer-analogous reaction was carried out until a desired number of B-type monomers was achieved, producing \( A_x B \), designed copolymers, where \( x \) is the mole number fraction of the B segments. The solubility of A- and B-type monomers was adjusted by varying the system reduced temperature, \( k_B T/|\varepsilon_{AB}| \), and the B–B interaction strength, \( \varepsilon_{BB} \).

The ratio of the B–B interaction strength to the A–A interaction strength \( (R_{BA} = |\varepsilon_{BB}|/|\varepsilon_{AA}|) \) was varied from 0.5 to 10. In this manner, the newly created B unit can be modeled as either more \( (|\varepsilon_{BB}| < |\varepsilon_{AA}|) \) or less \( (|\varepsilon_{BB}| > |\varepsilon_{AA}|) \) soluble than A. Interactions between nonbonded A and B units along the chain were modeled as atthermal, that is, \( \varepsilon_{AB} = 0 \). Typically, the reactions were run until the chain was composed of at least 60% B units. The distribution of A and B units along the polymer chain was calculated to determine the effect of solubility, chain length, surface density, and temperature on the comonomer distribution of the surface-tethered designed copolymers.

The simulations demonstrated that the monomer sequence distribution in the resultant copolymer depends on the interplay among (1) the length and the grafting density of the A-based homopolymer anchors, (2) the solubility of the parent homopolymer, and (3) the solubility of the B coloring units. In particular, the monolayer density profile, end-monomer probability profile, and the sequence distribution of A and B species along the designed copolymer chain were found to be functions of the parent homopolymer solubility and temperature, as measured by \( k_B T/|\varepsilon_{AA}| \) and modified monomer solubility, as measured by the \( R_{BA} \) ratio. The blockiness of short copolymers formed at low surface density (especially below the mushroom-to-brush overlap density) is similar to that of copolymers formed in the bulk due to the low-level of chain–chain interactions. When chain length was increased to the point where neighboring chains overlapped, the blockiness increased with increasing surface density. In general, increasing the system temperature reduced copolymer blockiness. High temperature promoted the extension of the chains away from the tethering surfaces, allowing for reactions along the length of the polymers. In contrast, low temperature promoted the collapse of the chains, concentrating the reactions at exposed portions of the polymer. It was shown that the effect of the surface density on blockiness depended on the chain length. In general, for a given increase in surface density, the blockiness of a longer-chained system increased more than that of a shorter-chained system. Increasing surface density at temperatures below the theta-temperature \( (T_\theta) \) had little effect on blockiness but had a large effect at temperatures above \( T_\theta \). In particular, at \( k_B T/|\varepsilon_{AA}| \gg T_\theta \) and high surface density, the distribution of the segments in the brush resembles closely that of diblock copolymers where the block close to the surface comprises primarily unmodified segments A and the block at the periphery of the tethered chain is made of a chemically modified chain section.

The simulations carried out by Genzer and coworkers also revealed how the coloring develops over time as a function of the solvent quality for both A and B units, as well as the grafting density of the parent homopolymer on the flat surface. It was shown that grafting density plays a pivotal role in determining the comonomer distribution along the designed copolymers, which can range from truly random to diblock-copolymer-like.

It should be noted that the present simulation work may provide design guidelines that go beyond those pertaining to the formation of grafted designed copolymers. Specifically, one can envision situations where the coloring species are replaced with nonpolymeric objects, such as metallic or oxide nanoparticles, thus producing polymer-based composites grafted to surfaces. Such systems have been realized experimentally and
studied theoretically.\textsuperscript{109} Results from those studies indeed reveal that the penetration depth of the nanoparticles inside polymeric grafts depends on the interplay between chain length, grafting density, and size of the nanoparticle.\textsuperscript{115,117–119}

6.18.2.3.1(vi) Adsorption-tuned copolymers and molecular dispenser

The idea of CDSD via PAT (postprocessing) can be generalized in many respects.\textsuperscript{72} Indeed, a special chemical sequence can be obtained not only from a globular conformation; any specific polymer chain conformation can play the role of a parent one. The simplest example of this kind is connected to the conformation of a homopolymer partly adsorbed onto a flat substrate (Figure 8, top panel). 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 catalyzes 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 \textit{et al.}\textsuperscript{120} compared the adsorption properties of copolymers with special ‘adsorption-tuned’ primary structures (adsorption-tuned copolymers, ATCs) with those of truly RCPs and random-block copolymers. Monte Carlo simulations revealed that specific features of the ATC primary structure promoted the adsorption of ATC chains, in comparison with their random and random-block counterparts under the same conditions. In other words, the resultant copolymer sequence ‘memorized’ the original state of the adsorbed homopolymer chain. Its statistical properties exhibit LRCs of the LF type similar to those found for PLCs obtained via PAT of a homopolymer globule.\textsuperscript{121}

Zheligovskaya \textit{et al.}\textsuperscript{120} have simulated the adsorption of ATCs. The critical adsorption energy and conformational properties were compared with those of RCPs with the same content of adsorbed segments and random-block copolymers with the same composition and the same average numbers of adsorbed and nonadsorbed blocks. It was found that the difference in the primary structure of the chains leads to the difference in the critical adsorption energy $\varepsilon^*$ and the characteristics of adsorbed chains. In particular, the ATC chains have the smallest (by the absolute value) critical adsorption energy $\varepsilon^*$. The RCPs are characterized by the largest $\varepsilon^*$, as compared to other copolymers. This fact is simply explained by the difference in block lengths: the RCPs have the shortest blocks. This is consistent with the analytical results\textsuperscript{122} for regular binary copolymers, according to which the absolute value of the critical adsorption energy decreases with increasing block length at the same fraction of adsorbed and nonadsorbed segments. At the same time, the difference in the critical adsorption energy for the random-block and ATC chains (which are characterized by the same average block lengths) can be explained only by the details of the ATC primary structure. Because the ATC chains have significantly longer end nonadsorbed blocks, as compared to their random-block counterparts, the adsorbed segments are placed more compactly in each ATC chain. This specific feature of the ATC primary structure promotes adsorption of ATCs. Thus, the difference in the adsorption behavior of ATC, random, and random-block chains can be rationally explained by taking into account the specific features of the primary structure of these chains.

The obtained results support the general idea of a CDSD: the generated ATC sequence ‘memorizes’ some features of the specific parent conformation of the adsorbed homopolymer. In particular, the position of adsorbed segments is tuned in the

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure8.png}
\caption{Top panel: Schematic representation of the sequence design procedure leading to an adsorption-tuned copolymer. (a) Initial partly adsorbed homopolymer, (b) chemical modification of adsorbed chain segments, (c) resultant copolymer. Modified segments are shown in red. Bottom panel: Stages of preparation of copolymer envelope. (a) Adsorption of homopolymer chain on a colloidal particle, (b) ‘coloring’ of the polymer chain (blue color corresponds to chemically modified monomer units and red color to adsorbed units) and introduction of cross-links (shown as green sticks) to stabilize hollow-spherical structure, and (c) elimination of the core particle.}
\end{figure}
best way for subsequent adsorption. Therefore, it is not surprising that this memorized hidden information became apparent as soon as we consider the adsorption of ATC chains. In other words, the polymer ‘learns to be adsorbed’ in the parent conformation, and this ‘experience’ is used in the subsequent ‘life’ of this copolymer.

Velichko et al. suggested the model of a so-called ‘molecular dispenser’. This idea is a further development in the direction of CDSD. Namely, they considered the conformation of a homopolymer chain adsorbed on a spherical colloidal particle (Figure 8, bottom panel) and performed design of sequence for this state of macromolecule. The motivation behind this design procedure is that if we eliminate the parent colloidal particle after the design is completed (e.g., by etching), the resultant copolymer will be tuned to adsorb selectively another colloidal particle of a parental size \(\sigma_p\). For instance, if such copolymer is exposed to a polydisperse colloidal solution of particles of different size, it will selectively choose to form a complex with the particle having the same radius as that in parental conditions. That is why such a macromolecular object can be called a molecular dispenser.

It should be noted that the development of such polymer systems is stimulated by existing experimental works. In particular, the experimental methods of preparation of nanometer-sized hollow sphere structures have been suggested because of their possible usage for encapsulation of molecules or colloidal particles. The preparation of hollow sphere structures, generally, is based on self-assembling properties of block copolymers in a selective solvent, that is, on the formation of polymer micelles with a nanometer-sized diameter. Further cross-linking of the shell of the micelle and photodegradation of the core part produce nanometer-sized hollow cross-linked micelles.

The sequence design procedure proposed in Reference can be described in more detail as follows. First, we consider a homopolymer chain attracting to a colloidal nanoparticle. Such chain is forming an adsorbed complex with the particle (its typical conformation is shown in Figure 8(a), bottom panel). Only a fraction of chain segments is in direct contact with colloidal particle, while other segments form flower-like loops. Then we ‘color’ the segments in the loops in ‘blue’, while the segments near colloidal particle remain ‘red’, that is, they are attracting to this particle. If the sequence design is stopped at this stage, the pronounced selectivity of the complex formation with another particle of parental size \(\sigma_p\) is not reached. However, if additional cross-links are introduced between adsorbed (red) units, thus fixing the cage structure of the central cavity (Figure 8(b), bottom panel), the macromolecule emerging after elimination of colloidal particle (Figure 8(c), bottom panel) does indeed exhibit the features of a molecular dispenser. This sequence design scheme was realized in the computer simulations.

To characterize the complexes formed by molecular dispenser and colloidal particles, the probability \(P(\sigma,T)\) of finding a complex made of the copolymer envelope and the particle of a given size, \(\sigma\), was calculated as a function of temperature \(T\). It was found that the selectivity of the complex formation strongly depends on the number of cross-links in the envelope. The particles of parent or smaller size \((\sigma \leq \sigma_p)\) are fully absorbed by the central cavity, because the corresponding fitting was ensured by the sequence design procedure. In contrast, particle of a larger size \((\sigma > \sigma_p)\) is too big for a central cavity, and thus the complex formed does not saturate all the possibilities for the attraction of adsorption-active (red) units to the surface of the particle. As to small particles, they easily penetrate inside the molecular dispenser, but the resultant complex is not stable (especially at high temperature) because of small surface of such particles. All these factors lead to a nonmonotonous behavior of \(P(\sigma,T)\): there is the peak in \(P(\sigma,T)\) at \(\sigma = \sigma_p\) for moderately cross-linked copolymer envelopes.

6.18.2.3.2 Copolymerization under heterogeneous conditions

6.18.2.3.2(i) How can we implement CDSD polymerization?

