

Formation of globules and aggregates of DNA chains in DNA/polyethylene glycol/monovalent salt aqueous solutions

H. Kawakita,¹ T. Uneyama,^{2,3} M. Kojima,⁴ K. Morishima,⁵ Y. Masubuchi,^{1,2,a)} and H. Watanabe²

¹*Department of Applied Chemistry, Tokyo University of Agriculture and Technology, Nakacho 2-24-16, Koganei, 184-8588, Japan*

²*Institute for Chemical Research, Kyoto University, Gokasho, Uji, 611-0011, Japan*

³*JST-CREST, Gokasho, Uji, 611-0011, Japan*

⁴*Department of Material Chemistry, Graduate School of Engineering, Kyoto University, Kyoto-Daigaku Katsura, Nishikyo-ku, 615-8510, Japan*

⁵*Graduate School of Bio-Applications and Systems Engineering, Tokyo University of Agriculture and Technology, Nakacho 2-24-16, Koganei, 184-8588, Japan*

(Received 16 April 2009; accepted 11 August 2009; published online 1 September 2009)

It has been known that giant DNA shows structural transitions in aqueous solutions under the existence of counterions and other polymers. However, the mechanism of these transitions has not been fully understood. In this study, we directly observed structures of probed (dye-labeled), dilute DNA chains in unprobed DNA/polyethylene glycol (PEG)/monovalent salt (NaCl) aqueous solutions with fluorescent microscopy to examine this mechanism. Specifically, we varied the PEG molecular weight and salt concentration to investigate the effect of competition between the depletion and electrostatic interactions on the coil-globule transition and the aggregate formation. It was found that the globules coexist with the aggregates when the unprobed DNA chains have a concentration higher than their overlap concentration. We discuss the stability of the observed structures on the basis of a free energy model incorporating the attractive depletion energy, the repulsive electrostatic energy, and the chain bending energy. This model suggested that both of the globules and aggregates are more stable than the random coil at high salt concentrations/under existence of PEG and the transition occurs when the depletion interaction overwhelms the electrostatic interaction. However, the coexistence of the globule and aggregate was not deduced from the thermodynamic model, suggesting a nonequilibrium aspect of the DNA solution and metastabilities of these structures. Thus, the population ratio of globules and aggregates was also analyzed on the basis of a kinetic model. The analysis suggested that the depletion interaction dominates this ratio, rationalizing the coexistence of globules and aggregates. © 2009 American Institute of Physics. [doi:10.1063/1.3216110]

I. INTRODUCTION

It has been reported that a giant DNA chain shows structural transitions in aqueous solutions in the presence of ions and other polymers.¹⁻⁵ In dilute DNA solutions containing polyethylene glycol (PEG) and monovalent salt, a random coiled DNA chain spontaneously changes into a compact globule form when the PEG and/or salt concentrations are larger than certain critical values.⁶⁻⁹ This phenomenon, known as the coil-globule transition,¹⁰ occurs in the presence of various condensing agents (besides PEG and monovalent salt) such as multivalent cations^{11,12} and cationic surfactants.¹³ On the other hand, in concentrated DNA solutions, DNA chains aggregate to form elongated structures in which the chains are packed in parallel.^{3,14-16} This structure is considered to be similar to the liquid crystalline structure, as suggested from a fact that short fragmented DNA chains form liquid crystalline phase under conditions of the aggregate formation of long DNA chains.^{3,17,18}

The mechanisms of the coil-globule transition and for the aggregate formation have been discussed in literature.^{2,5,10,19} For example, Vasilevskaya *et al.*⁸ proposed a free energy model for a single DNA chain incorporating the free energy of mixing of DNA and flexible polymers (e.g., PEG), the elastic free energy of the DNA chain, and the translational entropy of small ions. In this model, the transition is explained as a result of the competition of the destabilization due to the elastic free energy with the stabilization due to the translational entropy and the free energy of mixing. It should be noted that Vasilevskaya's model assumed screening of the electrostatic interaction for DNA by the counterions to neglect this interaction. In contrast, Ubbink and Odijk²⁰ proposed another free energy model incorporating the electrostatic interaction and showed that this interaction plays an important role in the coil-globule transition. Indeed Ou and Muthukumar²¹ performed Brownian dynamics simulations based on a similar model and showed that the electrostatic interaction is essential in particular to explain the effects of the valence and concentration of ions. Kojima *et al.*¹⁶ proposed that the transitions from random coils to

^{a)}Author to whom correspondence should be addressed. Electronic mail: mas@scl.kyoto-u.ac.jp.

globules and to aggregates occur as a result of competition between the depletion interaction^{21–25} and the chain bending elasticity. However, Kojima's model neglected the electrostatic interaction, as similar to the Vasilevskaya's model. Thus, no model incorporating both of the electrostatic and the depletion interactions has been proposed so far, though it is natural to consider a role of competition of these interactions in the DNA transitions.

