Core/Shell Structures of Proteinlike Copolymers: Are Finite Aggregates Thermodynamically Stable?

A. N. Semenov[†]

Physics Department, Moscow State University, Moscow 119992, Russia, and Institut Charles Sadron, 6 rue Boussingault, 67083 Strasbourg Cedex, France

Received June 27, 2003; Revised Manuscript Received October 20, 2003

ABSTRACT: We propose a theory of aggregation in solutions of amphiphilic copolymers consisting mostly of insoluble H (hydrophobic) units with a small fraction of soluble (P, polar) monomer units. When P units are arranged along the sequence in a periodic (regular) fashion, the resultant HP copolymers are essentially insoluble: they precipitate. The main result is that finite aggregates of HP copolymers can be made stable by an appropriate smart arrangement of the same number of soluble P units in the chemical sequence. An analytical approach yielding the thermodynamic quantities of H-core/P-shell copolymer structures in the weak stretch limit is developed. We show that different types of copolymer aggregates and microdomain structures can be thermodynamically stable depending on the copolymer chemical sequence. The relationship between the block-length distribution and copolymer aggregation is illustrated by a few phase diagrams. Copolymers that can form stable finite aggregates with H-core/P-shell structure (including single-chain globules) are often referred to as the *proteinlike* copolymers. Thus, the present theory sheds a new light as to the essential features of proteinlike copolymer sequences.

1. Introduction

Segregation and precipitation of polymers is one of the most common phenomena occurring in macromolecular systems. It often results in a macroscopic phase separation, e.g., between the polymer and the solvent. In some cases, the aggregation may be stopped at an intermediate stage yielding a system of finite aggregates (micelles) or even compact globules of individual polymer chains. It is the latter possibility that can provide the most promising basis for development of functional polymeric materials (just recall remarkable properties of block copolymer micellar solutions,[1] globular proteins,[2] or fine emulsions and dispersions in polymer blends[3]).

It is well-known that homopolymer globules are intrinsically unstable with respect to precipitation.[4] Hence, there is the necessity to resort to copolymers (heteropolymers) with two or more monomer unit types. For example, if hydrophilic side chains are grafted to a hydrophobic backbone, then the grafted chains can form stable starlike micelles (or globules) with a hydrophobic core protected by a hydrophilic shell. It is well-known, however, that such stars are thermodynamically stable only if the core is relatively small in comparison with the hydrophilic corona. It is likely that such a thick shell can provide undesirable screening of the (possibly functional) core. That might be one of the reasons why the typical structure of globular proteins is different: their hydrophobic cores are protected by relatively thin and transparent hydrophilic shells.

Hence the following question appears to be natural and relevant: Is it possible to make nonprecipitating copolymers (forming thermodynamically stable globules or finite aggregates) by incorporating a *small* fraction of soluble monomer units in the essentially insoluble polymer chain. This problem is not trivial: it is rather obvious that *regular* HP chains would readily precipitate from solution possibly forming an inverse microdomain structure (P cores, H coronas). Nevertheless, we show below that the desired stabilization can indeed be achieved with specially tuned irregular HP copolymer sequences.

The paper is organized as follows: In the next section we consider the general model, identify the main free energy contributions and discuss aggregation of regular HP copolymers. At the end of the section the notion of H−PH copolymers is introduced. Their aggregation behavior is analyzed in section 3. Next we turn to more general irregular sequences. Analytical theory treating equilibrium core/shell structures formed by such copolymers in the regime of weakly stretched blocks is developed in section 4. The resultant phase diagrams and other findings are discussed in the last section.

2. The Model

We consider a solution of HP copolymer chains mainly consisting of insoluble (hydrophobic, H) blocks separated by hydrophilic (polar, P) groups.[5] The solution is assumed to be dilute; hydrophobic blocks are much longer than soluble polar groups.[6] The mean length of an H block is $\bar{n} \gg 1$. It is assumed that long homopolymer H chains (without P units) would precipitate and form a homogeneous polymer phase with concentration (number of H units per unit volume) c_0. P units serve to stabilize either finite aggregates (globules) formed by HP copolymers or their microdomain structures in the precipitate. The anticipated structure of an aggregate involves a nearly homogeneous H core and a relatively thin surface layer decorated by P units. We assume that the core environment is so unfavorable for P units that all of them are located in the surface layer. Thus, the H blocks must be somewhat extended so as to span the aggregate's thickness D. Depending on the ratio of D to the Gaussian block size $b\bar{n}^{1/2}$, we distinguish the weak stretch ($D \ll b\bar{n}^{1/2}$) and strong stretch ($D \gg b\bar{n}^{1/2}$) regimes. The latter regime (SSL with polydisperse H blocks and in nonplanar geometry) does not seem to be tractable analytically. Therefore, it is the weak stretch limit (WSL) that is considered below (see section 4.1).

[†] Address correspondence to the Institut Charles Sadron.

10.1021/ma030357y CCC: $27.50

2.1. The Free Energy. Consider an aggregate formed by one or several HP chains each consisting of $\mathcal{N}_H$ hydrophobic monomer units and $\mathcal{N}_P$ polar (P) units, $\mathcal{N}_P \ll \mathcal{N}_H$. Its free energy F incorporates the following basic contributions:

$$F = F_{vol} + F_{surf} + F_{conf} + F_{bend}$$

where $F_{vol} = Vf_{vol}(c)$ is the free energy of the nearly uniform hydrophobic core (V is the core volume, $c = \mathcal{N}_H/V$ is the concentration in the core), $F_{surf} = A\gamma$ is the surface free energy (A is the surface area), F_{conf} is the conformational free energy of H blocks associated with the restriction that their ends must be located at the surface, and F_{bend} is the elastic energy due to surface curvature. The effective surface tension γ incorporates the interfacial tension γ_H between the reference hydrophobic (H) homopolymer and the solvent, and the contribution due to the presence of P units: $\gamma = \gamma(\sigma)$, where $\sigma = \mathcal{N}_P/A$ is the surface concentration of P units ($\gamma(\sigma)$ includes both the energy of affinity of P units to the surface and the energy of their interactions).

The bulk term F_{vol} is formally dominant, however it merely defines the core volume of the aggregate, $V = V_0 = \mathcal{N}_H/c_0$ and the concentration c_0 of H units in the core. Likewise, F_{surf} defines the optimal surface area $A = A_0$. (We assume that the typical surface energy is much larger than the conformational energy.) Thus, the aggregate's equilibrium structure must correspond to the minimum of the residual free energy per unit volume, $(F_{conf} + F_{bend})/V_0$, for a given surface-to-volume ratio $\alpha = A_0/V_0$. In the case of a deficit of P units, the optimum area A_0 may be too small to be accessible for a given V; this would only mean that larger aggregates (with larger V and smaller A/V ratio) are more favorable; i.e., the aggregates would coalesce.

The bending energy can be expanded in terms of the principal surface curvatures C_1, C_2:[7,8]

$$F_{bend} \simeq \int\left[\frac{1}{2}K(C_1 + C_2 - C_0)^2 + K_G C_1 C_2\right]dA \qquad (1)$$

where C_0 is the spontaneous curvature. Here we assume that the curvatures are small: the surface layer thickness h defined by the size (length) of P units is much smaller than the aggregate's size D. The mean (K) and Gaussian (K_G) bending moduli depend primarily on the structure of P units and their interactions. For example, we may assume that each P unit is actually a relatively short block of n_P subunits grafted to the main H chain. If the difference of chemical potentials that drives P blocks from inside the core to the surface is much higher than the thermal energy k_BT, then nearly all P units must accumulate at the surface. It is then natural to expect that P units form a surface brush facing the solvent. The relevant elastic constants predicted for swollen brushes are[9]

$$K = \frac{9}{64}\left(\frac{12}{\pi^2}\right)^{1/3} n_P^{\,3}\sigma^{7/3}w^{4/3}b'^{2/3}, \quad K_G = -\frac{64}{105}K \qquad (2)$$

where w is the second virial coefficient of interaction of primary P monomers and b' is the statistical segment length of a P block.[10]

Typically $K > 0$, $K_G < 0$.[8,9] The latter condition means that two separate aggregates have lower bending energy than a united aggregate obtained by connecting the two primary aggregates by a small neck (Figure 1).

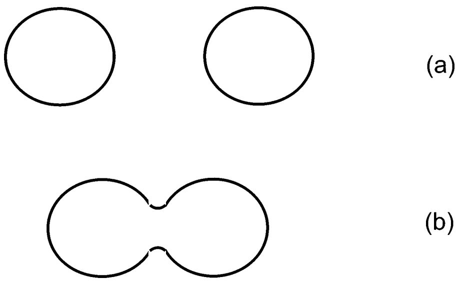

Figure 1. Two separate aggregates (a); united aggregate (b).

2.2. Regular Copolymer. In this section we consider the simplest case of regular copolymer (all H blocks of the same length). An estimate of the ratio of F_{bend} to F_{conf} can be done in a very simple way: The typical conformational energy per one H block is k_BT (in the intermediate regime at the crossover between WSL and SSL); the typical surface energy per P unit is somewhat larger than k_BT. Therefore, $F_{surf}/F_{conf} \sim 1$. On the other hand it seems likely that the bending energy F_{bend} scales as $F_{surf}(Ch)^2$, where $C \sim 1/D$ is the typical surface curvature, and h is the surface layer thickness (see eq 1). Therefore, $F_{bend}/F_{conf} \sim (h/D)^2$, i.e., $F_{bend} \ll F_{conf}$ for a regular copolymer since $h \ll D$. Hence, we may neglect F_{bend} in the first approximation.