Radical copolymerization at heterogeneous conditions offers additional opportunities not available in homogeneous (solution) copolymerization. These include the intrinsic possibilities of exploiting the heterogeneities of the reaction system to control the chemical microstructure of the synthesized copolymers, making possible new paradigms for synthesis and production of polymeric materials. We discuss some new CDSD-based synthetic strategies, which have been developed in recent years to provide effective control of the chemical sequences.

The CDSD copolymerization differs in principles and in results from the conformation-dependent PAT. The corresponding process can be considered as a variant of template polymerization (also called molecular imprinting) based on the noncovalent binding of polymerizing monomers to a template.

The essence of this technique consists of the copolymerization of monomers differing in their affinity to the template and therefore different distribution in the reaction system. In contrast to conventional types of template polymerization, in the CDSD copolymerization, all the monomers are bifunctional and thus form linear, not cross-linked, polymers. The sequence of the segments in the resultant molecule is determined by the template-controlled conformation of the propagating macroradical. Thus, the CDSD copolymerization regime is possible only in reaction systems with a strongly inhomogeneous spatial distribution of monomer concentrations. As an origin of inhomogeneities, one can mention interfaces, nanoparticles, molecular and supramolecular aggregates (e.g., for macroradicals capable of forming globular conformations), and so forth. They should selectively absorb one of the comonomers (Figure 9).

Moreover, the spatial scale of the concentration inhomogeneities should be comparable to the size of the growing macroradical. In addition, the polymerizing monomers should retain their properties after they are incorporated into a polymer chain. In the case of a copolymerization near an interface, one of the types of monomers and chain segments should be preferentially distributed near the adsorbing surface, whereas the other types of monomers and chain segments should be preferentially located in the solution. As a result, the chemical sequence and conformation of the growing macroradical become mutually dependent. Such interdependence determines the further way of chain growth through the comonomer concentrations in a small reaction volume (see Figure 9).

Finally, the rate of copolymerization should be slow enough. This guarantees that, during the chemical reaction,
the equilibrium concentration fields remain approximately constant, and the growing chain possesses an equilibrium conformation between successive attachments of the monomers. Therefore, the CDSD regime is realized when

$$\max \{ \tau_D, \tau_{rel} \} < \tau_R$$

Figure 9  Schematic representation of the CDSD copolymerization process in the cases when one of the comonomers is selectively absorbed by a polymer globule (top) or adsorbed on a surface (bottom). (a) Growing chains during the copolymerization (reaction zone around the growing chain end is marked with dashed line); resulting copolymers in (b) globular (adsorbed) and (c) coil-like states. Regions where absorbed (adsorbed) monomers dominate are shown in gray. The conditions in the reaction system compel an initially homopolymeric chain to take the desired template conformation, for example, a globular one. One part of the monomer units is screened from the solution in this conformation, while the other part is accessible for chemical modification.

6.18.2.3.2(ii) Copolymerization with simultaneous globule formation

Let us assume that we are performing a radical copolymerization of moderately hydrophobic (H) and hydrophilic (P) monomers in an aqueous medium. The conditions for copolymerization (e.g., the temperature and solvent composition) should be chosen in such a way that when the growing chain is long enough it can form a globule. As long as the current chain length $N_c$ is not too large, the growing hydrophobic–hydrophilic (HP) macroradical remains in a coil-like conformation. However, when its length becomes sufficiently large, the chain tends to form a two-layer globule with a hydrophobic core and a polar envelope. The origin of this effect is connected to the loss in translational entropy of covalently bonded hydrophobic species after their incorporation into the growing chain.

The presence of a hydrophobic–hydrophilic interface can dramatically change the reaction conditions. The hydrophobic core will selectively absorb hydrophobic species from the solution (Figure 9), and this will result in a redistribution of monomer concentrations between the core and bulk solution. Because the probability of attachment for each comonomer is determined by its concentration in a relatively small reaction volume near an active chain end, the active center inside the hydrophobic core will mainly attach more hydrophobic species; on the other hand, when the active center is located on the globule surface, it will mainly attach polar (soluble) monomers. In this way, the two-layer globule will grow, retaining its core–shell structure with a predominantly hydrophobic core and a hydrophilic outer envelope (cf. Figure 9).

Using Monte Carlo and molecular dynamics simulations, Berezkin et al. have studied the process of irreversible radical copolymerization of hydrophobic and hydrophilic monomers in a common solvent. The conformation-dependent polymerization was modeled as a step-by-step chemical reaction of the addition of H and P monomers to the growing copolymer chain, under the assumption that a depolymerization reaction was not allowed. To simplify the analysis, it was assumed that the copolymerization is ideal with equal reactivity ratios. The type of attached monomers and the probabilities of their addition were determined from the average concentrations of the reactive monomers in a reaction volume $V$ around the moving active end of the macroradical of a given length $N_c$, during a reaction time $\tau_g$. It was also assumed that the reaction bath was sufficiently large and the reactive species were sufficiently dilute so that there was significant time for the growing macroradicals to propagate independently. The preferential sorption of hydrophobic monomers in the core of the arising globule was explicitly taken into account.

The simulations showed that using the conformation-dependent polymerization mechanism, one can indeed end up with a designed copolymer having a dense hydrophobic core surrounded by a hydrophilic shell. Figure 10 presents a typical snapshot of the resultant copolymer globule and an example of the radial distribution that characterizes the distribution of H and P units with respect to the center of the globule. These findings prove explicitly the core–shell microstructure of the globule obtained via irreversible copolymerization in the solvent, which is moderately poor for hydrophobic species. In general, the same microsegregated structure is observed for core–shell globules obtained via ‘coloring’ procedure.

The average chemical composition of the resulting HP polymer is determined by the bulk concentrations of the monomers and by the solubility of hydrophobic monomers in a globular core. For approximately equal contents of H and P segments in the growing copolymer, the solution should contain an excess
of P monomers. In general, the chemical composition of the synthesized copolymer deviates strongly from the statistical one. Thus, the change in the solution concentration of polymerizing monomers allows the copolymer composition to vary over a wide range. Interestingly, although the reactivities of both monomers were equal, their copolymerization led to a copolymer that was preferentially enriched with hydrophobic segments. Therefore, a strong deviation from the principle of equal reactivity of Flory is observed for the CDSD polymerization regime. The type of monomer attached to the growth center during the simulation under kinetic control (at large $t_g$ values, see eqn [5]) is determined by the conformation and primary structure of the growing chain as a whole, not only by the local concentration of reactive monomers near the active end of the macroradical. As a result of such cooperativity, the formation of sequences with specific LRCs of the LF type was observed.

Because of the specific mechanism of chain propagation, the synthesized copolymers can have a gradient primary structure; that is, the HP composition can change along the growing chain during copolymerization. This result follows from the analysis of the distribution of monomers along the growing chain. Very short chains do not contain sufficiently large number of hydrophobic units to form a core–shell conformation, and the growth of these chains is similar to that observed for random free-radical copolymerization. The composition of these chains is also close to the monomer composition in solution. Then, as the macroradical becomes longer and the number of connected H units increases, a dense hydrophobic core can be formed. This core absorbs H monomers, and their fraction in the resultant copolymer increases. In this case, the probabilities of monomer addition are also not constant because of the changes in the ratio between the volumes of the hydrophobic core and the polar shell during the chain growth. Therefore, the gradient primary structure is formed because of a change in the chain conformation and a continuous redistribution of comonomers between the globule and the solution in the course of the polymerization; in this way, a compositional drift is produced along the chain.

Note that the statistical model developed by Berezkin et al.134,135 to predict sequence in free-radical polymerization is not applicable in living radical polymerization because of the reversible activation/deactivation procedure between radicals and capping agents.137

6.10.2.3(iii) Radical graft copolymerization near a selectively adsorbing surface

In a series of publications,136–141 a statistical model was developed for studying graft copolymerization in the vicinity of a surface. This approach represents an irreversible radical polymerization of selectively adsorbed A and B comonomers with different affinity to a chemically homogeneous impenetrable surface, allowing for a strong short-range monomer (A)-surface attraction. Thereby, one of the types of monomers and chain segments are preferentially distributed near the adsorbing surface, while other monomers and chain segments are located preferentially in the solution, as schematically depicted in Figure 9. As a result of such concentration inhomogeneities, the chemical sequence and conformation of growing macro-radical become mutually dependent that corresponds to the CDSD regime. It is found that under certain conditions, the adsorption copolymerization yields copolymers with gradient monomer sequence distribution.

Gradient copolymers (GCs) are characterized by local composition varying monotonically along the chain. They have attracted growing attention over recent years. The range of possible applications of these copolymers is quite wide.142–146 In the solutions or melts they can form diverse mesostructures. GCs can be used as surfactants and, particularly, as compatibilizers in polymer blends.145 Grafting of these copolymers yields ‘smart surfaces’, which can change the structure and properties, depending on external conditions.146

GCs are usually obtained via living polymerization.131,137,147–150 The gradient arises here due to the continuous variation of the monomer ratio during synthesis. However, this technique allows producing sequences with a relatively simple statistics only. The CDSD-based technique represents an alternative approach to the synthesis of copolymers with nontrivial statistics, including GCs, via free-radical
polymerization. As indicated above, the sequence of monomer units formed during synthesis and the equilibrium conformation of the reacting macromolecule should be interdependent. This interdependence results from the self-organization of monomers and copolymer macromolecules in the reaction system. We emphasize again that the conformation-dependent copolymerization is most effective when the chain propagation is slower than the diffusion of monomers and the formation of the equilibrium macroradical conformation.

The propagation of the grafted chain was considered as a sequential addition of monomer units of the type A or B to the free active end of macroradical. There are four possible ways in which monomer can be added:

\[ \sim A^* + A \xrightarrow{k_{AA}} \sim AA^* \]

\[ \sim A^* + B \xrightarrow{k_{AB}} \sim AB^* \]

\[ \sim B^* + A \xrightarrow{k_{BA}} \sim BA^* \]

\[ \sim B^* + B \xrightarrow{k_{BB}} \sim BB^* \]

where \( k_{XY} \) are reaction rate constants of the interaction between active center X and monomer Y. Relations between these constants are usually expressed through the reactivity ratios \( r_A = k_{AA}/k_{AB} \) and \( r_B = k_{BA}/k_{BB} \). For simplicity, it was assumed that the rate of propagation does not depend on the chemical structure of the active center \( (r_A = r_B = 1) \), so the ideal copolymerization was considered. The average concentrations depended on local monomer concentrations in the solution and near the surface. The concentration of the nonadsorbable monomer B was the same everywhere, while the concentration of monomer A in the adsorption layer exceeded that in the solution bulk by a factor \( q = \exp(-\varepsilon/k_B T) = \exp(u) \), where \( \varepsilon \) is the energy of a single adsorption contact and \( u \) denotes the dimensionless adsorption energy. Because chain propagation is usually a kinetically controlled process, the macroradical changes its conformation many times between the propagation acts, and the probability of finding the active center near the surface, \( f_{N} \), should be averaged over the macroradical conformations for a current chain length \( n \). This averaging is a significant theoretical problem. However, exact calculation of the statistical weight of each macroradical conformation is possible if the macroradical is considered as an ideal (Gaussian) chain on a lattice.