Focusing on this competition, this study performs direct, fluorescent microscopy observation for probed (dye-labeled), dilute DNA in the presence of unprobed DNA, PEG, and NaCl in buffer solutions. Specifically, we examine the coil-globule transition and the aggregate formation in a wide range of the unprobed DNA concentration by utilizing two PEG samples of different molecular weights (to tune the depletion interaction) and also varying the NaCl concentration (to tune the electrostatic interaction). We propose a new thermodynamic model to show that the stabilities of the structures are determined by the competition of the depletion and the electrostatic interactions. We also discuss the population of each structure on the basis of a kinetic model (because the full equilibration of the population did not appear to be achieved).

II. EXPERIMENTS

A. Preparation of probed DNA

Probed DNAs, labeled with YOYO-1 (quinolinium, 1, 19-(1,3-propanediylbis-[(dimethyliminio)-3,1-propanediyl]) bis [4-([3-methyl-2(3H)-benzoxazolylidene]-methyl)]-tetraiodide obtained from Molecular Probes), were prepared by dropwisely adding a dimethylsulfoxide (DMSO) solution of YOYO-1 into a DNA aqueous solution (Bacteriophage T4 GT7 DNA, Nippon Gene). The labeling reaction was completed after overnight settlement at 4 °C. The molar ratio of the probed DNA base pairs to the fluorescent YOYO-1 dye was about 5/1.^{26,27} 2-mercaptoethanol (Nakarai Tesuque) and antioxidant (Glucose Oxidase, D-(+)-glucose and catalase, Tokyo Chemical) were added to the solutions as antifaders that eliminate oxygen radicals to suppress the photobleaching of the fluorescent dye.²⁸ In the stock solution of the probed (YOYO-1 labeled) DNA, the DNA concentration was 3.5 μ Mbp and the concentrations of the antifader components were 3 vol % for 2-mercaptoethanol, 50 μ g/ml for glucose oxidase, and 5 μ g/ml for D-(+)-glucose, and 5 μ g/ml for catalase.

B. Preparation of sample solutions for fluorescent microscopy

The DNA solutions for fluorescent microscopy observation were prepared by gently mixing a small volume of the stock solution of probed (dye-labeled) DNA with solutions of unprobed DNA of known concentration in TE buffer (containing tris(hydroxymethyl)aminomethane 100mM/ ethylenediaminetetraacetic acid 10mM; Wako Pure Chemical). Microtubes were used for this mixing. Separately prepared aqueous solutions of two PEG samples, PEG6000 and PEG20000 (average molecular weights $M_{\text{PEG}}=8.2$ and 21 k;

Kishida Chemical), and NaCl (Kanto Chemical) were also added to these solutions of probed and unprobed DNA.

In all solutions subjected to fluorescent observation, the concentrations of the probed DNA and PEG were fixed at 7.0×10^{-2} μ Mbp and 230 mg/ml, respectively. These probed DNA chains were dilute and not mutually overlapping. It should be noted that our antifader concentration was lower than that in the earlier studies^{29,30} but gave a sufficient antifading effect. The salt and unprobed DNA concentrations were varied independently to examine the effects of the electrostatic interaction and overlapping of DNA chains on the structural transitions. In one series of the solutions, the salt concentration (equal to the sum of the concentrations of NaCl and salt in the TE buffer solution⁸) was varied in the range from 9.6×10^{-1} to 1.0×10^3 mM while keeping the unprobed DNA concentration at 31 μ Mbp (i.e., in the semi-dilute overlapping regime). In the other series of solutions, the unprobed DNA concentration was varied in the range from 7.0×10^{-2} to 2.5×10^2 μ Mbp while keeping the salt concentration at 120 mM where the electrostatic interaction was well screened.

C. Fluorescent microscopy observation

The sample solutions prepared as above were injected into an observation cell made of polydimethylsiloxane (PDMS:SILPOT 184, Dow Corning Toray) and microcover glass (Matsunami Glass, size: 25×60 mm², thickness: 0.12–17 mm). The cell height was 0.34 mm and the sample solution volume was 20 μ l. Prior to the sample injection, the PDMS cells were conditioned/equilibrated with distilled water in order to avoid adsorption and/or evaporation of water from the solution. In addition the observation was performed under saturated water vapor.

The probed DNA chains in the solutions were observed at 25 °C (room temperature) with a fluorescent microscope, (ix-71, Olympus) with UPLSAP 100 \times (numerical aperture = 1.4) oil immersed objective lens. Olympus U-IBW3 mirror unit and the high-resolution electron bombarded CCD camera (C7190-43, Hamamatsu) were used.

The fluorescent images of probed DNAs thus obtained were processed by an image processor, Hamamatsu Argus 10. For each sample solution, the observation was performed over 10 min at least. (10 min were sufficient for the DNA chains to form steady structures after the injection of sample solutions.) The observation was focused on DNA chains located at ~ 50 μ m away from the glass surface in order to avoid any effect of the surface-DNA interaction on the DNA conformation/structure. Image-J developed by National Institutes of Health was utilized to achieve the conformational measurement explained below.