A qualitative argument pointing to the equilibrium structure of regular copolymers is presented below: The conformational energy per H block can be crudely estimated as $\mathcal{F}_{conf} \sim D^2/(\bar{n}b^2)$, where b is the statistical segment of the block. We expect that HP copolymers either form low-dimensional aggregates (spheres, cylinders or layers) or precipitate forming an inverse structure (e.g., spherical domains containing solvent in the copolymer matrix). The size of low-dimensional aggregates is defined by the ratio $\alpha = A/V$ of the (optimal) area of H/solvent interfaces to the volume of hydrophobic cores: $D \sim 1/\alpha$. On the other hand, D for an inverse structure can be made arbitrarily small for any given α.

As a simple example consider inverse spheres of radius R (inverse cylindrical structure is less favorable; see section 4.1.1). The volume fraction ϕ_s of inverse spheres (the solvent domains) is then defined geometrically:

$$\frac{1 - \phi_s}{\phi_s} = \frac{3}{R\alpha}$$

It is obvious that both R and intermicellar distance ($\sim R/\phi_s^{1/3} \propto \phi_s^{2/3}$) tend to 0 as ϕ_s is decreased (the optimum R is still finite of course: it is determined by the relevant cutoff, the size of polar groups).

Therefore, the lowest conformation free energy is achieved with an inverse structure (a more careful analysis shows that it is the inverse spherical structure that is favored, see section 4.1.1); i.e., precipitation of regular HP copolymers is predicted.

Below we address the following basic question: is it possible to make low-dimensional (in particular, finite) copolymer aggregates thermodynamically stable by resorting to an irregular arrangement of P units along the hydrophobic backbone. The answer is positive, but the general problem is rather complicated. That is why we start analyzing it with the simplest example: aggregation of the so-called H−PH copolymers[11] consisting

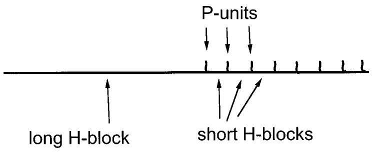

Figure 2. H−PH copolymer.

of one long H block and many short H blocks (Figure 2).

3. H−PH Copolymers

The very copolymer structure shown in Figure 2 implies that almost every P unit is conjugate with a short H block. As all P units are at the surface, the HP section of the chain (with short H blocks) must form interfacial (surface) layer with an optimum surface area A_0. We assume that this layer is relatively thin: its thickness h is much smaller than the characteristic size of the aggregate, D (the thickness h is comparable to the size of P units). This is equivalent to assuming that the total length of all short H blocks is smaller than that of long H blocks (we also assume that the size of short H blocks is comparable to the size of P units). The core(s) of the aggregated structure must be then formed primarily by long H blocks. We assume that the core H blocks are so long that their conformational energy (which is due to their elastic elongation and localization of their ends) is negligible (apart from direct contribution of long blocks to the interfacial tension due to the inhomogeneous distribution of H units near the surface). The necessary condition for that is $KC_0^2 \gg 1/(Nb^2)$, where N is the number of H units in a long block (i.e., approximately, the total copolymer polymerization degree).

Therefore, the equilibrium structure of the system is defined by minimization of F_{bend}/V (see eq 1) for a given $\alpha = A/V$. There are two essential nondimensional parameters involved: the ratio of elastic constants K_G/K and the reduced spontaneous curvature $\tilde{C} = C_0/\alpha$. The surface layer can be viewed as an inner sublayer of short H loops and outer sublayer of P units. If each P unit is actually a short grafted soluble block, then the P sublayer may be considered as a swollen brush. On the other hand, the H sublayer is more like a molten brush. The theoretical ratio K_G/K for a swollen brush is defined in eq 2: $K_G/K \simeq -64/105$. The similar ratio for a molten brush is $K_G/K \simeq -4/15$.[9] It seems reasonable to assume that K_G/K for H−PH copolymers is between the two values mentioned above which we may consider as the reference values. As for C_0, it is determined by a competition between the inner and the outer sublayers, so C_0 can take any value, either positive or negative, and the same is true for the reduced curvature $\tilde{C}$. In what follows we typically assume that $C_0 \geq 0$; i.e., the surface tends to be convex.[12]

Minimization of the bending energy may seem a straightforward task, however it is in fact rather involved technically and often requires an extensive numerical work.[13] To simplify the problem, we restrict the consideration to the following basic shapes of the copolymer aggregates: infinite plain layer, infinite cylinder, sphere, finite cylinder with semispherical caps, disk with round edge, vesicle (spherical layer), torus, and a "swollen star" consisting of a central large sphere and a number of smaller spheres (Figure 3).[14]

Inverse structures (inverse cylinders or spheres) involve concave interfaces (with negative mean curvature), and hence, they are always less favorable than, e.g., plain layers (since $C_0 > 0$ and $-2 < K_G/K < 0$). An analysis shows that other more complicated structures like modulated cylinder or bicontinuous phases with triply periodic surfaces[15] can be competitive if both $|K_G/K|$ and $\tilde{C}$ are small enough. For example, modulated

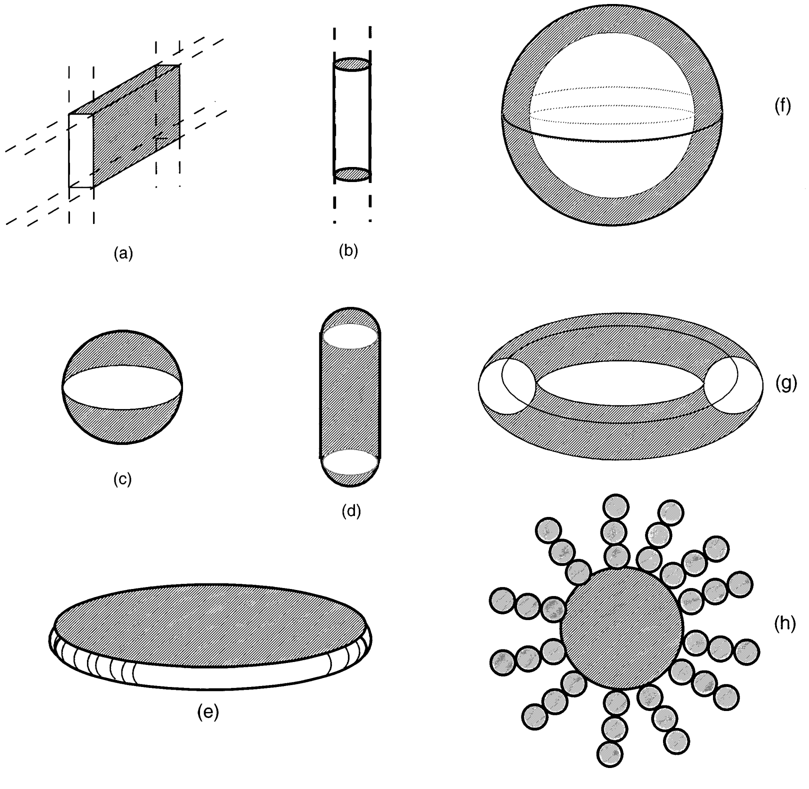

Figure 3. Infinite plain layer (a); infinite cylinder (b); spherical aggregate (c); finite cylinder with semi-spherical caps (d); finite disk with round edge (e); vesicle (f); torus (g); "swollen star" (h).

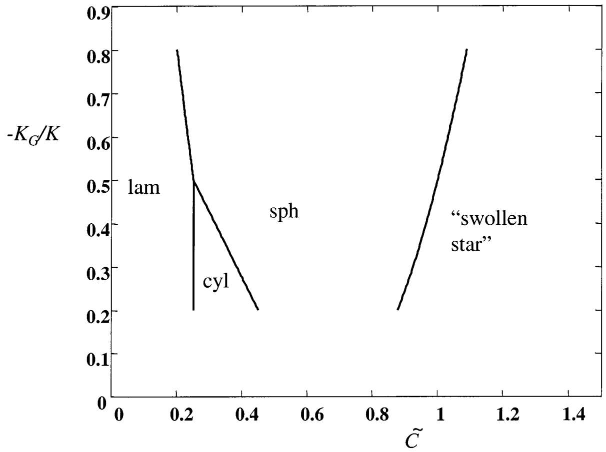

Figure 4. Phase diagram for H−PH copolymers; $\tilde{K} = K_G/K$, $\tilde{C} = C_0/\alpha$; lam = infinite plain layers; cyl = cylinders; sph = spheres.

cylinder may become favorable for $|K_G/K| < {}^1/_8$. Thus, we disregard these structures here since the reference values of $|K_G/K|$ are higher.

The bending free energy of a spherical aggregate (radius R) is

$$F_s = 2\pi K(2 - 3\tilde{C})^2 + 4\pi K_G \tag{3}$$

where $\tilde{C} = C_0/\alpha$ and $\alpha = 3/R$. Similar equations can be obtained for other shapes of the aggregates. By comparing the bending free energy per unit area for different shapes, we found that the equilibrium structure never involve finite cylinders or disks, etc: it is only infinite layers, infinite cylinders, spheres, and "swollen stars" that compete with each other. The phase diagram is shown in Figure 4. Infinite plain layers are always favorable at low positive $\tilde{C}$; "swollen stars" are formed if the reduced curvature is large enough.