The computational procedure was realized as follows. The first monomer unit of the grafted chain was placed at the origin. The surface coincided with the XY plane; the z axis was normal to the surface and directed to the bulk of the solution (cf. Figure 11, top left panel). If the distance between the nth monomer unit and the surface \( z_n = 0 \), then the unit contacted with the surface, and if this unit belonged to the A type, then it was adsorbed.

In order to find the partition function and statistical properties of the growing macroradical, Berezkin et al. introduced a variable \( \alpha_n(h) \) that was defined as follows: \( \alpha_n(h) = 1 \) if the nth unit located at the point \( r = (x,y,z) \) is adsorbed, and \( \alpha_n(h) = 0 \) if the unit is not adsorbed. If the macroradical of length \( n \) has in some conformation \( h \) adsorption contacts with the surface, and its active center is located at the point \( r \), the possible number of such conformations is denoted as \( \beta_n(k,h) \). The following recurrence equation determines the value of \( \beta_n(k,h) \):

\[ \beta_n(k,h) = \sum_{R(r)} \beta_{n-1}(k-a_n(h),r') \]

where \( R(r) \) is the set of all possible locations of the monomer unit \((n-1)\) if the nth unit is located at \( r \); \( r' \) is the point belonging to the set \( R(r) \). Then, the partition function of an \( N \)-unit chain is given by

\[ Z_N = \sum_{V=0}^{n} \sum_{k=0}^{n} \beta_N(k,h) q_k \]

where \( V \geq 0 \) is a subset of lattice points with \( z = 0 \). The attachment of comonomers is determined by their average local concentrations, which, in turn, depend on the corresponding bulk concentrations, \( \phi_A \) and \( \phi_B \), and on the energy parameter \( u \). The technique developed in Reference 141 allows one to obtain an exact analytical solution of the problem for the Gaussian chain model.

Figure 11 (top right panel) shows that at a fixed solution composition, the adsorption leads to an increase in the fraction \( \phi_A \) of adsorbing monomer (A) units in the copolymer, as expected. This is explained by the growth of monomer A concentration in the adsorption layer.

To analyze the local compositional inhomogeneities along the chain of synthesized copolymers, one can use the probabilities of finding a monomer unit of type A at the nth position from the beginning of a growing macromolecule, \( \phi_A(n) \). It is clear that for a truly RCP, in which chemically different segments follow each other in statistically random fashion, the \( \phi_A(n) \) function should coincide with the average fraction of A segments for any \( n \). For a random-block copolymer, the fraction of one component averaged over many sequences should also be uniform along the chain.

Figure 11 (bottom left panel) presents the local copolymer composition \( \phi_A(n) \). It is seen that when \( u > 0 \), the simulated polymerization process yields copolymers with a well-defined gradient structure. For such copolymers, the \( \phi_A(n) \) function smoothly decreases with \( n \). It should be kept in mind that speaking about GCs, we have to consider their ensemble generated by the same synthetic process, not a single chain which can, in principle, have arbitrary statistics. Note that in the range \( u = 0.3 \)–0.4, the composition gradient is most pronounced. Interestingly, this adsorption energy is very close to the critical adsorption energy \( u_c \) estimated for the same model.

To examine the influence of the adsorption energy on the gradient in more specific way, one can introduce a quantitative measure of the gradient as the variance of the preaveraged local copolymer composition:

\[ D = N^{-1} \sum_{n=1}^{N} [f_A(n)]^2 - [N^{-1} \sum_{n=1}^{N} f_A(n)]^2 \]
The value of $D$ is shown in Figure 11 (bottom right panel) as a function of $u$. It is seen that the dependence of $D$ on the adsorption energy goes through a well-pronounced maximum located at $u = u_c$, when the largest compositional nonuniformity is observed (cf. Figure 11, bottom left panel).

A detailed analysis of the model showed that the key factor determining the microstructure of the resultant copolymer is the dependence of the probability $f_S(n)$ on the current macroradical length $n$. This could be considered as a ‘memory function’. Its behavior is strongly influenced by the adsorption energy $u$. There are three regimes of the adsorption copolymerization that correspond to different adsorption energies:

1. In the nonadsorbing regime ($u \ll u_c$), the probability for the active center to be located near the surface scales as $f_S(n) \propto n^{-1}$, where $n$ is the current macroradical length. Therefore, in this regime, the chain propagation leads to asymptotical convergence of the local copolymer composition to that of RCPs synthesized in a solution under the same conditions. For finite values of $n$, a fast change in the

![Figure 11](image-url)
local chemical composition is observed for initial sections of the growing macroradical.

2. In the weak adsorption regime \((u = u_c)\), the probability \(f_n\) is nonzero for any \(n\) and \(u\). As a result, the local copolymer composition in the long-chain limit differs dramatically from that of a RCP synthesized in the solution bulk. In the vicinity of \(u_c\), the magnitudes of the compositional nonuniformity and gradient are maximal, and the gradient extends along the entire chain for any chain length.

3. For the strong adsorption regime \((u \gg u_c\) or \(u > u^*\)), the copolymer statistics does not depend on \(u\) and is determined solely by the solution monomer concentrations.

Thus, the change in the adsorption energy can provide versatile and accurate control of the statistics of the resulting copolymers obtained via the conformation-dependent adsorption copolymerization. It is known that the gradient (tapered) nature of copolymers, which can be synthesized in free-radical polymerization processes, is due to a drift in the free monomer composition during solution polymerization. Such copolymers can be considered a special type of block copolymers in which the composition of one component varies along the chain. With a decreasing difference in the monomer reactivity ratios, the formation of gradient statistical copolymers rather than gradient (tapered) block copolymers occurs. However, in the model polymerization process considered in References 138–141, all the polymerizing species had the same reactivities, and the monomer concentrations remained unchanged during synthesis. Therefore, the change in the probabilities of the addition of components to the growing macroradical are due to the evolution of its chemical composition; this result is typical for the CDSD regime.

### 6.18.2.3.2(iv) Radical copolymerization near a patterned surface

Because the spatial scale of the monomer concentration gradient in the reaction system discussed in the previous sections was comparable with the size of the macroradical, the chemical inhomogeneity has been observed for the generated copolymer sequence as a whole. To limit and control the size of the gradient regions, the spatial scale of the concentration inhomogeneities in the reaction system should be limited. The easiest way to achieve this is to perform the copolymerization near a solid patterned surface with discrete adsorption sites. In this case, each of the sites represents a small independent source of concentration disturbances lying at a certain distance from other sites. Surfaces with a regular distribution of adsorption sites are of most interest because they can allow fine control of the primary structure of the copolymers obtained with the CDSD technique.

Another motivation for this work is the development of copolymers that are tuned to a certain surface, that is, copolymers that have a 'memory' of the preparation conditions and are able to reproduce their specific conformation in the vicinity of the surface with predefined chemical heterogeneity. It is believed that copolymers designed in this way could have potential for the recognition of patterned substrates through the formation of stable adsorption complexes with planar or spherical substrates composed of two chemically distinct sites, one of which has a preferential affinity for one of the cocomomers. Obviously, these copolymers could have a great potential in molecular technology and biotechnology.

A Monte Carlo simulation of irreversible template copolymerization near a chemically heterogeneous surface with a regular (hexagonal) distribution of discrete adsorption sites was performed and reported in References 151 and 152. The sites could selectively adsorb one of the two polymerizing monomers and the corresponding chain segments from solution. The focus of this work was on the influence of the polymerization rate, adsorption energy \(v\), and distance between adsorption sites \(r_s\) on the chain conformation and chemical sequence of the resulting A/B copolymers and, specifically, on the coupling between polymerization and selective adsorption.

Under the preparation conditions corresponding to the CDSD regime, the formation of quasi-regular copolymers with blocky primary structure was observed. In such copolymers, there are two types of alternating sections. One of them contains randomly distributed A and B segments. The second one consists mainly of strongly adsorbed A segments. The average length of the random sections is proportional to the distance separating the nearest neighbor adsorption sites \(r_s\). The average length of the A-rich sections is determined by the adsorption capacity of the adsorption sites. Therefore, by varying the interaction parameters and the distribution of adsorption sites on the substrate, one can design and synthesize copolymers with different surface-induced chemical sequences in a controlled fashion.

Under the preparation conditions corresponding to the CDSD regime, the formation of quasi-regular copolymers with blocky primary structure was observed. A typical snapshot from the simulation is presented in Figure 12(a). In such copolymers, there are two types of alternating sections. One of them contains randomly distributed A and B segments, the second one consists primarily of strongly adsorbed A segments. The average length of the random sections is proportional to the distance separating the nearest neighbor adsorption sites \(r_s\). The average length of the A-rich sections is determined by the adsorption capacity of the adsorption sites. Therefore, by varying the interaction parameters and the distribution of adsorption sites on the substrate, one can design and synthesize copolymers with different surface-induced chemical sequences in a controlled fashion.

In particular, the variation of the strength of the effective monomer (A)–substrate interaction \(v\) allows us to determine the asymptotic regimes corresponding to random copolymerization or CDSD. The average intramolecular chemical composition \(\phi_A\) emerging in the simulation is shown as a function of \(v\) in Figure 12(b) for the case in which the solution concentrations of A and B monomers are equal, \(C_A^S = C_B^S\). In the weak adsorption regime, \(v/k_BT \ll 1\), random copolymerization dominates \((\phi_A = 0.5)\); in the region \(v/k_BT \gg 1\), CDSD dominates. The threshold value of the energy parameter \(v\) when random copolymerization and CDSD have about the same probability is \(v/k_BT = 8\) for a given choice of the polymerization parameters and simulation model.