Ten snapshots were taken for individual, probed DNA chains to calculate their average dimension R_L , whereas the data obtained for 50 chains were utilized to evaluate a distribution of the DNA dimension. For the DNA chains in the random coil and globule states, the dimension was estimated as the length of the longer principal axis of the two-dimensional (2D) inertia tensor obtained from imaging. For the elongated DNA aggregates, the dimension was evaluated

as the distance between the most separated imaging pixels having the signal from DNA. Although the actual size of DNA should be somewhat smaller than the size evaluated from a fluorescent image due to the blurring effect,³¹ we simply use the latter size in this study.

III. RESULTS

A. Fluorescent images of DNA

Figures 1(a)–1(c), respectively, show typical snapshots of fluorescent images of the probed DNA in the random coiled, globular, and elongated/aggregated states, and Figs. 1(d)–1(f) show illustrations of respective states. The salt concentration C_{Na^+} was 0.96 mM for (a) and 120 mM for (b) and (c). For all of (a)–(c), the concentration of the unprobed DNA was 31 μMbp , which corresponded to a number ratio of 1/440 between the probed and unprobed DNA chains. Images (b) and (c) were obtained from the same solution; namely, the globules and aggregates coexisted in the solution. Minagawa *et al.*⁶ evaluated the dimensions of random coiled and globular DNA as the length of the longer principal axis of the 2D inertia tensor of the fluorescent images (i.e., with the same method as we used) and reported that the dimension of T4 DNA is 2.0–5.5 and 0.4–1.5 μm in the random coil and globule states, respectively. For the elongated aggregates, Kojima *et al.*¹⁶ reported the dimension ~ 40 μm . Our observation is consistent with these earlier results.

B. Dependence of DNA dimension on salt concentration

Figure 2 shows dependence of the average DNA dimension R_L on the salt concentration in the solutions containing (a) PEG6000 and (b) PEG20000. These solutions commonly had the PEG concentration of 230 mg/ml and the unprobed DNA concentration of 31 μMbp .

We note that R_L is about 3 μm and thus the probed DNA in the semidilute matrix of unprobed DNA has the random coil conformation at low salt concentrations, below

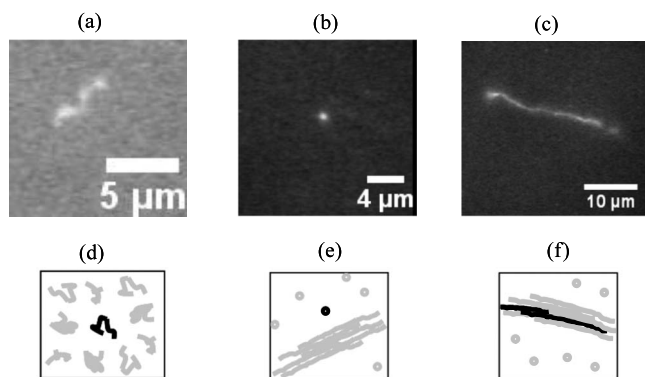


FIG. 1. Typical snapshots of the probed DNA in the solutions containing unprobed DNA at 31 μMbp with various salt concentrations. (a) A random coiled DNA at $C_{\text{Na}^+}=0.96$ mM, (b) a globular DNA at $C_{\text{Na}^+}=120$ mM, and (c) an elongated DNA at $C_{\text{Na}^+}=120$ mM. (d)–(f) show illustrations of (a)–(c).

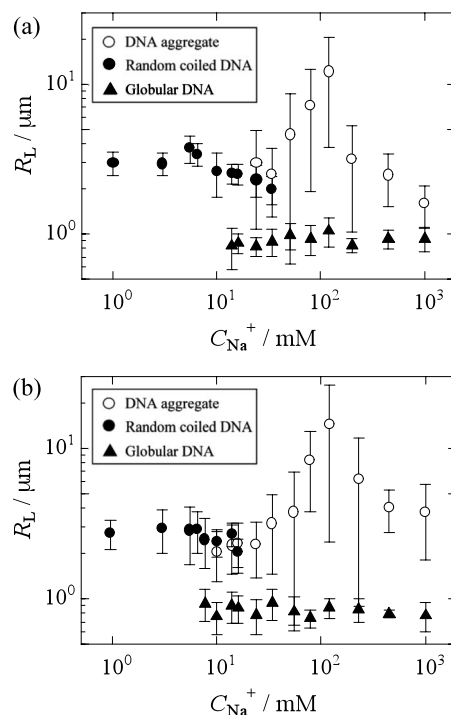


FIG. 2. The average dimension R_L of DNA plotted against the salt concentration C_{Na^+} for the solutions containing (a) PEG6000 and (b) PEG20000. Filled circles, filled triangles, and open circles represent random coiled, globular, and elongated DNA, respectively.