Each "swollen star" consists of a large central sphere of radius R_c and many much smaller spheres forming "arms" (see Figure 3). The radius R of the peripheral spheres is defined by the spontaneous curvature: $R \simeq 2/C_0$; $R_c \gg R$. The volume of the central sphere (V_c) and the total volume of all smaller spheres (V_{small}) are comparable: a simple geometrical argument shows that $V_{small}/V \simeq 2/3\tilde{C}$ (here $V = V_c + V_{small}$ is the total volume). On the other hand the total surface area of the smaller spheres is dominating. Thus, the overall size of the "swollen star" is nearly equal to R_c if $\tilde{C} \gg 1$. To estimate this size, we have to invoke the elastic free energy of the long blocks (which is negligible for other morphologies as discussed above). Assuming that the "star" is formed by many H−PH copolymers (and that $\tilde{C} \gg 1$), we write its free energy as

$$F_{star} \simeq \frac{1}{2}KC_0^{\ 2}4\pi R_c^{\ 2} + \mathcal{N}\frac{R_c^{\ 2}}{10Nb^2}$$

where

$$\mathcal{N} \simeq \frac{4\pi R_c^{\ 3}}{3Nv_0}$$

is the number of chains in the aggregate, and $v_0 = 1/c_0$ is volume per H unit. Here the first term is the surface bending energy of the central sphere (the bending energy of the small spheres is negligible because their surface curvature is C_0; the contribution of the Gaussian curvature term amounts to $4\pi K_G$ for the whole star; i.e., it is also negligible in comparison with the first term). The second term is the conformational (elastic) energy of long H blocks (the numerical factor 1/10 corresponds to the "weak stretch limit" considered in section 4.1.1, see eq 15). Minimizing the free energy per copolymer, $F_{star}/\mathcal{N}$, we get $R_c \simeq (15Kv_0b^2\alpha^2\tilde{C}^2N^2/2)^{1/3}$. The number of peripheral spheres forming the "arms" is

$$\mathcal{M} \simeq \frac{2}{3\tilde{C}}\frac{R_c^{\ 3}}{R^3} \simeq \frac{5}{8}Kv_0b^2\alpha^5\tilde{C}^4N^2$$

It should be expected that $\tilde{C} = C_0/\alpha$ is rather sensitive to external conditions like solvent quality (or temperature) because the spontaneous curvature C_0 is controlled by a subtle balance of bending elasticities of the two (P and H) surface sublayers. Therefore, it is reasonable to assume that $\tilde{C}$ varies while other parameters like α, K_G/K, and c_0 stay nearly constant. The following morphological changes are then predicted as $\tilde{C}$ increases from $-\infty$ to $+\infty$, for example, at $K_G/K = -4/15$: the H−PH copolymers precipitate for $\tilde{C} < 0$, form infinite layers for $0 < \tilde{C} < 0.25$, infinite cylinders for $0.25 < \tilde{C} < 0.406$, spherical aggregates for $0.406 < \tilde{C} < 0.91$, and "swollen stars" for $\tilde{C} > 0.91$.[16]

The thickness of layers and cylinders, and the diameter of spheres are nearly constant in the respective windows of their stability (since $\alpha \approx$ const). On the other hand, the size of "swollen stars" is increasing with $\tilde{C}$ being approximately proportional to $\tilde{C}^{2/3}$. Similar changes are predicted for $K_G/K = -64/105$, except that the window of stability of cylindrical aggregates disappears in this case.

Thus far, we assumed many copolymer chains per aggregate, i.e., that $c_0/(N\alpha^3) \gg 1$. If this is not the case, then the mass of an aggregate has to be quantized. This effect may destabilize simple spherical aggregates in favor of "swollen stars" , infinite or finite modulated cylinders and other nonspherical shapes. For example, a simple analysis of the opposite limit of very long copolymers, $N\alpha^3/c_0 \gg 1$ (many "optimal spheres" per copolymer) yields the following transitions as $\tilde{C}$ is increased: plain layers (see Figure 3a) $\xrightarrow{0.25}$ infinite cylinders (see Figure 3b) $\xrightarrow{0.5}$ infinite modulated cylinders (Figure 5a) → finite modulated cylinders (Figure 5b) $\xrightarrow{2/3}$ "swollen stars" (Figure 3h); the numbers above the arrows indicate approximate values of $\tilde{C}$ at the transitions.

4. Irregular Sequences

Thus far, we considered a rather special case of H−PH copolymers where formation of finite aggregates is induced by spontaneous curvature of their surface which in turn is defined, in particular, by sufficiently strong interactions of polar units at the surface. In this regard, the very mechanism selecting the shape of the aggregates is rather similar to that defining the geometry of block copolymer micelles and the related microdomain structures.

In this section, we consider a much more subtle case, assuming that polar units are structureless (material points), so they do not prefer any surface curvature, and their fractions are still small. Thus, to obtain the equilibrium structure of the system we minimize its free energy disregarding any explicit bending energy. This

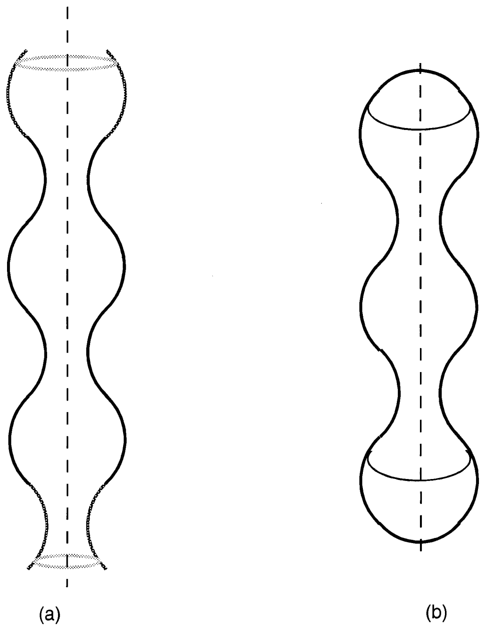

Figure 5. Infinite modulated cylinder (a); finite modulated cylinder (b).

is equivalent to minimizing the conformational free energy (per copolymer chain) for a given $\alpha = A/V$.

The same model was considered for the regular HP copolymers in section 2.2: we argued that regular copolymers always precipitate under such conditions. Below we show that it is possible to stabilize finite copolymer aggregates by playing with molecular weight distribution of H blocks.

The general task of free energy calculation is considered in the next section. To simplify the theoretical treatment we restrict the analysis to the most symmetric morphologies: inverse spheres, inverse cylinders, plain layers, cylinders, and spheres. This choice is inspired by the results of the previous section showing that other less symmetric shapes of the aggregates are likely to be less stable (we disregard the "swollen star" shape though).

4.1. Conformational Free Energy. The core of an aggregate can be considered as a system of H loops with ends at the surface. Analytical calculation of their conformational free energy (with the additional condition that the overall density of H units is nearly uniform inside the core) in the general case seems to be an untractable task: a similar problem with *polydisperse* end-grafted chains has not yet been solved analytically for nonplanar geometries even in the asymptotic strong segregation regime. That is why below we assume the regime of weakly stretched blocks, which has been proven to be tractable analytically. We prefer to call it the weak stretch limit (WSL) rather than weak segregation because segregation (of P and H units) is actually extremely strong: P at the surface and H in the core. What is assumed is that H blocks in the core are not significantly stretched, i.e., that the corresponding molecular field in the core is weak enough and can be considered as a perturbation. Clearly this regime must correspond to the situation when the mean Gaussian size of the H blocks is much larger than the aggregate's thickness: $\bar{n} \gg D^2/b^2$. However this does not mean that every block is long: a broadly polydisperse distribution of block length including a small fraction of short blocks is allowed.

The conformational free energy can be derived in two steps. First we assume no molecular field in the core, i.e., ideal noninteracting loops. The corresponding free energy term, F_{loc}, is entirely due to localization of the loop ends at the surface:[11]

$$F_{\text{loc}} = -\mathcal{N}\int p(n)\ln G_{\text{n}}\,\mathrm{d}n \tag{4}$$

where $\mathcal{N}$ is the total number of H blocks (thus $\mathcal{N} \simeq \mathcal{N}_{\text{P}}$), $p(n)$ is the fraction of H blocks with exactly n units, and G_n is the statistical weight of a loop of n units with one end fixed at the surface and another end anywhere at the surface:[11]

$$G_{\text{n}} = \frac{1}{A}\oint\oint \mathrm{d}^2r\,\mathrm{d}^2r_0\,G(\underline{r}, n;\underline{r}_0) \tag{5}$$

Here $G(\underline{r}, n;\underline{r}_0)$ is the statistical weight of an H block (n) with ends at $\underline{r}_0$ and $\underline{r}$.