To analyze the correlations in the copolymer sequences, one can use the so-called two-point chemical correlators:

\[
\Theta_{ij}^{(1)} = n_{ij}^{(1)} \sum_{\alpha \neq \beta} n_{\alpha \beta}^{(1)}
\]  

[11]
Figure 12 (a) Snapshot illustrating a typical conformation of a 512-unit copolymer chain with excluded volume interaction synthesized near a patterned surface in the strong adsorption regime. Chain segments of A and B type are shown as red and green sticks, respectively. Spheres depict adsorption sites. (b) Average chemical composition as a function of the adsorption energy parameter for the case when the solution concentrations of A and B monomers are equal. (c) Chemical correlators for AA diads for different values of the adsorption energy parameter $\varepsilon$.

where $n_{\alpha\beta}^{(i)}$ is the number of $\alpha\beta$ (AA, BB, and AB) pairs and $i$ is the 'chemical distance' between these pairs along the chain. Chemical correlator $\Theta_{\alpha\beta}^{(i)}$ represents the probability of finding a pair of $\alpha$ and $\beta$ segments ($\alpha, \beta = A, B$) among all the possible pairs separated by $i$ segments along the chain in a given copolymer sequence. Some of the results of the calculation are shown in Figure 12(c) for a few different values of $\varepsilon$.

The dependences of $\Theta_{\alpha\beta}^{(i)}$ on $i$ indicate that quasi-regular copolymers are synthesized in the CDSD regime when the adsorption interaction is sufficiently strong. Indeed, we see that these copolymers are characterized by a periodic variation of the composition along the chain. For the system simulated, the period of this variation is about 20 segments. Therefore, the formation of copolymers with blocky primary structures can be observed.

The following mechanism of the formation of quasi-regular copolymers was suggested. Chain growth begins near an adsorption site at which the concentration of strongly adsorbed A monomers is high. These monomers are attached preferably to the end of the growing macroradical and form an initial chain section comprising mainly A segments. This section gradually covers the nearest adsorption site. Because of the limited adsorption capacity, the screened adsorption site looses the ability to attract the active end group of the macroradical. Afterwards, the reaction volume moves into solution away from the substrate in which random copolymerization occurs. The random chain section grows until it reaches the nearest free adsorption site. Subsequently, the formation of a new adsorbed section commences again. Such cycles are repeated many times.

As a result, the formation of a quasi-regular copolymer with alternating A-rich and random AB sections is observed. Thus, the copolymer sequence consists of repeating blocks. Each of these blocks comprises a short 'gradient sequence' formed by strongly adsorbed sections and random weakly adsorbed sections. The average length of the adsorbed section is nearly constant and is related to the adsorption capacity of the adsorption center. In contrast, the average length of the random section depends on its conformation between the adsorption centers, and in principle, it can vary over a wide range. It should be emphasized that the quasi-regular copolymer is formed only when the random bridges connecting neighboring adsorption sites are strongly stretched and the variation of their length is sufficiently small. The strongly stretched regime leads to the periodical composition variations. Figure 12(c) shows that this regime is realized for $\varepsilon/k_BT > 15$.

The distribution of B blocks, which are included mostly in nonadsorbed chain sections, decays exponentially and thus should obey Bernoullian statistics that correspond to a zeroth-order Markov process. The average length of such blocks is close to 2, which is the same as that of a RCP. In the case of A blocks, the distribution function $f_\alpha(\ell)$ also decays exponentially in the initial region, which corresponds to short blocks included in the random chain sections. For longer A blocks, however, the distribution becomes significantly broader and has a local maximum at $\ell \sim 10$. Hence, one
can conclude that the distribution of A blocks strongly deviates from that known for random sequences.

By varying the distance between nearest adsorption sites, \( r_p \), one can control the composition variation period of the synthesized copolymer. From the chemical correlators defined by eqn [11], it is easy to find the average number of segments in the repeating chain sections, \( N' \), for different \( r_p \) values. It is instructive to analyze the relation between \( N' \) and \( r_p \). As expected, a power law \( N' \propto r_p^\mu \) is observed. It is clear that the value of the exponent \( \mu \) in this dependence should range be between 1 (for a completely stretched chain) and \( v^{-1} \) with \( v = 0.6 \) (for a random coil with excluded volume). The calculation yields \( \mu = 1.33 \) for \( N' \gtrsim 15 \). This supports the aforementioned assumption that the repeating chain sections are strongly stretched between the adsorption sites. The same conclusion can be drawn from the visual analysis of typical snapshots similar to that presented in Figure 12(a).

6.18.2.3.3 Evolutionary approach and theoretically informed sequence design

6.18.2.3.3(i) Design as a simulation of evolutionary process

The concept of evolution of primary sequences of biopolymers has attracted significant interest from biologists, chemists, and physicists for a long time. As has been discussed, it is natural to expect that the content of information in the sequences of biopolymers (proteins, DNA, RNA) is relatively high in comparison with random sequences where it should be almost zero. Presumably, the information complexity of early ancestors of present day biopolymers has been increased in the course of molecular evolution when the copolymer sequences became more and more complicated. The study of various possibilities of this evolution of copolymer sequences is just the area where the concept of evolution can be used in the context of polymer science.

It is worthwhile to note that since the information content of sequences can be represented as a mathematically defined quantity, the whole process of evolution of biopolymer sequences can be specified in exact mathematical terms. The formulated fundamental problem is extremely difficult to handle because of the absence of direct information on the early prebiological evolution. Therefore, of particular interest are `toy models’ of evolution of sequences that show different possibilities for appearance of statistical complexity and of LRCs in the sequences.

Evolutionary computation approaches represent optimization methods. They are conveniently presented using the metaphor of natural evolution: a randomly initialized population of individuals evolves following a crude parody of the Darwinian principle of the survival of the fittest. New individuals are generated using simulated evolutionary operations such as mutations. The probability of survival of the newly generated solutions depends on their fitness (how well they perform with respect to the optimization problem at hand); the ‘best’ are kept with a high probability, and the ‘worst’ are rapidly discarded.

Some computer models describing the evolution of copolymer sequences have been proposed previously. Most of them are based on a stochastic Monte Carlo optimization principle (Metropolis scheme). Such optimization algorithms start with arbitrary sequences and proceed by making random substitutions biased to minimize relative potential energy of the initial sequence and/or to maximize folding rate of the target structure.

As has been emphasized, the problem that we address here is somewhat different from that usually discussed in the context of protein physics. We do not intend to search for unique three-dimensional (native) conformations with fast folding rate. On the contrary, we are interested in the state with large entropy. In general, our aim is to learn whether it is possible to make with synthetic copolymers a step along the same line as molecular evolution.

Ascending and descending branches of sequence evolution. The aim of the study was to introduce explicitly the concept of sequence evolution into the CDSSD scheme. Using the method of molecular dynamics, the conformation-dependent evolution of model HP copolymer sequences was simulated. The sequence evolution mechanism involved the generation of an initial protein-like sequence by inspecting a homopolymer globule and by attributing H type to the monomeric units in the core of this globule and P type to the units on the surface of the globule. The resulting copolymer was then transferred to a coil conformation and then refolded by means of strong attractions acting among the H units. The HP sequence was further modified, depending on the position of a monomer in the core or on the surface of a newly formed globule. Such modifications, leading to changing the primary HP sequence, are repeated many times (~10^3). With this evolutionary process that can be called ‘repeated coloring’, structures and sequences are formed self-consistently.

A 128-unit flexible-chain heteropolymer with an HP composition fixed at 1:1 was simulated for the condition when hydrophobic H monomers strongly attract each other, thus stabilizing a dense globular core. Concurrently, the affinity energy \( \varepsilon_{HP} \) among two hydrophilic P monomers was considered as a parameter (the interaction between H and P monomers is given by \( \varepsilon_{HH} = \varepsilon_{PP} \)). For this model system, various conformation-dependent and sequence-dependent properties, including information-theoretic-based quantities, can be calculated.

Depending on the attraction energy between polar segments \( \varepsilon_{PP} \), it is possible to find two regimes (branches) of evolution (regimes I and II). If \( \varepsilon_{PP} \) is smaller than some crossover energy \( \varepsilon_{PP}^* \) (regime I), the evolution can lead to a second-order-like transition in sequence space from the sequences with a protein-like primary structure capable of forming a core–shell globule to the degenerated (nonprotein-like) sequences having long uniform H and P blocks (cf. Figure 13, left panel). This transition is also accompanied by large changes in the conformational properties of copolymer.

Figure 13(a) (right panel) shows the mean square gyration radius, \( R_g^2 \), plotted versus \( \varepsilon_{PP} \). As seen, \( R_g^2 \) is a weakly decreasing function of \( \varepsilon_{PP} \) in the range \( \varepsilon_{PP} > \varepsilon_{PP}^* \) and demonstrates a rapid growth when \( \varepsilon_{PP} \) decreases and becomes less than \( \varepsilon_{PP}^* \). The critical value \( \varepsilon_{PP}^* \) is found to be smaller than the critical energy at which coil-to-globule transition takes place in a homopolymer chain of the same length.

The degenerated primary structure resembles a di- or triblock sequence (‘core–tail’ or ‘tadpole-like’ conformation), as seen in Figure 13.

Thus, when the attraction between the hydrophilic segments is not sufficiently strong, we deal with the ‘descending branch’ of the evolution, which leads to nonprotein-like
sequences having low information content and low complexity. In contrast, in the second regime (at \( \varepsilon_{pp} \geq \varepsilon^*_{pp} \)), the complexity of protein-like structures is found to increase and therefore we have the ‘ascending branch’ of the evolution.

Using the Jensen–Shannon divergence, \( JS \) (cf. eqn [1]), as a measure of complexity for the generated sequences, one can obtain an interesting result (cf. Figure 13(b), right panel). The most important feature is that the \( JS \) value is a nonmonotonic function of \( \varepsilon_{pp} \), whereas Shannon’s entropy, Shannon’s index, and many other sequence-dependent parameters change always gradually.\(^{10}\)

For the sequences generated in the evolutionary process described above, it was shown that at \( \varepsilon_{pp} \geq \varepsilon^*_{pp} \) (regime II), the degree of complexity, as measured by \( JS \) divergence, can be considerably higher compared to that observed for the regime I, at \( \varepsilon^*_{pp} < \varepsilon_{pp} \). The complexity increases slightly with decreasing \( \varepsilon_{pp} \), reaches a maximum just on the boundary of regimes I and II, and then drops sharply (Figure 13(b), right panel). Therefore, in the regime II, the evolution preserved the copolymer sequence of high complexity, whereas in regime I, the information content of the sequence has degenerated in the course of evolution.