14 mM in Fig. 2(a) and below 7.7 mM in Fig. 2(b). In a certain range of salt concentration, between 14–34 mM in Fig. 2(a) and between 7.7–16 mM in Fig. 2(b), the globule and random coil coexisted to give the double-valued dimension, $R_L \cong 0.8$ and 3 μm . At the salt concentration slightly higher than this coil-globule coexistence regime, the elongated aggregates also coexisted and R_L was triple valued. However, this triple-coexistence window was narrow and the random coil state vanished, and R_L became double valued (for the globules and aggregates) at higher salt concentrations, >34 mM in Fig. 2(a) and >16 mM in Fig. 2(b). In the fluorescent image traced for at least 10 min, no structural evolution (from aggregate to globule and *vice versa*) was observed, suggesting that both of these structures are considerably stable under the experimental conditions examined. The observed structural behavior is similar to that reported for DNA in spermidine solutions.¹⁵

Figure 2 gives an estimate of the critical salt concentrations for the coil-globule transition, 14 and 7.7 mM in solutions (a) and (b) containing PEG6000 and PEG20000, respectively. These estimates may not be very accurate and somewhat smaller than the real critical concentrations because of a difficulty of observing a random coiled DNA in a certain range of salt concentration (10–100 mM) due to contamination of unbound dye (that reduces fluorescent contrast).³² However, we are only interested in the changes in the structural transition point with the experimental parameters such as the salt and DNA concentrations and PEG molecular weight. Thus, the rough estimates mentioned above are still fine for our purpose.

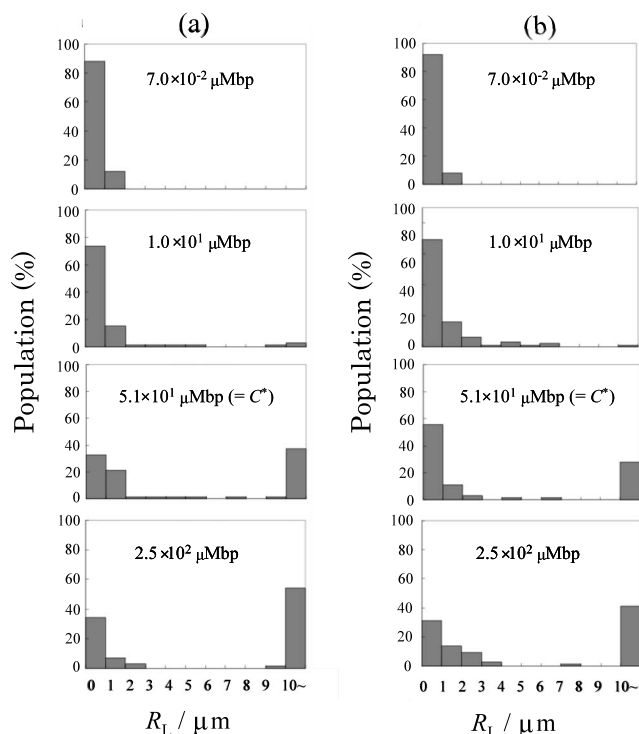


FIG. 3. Distribution of the DNA dimension at various DNA concentrations C_{DNA} (indicated in the figure) for the solutions containing (a) PEG6000 and (b) PEG20000.

C. Dependence of DNA dimension on DNA concentration

Figure 3 shows the distribution of the probed DNA dimension at various unprobed DNA concentrations in the high salt concentration regime (at $C_{\text{Na}^+} = 120$ mM) where the globules and aggregates coexisted (see Fig. 2). Up to a certain critical DNA concentration, the DNA chains are mostly in the globular state and hardly form the aggregates. We note that this critical concentration of aggregate formation is close to the overlapping concentration of T4 DNA in aqueous solutions, $C^* = 51 \mu\text{Mbp}$.³³ Interestingly, globules do not disappear even if we increase the DNA concentration up to $250 \mu\text{Mbp}$ ($\cong 5C^*$). From Figs. 3(a) and 3(b), we also note that the globules are more populated in the solution (b) containing PEG of the higher molecular weight.

IV. DISCUSSION

A. Stability of observed structures: Thermodynamic argument

In this section, we discuss the stabilities and coexistence of the globules and aggregates of the DNA chains on the basis of a simple thermodynamic model based on semiflexible chain description, where DNA chain is regarded as consecutive persistent rods rather than freely jointed segments. Since our discussion is for the regime where the random coil is not observed, we shall focus on the globules and aggregates and consider the random coil just as a reference for the globules/aggregates. It is expected that the conformational entropy is the most dominant contribution to the free energy where the random coil exists as a thermodynamically stable

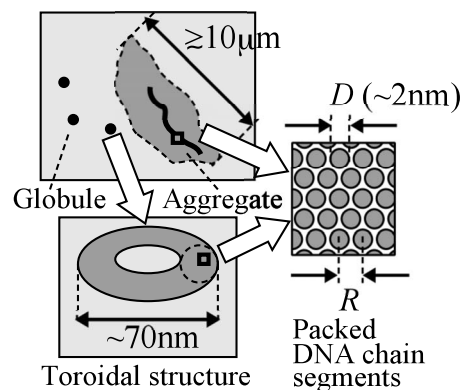


FIG. 4. Illustration of microscopic structures for globular DNAs and DNA aggregates considered in the free energy model.

structure. However, it will be less important in the current situation where the random coil is not stable because of the other contributions. For example, the entropic contribution *per unit length of DNA* (1 nm) is of the order of $10^{-2} k_B T$ (k_B is the Boltzmann constant and T is the absolute temperature) as estimated from the Grosberg–Khokhlov type elastic energy model,¹⁰ whereas the depletion interaction and the electrostatic interaction per unit length are estimated to be of the order of several $k_B T$ as shown later. Thus the contribution of the conformational entropy should be sufficiently small compared to others and is neglected in the following argument. The translational entropies of a DNA chain are of the order of $10^{-3} k_B T$ *per unit length* (order of $k_B T$ per chain) and also neglected in our analysis. The effect of the fluorescent dye, YOYO-1, is taken account only as a decrease in effective charge density of DNA (see Appendix A). We assume that the effect of solvent^{34,35} is similar for globules and aggregates and do not consider it explicitly.