It is important to note that the overall distribution of H units in the core is not necessarily uniform at the first step: $c_1 = c_1(\underline{r}) = c_0 + \delta c_1(\underline{r})$. That is why a molecular field is required. The molecular potential $U(\underline{r})$ (=potential energy of an H monomer unit located at $\underline{r}$) is introduced at the second step. We find the corresponding contribution to the conformational free energy (which we somewhat arbitrarily refer to as *elastic*) by assuming that $U(\underline{r})$ is being turned on gradually:

$$F_{\text{el}} = \int \delta F, \quad \delta F = -\int \mathrm{d}^3r\; U(\underline{r})\delta c(\underline{r})$$

where $\delta c(\underline{r})$ is the variation of concentration of H units induced by a variation $\delta U(\underline{r})$ of the molecular field (such that $\int \delta c(\underline{r}) = -\,\delta c_1(\underline{r})$). A great simplification is coming now from the fact that δc is dominated by the response of long blocks (since the weight fraction of short blocks is small, and any short block produces a weaker response to δU as compared to a long block). Therefore, the ground-state approximation is applicable; i.e., we may calculate δc assuming that the loops are infinite. The result

$$F_{\text{el}} \simeq \frac{b^2}{24c_0}\int (\nabla c_1)^2\,\mathrm{d}^3r \tag{6}$$

coincides with the well-known square-gradient conformational energy. To use this equation we have to find $\delta c_1(\underline{r})$. Neglecting a contribution of short blocks, we write

$$\delta c_1(\underline{r}) \simeq 2\mathcal{N}\int \mathrm{d}n\;\psi_{\text{n}}(\underline{r}) \tag{7}$$

where $2\mathcal{N}$ is the total number of ends of all long blocks and $\psi_n(\underline{r})$ is the distribution density of the n th monomer unit of an infinite ideal H block that is end-grafted to the surface (the grafted end is $n = 0$).[17]

The ψ-functions satisfy the well-known Edwards equation

$$\frac{\partial \psi_{\text{n}}}{\partial n} = \frac{b^2}{6}\nabla^2\psi_{\text{n}} \tag{8}$$

with the initial condition $\psi_0(\underline{r}) = \text{const}\;\delta(h(\underline{r}))$ stemming from the fact that the grafted unit ($n = 0$) can be anywhere on the surface. Here $\delta(\cdot)$ is the δ function, $h(\underline{r})$ is the distance from $\underline{r}$ to the surface, and const $= 1/A$ is defined by the normalization condition: $\int \psi_n(\underline{r})\,\mathrm{d}^3r = 1$. The boundary condition

$$\nabla_{\perp}\psi_{\mathrm{n}} = 0, \quad \text{for } \underline{r} \text{ on } A \tag{9}$$

($\nabla_{\perp}$ is the gradient in the direction normal to the surface) reflects the desired uniform distribution of H units in the core.

Thus, we get using eqs 6−8

$$F_{\mathrm{el}} \simeq \frac{\mathcal{N}^2}{c_0 A}\int_0^{\infty} \mathrm{d}n \oint (\psi_{\mathrm{n}}(\underline{r}) - 1/V)\,\mathrm{d}^2 r \simeq \frac{\mathcal{N}}{\bar{n}A}\oint \Phi(\underline{r})\,\mathrm{d}^2 r \tag{10}$$

where $\bar{n} \simeq c_0V/\mathcal{N}$ is the mean length of an H block; $\Phi(\underline{r})$ is defined by the following equations:

$$\frac{b^2}{6}\nabla^2\Phi = 1 - \frac{V}{A}\delta(h(\underline{r})), \quad \text{for } \underline{r} \text{ in } V;$$

$$\nabla_{\perp}\Phi = 0, \quad \text{for } r \text{ on } A; \quad \int \Phi\,\mathrm{d}^3 r = 0$$

Noting an obvious relation between $\psi_n(\underline{r})$ as defined by eq 8 and the partition function $G(\underline{r}, n;\underline{r}_0)$

$$\psi_{\mathrm{n}}(r) = \frac{1}{A}\oint \mathrm{d}^2 r_0 G(\underline{r}, n;\underline{r}_0)$$

and using eqs 4 and 5, we get an alternative representation of F_{loc}

$$F_{\mathrm{loc}} \simeq -\mathcal{N}\ln\alpha - \mathcal{N}\int p(n)\ln\Psi_n\,\mathrm{d}n \tag{11}$$

where

$$\Psi_n = \frac{1}{\alpha}\oint \mathrm{d}^2 r\,\psi_{\mathrm{n}}(\underline{r}) \tag{12}$$

The constant term $-\mathcal{N}\ln\alpha$ does not depend on the aggregate's shape.

Employing the analogy between eq 8 and the simple diffusion equation we note that Ψ_n is proportional to the probability that a Brownian trajectory that starts at the surface would return to it after n time steps. Noting also that eq 9 corresponds to the reflective boundary conditions for the diffusion problem, we find that $\Psi_n \to 1$ as $n \to \infty$ since $\psi_n(\underline{r})$ tends to a uniform function of $\underline{r}$ for large n. Therefore, long blocks do not contribute to the second term in F_{loc} (since $\ln\Psi_n \to 0$ for $n \gg D^2/b^2$, where $D \sim 1/\alpha$ is the aggregate's size). The "elastic" free energy can be also expressed in terms of Ψ_n:

$$F_{\mathrm{el}} \simeq \frac{\mathcal{N}}{\bar{n}A}\oint \Phi(r)\,\mathrm{d}^2 r = \frac{\mathcal{N}}{\bar{n}}\int_0^{\infty}(\Psi_n - 1)\,\mathrm{d}n \tag{13}$$

Thus, the total conformational energy is

$$F_{\mathrm{conf}} = F_{\mathrm{loc}} + F_{\mathrm{el}} \tag{14}$$

where F_{el}, F_{loc} are defined in eqs 11 and 13.

4.1.1. Regular Copolymers. We start the detailed consideration of the WSL with the simplest case of regular copolymers. Here the localization term (per block) is just a constant: $F_{\mathrm{loc}}/\mathcal{N} \simeq -\ln\alpha$ since all blocks are long, $n = \bar{n} \gg D^2/b^2$. The elastic term can be obtained using eq 10. The result is

$$F_{\mathrm{el}}/\mathcal{N} \simeq K_{\mathrm{el}}\frac{6}{\alpha^2\bar{n}b^2} \tag{15}$$

where $K_{\mathrm{el}} = d/(d+2)$ with $d = 1$ for layers, $d = 2$ for cylinders, and $d = 3$ for spheres. Note that the energy per block is proportional to D^2/R_{n}^2, where $D \sim 1/\alpha$ is the aggregate's size, and R_n is the Gaussian gyration radius of the block. Thus, F_{el} is formally nearly the same as the energy of block elongation in the strong stretch limit (apart from a numerical factor). This perhaps justifies the name "elastic" attributed to this energy.

For inverse structures, we use the spherical (cylindrical) approximation for the Wigner−Seitz cell: the H domain within one cell is represented as a layer between two concentric spheres (cylinders) of radii R and rR, respectively, $r > 1$. Thus, we get

$$K_{\mathrm{el}} = \frac{1}{(r^2-1)^2}\left[\frac{r^4\ln r^2}{r^2-1} - \frac{1}{2}(3r^2-1)\right]$$

for inverse cylinders (i.cyl) and

$$K_{\mathrm{el}} = \frac{3}{5}\,\frac{5r^3+6r^2+3r+1}{(r^2+r+1)^3}$$

for inverse spheres (i.sph). The reduced surface curvature (defined after eq 3) is related to r:

$$\tilde{C} = \frac{-1}{R\alpha} = \frac{1}{2}(1-r^2), \quad \text{for i.cyl;}$$

$$\tilde{C} = \frac{-2}{R\alpha} = \frac{2}{3}(1 - r^3), \quad \text{for i.sph.}$$

Obviously K_{el} for inverse structures is vanishing as r is increased. Therefore, inverse structures with large $|\tilde{C}|$ (small R) are more favorable than other structures as already mentioned in section 2.2. The asymptotic behavior of K_{el} for large r (large $|\tilde{C}|$) is

$$K_{\mathrm{el}} \sim \frac{\ln|\tilde{C}|}{|\tilde{C}|} \text{ for i.cyl;} \quad K_{\mathrm{el}} \sim \frac{1}{|\tilde{C}|} \text{ for i.sph}$$

Obviously the lowest F_{el} for a given large $|\tilde{C}|$ corresponds to a structure of inverse spheres as stated in section 2.2.

4.1.2. Irregular Copolymers. We now turn to the effect of polydispersity of H blocks assuming that there is a fraction of relatively short H blocks. Therefore, the second term in F_{loc}, eq 11, must be taken into account. As for the elastic term, F_{el}, it stays the same as for regular copolymers, i.e., eq 15 is applicable. Therefore, the only missing quantity required for calculation of the conformational free energy is the reduced "return probability" , Ψ_n. It can be obtained using eqs 8 and 12. Skipping the details, we write the result:

$$\Psi_n = \Psi(z) = 1 + \sum_{q>0} K\exp(-z\lambda(q)) \tag{16}$$

where

$$z = \frac{nb^2\alpha^2}{6}$$

$K = 2/d$ (i.e., $K = 2$ for layers, $K = 1$ for cylinders, and $K = {}^2/_3$ for spheres)

$$\lambda(q) = \frac{q^2}{d^2}$$

and q is taking the values of all roots of the characteristic equations:

$$\sin q = 0, \quad \text{for layers}$$

$$J_1(q) = 0, \quad \text{for cylinders}$$

$$\tan q - q = 0, \quad \text{for spheres}$$

In the case of inverse structures Ψ_n depend also on the ratio r of the outer radius to the inner radius of the primitive cell:

$$\Psi_n = \Psi(z) = 1 + \sum_{q>0} K(q, r) \exp(-z\lambda(q, r))$$

where

$$K(q, r) = (r^2 - 1)/(r^2\Xi^2 - 1), \quad \Xi = \frac{Y_1(q)J_0(qr) - J_1(q)Y_0(qr)}{Y_1(q)J_0(q) - J_1(q)Y_0(q)}$$

$$\lambda(q, r) = \frac{q^2}{4}(r^2 - 1)^2, \quad J_1(q)Y_1(qr) - Y_1(q)J_1(qr) = 0$$

for inverse cylinders, and

$$K(q, r) = \frac{2}{3}(r^2 + r + 1)q^2 \Bigg/ \left[q^2 + 1 + \frac{q^2 - 1 + \frac{2q^2(r-1)}{1+rq^2}}{1 + rq^2 + \frac{q^2(r-1)^2}{1+rq^2}} \right]$$

$$\lambda(q, r) = \frac{q^2}{9}(r^3 - 1)^2, \quad \tan[q(r - 1)] - \frac{q(r-1)}{1 + rq^2} = 0$$

for inverse spheres. Here $J_\nu(\cdot)$, $Y_\nu(\cdot)$ are the Bessel functions of the first and second kinds, respectively.