**Simultaneous evolution of sequences and conformations.** Chertovich et al.\(^{157}\) reported on simultaneous evolution of sequences and conformations. This design procedure leads to the final state that depends on the set of interaction parameters and on the rearrangements both in conformational space and in sequence space. These rearrangements are characterized by the usual thermodynamic temperature, \( T \), for conformational space as well as effective sequence rearrangement temperature, \( T_{c} \). Namely, after certain number of Monte Carlo steps in conformational space (in the course of this process the system equilibrates at temperature \( T \)), the possibility of mutation of two randomly chosen monomeric units is attempted: the monomeric unit H converts into P and vice versa. The move is accepted provided this process leads to a decrease in the energy of a protein-like globule. If the globular energy is increased by the amount \( \Delta E \), this move is accepted with the probability defined by \( \exp(-\Delta E/k_B T) \).

In general, we can define the evolution of sequences for any value of \( T_{c} \). However, the simulation for three characteristic cases can be most easily understood.

1. The inequality \( T \ll T_{c} \) means that all the moves in sequence space are accepted independent of conformation. This corresponds to random mutations, and the final sequence (after long evolution) will be that of a random HP copolymer (no information complexity).

2. The inequality \( T \gg T_{c} \) means that only the moves leading to a decrease in the globular energy \( E \) are accepted. Such an evolution should lead to a sequence corresponding to a minimum of globular free energy in conformational space. It was shown\(^{58}\) that in the absence of any attractive interactions between P units the final sequence after evolution should have hydrophobic core with very few hydrophilic (or polar) loops and a long hydrophilic tail. This sequence is close to that of HP diblock copolymer, and should not exhibit any information complexity, as has been stated above.

3. The case \( T = T_{c} \) corresponds to annealed HP sequence. This case is equivalent to the situation when monomeric unit H can be converted into P by attaching some ligand L:P \( \neq H + L \). We assume that the number of ligands is fixed to maintain 1:1 HP composition; however, they can choose which monomeric unit to bind to. This defines, in particular, the chain

**Figure 13** Left panel: Snapshots of two typical conformations of designed copolymers obtained after long evolution (3.63 \( \times \) 10\(^6\) MD time steps) of sequences. (a) Core–tail (tadpole-like) structure at \( \varepsilon_{pp} = 0 \), (b) Core–shell structure at \( \varepsilon_{pp} = 0.3 \ k_BT \). Hydrophobic and hydrophilic units are colored red and green, respectively. Right panel: (a) mean square gyration radius and (b) Jensen–Shannon divergence measure as a function of the attraction energy \( \varepsilon_{pp} \) between hydrophilic segments, after the sequence evolution procedure. The characteristic energy of H–H interactions is fixed at \( \varepsilon_{HH} = 2k_BT \), thus stabilizing a dense globular core.
sequence. Annealed HP sequences in the context of polymer globules were first considered by Grosberg.\(^\text{156}\) When \(T_c\) does not correspond to any of the characteristic cases described above, the evolution of sequences coupled with conformations can be still defined in the same way. One has to only remember that if \(T \neq T_c\), and both \(T\) and \(T_c\) are finite, there is a flow of heat between conformational space and sequence space, so that full thermodynamic equilibrium is impossible. Still, we can be in a stationary regime corresponding to the sequences tending to a certain fixed point, and possessing (or not possessing) information complexity.

The simulations and theoretical arguments\(^\text{157}\) predict that at high \(T_c\), the sequence free energy \(F_s\) dominates and the sequence tends to be completely random, corresponding to the minimum of \(F_s\). At low \(T_c\), the evolution selects those sequences, which correspond to the low value of the conformation free energy (structure of the core–tail type). In the intermediate regime, there is an interplay between the conformational thermodynamic force and the sequence contribution (evolution pressure). Therefore, the formation of nontrivial structures and sequences is possible. Even in the absence of attraction between \(P\) units, the final sequences remain protein-like (i.e., the final copolymer has the core–shell structure of a globule) and therefore maintain certain information complexity.

6.18.2.3.3(ii) Theoretically assisted sequence design

The examples of sequence design described above belong to a spectrum of \textit{in silico} methodologies that can be called knowledge-based model approach. A knowledge-based approach requires some form of \textit{a priori} knowledge about a particular target model (e.g., core–shell globule) and is therefore limited in its applicability by the models that are available. As has been noted, there is another sequence-design strategy where existing theoretical concepts are employed. In this section, we will discuss a simple theoretically assisted algorithm that introduces a ‘selection pressure’ under which a random two-letter (AB) copolymer sequence can mutate and transform into the sequence tuned to microphase separation transition (MIST). In particular, we are interested in determining how a sequence of A’s and B’s should be organized in order to reach maximum characteristic length scale for MIST at a given temperature, AB composition, and polymer chain length \(N\). It is natural to call such a sequence ‘MIST-tuned sequence’.

It is well known that heteropolymers can self-assemble into highly ordered patterns of microstructures, both in solution and in bulk. This subject has been reviewed extensively.\(^\text{1–4,159–161}\) The driving force for structure formation is competing interactions, that is, immiscibility of chemically different monomer species, on the one hand, and covalent bonding of units within the same macromolecule, on the other hand. The latter factor prevents the separation of the system into homogeneous macroscopic phases. Spontaneous ordering of copolymers, which occurs with changing (typically decreasing) temperature and is accompanied by an explicit symmetry change, is treated as MIST or order-disorder transition (ODT) and the respective temperature is the ODT temperature. With further decrease of temperature, the ODT is often followed by various order–order transitions between the different ordered morphologies.

Khalatur et al.\(^\text{162}\) proposed a theoretically assisted design scheme that optimizes monomer distributions and leads to the MIST-tuned sequence capable of forming microphase-separated structures with a required (micro) domain spacing \(r^*\). In this evolutionary approach, a simulation is performed to search for ‘point mutations’ that favor the microphase separation of copolymer melt. A random AB sequence is taken as an initial generation \((G = 0)\), which is considered as a ‘common ancestor’ of a given run. Then a procedure of the evolution (annealing) of the sequence starts. The iterative procedure consists of many mutation steps. At each step – with every ‘click of the evolutionary clock’ – two monomers are chosen randomly and, if they happen to be of different types, an attempt is made to exchange their types \((A \leftrightarrow B)\). This point mutation changes the copolymer sequence \((S)\) and the corresponding sequence-dependent intrachain correlation function \(\omega(q,S)\) calculated in the reciprocal \(q\)-space. The \(\omega(q,S)\) function is used as an input for some theory (see below) that gives the transition temperature \(T^*\) (or the critical value of the Flory–Huggins parameter \(\chi^* \approx 1/T^*\)) and the wave number of maximum instability \(q^* \approx (2n/r^*)^2\). These quantities can be viewed as a scoring (fitness) function \(f\) that describes how well a particular sequence is optimized to match the target property. There should be a feedback in the system. In other words, to accept or reject current mutation, one should compare two fitness functions, old and new ones. This is done following the Metropolis ideology, using some design parameter, which is called design (or sequence) temperature. This parameter characterizes the tolerance to mutations in sequence space or, in other words, an ‘evolution pressure’. In fact, this is the same parameter as that discussed in Section 6.18.2.3.3(i).

If the ODT temperature \(T^*\) is viewed as a fitness function \(f = T^*(S)\), the resulting change in the transition temperature \(\Delta T^*\) after an attempted mutation is found and the probability \(p\) to fix the mutation is guided by the standard Metropolis algorithm: \(p = 1\), if \(\Delta T^* > 0\), otherwise \(p = \exp(\Delta T^*/T_c)\), where \(T_c\) is the fictitious temperature referred to as sequence design temperature. It is assumed that after \(N\) point mutations, the \(N\)-unit sequence is passed on to the next generation, \(G = G + 1\). Such modifications, leading to changing in copolymer sequence, are repeated many times \((n_c \sim 10^5)\). For each trajectory corresponding to a stationary regime, one may interpret the set of sequences generated in the course of the evolutionary process as \(n_c\) different ‘species’ originating from a common ancestor. It is reasonable to constrain the sequence composition so that there are \(f_0\) units of type A and \((1-f_0)\) units of type B.

In the weak-segregation regime, the phase behavior of a polymer melt composed of flexible-chain macromolecules can be described on the basis of the random-phase approximation (RPA)\(^\text{159}\) or the polymer integral equation reference interaction site model (pRISM)\(^\text{163,164}\) that allow finding the conditions under which the spatially homogeneous state of the system becomes unstable.

Using the simplest incompressible random phase approximation, the critical value of the Flory–Huggins parameter, \(\chi^*\), and the corresponding transition temperature \(T^* \approx 1/\chi^*\), are determined by the condition that the scattering intensity \(S(q)\) reaches its maximum value at a particular nonzero wave number \(q^*\). Within the RPA, the scattering intensity is given by
\[ S_{\text{BA}}^2(q) = \Re(q) - 2\chi \]  

with

\[ \Re(q) = \frac{1}{\delta\omega(q)} \left( \frac{\omega_{\text{AA}}(q)}{\phi_B} + \frac{\omega_{\text{BB}}(q)}{\phi_A} + \frac{2\omega_{\text{AB}}(q)}{f_A f_B} \right) \]

where the intramolecular correlation functions \( \omega_{ij} \) characterize the conformation of a macromolecule and its sequence distribution, \( f_A = N_A/N \) and \( f_B = N_B/N, \phi_A \) and \( \phi_B \) are the volume fractions of the corresponding monomer species, and \( \delta\omega \) is defined as

\[ \delta\omega(q) = \omega_{\text{AA}}(q)\omega_{\text{BB}}(q) - f_A^2 f_B^{-1} \omega_{\text{AB}}^2(q) \]

To a first approximation, one can treat polymer chains on the basis of the highly simplified unperturbed model without intramolecular excluded volume interactions. This allows us to considerably simplify the problem by calculating the \( \omega_{ij} \) using the Gaussian function \( \exp(-q^2\sigma^2|j-j|/6) \) which describes the distribution of segments \( i \) and \( j \) belonging to the species \( A \) and \( B \) inside an \( N \)-unit polymer.