Thus, we model the free energy per DNA chain in the system as

$$F_{\text{tot}} = F_{\text{dep}} + F_{\text{elec}} + F_{\text{bend}}, \quad (1)$$

where F_{dep} , F_{elec} , and F_{bend} are the depletion interaction energy, the electrostatic interaction energy, and the bending energy,

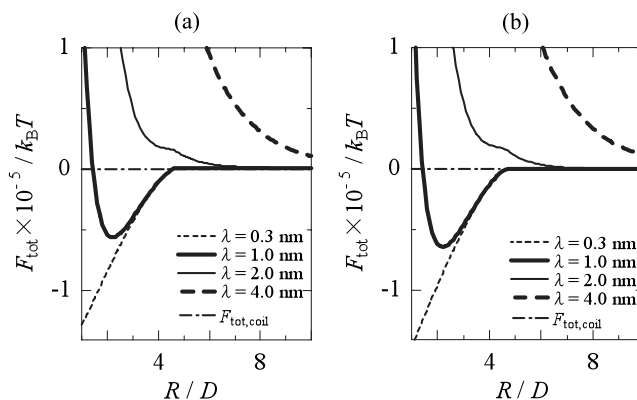


FIG. 5. The total free energy per chain F_{tot} in the presence of PEG6000 as a function of the distance between DNA segments for (a) globular DNA (isolated) and (b) elongated DNA (in the aggregate) for various Debye lengths (λ). Horizontal broken line shows the free energy of random coiled DNA for comparison.

ergy, respectively. For calculation of each energy term, we further model a globular DNA as a toroid in which a single DNA chain is wound and its sections are packed hexagonally in the toroidal cross section, as illustrated in Fig. 4.³⁶ We also model the DNA aggregates as a liquid-crystalline like bundle of stretched DNA chains.¹⁷

On the basis of the above structural modeling, we can calculate depletion energy F_{dep} with the aid of the Asakura–

Oosawa theory.^{22–24} We regard each persistence section of the DNA chain as a rigid rod of a diameter D and the PEG chains as rigid spheres having a diameter $d=2R_{G,\text{PEG}}$, with $R_{G,\text{PEG}}$ being the root-mean-square radius of gyration of PEG. (In other words, for calculation of F_{dep} , we regard even the globule of DNA locally as a sequence of rods.) Then, as shown in Appendix B, F_{dep} can be written as a function of distance between the nearest neighbor DNA chains, R :

$$\frac{F_{\text{dep}}}{k_B T} = \begin{cases} -\frac{L}{2}\rho(D+d)^2 \left[\frac{\pi}{2} - \frac{R}{D+d} \sqrt{1 - \left(\frac{R}{D+d}\right)^2} - \sin^{-1}\left(\frac{R}{D+d}\right) \right] & (D < R \leq D+d) \\ 0 & (D+d < R) \end{cases}, \quad (2)$$

Here, L is the contour length of DNA and ρ is the number density of PEG. In Eq. (2), we find that the depletion energy is always negative for $R < D+d$ and it thus stabilizes the compact structures (globules and aggregates). The strength of the depletion interaction is dependent on ρ and d and thus determined by the concentration and molecular weight of PEG.

The electrostatic interaction can be calculated from the linearized Poisson–Boltzmann equation. We again regard local sections of DNA chain as rigid rods having the diameter D (less than the distance R between the DNA chains) and the interaction is considered only between two neighboring persistent rods although rods are packed rather closely in the globule or the aggregate. This assumption is reasonable at the salt concentrations examined; the estimated Debye length is from 0.30 to 4.0 nm, which is comparable to the diameter of the DNA double helix (2 nm) and much smaller than the diameter of the toroidal DNA globule (70 nm).³⁷