Note that Ψ_n actually depend on just one parameter, $z = b^2\alpha^2 n/6$ (apart from the geometry of the aggregates): $\Psi_n = \Psi(z)$. Plotting $\Psi(z)$ for different morphologies one can see that (for a given z) the “return probability” is the highest for spheres. The same tendency can be observed by analyzing $\Psi(z)$ for small z. After some algebra we get

$$\Psi(z) \simeq \frac{1}{\sqrt{\pi z}} + 0.5\,\tilde{C} + \mathcal{O}(\sqrt{z}) \tag{17}$$

where $\tilde{C} = (C_1 + C_2)/\alpha$ is the reduced surface curvature (see section 3), $\tilde{C} = 0$ for plain layers, $= {}^1/_2$ for cylinders, $= {}^2/_3$ for spheres, and $\tilde{C} < 0$ for inverse structures. Qualitatively eq 17 can be understood as follows: the free end of a block attached to the surface can explore a region of size $s \sim \sqrt{z}$, Ψ is nearly proportional to the ratio of accessible area to the accessible volume, and the latter is smaller if the surface is bent toward the H region; hence, higher Ψ is seen for spherical aggregates than for plain layers or for inverse structures.

Omitting the first constant term in eq 11, we write

$$F_{\text{loc}}/\mathcal{N} \simeq -\sum_n p(n) \ln\Psi(b^2\alpha^2 n/6) \tag{18}$$

Thus, F_{loc} is always decreasing when the morphology changes from an inverse structure to plain layers, then to cylinders, then to spheres, i.e., behavior which is exactly opposite to the behavior of F_{el}. Therefore, the localization free energy, F_{loc} (that is effected for irregular sequences) may give rise to stabilization of spherical (or cylindrical) aggregates if the ratio $|F_{\text{loc}}|/F_{\text{el}}$ is not too small. This ratio can be estimated as (see eqs 15 and 18)

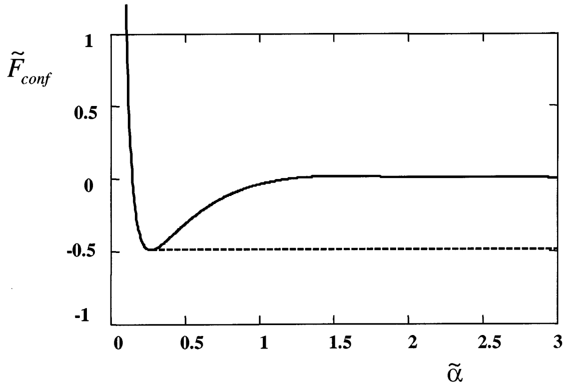

Figure 6. Typical dependence of the conformational free energy $\tilde{F}_{\text{conf}} = F_{\text{conf}}/(\mathcal{N}\hat{p}_s)$ on the reduced surface area $\tilde{\alpha} = \alpha b\sqrt{n_s/6}$ (for spherical aggregates); $\mathcal{N}\hat{p}_s$ = number of short blocks; $B \equiv n_s/\bar{n}p_s = 0.05$. The dashed line corresponds to spherical aggregates with crumpled surface.

$$\frac{|F_{\text{loc}}|}{F_{\text{el}}} \sim p_s \frac{\bar{n}b^2}{D^2} \tag{19}$$

where p_s is the fraction of relatively short blocks with Gaussian size $\lesssim D \sim 1/\alpha$, the aggregate's size. Clearly, this ratio may be however high provided that $\bar{n}$ is long enough (or $\alpha \sim 1/D$ is large enough).

Another important dissimilarity between the elastic and the localization terms concerns their dependence on the reduced surface area α. Obviously F_{el} is a decreasing function of α (see eq 15). On the other hand F_{loc} is always increasing as α is increased: note that $\Psi_n = \Psi(z)$ is a strictly decreasing function of $z = nb^2\alpha^2/6$.

As a simple example let us consider the total conformational energy $F_{\text{conf}} = F_{\text{el}} + F_{\text{loc}}$ for a bimodal block-length distribution (just short and long blocks)

$$p(n) = p_s\delta(n - n_s) + (1 - p_s)\delta(n - n_l) \tag{20}$$

assuming the spherical geometry ($p_s \ll 1$, $\bar{n} \simeq n_l \gg n_s$). Using eqs 15 and 18, we get

$$F_{\text{conf}}/\mathcal{N} = F_{\text{el}}/\mathcal{N} + F_{\text{loc}}/\mathcal{N} \simeq \frac{3}{5}\frac{6}{\alpha^2 b^2 \bar{n}} - p_s \ln\Psi(\alpha^2 b^2 n_s/6) \tag{21}$$

where $\Psi(\cdot)$ is defined in eq 16.

A typical plot of the total conformational energy vs α (for the spherical morphology) is shown in Figure 6. $F_{\text{conf}}/(\mathcal{N}\hat{p}_s)$ depends on just two parameters: the dimensionless surface area

$$\tilde{\alpha} \equiv \alpha b\sqrt{n_s/6}$$

and

$$B \equiv \frac{n_s}{\bar{n}p_s}; \quad B = 0.05,\ p_s \ll 1$$

Note that F_{conf} is decreasing at both small and large α's, while it is increasing with α in the intermediate region when F_{loc} is dominating. Thus, the plot shows a

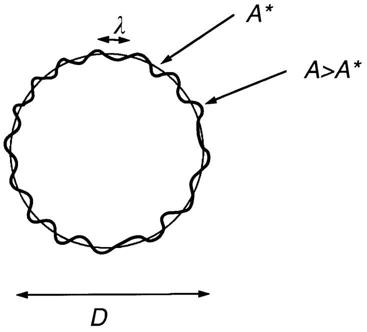

Figure 7. Crumpled interface with $A > A^*$; the apparent area of the smoothed surface is A^*. The length-scale of the surface pattern, λ, is much shorter than D.

minimum at $\alpha = \alpha_{\rm min}$ and a broad maximum at a larger α. Note that

$$\alpha_{\rm min} = \tilde{\alpha}_{\rm min}/R_{\rm s},\ R_{\rm s} = b\sqrt{n_{\rm s}/6} \qquad (22)$$

where $\tilde{\alpha}_{\rm min}$ is a coefficient that depends only on

$$B = \frac{n_{\rm s}}{\bar{n}p_{\rm s}}$$

(and on the shape of the aggregates) and $R_{\rm s}$ is the unperturbed gyration radius of short blocks.

Now we argue that the increasing part of the dependence $F_{\rm conf}$ vs α is unstable: In fact it is always possible to reduce the *apparent* area of the smoothed surface, A^*, by introducing an appropriate short-range waviness of the core/solvent interface (Figure 7). It is important that both $F_{\rm el}$ and $F_{\rm loc}$ (defined in eq 18) depend on the apparent area parameter

$$\alpha^* = \frac{A^*}{V}$$

i.e., it is α^* that must be substituted for α in eqs 15 and 18. If $\alpha > \alpha_{\rm min}$, then aggregates with lower apparent surface area ($\alpha^* = \alpha_{\rm min}$) are favorable; hence, the actual core/solvent interface has to become crumpled. The degree of crumpling is

$$\frac{A}{A^*} = \frac{\alpha}{\alpha^*} = \frac{\alpha}{\alpha_{\rm min}}$$

(see Figure 7; here we assume that the length-scale of the surface pattern is much shorter than both the typical size of the aggregate, D, and the size of a short block).

Note that D keeps constant in the region $\alpha > \alpha_{\rm min}$: $D \sim 1/\alpha_{\rm min}$. Let us estimate D for the bimodal case, eq 20, assuming that

$$B = \frac{n_{\rm s}}{\bar{n}p_{\rm s}} \ll 1.$$

As verified below this condition is equivalent to assuming that D is much larger than the size of short blocks, i.e., that $\tilde{\alpha}^2 = \alpha^2 b^2 n_{\rm s}/6 \ll 1$. Therefore, the asymptotic eq 17 is applicable here, i.e., (see also eq 21):

$$F_{\rm conf}/\mathcal{N} \simeq \frac{3}{5}\,\frac{6}{\alpha^2\bar{n}b^2} + \frac{1}{2}p_{\rm s}\ln(\pi\alpha^2 b^2 n_{\rm s}/6) \qquad (23)$$

Minimizing $F_{\rm conf}$ with respect to α we get

$${\alpha_{\rm min}}^2 \simeq \frac{36}{5}\,\frac{1}{p_{\rm s}\bar{n}b^2},$$

i.e., the aggregate's diameter[18]

$$D = 6/\alpha \simeq \sqrt{5}\sqrt{p_{\rm s}}b\sqrt{\bar{n}}$$

Note that D is proportional to the size of long blocks $R_{\rm l} \sim b\sqrt{\bar{n}}$ and to the square root of the fraction of short blocks. We can now verify that $\tilde{\alpha}^2 = \alpha^2 b^2 n_{\rm s}/6 \sim B \ll 1$ and that $R_{\rm s} \ll D \ll R_{\rm l}$.