In pRISM theory, the MIST process is directly reflected in the normalized static structure factor, \( S(q) \). As a copolymer system is cooled and microdomains are forming, the peak scattering intensity grows in a mean-field manner corresponding to the linear portion of the \( S(q) \) curve in the coordinates \( S^2(q) - \chi \). Extrapolation of this linear portion to divergent intensity defines an apparent mean-field spinodal transition.\(^{164}\)

For the \( T^*(S) \) fitness function used in the sequence-selection process,\(^{161,162}\) it was found that if the sequence design temperature \( T_s \) is too high, the design can be viewed as a random walk in sequence space, and it yields random sequences, as expected. In contrast, when \( T_s \) is too low, \( A \) and \( B \) units tend to be separated within the chain and the evolutionary algorithm leads to trivial symmetric diblocks for \( f_A = f_B = \frac{1}{2} \). In the \( T_s \rightarrow 0 \) limit, one observes a typical mean-field behavior: \( T^* \propto N \) and \( r^* \propto N^{1/2} \) (for long symmetric diblocks, \( T^*/N = 0.096 \) or \( r^*/N = 10.4 \)). This is well seen in Figure 14 where the transition temperature \( T^* \) and the characteristic length scale \( r^* = (2\pi/q^*) \) are shown as a function of \( T_s \) for different \( N \). In between the \( T_s \rightarrow 0 \) and \( T_s \rightarrow \infty \) regimes, there is a certain critical design temperature \( T_s^{(c)} = 0.78 \) at which the transition from random to nonrandom sequences occurs.

What is rather unexpected here is that we find nonzero design temperature near which the \( r^* \) value has a maximum for sufficiently long chains, \( N \geq 200 \). In other words, there are such sequences, which produce microphase-separated structures with larger characteristic length scales than simple symmetric diblocks (at least, in the weak segregation regime).

Another important observation is that the transition in sequence space from random to nonrandom sequences occurs as a first-order-like transition. Indeed, the calculations predicted that near \( T_s^{(c)} \), the distribution function \( W(r^*) \) for the ensemble of generated sequences demonstrates a clear bimodality.\(^{162}\)

The most intriguing observation that can be done for this transitory regime is that sufficiently long sequences providing the maximum of \( r^* \) do not correspond to simple diblocks but rather they have a gradient shape, or it maybe better to say, an S-like shape (Figure 15(a)). Of course, among all generated sequences, symmetric diblocks have the highest value of \( T^* \) (i.e., their critical \( \chi \) parameter is lowest). However, S-like sequences can show larger characteristic length scales.

We now analyze the sequence design process with the value of \( r^* \) used as a fitness function. The maximization of the fitness

---

**Figure 14** (a) Transition temperature \( T^* \) and (b) characteristic length scale \( r^* = (2\pi/q^*) \) as a function of the sequence design temperature for different chain lengths \( N \) at 1:1 AB composition. Both \( T^* \) and \( r^* \) are averaged over the ensemble of generated sequences (\( \sim 10^5 \)).
The Jensen–Shannon divergence, JS, as a measure of complexity for the generated sequences, one can obtain an interesting result (cf. Figure 16). The most important feature is that JS value is a nonmonotonous function of the sequence design temperature Ts. Another intriguing result is that the Jensen–Shannon divergence measure shows a maximum in the vicinity of Ts(c). For the sequences generated in the sequence design process described above, it was found that at Ts = Ts(c), the degree of complexity, as measured by JS divergence, can be considerably higher as compared to that observed for very low and high Ts. The complexity increases with Ts decreasing, reaches its maximum just near Ts(c), and then sharply drops (Figure 16). Therefore, in the vicinity of Ts(c), the copolymer sequences of high complexity are generated, whereas at low Ts, the information content of the sequence is degenerated in the course of the design. One can say that the most interesting and unusual

Figure 15 Composition profiles shown as a function of reduced monomer number i/N for 512-unit ‘MIST-tuned’ copolymer sequences having 1:1 AB composition. The present definition assumes that the A-type monomers are coded by symbol +1, whereas symbol −1 is assigned to the B-type monomers. For an ideal RCP in which chemically different units follow each other in a statistically random fashion, the probability ρA that monomer A is located at the i-th position in the chain is ½ for any i. (a) Optimization of the transition temperature T* at different Ts leads to sequences that have a gradient (or S-like) distribution of A and B monomers, as schematically depicted in the right panel. (b) Optimization of the characteristic length scale r* at different rs leads to sequences that have a tapered (or ‘bolas-like’) distribution of A and B monomers, as schematically depicted in the right panel.
sequences appear at the ‘edge of chaos’ when \( T_\text{i} \) is close to the transition temperature \( T_\text{c} \).

### 6.18.3 Synthesis of Designed Copolymers

Three experimental approaches have been employed to synthesize PLCs. One of them utilizes synthesis of copolymers from monomers having various solubilities in a common solvent. Changing the solution conditions leads to local agglomeration of monomers; the synthesis performed under these conditions produces PLCs. One drawback of this approach is that copolymers thus formed comprise multiple fractions having varying chemical composition and comonomer sequences. Various separation methods have to be employed to fractionate the resulting macromolecules into monodisperse (composition and monomer sequence) fractions. Another method reported in the literature involves ‘grafting through’ approach of (typically hydrophilic) oligomeric or polymeric monomers onto a backbone of a parent macromolecule. The parent polymer is a ‘truly random’ copolymer containing responsive segments (i.e., those that respond to variation in temperature, i.e., \( N \)-isopropylacrylamide) and hydrophilic reactive units (i.e., glycidyl methacrylate) that will serve as attachment points for the macrografts. The characteristics of these PLCs will depend on the density and length of the grafts, the chemical composition of the grafts and the spatial distribution of the grafts along the parent macromolecule. In case that thermoresponsive segments are present in the parent RCP, the latter can be adjusted by varying the solution temperature during the ‘grafting through’ reaction. Finally, the last approach leading to PLCs reported in the literature follows the original sequence design scheme suggested by Khokhlov and Khalatur.59-77 Specifically, a parent homopolymer is placed in a selective solvent, where it adopts either an expanded (good solvent), Gaussian (theta solvent), or collapsed (poor solvent) conformation. A ‘chemical coloring’ species capable of modifying chemically monomers inside the parent homopolymer is added to the solution and it reacts with those monomers in the parent homopolymer that are accessible to the reaction thus converting them into another species. As will be discussed in detail below, this chemical modification can follow either chemical addition (A \( \rightarrow \text{C} \)) or elimination (A \( \rightarrow \text{C} \)). If no conformational change takes place during the reaction, the distribution of the newly formed ‘coloring’ segments in the PLC will be governed by the original conformation of the parent macromolecule. In reality, the situation may be more complex because the solubility and thus the overall conformation of the PLC changes with increasing ‘degree of coloring’. Depending on the interplay between the solubility of the original and newly created units in the copolymer, the chain may either expand (solubility of the newly added species is better than that of the original monomeric units) or can further collapse or even invert (solubility of the newly added species is worse than that of the original monomeric units). This behavior was predicted in computer simulations101 (see Section 6.18.2.3.1(iv)).

Below we provide more detailed description of the experimental approaches employed in the synthesis of PLCs.

#### 6.18.3.1 ‘Grafting through’ Approach

Early experimental realizations pertaining to the formation of PLCs appeared in the early 2000s. Virtanen and Tenhu165 prepared RCPs comprising \( N \)-isopropylacrylamide (NIPAAm) and glycidyl methacrylate (GMA) or \( N \)-acyrloylsuccinimide (NASI) units and grafted amine-terminated poly(ethyleneoxide) (PEO) grafts to the GMA (or NASI) moieties.166,162 Light scattering studies on those systems have revealed that the conformation of P(NIPAAm-g-EO) copolymers depends on the interplay between the solubility of P(NIPAAm, a thermo-responsive polymer, which exhibits the lower critical solution temperature (LCST) behavior in aqueous solutions at \( \approx 32 \) °C, and PEO and the conditions of the conformation of the parent P(NIPAAm-co-GMA) chain to which the PEO grafts were added. Although some of the observed properties (i.e., the critical solution temperature or viscosity) could be associated with varying number of PEO grafts in systems prepared under good and nearly critical conditions (15 and 29 °C), other characteristics (i.e., the sizes of aggregates under poor solvent conditions, fluorescence electron paramagnetic resonance) reflect various location of the PEO grafts along the copolymer.166-168 The authors noted that this suggest ‘conformational memory’ of the parent polymer.

#### 6.18.3.2 Conformation-Dependent Synthesis: Direct Copolymerization

Another approach toward preparing PLCs involves direct copolymerization of monomers or macromonomers under variable solubility conditions. This method is based on selective solubility of the two monomers leading to aggregation of the monomers and resulting in more blocky character of the resultant RCP. In early studies, Lozinsky et al.169-172 reported on radical copolymerization of mixture of \( N \)-vinylcaprolactam (NVCL)/\( N \)-vinylimidazole (1-VIA) mixtures in 10% dimethyl sulfoxide at temperatures both below and above phase separation. The researchers isolated three fractions of NVCL/1-VIA copolymers. Differential scanning calorimetry studies revealed that one fraction underwent conformational transition predicted by the theory.134,135 Siu et al.173 later reported on copolymerization of NIPAAm and \( N \)-vinylpyrrolidone (VP) at various temperatures. While PNIPAAm/VP copolymers synthesized at temperature below the LCST, copolymerization at temperatures higher than LCST resulted in a globular ‘protein-like’ comonomer sequences. Wahlund et al.174 copolymerized 1-VIA with NIPAAm and employed affinity chromatography to separate the resultant copolymers into two fractions with different physico-chemical characteristics. While copolymers with truly random distributions of the two monomers did not interact with the metal chelate iminodiacetic acid Sepharose CL 6B (\( \text{Cu}^{2+} \)-ID Asepharose), the other fraction comprising PLCs adsorbed readily. The authors reconciled the observed behavior by the accumulation of the hydrophilic pendant 1-VIA groups responsible for complexing at the hydrophilic periphery of the resultant copolymer.

One of the major drawbacks of the ‘direct copolymerization’ method is that one typically ends up with multiple fractions of the monomers and thus resulting comonomer distributions of the resulting copolymers. The authors noted
that the various fractions have to be isolated via tedious separations, however.

In a subsequent study, Kumar et al. used the P(NIPAm-co-1-VIA) copolymers to purify histidine-tagged green fluorescent protein (His-tag GFP) from recombinant *E. coli* by copper-chelate affinity precipitation. The authors noted that the comonomer distribution in the copolymer played a critical role in the interaction between the copolymer and the His-tag GFP, while complete elution of the affinity-bound Cu P(NIPAm-co-1-VIA) was achieved with PLC comonomer sequences (imidazole groups exposed to the outer solution), no recovery was obtained with IMAC non-bound copolymer fraction (imidazole groups unexposed). More details pertaining to the behavior of these systems can be found in a review by Lozinsky.