Then, F_{elec} is expressed as^{25,38}

$$\frac{F_{\text{elec}}}{k_B T} = \frac{Q^2 L}{2\pi\epsilon_r\epsilon_0 k_B T} K_0\left(\frac{R}{\lambda}\right) \quad (3)$$

where Q is the effective line charge density of DNA, ϵ_r is the relative dielectric constant of the solvent ($\epsilon_r=78.54$ for water), ϵ_0 ($=8.85 \times 10^{-12} \text{ J}^{-1} \text{ C}^2 \text{ m}^{-1}$) is the absolute dielectric permittivity of vacuum, λ is the Debye length, K_0 is the zeroth order modified Bessel function of the second kind. As seen from Eq. (3), the electrostatic interaction energy is always positive. (Note that in the counterion condensation theory an attractive interaction is derived for small R region.³⁹ However the effect is not essential in the conditions of our interest for R range³⁹ and DNA concentration.⁴⁰) Thus the electrostatic interaction stabilizes the isolated random coils but destabilizes the globules and aggregates that involve crowded charges. The Debye length λ appearing in Eq. (3) changes with the ionic strength I ,

$$I = \frac{1}{2} \sum_i C_i Z_i^2, \quad (4)$$

where C_i and Z_i are the concentration and valence number of the i th ion species in the solution. Specifically, the Debye length is written in terms of I as

$$\lambda = \sqrt{\frac{\epsilon_r \epsilon_0 k_B T}{2N_A e^2 I}}, \quad (5)$$

with e being the elementary electric charge.

The bending energy F_{bend} appearing in Eq. (3) can be calculated on the basis of a semiflexible chain model⁴¹ as

$$\frac{F_{\text{bend}}}{k_B T} = \frac{1}{2} l_p \frac{L}{r^2}, \quad (6)$$

Here, l_p is the persistence length of the DNA and r is the average radius of curvature of DNA. The bending energy increases with decreasing r and thus prefers the aggregates containing elongated DNA chains.

All parameters appearing in Eqs. (1)–(6) can be evaluated/estimated experimentally. $R_{G,\text{PEG}}$ is evaluated from an empirical relationship between $R_{G,\text{PEG}}$ and PEG molecular weight⁴² as $R_{G,\text{PEG}6000}=3.7$ nm and $R_{G,\text{PEG}20000}=6.4$ nm. The Debye length λ is evaluated from the known ionic strength [cf. Eqs. (4) and (5)]. For the DNA solutions examined in this study, λ ranged from 0.30 to 4.0 nm. Considering the electrostatic screening by YOYO-1 explained earlier, we estimated the effective charge density Q on DNA as 60% of the bare charge, i.e., $Q=5.6 \times 10^{-19} \text{ C/nm}$. For the random coil of DNA utilized as a reference for the globules/aggregates, the average radius of curvature, r , is estimated to be 53 nm from atomic force microscope images of T4 DNA.⁴³ The globular DNA has $r=35$ nm, as evaluated from the electron microscopy images of T4 DNA,³⁷ and the elongated DNA chains in the aggregates should have $r \approx 0$. The contour length, persistence length, and diameter of DNA are known to be $L=57 \mu\text{m}$, $l_p=50$ nm, and $D=2$ nm, respectively.

Utilizing the above parameter values in Eqs.(1)–(6) (together with $T=298$ K), we calculated the free energy per

chain. For simplicity, we assume that a random coil DNA has an expanded conformation and its segments are locally so scarce that the depletion and the electrostatic interaction are negligible. Then, the free energy per random coil can be simply given by $F_{\text{tot,coil}} = F_{\text{bend,coil}}$, which does not change with the interchain distance R . For the solution containing PEG6000, Fig. 5 shows plots of the free energy thus obtained for (a) single globular DNA and (b) single DNA in the aggregate. The free energy is plotted against the distance between DNA chains and the Debye length is varied in the experimental range explained earlier. As noted in Fig. 5, F_{tot} has the minima if the Debye length λ is smaller than a critical value λ^* ($=1.6$ nm found from comparison of the F_{tot} curves for finely varied λ not shown here). This result suggests that the globule and/or the aggregate are the thermodynamically stable for $\lambda < \lambda^*$. The minima of F_{tot} for the globule and aggregates are much deeper than $F_{\text{tot,coil}}$ for the random coil. Namely, $|F_{\text{tot,min}} - F_{\text{tot,coil}}|$ is of the order of $10^4 k_B T$, indicating that the random coil state is never achieved through thermal activation when the globule and aggregate having such a deep minimum are formed.

The above result suggests the structural transitions from the random coil to the globule and aggregate are mainly driven by the competition between the depletion and electrostatic interactions. Qualitatively speaking, the transition occurs when the depletion interaction overwhelms the electrostatic interaction and this situation can be realized through the electrostatic screening due to salt (and partly to YOYO-1). The change in the transition point with the PEG molecular weight M_{PEG} can be understood similarly. Our model shows that an increase in M_{PEG} enlarges the depletion energy F_{dep} thereby allowing the transition to occur at a lower salt concentration. This molecular scenario is consistent with the experimental observation in Fig. 2.

Despite the above success of our model, we note a serious problem. As noted in Fig. 5, the minimum of F_{tot} is much deeper, by a factor of $10^3 k_B T$, for the single chain in the aggregates than the single-chain globule. This result means that the DNA aggregate is the most stable structure deduced from our model and the globules should be converted to the aggregates at equilibrium. (Note that the coil state is unstable at high salt and/or PEG concentrations in our model.) Since we observed the coexistence of the globules and aggregates, the equilibrium has not been achieved in our solutions (i.e., the globules survive as a metastable structure) because of a high activation barrier for the transition from the globule to the aggregate, or, some important free energy term is still missed in our model. We adopt the former viewpoint and further examine the kinetics of the globule/aggregate formation in the following section.