It is interesting that the localization free energy generates an apparent surface tension in the regime $\alpha > \alpha_{\rm min}$ considered above. In fact

$$F_{\rm loc} \simeq \mathcal{N}\frac{p_{\rm s}}{2}\ln(\pi(\alpha^*)^2 b^2 n_{\rm s}/6)$$

in this regime, where $\alpha^* = A^*/V$. Hence the apparent surface tension

$$\gamma^* = \frac{\partial F_{\rm loc}}{\partial A^*} \simeq \frac{\mathcal{N}p_{\rm s}}{A^*},$$

i.e., γ^* is proportional to the number of short blocks per unit area.

4.2. Phase Diagrams for Irregular Copolymers. Having considered a few simple examples illustrating the typical features of aggregation of irregular copolymers, we now turn to a systematic description of the aggregation for different molecular-weight distribution of H blocks. We first consider the simple bimodal distribution, eq 20. The phase diagram is shown in Figure 8. The relevant dimensionless parameters are

$$\tilde{\alpha} \equiv \alpha b\sqrt{n_{\rm s}/6} \quad \text{and} \quad B \equiv \frac{n_{\rm s}}{\bar{n}p_{\rm s}}$$

Note that the ratio

$$|F_{\rm loc}|/F_{\rm el} \sim \frac{\tilde{\alpha}^2}{B}$$

Since irregularity affects the localization free energy only, it is not surprising that for small $\tilde{\alpha}$ or for larger B the structure is the same as predicted for regular copolymers, i.e., an inverse structure of precipitated copolymer matrix with an array of relatively small solvent spheres in it (0 precipitate); the size D of the spheres is defined by the competition of the elastic energy and the surface bending energy; the resultant D is much smaller than the size of H blocks.

For small B the main effect of the localization free energy is just a renormalization of the effective surface energy: note that the second term in the rhs of eq 23 depends on α, but it does not depend on the morphology. Therefore, in analogy with regular copolymers, separate aggregates are less favorable than inverse structures in this regime. However, the competition between inverse structures is more subtle here: the asymptotic eq 17 is not applicable for inverse structures since typically the interfacial curvature is too high. As a result, inverse cylindrical structures are stable in a relatively narrow region $\alpha_1 < \alpha < \alpha_2$, where

$$\tilde{\alpha}_1 \sim \exp\left(-\frac{\rm const}{B}\right) \quad \text{and} \quad \tilde{\alpha}_2 \sim \frac{B}{|\ln B|}$$

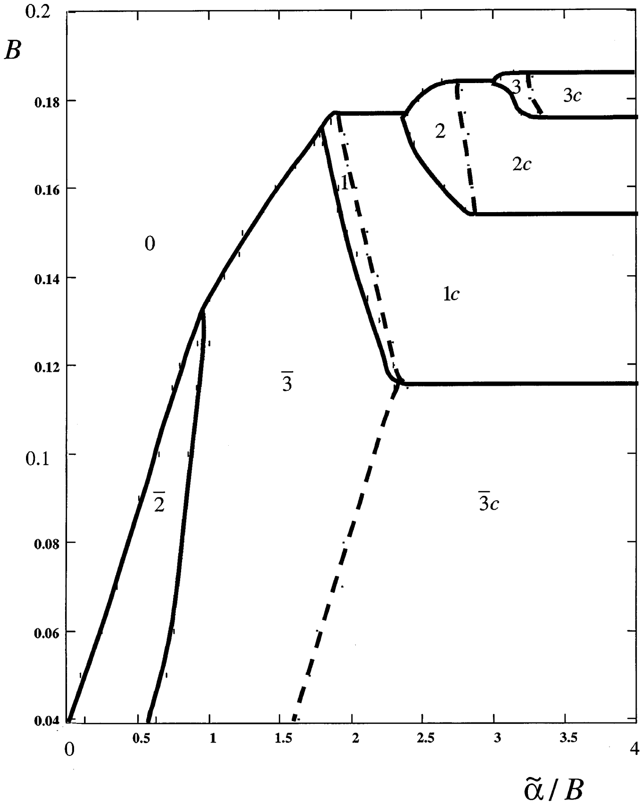

Figure 8. Phase diagram for bidisperse distribution of H block length; $\tilde{\alpha} = \alpha b\sqrt{n_s/6}$, $B = n_s/\bar{n}p_s$. 0 = 0 precipitate (characteristic of regular HP copolymers); 1 = plain layers; 2 = cylinders; 3 = spheres; $\bar{2}$ = inverse cylinders; $\bar{3}$ = inverse spheres; *c* stands for "with crumpled interfaces".

The reason is that the radius of curvature of the copolymer/solvent interface, $R_{\rm in}$, is much smaller than the size of short blocks, $R_{\rm s}$, in this regime. Therefore, the probability that a short block returns to the same interface forming a loop can be much higher for an inverse cylinder than for an inverse sphere of a comparable radius; hence, an inverse cylindrical morphology can be favored. The ratio $R_{\rm in}/R_{\rm s}$ increases with α, thus shifting the equilibrium to a structure of inverse spheres. For $\alpha > \alpha_3$ inverse spheres with crumpled interfaces are favorable; α_3 corresponds to the minimum of $F_{\rm conf}(\alpha)$ as explained above.[19]

For not too small (and not too large) B's the localization term favoring finite aggregates is still competitive with the elastic term, and, at the same time, the regime of crumpled interfaces is shifted to higher $\tilde{\alpha}$'s. Therefore, plain layers, or cylindrical (spherical) aggregates are becoming favorable in this regime (see Figure 8). Note that finite spherical aggregates (including spheres with crumpled surfaces) are stable in a relatively narrow range of B's: $0.175 < B < 0.187$.

Next we turn to the case of polydisperse short blocks:

$$p(n) = \frac{p_{\rm s}}{n_{\rm s}} e^{-n/n_{\rm s}} + (1 - p_{\rm s})\delta(n - n_{\rm l})$$

i.e., an exponential (Flory) molecular-weight distribution for short blocks. This distribution is characteristic of correlated random copolymers[20] (the assumption of monodisperse long blocks is inessential: a moderate polydispersity of long blocks does not affect either the free energy, or the morphology). The corresponding phase diagram (Figure 9) is qualitatively similar to that for the bidisperse case. The main quantitative difference: the window of stable finite aggregates (spheres) is much wider in Figure 9 as compared with the bidisperse case.

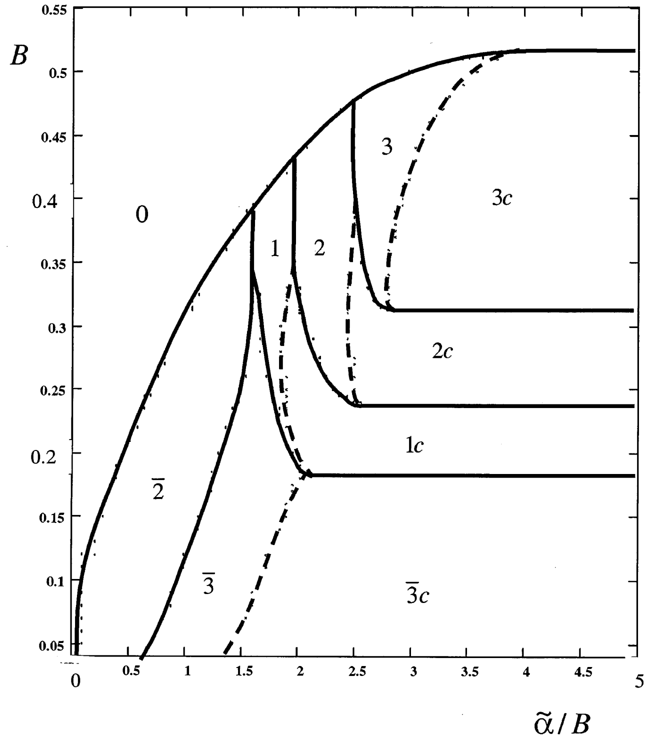

Figure 9. Phase diagram for exponential length distribution of short H blocks. All notations are as in Figure 8.