There were other reports in the literature that reported on direct copolymerization of various monomers that may lead to PLC.

With some exceptions, one of the monomers in the copolymer was a thermoresponse unit, that is, NIPAm, whose presence endowed the resultant copolymer with thermoresponse nature. Upon heating in water, PNIPAm undergoes coil-to-globule transition due to breaking hydrogen bonds between PNIPAm units and water and forming intrachain hydrogen bonds. While for PNIPAm, this transition occurs at \( \approx 32°C \), copolymerization of PNIPAm with another unit (say, U) shifts the transition point up or down, depending on the physico-chemical nature of U, chemical composition of the resultant copolymer, and comonomer distribution of NIPAm and U.

Works pertaining to studying the formation and properties of protein-like behavior included copolymers made by copolymerizing NIPAm with hydrophilic monomers, including, methacrylic acid (MAA), acrylic acid (AA), and sodium styrene sulfide (NaSS). Inter- and intramolecular association was observed that was governed by the interplay between the solubility of the hydrophilic units and temperature-dependent solubility of NIPAm. In most instances, the monomer sequence distribution in the resultant copolymer was not measured; the presence of blockiness was inferred indirectly.

Zhang and Wu performed an extensive series of experiments aiming at comprehending the phase behavior of RCPs bearing a NIPAm unit. For instance, they noted that copolymerization of NIPAm with hydrophilic VP at temperatures higher and lower than its LCST, respectively, resulted in segmented and random VP distributions in the resultant P(NIPAm-co-VP) copolymers. The authors reported that P(NIPAm-co-VP) chains with a random-block distribution of the VP units aggregated readily to form larger mesoglobules relative to their counterparts having a random VP distribution. The formation of mesoglobules was attributed to the competition between intrachain contraction and interchain association.

In another study, the researchers copolymerized NIPAm and EO at 45 °C, that is, above the LCST of NIPAm. Wu and coworkers used a palette of experimental tools including light scattering and differential scanning calorimetry and observed that the copolymer with a higher PEO content involved two transitions; one sharp transition at 33 °C and another broader one around 45 °C (both in heating and cooling processes). The second transition disappeared gradually with lowering the PEO content in the copolymer. The researchers attributed the first transition to the collapse of the PNIPAm segments. The second, higher-temperature transition, was associated with steric repulsion-induced stretching and the 3-clustering-induced collapse of PEO chains on the periphery of the collapsed PNIPAm cores.

Cheng et al. also reported on phase behavior of NIPAm and sodium acrylate (NaA) copolymers grafted on solid substrates. The researchers noted that the brushes can refold to the state at which the chains were prepared, a behavior that was consistent with the theoretical prediction for the folding of PLCs. They also commented on the formation of hydrogen bonds in the collapsed state of the copolymer chains, indicating that the hysteresis in the cooling process was due to the incomplete removal of additional hydrogen bonds.

Tang et al. reported on the formation of ‘macromolecular thermometers’ by direct copolymerization of NIPAm and a tetraphenylene (TPE)-based unit; TPE is nonemissive when dissolved but becomes highly emissive when aggregated. While no direct characterization of the NIPAm and TPE sequences was shown, the authors noted that increasing the content of the TPE units resulted in polymer aggregation due to hydrophobic interactions among TPE units, as deduced from increased fluorescence signals. While increasing the solution temperature collapses NIPAm units, it leads to breaking up TPE complexes due to increased molecular motion that ultimately leads to decrease of fluorescence signal. One can only speculate that variation of comonomer sequences in P(NIPAm-co-TPE) can be observed by monitoring the fluorescence emission as a function of temperature.

Similarly, Liu et al. prepared graft copolymers comprising poly(acrylic acid) (PAA) backbone and side-chain PNIPAm grafts that exhibited pH-responsive (due to PAA) and thermo-responsive (due to PNIPAm) behavior. The macromolecules were formed by AIBN-initiated direct copolymerization in solution. No effect of comonomer sequences on the phase behavior was reported, although the system would, in principle, allow for probing this effect efficiently.

Shimori and coworkers studied the formation of a copolymer of a RCP P(NIPAm-co-AAPBA) made by direct copolymerization of NIPAm and 3-(acrylamido)phenylboronic acid (AAPBA) and its interaction with glucose. It was reported that the temperature-induced precipitation of P(NIPAm-co-AAPBA) increased by 1.5 and 4 °C in the presence of a small concentration of glucose and fructose, respectively. While the authors did not comment on the effect of comonomer sequences in their paper, it is obvious that tailoring the distribution of the comonomer units in the copolymer will have a profound effect on the bind efficacy to sugars.

Multiblock copolymers comprising NIPAm and N,N-dimethylacrylamide (DMAAm) with variable length of blocks of both components prepared by reversible addition–fragmentation chain transfer polymerization were also investigated and their phase behavior was found to exhibit ‘protein-like behavior’ exhibiting intramolecular collapse upon heating, forming unimolecular flower-like micelles above the thermal phase transition temperature.

### 6.18.3.3 Polymer-Analogous Reactions

A few groups have developed experimental protocols for generating PLCs that follow very closely the scheme suggested by
early computer simulations of Khokhlov and Khalatur.69—71 Starting from a parent homopolymer made of chemically equivalent segments, PLCs can be generated by chemically modifying those segments that are accessible for reaction. In the initial stages of the reaction, the process is governed by the formation of the parent macromolecule; as the reaction proceeds and the solubility of the newly created copolymer changes, the formation of the copolymer and thus accessibility of the modifiable segments get altered, which, in turn, affects the accessibility of the reaction sites.90,101

Manokrung and Manias194 formed RCP comprising methyl methacrylate (MMA) and MAA units by alkaline hydrolysis of parent PMMA homopolymers with NaOH, which converted MMA into MAA. By performing the reaction under poor solvent conditions (isopropyl alcohol with small amounts of deionized water), the researchers speculated that the resulting P(MMA-co-MAA) copolymers possessed blocky distributions of the two monomers. While no attempts were made to characterize the comonomer distribution, the authors reported on sharp phase transitions in aqueous solutions as a function of pH and a corresponding hysteresis associated with increasing and decreasing the pH. The latter was attributed to inter- and intrachain interactions and stabilization of the collapsed state due to strong hydrophobic interactions acting among MMA units.

One of the most comprehensive studies pertaining to the formation of PLCs has been carried out by Semler et al.100 They utilized electrophilic substitution of Br onto PS135 in selective solvents as a means of preparing poly(styrene-co-4-bromostyrene) (PBr,S) with tunable distribution of the 4-bromostyrene units whose mole fraction (x) can be tune smoothly from 0 to 1. Jhon et al.196 have studied the kinetics of bromination of PS in selective solvents and established that regardless of the solvent type, the reaction is second order in bromine. The researchers have also carried out a series of experiments examining the role of solvent in the bromination reaction on the comonomer sequence distribution in PBr,S. A series of PBr,S copolymers was prepared by brominating parent PS (Mw = 30 kDa) in 1-chlorodecane (CD), 1-chloroundecane (CUD), or 1-chlorododecane (CDD) at various temperatures. Because the solvents have similar dipole moments but varying different theta temperatures (Tθ), (TθCD = 6.6 °C, TθCUD = 32.8 °C, TθCDD = 58.6 °C), the kinetics of bromination is similar at a given temperature for all three solvents.196 Because of the temperature- and solvent-dependent conformational changes of PS, bromination was expected to produce PBr,S with different sequence distributions of 4-BrS. For instance, at 32.8 °C, PS adopts a swollen coil conformation in CD, Gaussian coil in CUD, and a collapsed conformation when dissolved in CDD. In the following, we refer to the samples as PBr,S-SOLCD, where x denotes the mole fraction of 4-BrS, SOL stands for solvent (CD, CUD, or CDD), and the symbol ‘t’ denotes the bromination temperature in °C.

In order to establish the degree of blockiness of 4-BrS in PBr,S-SOLCD, Semler et al.100 utilized an electro-optical Kerr effect measurement in order to determine the molar Kerr constant (mK), which carries information about both the monomer stereoregularity and PS/4-BrS sequencing in PBr,S (cf. Figure 17). To gain a more quantitative insight into the results of the electrical birefringence measurements, the researchers calculated mK using the rotational isomeric state (RIS) model and matrix multiplication methods of Flory for a few selected periodic sequences having the same degree of bromination (~60%) and degree of polymerization (~300) as the experimental system. Comparison of the mKs for the RIS periodic sequences and the experimental sequences provided clear evidence that PBr0.58S-CDD33 and PBr0.63S-CUD33 have a more blocky character relative to PBr0.59S-CD33; the average block length of the 4-BrS units within PBr0.58S-CDD33 and PBr0.63S-CUD33 is in the range of 20 and 10 monomers, respectively.

The differences in monomer sequence distribution influence significantly the partitioning of copolymers at interfaces. Specifically, macromolecules with longer consecutive sequences of modified monomers that interact preferentially with the substrate adhere more strongly to the substrate relative to the polymers with random distribution of the modified units. Therefore, the mobility of RCPs on substrates, as reflected by the rate of dewetting, is also expected to decrease with increasing degree of blockiness of the modified monomers.

To confirm the differences in monomer sequence distributions between PBr0.59S-CD33 and PBr0.63S-CDD33, Genzer and coworkers studied dewetting of thin films of those polymers from solid substrates. The effect of the monomer sequence distribution in PBr,S was elucidated by depositing films of the same thickness onto substrates covered with semi-fluorinated organosilane (SiOS) self-assembled monolayers (SAMs) and annealing the specimens. The growth and coalescence of holes in the films of various thicknesses were recorded using optical microscopy as a function of time revealed that the polymer with a higher degree of blockiness dewetted more slowly relative to the specimen having a ‘truly random’ monomer sequence distribution. The sample set up and the results of the experimental findings are summarized in Figure 18.

Jhon et al.197 studied the effect of polymer composition, blockiness, solvent quality, polymer concentration, and temperature on adsorption of PBr,S PLCs onto silica substrates. They reported that among the solvents studied, cyclohexane showed the highest PBr,S adsorption to substrates because of high affinity of the 4-BrS unit to silica, strong repulsion of 4-BrS from the solvent, and weak adsorption of cyclohexane on the substrate. Under these conditions, the amount of PBr,S adsorbed onto silica substrates increased with increasing 4-BrS content in the copolymer. Upon increasing the temperature from 25 to 50 °C, the amount of PBr,S on silica decreased because styrene solubility in cyclohexane increased (the system crossed the theta-to-good solvent quality for PS), as shown by the data presented in Figure 19. While at low degrees of bromination, Jhon et al. could not detect any dependence of comonomer sequences in PBr,S on adsorption, increasing the degree of bromination to 35% produced adsorption isotherms that depended on the comonomer sequences. Specifically, samples with higher degree of blockiness exhibited higher adsorption on silica relative to PBr,S that possessed a random distribution of PS and 4-BrS units.