B. Structure formation kinetics

Here, the coexistence of globules and aggregates observed in our experiments is discussed on the basis of a simple kinetic model. We assume that a DNA chain can take only three structures, the random coil, the globule, and an elongated chain in the aggregate. We further assume that the structural transition is limited to the following three pro-

cesses; (a) formation of a globule from a random coil, (b) formation of an aggregate from two random coils, and (c) growth of an aggregate through incorporation of a random coil into the aggregate. We consider no reverse processes for these (a)–(c) processes because of the very large free energy difference between the coil and globule/aggregate discussed for Fig. 5. We also consider that no aggregate formation/growth occurs through either fusion of two globules or fusion of a globule into an aggregate, because of a high activation barrier for the fusion process that is naturally expected from the large surface charge⁴⁴ and thick hydration layer⁷ of the globule (as well as the aggregate). Namely, we consider that the aggregates grow only through the coil-aggregate collision and successive fusion. We do not distinguish large and small aggregates and regard them as equivalent aggregates in the kinetic analysis, as in the case of usual analysis of propagation kinetics of anionic polymerization.⁴⁵

From the above assumptions/considerations, the kinetic equations for the molar concentrations of the random coil, globule, and aggregate, $n_c(t)$, $n_g(t)$, and $n_a(t)$, can be written as

$$\frac{dn_c(t)}{dt} = -k_g n_c(t) - k_a n_c^2(t) - k'_a(m_a(t)) n_c(t) n_a(t), \quad (7)$$

$$\frac{dn_g(t)}{dt} = k_g n_c(t), \quad (8)$$

$$\frac{dn_a(t)}{dt} = k_a n_c^2(t), \quad (9)$$

where $m_a(t)$ is the average aggregation number of DNA aggregates at time t , k_g , k_a , and k'_a are the rate constant for the coil to globule, coils to aggregate, and aggregate growth processes, respectively. [In relation to Eq. (9), we should note that $n_a(t)$ does not change on fusion of a coil into an aggregate.] In principle, the rate constants are determined from molecular level transition kinetics. Although there have been many studies on the transition kinetics especially on coil-globule transition,^{46–49} we do not discuss the transition process in detail and regard the rate constants as phenomenological parameters. We do not need molecular expressions of the rate constant for our current purpose, because we focus only on steady state mass fractions as shown later. In addition, we do not consider the two-stage process in coil-globule transition⁵⁰ for simplicity. The average aggregation number, $m_a(t)$, is determined according to the mass conservation,

$$n_c(t) + n_g(t) + m_a(t) n_a(t) = n_{\text{DNA}}, \quad (10)$$

with n_{DNA} being the time-independent, total molar concentration of DNA chains in the system.

It should be emphasized that the rate constant $k'_a(m_a(t))$ appearing in Eq. (7) increases with $m_a(t)$ because the collision of the coil and aggregate occurs more frequently to accelerate the aggregate growth at longer t where the aggregate grows to have a larger surface area available the collision/fusion. However, an explicit functional form of $k'_a(m_a(t))$ is not known. For simplicity, we assume

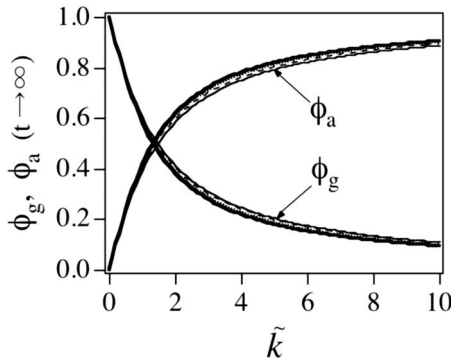


FIG. 6. The steady state solutions of ϕ_g and ϕ_a ($\bar{t} \rightarrow \infty$) plotted against \tilde{k} . Bold, dotted, broken, and thin lines indicate the solutions for $\alpha=1$, $2/3$, $1/3$, and 0 , respectively.

$$k'_a(m_a) = m_a^{-\alpha} k_a, \quad (11)$$

where α is a parameter determined from the ratio of the number of coil-fusion sites per aggregate to the number of DNA chains per aggregate. If this number of the coil-fusion sites is constant, we have $\alpha=0$, which is the lowest limit. Another extreme is $\alpha=1$, which corresponds to a situation that the surface of *all* stretched DNA chains in the aggregate, including the surface of an *internal* chain not located at the aggregate surface, is available for fusion of the random-coil DNA into the aggregate. In realistic cases (for which the internal DNA chains do not behave as the fusion site), we expect $0 < \alpha < 1$. For instance, if the aggregate isotropically grows through the coil-fusion only at the aggregate surface, we have $\alpha=2/3$.