Molecular-weight distributions of polymers often obey power-laws for short chains (Schultz–Zimm distributions[21]). On the other hand, the power-law distributions of block lengths had been discussed for proteins[22,23] and "proteinlike copolymers" [24] (Levy flight statistics of copolymer sequence). These notions inspire consideration of the following distribution

$$p(n) = C_{\rm s}\left(\frac{\bar{n}}{n}\right)^{\theta} H(\bar{n} - n) + (1 - p_{\rm s})\delta(n - \bar{n})$$

where

$$C_{\rm s} = \frac{p_{\rm s}(1 - \theta)}{\bar{n}},$$

$H(\cdot)$ is the Heaviside function, $p_{\rm s} \ll 1$. A simple scaling analysis of eq 18 yields

$$F_{\rm loc}/\mathcal{N} \sim -\frac{p_{\rm s}}{(\alpha^2 \bar{n} b^2)^{1-\theta}}$$

On the other hand

$$F_{\rm el}/\mathcal{N} \sim \frac{1}{\bar{n}b^2}\frac{1}{\alpha^2}$$

(see eq 15). The structure of the aggregates depends on the ratio $|F_{\rm loc}|/F_{\rm el}$ defining the following essential parameter:

$$\tilde{\alpha} = \text{const}\ \alpha b \bar{n}^{1/2} p_{\rm s}^{\ 1/2\theta}$$

Thus, we predict a universal phase diagram in terms

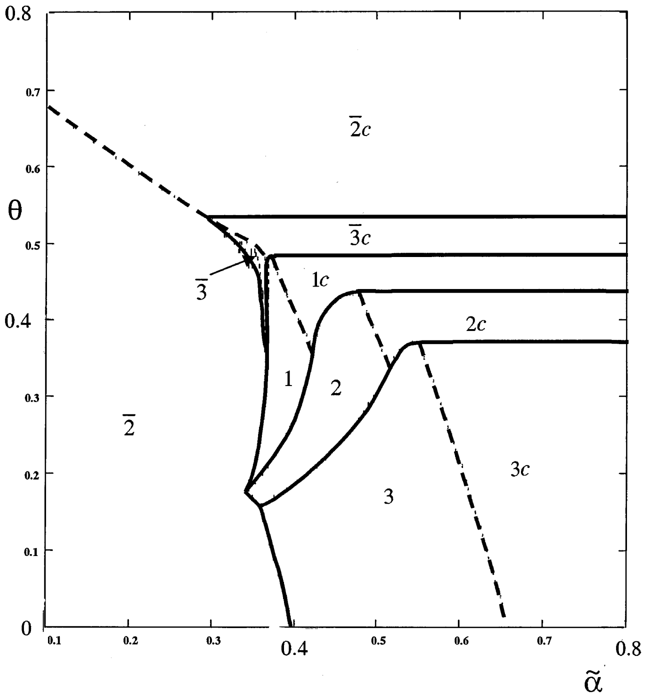

Figure 10. Phase diagram for power law length distribution of short H blocks ($p(n) \propto n^{-\theta}$). $\tilde{\alpha} = \text{const}\ \alpha b \bar{n}^{1/2} p_s^{1/2\theta}$. All other notations are as in Figure 8.

of $\tilde{\alpha}$ and the exponent θ (provided that $p_s \ll 1$ and that the weak stretch conditions are satisfied). We have considered the range $0 < \theta < 1$ in order to avoid divergencies of the free energy terms for either large or small n's. The diagram calculated using the general equations obtained in the previous section is shown in Figure 10. If $\theta > 0.48$, then only inverse morphologies are thermodynamically favorable. We also observe that spherical aggregates are not stable if $\theta > 0.37$. On the other hand, if $\theta < 0.37$, then finite aggregates (globules) are always stable (perhaps with crumpled surfaces), provided that the optimum surface area is large enough: $\tilde{\alpha} \gtrsim 0.4$. The size of the spheres with crumpled surface is

$$D \sim b\bar{n}^{1/2} p_s^{1/2\theta}$$

i.e., this size is quite sensitive to the fraction of short blocks, in particular, if θ is small.[25]

5. Discussion and Conclusions

(1) In this paper we proposed a theory of aggregation in solutions of amphiphilic copolymers (so-called HP copolymers). In particular, we consider the case when the fraction of soluble (P) monomer units in the sequence is small, therefore regular copolymers are essentially insoluble: they are precipitating. The main conclusion is that finite aggregates (including single-chain globules) of the copolymers can be made stable by an appropriate smart arrangement of the same soluble P units along the chemical sequence. That sort of copolymers that can form finite aggregates with H-core/P-shell structure, can be referred to as *proteinlike* copolymers.[26] Thus, the present theory elucidates certain essential features of proteinlike copolymer sequences.

More generically, we refer to such sequences as "irregular copolymers" thus stressing that they are neither regular nor totally random. An irregular HP copolymer considered in the paper is a sequence of insoluble H blocks separated by identical P units. The irregularity of the copolymer sequence is thus related to the polydispersity of H blocks.

(2) We considered two qualitatively different cases: (i) one or a few long H blocks and a lot of very short H blocks, the size of short blocks being comparable to the size of P units (each P unit can be viewed as a short monodisperse soluble block); (ii) a lot of long H blocks with a small fraction of shorter H blocks, the size of short H blocks being still much larger than the size of P units. In all cases, the copolymers tend to form a core/shell structure: interactions of H units in the core define its (optimum) volume V, and interactions of P units at the surface define the optimum surface area A; the relevant important parameter is $\alpha = A/V$. Finite aggregates can be stable in both cases i and ii, yet for rather different physical reasons.

(3) In case i finite copolymer aggregates (spheres and "swollen stars") are stabilized by a spontaneous interfacial curvature favoring convex surfaces (when observed from the side facing the solvent). The spontaneous curvature can be induced by, e.g., repulsion of the soluble P units at the surface. The relevant phase diagram indicating the regions of stability of different aggregate kinds is shown in Figure 4. Finite aggregates are stable if the reduced preferred curvature $\tilde{C}$ is high enough. Note that the aggregates are thermodynamically stable: they do not precipitate.

(4) For the copolymer model ii, the driving force destabilizing inverse structures in favor of finite aggregates is the localization entropy related to the requirement that both ends of a short H block must be located at the surface. The entropy is higher if the surface is curved toward the "hydrophobic" side thus increasing the probability of an H loop (an H block of a given length starting and terminating at the surface).

Another effect predicted in the paper is the formation of short-scale surface patterns in the case when the (optimum) interfacial area is large enough, i.e., for high α's (this effect is referred to as surface "crumpling"). The length-scale of the pattern is much smaller than the size of H blocks. We did not investigate the precise short-scale structure of the pattern: the only essential parameter is the ratio of the actual area of the "crumpled" interface to the apparent area of the smoothed surface. The driving force for "crumpling" is the same entropy of short H loops (with both ends "attached" to the interface) that favors structures with higher accessible area of the interfacial region.

The phase behavior for the copolymer model ii is extremely sensitive to the length distribution of short H blocks. The three illustrative examples are shown in Figures 8−10 for monodisperse, exponential, and power-law ($\sim n^{-\theta}$) distributions, respectively.[28] Finite aggregates are stable in wide α ranges for the exponential distribution, and for the power-law distributions with a low exponent θ. For example, for exponential distribution with $B = 0.33$, the following transitions are predicted as α is increased: 0 precipitate → precipitate with inverse cylindrical structure → inverse spherical structure → unbound plain copolymer layers → layers with crumpled surface → cylindrical aggregates → cylinders with crumpled surface → spheres → spheres with crumpled surface. For the power-law distribution with $\theta < 0.16$ the following morphologies are obtained:

0 precipitate → inverse cylinders → finite spheres → spheres with crumpled surface.

The diameter (thickness) D of the aggregates is inversely proportional to α for $\alpha < \alpha_{\min}$, and $D = \text{const}$ in the regime of crumpled interfaces, $\alpha > \alpha_{\min}$; in all cases, D is roughly proportional to the typical Gaussian size of short blocks (except the case of power-law molecular-weight distribution when the typical size of short blocks just cannot be defined).

(5) To illustrate how the predictions (for the second model) work for a realistic situation, let us consider HP-copolymers in a slightly poor solvent for H blocks, i.e., close to the coil−globule transition for the reference H homopolymers (still we assume that H/H attraction and H/P repulsion are strong enough, so that a core/shell structure is formed). As we showed above, the aggregated structure for the second model is defined by the surface area parameter α (apart from the block-length distribution). Therefore, it would be instructive to find out how α depends on the "distance" to the θ-point

$$\tau = \frac{T_\theta}{T} - 1$$

It is well-known that homopolymer in a slightly poor solvent normally precipitate; concentration in the polymer phase scales as $c_0 \propto \tau$ for $\tau \ll 1$, and polymer/solvent interfacial tension, $\gamma_H \propto \tau^2$.[4] It is likely that incorporation of a small fraction of P units in the polymer chains does not affect c_0 much; i.e., the core volume

$$V = \frac{\mathcal{N}_H}{c_0} \propto 1/\tau.$$

On the other hand, it is interactions of P units that define the optimum surface area A. Let us assume that P units are actually relatively short grafted chains of n_P subunits, forming a brush facing a marginal solvent. Then the energy of their repulsion (per unit area) scales as[29] $\gamma_P \propto \sigma^{5/3}$, where

$$\sigma = \frac{\mathcal{N}_P}{A}$$

is surface density of P units. Minimizing the total surface energy $A(\gamma_H + \gamma_P)$ with respect to A, we find that the optimum density $\sigma \propto \tau^{6/5}$ and the optimum area

$$A \propto \frac{1}{\sigma} \propto \tau^{-6/5}$$

Hence $\alpha = A/V \propto \tau^{-1/5}$.