Solubility of PBr,S copolymers in cyclohexane was investigated in great detail. Specifically, Jhon et al.198 measured the turbidity of the solution as a function of solute concentration and temperature. The cloud point was found to depend on the comonomer distribution; it is always higher for random-blocky samples, relative to truly random comonomer distributions. By performing the experiments at various PBr,S concentrations in
cyclohexane and various amounts of 4-BrS in the sample, Jhon et al. began generating phase diagrams for solubility of RCPs as a function of comonomer sequences. During the course of these experiments, they have also discovered a strong isotope effect on solubility. These experiments were performed with all possible combinations of deuterium labeling: (1) PBr$_{1.0}$S in cyclohexane, (2) d-PBr$_{1.0}$S in cyclohexane, (3) PBr$_{1.0}$S in d-cyclohexane, and (4) d-PBr$_{1.0}$S in d-cyclohexane. In addition, the transition from coil to globule in random-blocky samples suggested that one may encounter multiple-chain collapse, as supposed to single-chain collapse present in the coil-to-globule transition of random PBr$_{1.0}$S.

Recently, Han et al.$^{199}$ carried out analysis of selective adsorption of PBr$_{1.0}$S copolymers using interaction chromatography (IC). Their IC experiments performed with a series of PBr$_{1.0}$S samples reveal that (1) IC is capable of determining the extent of bromination in PBr$_{1.0}$S samples regardless of the solvent used, and (2) for some solvents, one can, in addition to measuring the chemical composition of the random PBr$_{1.0}$S copolymer, also detect differences in monomer sequence distribution. The data in Figure 20 represent a summary of the IC experiments. As discussed by Han et al., the interplay between solvent-induced conformational changes and surface adsorption on chromatographic stationary phases offers new insights in chromatographic separation for the characterization of the synthetic copolymers and biomacromolecules that embody comonomer sequence distribution.

6.18.4 Concluding Remarks

Under the general heading of ‘sequence design’, there is an ever-increasing body of methodology that has been rapidly evolving over the past few years in polymer chemistry. In this
review, we have discussed advances that have recently been
demonstrated in the computer simulation, theoretical under-
standing, and synthesis of designed functional copolymers.
The fundamental principle of these approaches was con-
formational-dependent sequence design (CDS), which
takes into account a strong coupling between the conformation
and primary structure of copolymers during their synthesis. The
focus was on PLCs with a nontrivial primary sequence
exhibiting large-scale compositional heterogeneities and
long-range statistical correlations between monomeric units.
These features are intrinsically related to the CDS scheme and
cannot be explained by the basic stochastic processes such as random sequence or Markov chain; the first has no correlations and the second only has SRCs. We have tried to demonstrate how the preparation conditions dictate the resulting copolymer sequences. In many cases, the presence of LRCs...
in designed copolymers can bring about dramatic changes in their physical properties with respect to the corresponding copolymers whose sequence has only minor correlation between adjacent monomeric units.

While PLCs have been realized experimentally\(^\text{165–199}\) and their unique physical characteristics have been documented in multiple studies,\(^\text{76,77,200–210}\) much more remains to be understood about these unique macromolecular systems. First and foremost, our understanding of the coloring process in real experimental systems is limited. This stems primarily from technical difficulties associated with monitoring the sequence design (coloring) process with a sufficient spatio-temporal resolution. It is likely that, in order to gain insight into this process, multiple experimental probes would have to be employed concurrently. While some initial attempts have been recently for the \(\text{PBr}_S\) using small-angle neutron scattering in combination with Kerr effect and elemental analysis,\(^\text{211}\) much more work needs to be done before one can describe the experimental coloring pathways conclusively.

As discussed earlier in the document, the actual coloring pathway may follow multiple paths depending on the solubility of the parent homopolymer and that of the coloring units. While it appears from the preliminary data that during the formation of \(\text{PBr}_S\) a conformational rearrangement of the original PS coil takes place, other experimental situations may be realized that follow a different route. For instance, one can envision situations, also discussed earlier, where a compact globule opens up progressively during the coloring chemical reaction because the newly added monomer increases the overall solubility of the mother macromolecule. An example of such a system can be quaternization of poly(dimethylaminoethyl methacrylate) (PDMAEMA) in water at high temperatures, where PDMAEMA is originally collapsed.\(^\text{212}\)

Another area that merits further work involves detailed investigation of the effect of the molecular weight of the parent homopolymer on the comonomer sequence distribution in PLC. While it is desired that macromolecules with larger molecular weight are used in this process because they form a more compact globule and are thus expected to result in PLCs with more pronounced degrees of blockiness, technical limitations exist that set an upper limit on the molecular weight. To explain, in order to perform the coloring reaction under poor solvent conditions, one has to work with sufficiently low polymer concentrations (typically much below the overlap concentration, \(C^*\)) in order to avoid agglomeration of the collapsed homopolymer as well as the newly formed copolymer globules. Because \(C^*\) decreased dramatically with increasing molecular weight of the polymer, one may not produce a sufficient amount of the PLC when working with high-molecular-weight macromolecules. One possible solution to this problem could involve grafting the parent macromolecules onto a solid substrate in order to decrease their tendency to overlap with their neighbors. This situation can be realized technically by growing the parent polymers by ‘grafting from’ methodology from supports comprising inorganic or metallic (nano)particles decorated with polymerization initiators. The density of the initiators should be sufficiently low that the growing chains remain in the ‘mushroom conformation’ and

![Diagram](image-url)
are thus isolated. Sufficient amounts of polymer can be produced by employing small particles whose overall surface area-to-volume is high. Using small particles also provides the advantages of minimizing any confinement effect both during the ‘grafting from’ homopolymerization as well as the coloring reaction itself. Alternatively, end-functionalized polymers can be grafted chemically by the ‘grafting onto’ methodology from solution; this technique has the advantage of avoiding any confinement effects during the ‘grafting from’ technique as well as the ability to work with polymers that have been characterized previously in bulk. One has to stress, however, that avoiding confinement effect completely during the coloring process is likely not possible because the collapsed macromolecule may (and likely will) interact strongly with the particle support.

Finally, adequate analytical methods should be employed in order to characterize the structure of the copolymer. Characterization of the overall molecular weight and chemical composition of the copolymer is determined by utilizing well-established analytical methodologies. However, determination of comonomer sequence distribution is more challenging. While in some situations one can use nuclear magnetic resonance (NMR), information thus obtained is limited because NMR cannot distinguish sequences beyond the length of triads. As recently demonstrated by Semler et al.,\(^{100}\) Kerr effect, that is, measurement of optical birefringence of solutions under electric field, may provide a convenient means of determining comonomer sequence distribution (in addition to chemical composition and tacticity) when there is sufficient contrast between the dipole moments of the monomers present. Kerr effect does not suffer from the same limitation of NMR.\(^{211}\) However, in order to interpret the experimental data by complementing the measurements with modeling using RIS model,\(^{214,215}\) and calculating the so-called molar Kerr constant by employing the matrix formalism developed by Flory\(^{216,217}\) and others\(^ {217–222}\)

It should be noted that PLCs prepared by either of the aforementioned methodology are not monodisperse in their length, chemical composition, and comonomer distribution. The degree of polydispersity in any properties varies from method to method and even within each method, it may change depending on the actual system synthesized. Characterizing these polydispersities is a daunting task. Even more complex is fractionation of the resultant copolymers into monodisperse fractions.

References

Biographical Sketches

Jan Genzer received his 'Diploma-engineer' degree (Dipl.-Ing.) in Chemical and Materials Engineering from the Institute of Chemical Technology in Prague, Czech Republic, in 1989. In 1991, he moved to the United States to pursue graduate studies at the University of Pennsylvania under the direction of Professor Russ Composto, receiving his PhD degree in Materials Science and Engineering in 1996. After two postdoctoral stints with Professor Ed Kramer first at Cornell University (1996–97) and later at University of California at Santa Barbara (1997–98), Genzer joined the faculty of chemical engineering at the NC State University as an assistant professor in fall 1998. He is currently the Celanese Professor of Chemical and Biomolecular Engineering at NC State University. His honors include Camille Dreyfus Teacher-Scholar Award, National Science Foundation CAREER award, John H. Dillon Award of the American Physical Society, National Science Foundation Award for Special Creativity, and others. He is a Fellow of the American Physical Society. His group at NC State University is actively involved in research related to the behavior of polymers at interfaces and in confined geometries, with particular emphasis on self-assembly and forced assembly and combinatorial methods.

Pavel G. Khalatur received a Ph.D. degree in chemistry from the Tver State University and a Dr.Sc. degree in physics and mathematics from the Institute of Macromolecular Compounds of the Russian Academy of Sciences. In 1990, he became a full professor of the Department of Physical Chemistry of Tver State University. In 2002, he moved to the University of Ulm to pursue his research with A. R. Khokhlov, dedicated to computer-aided molecular design of bioinspired polymers. He has been awarded four Distinguished University Professor Awards from the International Science Foundation for excellence in science and teaching. Prof. Khalatur has authored and coauthored over 300 articles and four books in the field of polymer science and advised more than twenty Ph.D. candidates. The major foci of his research are computer modeling of polymer systems, computer-assisted molecular design, nanoscience and the integral equation theory of simple and polymeric liquids.

Alexei R. Khokhlov is the head of a scientific school, and the author of textbooks widely known both in Russia and abroad. The range of his interests covers polymer science, statistical physics of macromolecules, physical chemistry of polyelectrolytes and ionomers, microphase separation in polymer systems, polymer liquid crystals, polyelectrolyte responsive gels, topological restrictions in polymer systems, dynamics of concentrated polymer solutions and melts, coil–globule transitions, associating polymers, computer simulation of polymer systems, biomimetic polymers, and proton-conducting polymer membranes. His publications include works on both theoretical and experimental science, some of the former being classical in the field. Among others, he has proposed methods to synthesize ‘protein-like’ copolymer sequences as well as macromolecules whose functional properties can be preset. The theoretical principles he developed have been successfully implemented in the field of oil extraction, in the production of durable biocompatible polymers and new efficient catalysts to synthesize vitamins.