We are now in a position to describe the theoretically anticipated sequence of morphological transitions that should occur as the solvent quality is decreased. For example, let us consider the case of power-law length distribution for shorter H blocks with the exponent $\theta = {}^1/_4$. An initial increase of τ will result in formation of finite copolymer aggregates (or single-chain globules) at $\tau = \tau_0$: nearly spherical aggregates with "crumpled" surfaces are predicted here. The diameter D of the aggregates is defined by an average of the Gaussian sizes of H blocks. A further increase of τ nearly does not affect D. On the other hand, the amplitude of surface pattern decreases with τ, and it disappears at $\tau_1 > \tau_0$ (transition from "crumpled" spheres to normal spheres). For $\tau > \tau_1$, the size D weakly increases with τ (as $\tau^{1/5}$). The following transitions are predicted as τ is further increased: spheres → cylinders → unbound plain layers → polymer precipitate with inverse cylindrical structure. The first transition is accompanied by an abrupt decrease of D by a factor of 1.5 (the diameter of cylinders is two-thirds of the diameter of "coexisting" spheres); at the second transition D decreases by a factor of 2.[30]

(6) It is important to distinguish between dynamical and thermodynamical stabilization of finite aggregates. The dynamical stabilization can be always provided by thick enough protecting surface layers of soluble P units (P blocks). However, these layers as such do not ensure at all that finite aggregates are thermodynamically more favorable than other structures (for example, infinite plain layers, or inverse morphologies). Hence, finite aggregates could be dynamically stable, but unstable thermodynamically. It is the thermodynamical stability of different aggregated structures that was considered in the present paper; that sort of stability is controlled by a subtle competition between conformational preferences of different insoluble H blocks.

(7) Turning to a biological motif (already hinted at by the term "proteinlike") we first note that protein chemical structures are normally much more complicated than those of "two-letter" irregular copolymers. Therefore, the present theory is not applicable (at least not directly) to compact protein structures. Nevertheless certain analogies are not impossible. For example, some proteins can form extended collective fibrillar structures that are possibly more favorable thermodynamically than protein globules[31] (similar behavior had been observed for synthetic peptides[32]). This phenomenon resembles transitions between finite spherical and cylindrical aggregates considered above for irregular (proteinlike) copolymers.

Acknowledgment. I acknowledge stimulating discussions with A. R. Khokhlov and P. G. Khalatur. This work was supported in part by the Alexander von Humboldt Foundation, Program for Investment in the Future (ZIP, W. Paul Award Project "Sequence Design of Functional Copolymers"), and by NATO SfP Grant 974173.

References and Notes

(1) Kinning, D. J.; Winey, K. I.; Thomas, E. L. *Macromolecules* **1988**, *21*, 3502. Halperin, A. *Macromolecules* **1991**, *24*, 1418. Lairez, D.; Adam, M.; Carton, J. P. et al. *Macromolecules* **1997**, *30*, 6798. Hamley, I. W. *The Physics of Block Copolymers*; Oxford University Press: Oxford, England, 1998. Luo, L. B.; Eisenberg, A. *Langmuir* **2001**, *17*, 6804. Bartel, C.; Shen, H.; Eisenberg, A. In *Encyclopedia of Materials: Science and Technology*; Elsevier Science: Oxford, England, 2001; p 718. Hillmyer, M. A.; Lodge, T. P. *J. Polym. Sci., Part A: Polym. Chem.* **2002**, *40*, 1. Lodge, T. P. *Macromol. Chem. Phys.* **2003**, *204*, 265.
(2) Anfinsen, C. *Science* **1973**, *181*, 223. Hecht, M. H. *Science* **1990**, *249*, 884.
(3) Leibler, L. *Makromol. Chem., Macromol. Symp.* **1988**, *16*, 1.
(4) Lifshitz, I. M.; Grosberg, A. Yu.; Khokhlov, A. R. *Rev. Mod. Phys.* **1978**, *50*, 683.
(5) Khokhlov, A. R.; Khalatur, P. G. *Physica A* **1998**, *249*, 253. Khokhlov, A. R.; Khalatur, P. G. *Phys. Rev. Lett.* **1999**, *82*, 3456.
(6) To be definite we assume an aqueous solution; in the general case H units are just poorly soluble or insoluble, while P units are soluble.
(7) Helfrich, W. *Z. Naturforsch.* **1973**, *28c*, 693.
(8) Lipowski, R. *Nature (London)* **1991**, *349*, 475.
(9) Milner, S. T.; Witten, T. A. *J. Phys. Fr.* **1988**, *49*, 1951.
(10) Here and below the thermal energy k_BT is considered as the energy unit.

(11) Govorun, E. N.; Khokhlov, A. R.; Semenov, A. N. *EPJ E*, to be published.

(12) Note that positive C_0 stems from the surface repulsion of P units which can be partially released by convex bending of the surface.

(13) Wintz, W.; Döbereiner, H.-G.; Seifert, U. *Europhysics Lett.* **1996**, *33*, 403. Seifert, U.; Berndl, K.; Lipowsky, R. *Phys. Rev. A* **1991**, *44*, 1182. Seifert, U. *Phys. Rev. Lett.* **1993**, *70*, 1335.

(14) The following idea behind this structure: for a given mass (volume, surface area) of the aggregate its surface bending energy consists of two terms corresponding to the two terms in eq 1. The second term is constant $= 4\pi K_G$ (or larger for multiconnected topologies with $K_G < 0$). The minimum of the first term points to any surface with constant optimum mean curvature, $C_1 + C_2 = C_0$. The largest possible $\tilde{C}$ compatible with this condition is $\tilde{C} = {}^2/_3$, corresponding to a system of connected spheres of "optimal" radius $R = 2/C_0$. If $\tilde{C} > {}^2/_3$, then a new element of the aggregate accommodating the excess volume must emerge; minimization of its surface energy is virtually equivalent to minimization of its surface area and hence the nearly spherical shape of the new element (with radius R_c). For a given $\tilde{C}$, the total bending energy per unit volume is then proportional to $({}^1/_2)KC_0{}^2/R_c + K_G/R_c{}^3$ if $R_c \gg 1/C_0$, suggesting a minimum at $R_c \to \infty$; the mutual arrangement of the spheres, the actually finite optimal R_c, and the overall "swollen star" structure are then defined by invoking the elastic energy of long H blocks; see below.

(15) Dubois-Violette, E., Pansu, B., Eds. *Geometry and Interfaces*; Proc. Aussois, 1990 (*J. Phys. (Paris)* **1990**, *51* (23)). Schoen, A. H. *NASA Tech. Rep.* **1970**, No. 05541.

(16) Precipitation at $\tilde{C} = 0$ is a result of the spherical (cylindrical) approximation for the Wigner−Seitz cell of an inverse structure.

(17) We assume that the *number* fraction of short blocks is small. Therefore $\mathcal{N}$, the total number of all H blocks, nearly coincides with the number of long blocks. With this condition the expression for the conformational free energy obtained below, eq 14, is asymptotically exact.

(18) A spherical aggregate is assumed; the result for D, apart from a numerical prefactor, is also applicable to other morphologies.

(19) Note that the apparent reduced surface area $\alpha^* = \alpha_3$ for high α's. Thus the region of high apparent α's is just not accessible for small B's because of the effect of surface crumpling.

(20) Angerman, H.; Ten Brinke, G.; Erukhimovich, I. *Macromolecules* **1996**, *29*, 3255.

(21) Leibler, L.; Benoit, H. *Polymer* **1981**, *22*, 195.

(22) Pande, V. S.; Grosberg, A. Y.; Tanaka, T. *PNAS* **1994**, *91*, 12972. Pande, V. S.; Grosberg, A. Y.; Tanaka, T. *PNAS* **1994**, *91*, 12976.

(23) Peng, C.-K.; Buldyrev, S. V.; Goldberger, A. L.; Havlin, S.; Sciortino, F.; Simons, M.; Stanley, H. E. *Nature (London)* **1992**, *356*, 168.

(24) Govorun, E. N.; Ivanov, V. A.; Khokhlov, A. R.; Khalatur, P. G.; Borovinsky, A. L.; Grosberg, A. Yu. *Phys. Rev. E* **2001**, *64*, R40903.

(25) The phase diagram is discontinuous at $\theta = 0$: note a singularity in the equation defining $\tilde{\alpha}$ at $\theta = 0$. That is why there is no matching similarity between the diagrams shown in Figures 9 and 10.

(26) The term "proteinlike copolymer" is also widely used in a more specific sense assuming a certain sequence design procedure[5] and/or a certain Levy-flight sequence statistics.[24] The same term ("proteinlike") was also used in a different sense (not applicable in this paper) marking the ability of the copolymer to fold in a specific conformation (see, for example, ref. 27).

(27) Pande, V. S.; Grosberg, A. Yu.; Tanaka, T. *J. Phys. A*: *Math. Gen.* **1995**, *28*, 3657.

(28) Note that in all the cases the complete distribution is bimodal involving a lot of long blocks and some short blocks. So, the mean block length $\bar{n}$ always nearly coincides with the length of long blocks.

(29) Milner, S. T.; Witten, T. A.; Cates, M. E. *Macromolecules* **1988**, *21*, 2610.

(30) A similar analysis for H−PH copolymers shows that the spontaneous surface curvature decreases as solvent quality worsens (i.e., as τ is increased). Hence the following possible sequence of morphological transitions: homogeneous solution → "swollen star" aggregates → spheres → cylinders → layers → precipitate (see section 3).

(31) Tan, S. Y.; Pepys, M. B. *Histopathology* **1994**, *25*, 403. Cohen, F. E. *J. Mol. Biol.* **1999**, *293*, 313. Fandrich, M.; Fletcher, M. A.; Dobson, C. M. *Nature (London)* **2001**, *410*, 165.

(32) Aggeli, A.; Nyrkova, I. A.; Bell, M.; Harding, R.; Carrick, L.; McLeish, T. C. B.; Semenov, A. N.; Boden, N. *PNAS* **2001**, *98*, 11857.

MA030357Y