All DNA chains are considered to be in the random coil state at $t=0$, giving the initial condition for Eqs. (7)–(10):

$$n_c(0) = n_{\text{DNA}}, \quad n_g(0) = 0, \quad n_a(0) = 0. \quad (12)$$

It is useful to introduce mass fractions of the coil, globule, and aggregate, $\phi_c \equiv n_c/n_{\text{DNA}}$, $\phi_g \equiv n_g/n_{\text{DNA}}$, and $\phi_a \equiv n_a/n_{\text{DNA}}$, and rewrite Eqs. (7)–(10) as

$$\frac{d\phi_c(\bar{t})}{d\bar{t}} = -\phi_c(\bar{t}) - \tilde{k}\phi_c^2(\bar{t}) - \tilde{k}\phi_c(\bar{t})\phi_a^\alpha(\bar{t})\psi_a^{1-\alpha}(\bar{t}), \quad (13)$$

$$\frac{d\phi_g(\bar{t})}{d\bar{t}} = \phi_c(\bar{t}), \quad (14)$$

$$\frac{d\psi_a(\bar{t})}{d\bar{t}} = \tilde{k}\phi_c^2(\bar{t}), \quad (15)$$

$$\phi_c(\bar{t}) + \phi_g(\bar{t}) + \phi_a(\bar{t}) = 1. \quad (16)$$

Here, $\bar{t} \equiv k_g t$ is the dimensionless time, $\tilde{k} \equiv n_{\text{DNA}} k_a / k_g = C_{\text{DNA}} k_a / M_{\text{DNA}} k_g$ (M_{DNA} is the number of base pairs per DNA chain) is a dimensionless rate constant, and $\psi_a \equiv n_a / n_{\text{DNA}}$ is an auxiliary variable. Figure 6 shows numerical solutions of Eqs. (13)–(16) for the limit of $t \rightarrow \infty$ (steady state solution) as functions of \tilde{k} for some α values. We note that α does not significantly affect the calculated ϕ_g and ϕ_a . Thus, we hereafter limit ourselves for the case of $\alpha=1$ where we

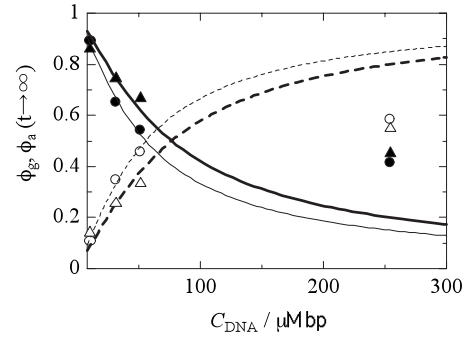


FIG. 7. The number fraction of DNA globules (ϕ_g) and aggregates (ϕ_a) plotted against the DNA concentration C_{DNA} . Filled and unfilled symbols are experimental values of ϕ_g and ϕ_a for PEG6000 (circles) and PEG20000 (triangles), respectively. Solid and broken curves are the steady state solution for ϕ_g and ϕ_a obtained by the kinetic model at $\alpha=1$ for PEG6000 (thin curves) and PEG20000 (bold curves).

find the following analytical expressions of ϕ (cf. Appendix C):

$$\phi_c(\bar{t} \rightarrow \infty) = 0, \quad (17)$$

$$\phi_g(\bar{t} \rightarrow \infty) = \frac{1 + \tilde{k} - \sqrt{1 + \tilde{k}^2}}{\tilde{k}}, \quad (18)$$

$$\phi_a(\bar{t} \rightarrow \infty) = \frac{-1 + \sqrt{1 + \tilde{k}^2}}{\tilde{k}}. \quad (19)$$

As noted from Eqs. (17)–(19), the steady state mass fractions ϕ depend only on $\tilde{k} = C_{\text{DNA}} k_a / M_{\text{DNA}} k_g$ and thus change only with the DNA concentration C_{DNA} if the salt concentration is kept constant. Specifically, ϕ_a of the aggregates increases with increasing C_{DNA} .

In Fig. 7, we attempted to fit the experimental ϕ data (symbols) with Eqs. (18) and (19) to estimate the relative rate constant, $\tilde{K} \equiv \tilde{k} / C_{\text{DNA}} = k_a / M_{\text{DNA}} k_g$. The fitting curves, with $\tilde{K} = 4.0 \times 10^9$ and 2.9×10^9 Mbp^{-1} for the DNA solutions containing PEG6000 and PEG20000, respectively, are reasonably close to the data except for the highest C_{DNA} , suggesting that the kinetic model formulated here captures an essential feature of the globule/aggregate formation.

The aggregate formation would be affected by the depletion interaction as well as the coil diffusion toward the aggregate, while the globule formation is affected only by this interaction. Thus, we expect that the depletion interaction accelerates the globule formation more significantly than the aggregate formation (in particular when the coil diffusion is slow). For this case, the rate constant k_g for the globule formation increases with the PEG molecular weight M_{PEG} more significantly compared to the rate constant for the aggregate formation k_a , which leads us to expect that $\tilde{K} = k_a / M_{\text{DNA}} k_g$ is larger for smaller M_{PEG} . This expectation, being in harmony with the thermodynamic model [Eqs. (1)–(6)], is in accord with the above observation ($\tilde{K}_{\text{PEG6000}} > \tilde{K}_{\text{PEG20000}}$), suggesting again that our kinetic model captures an essential feature of the globule/aggregate formation. At the same time, we should note that the model has been formulated under several

assumptions explained earlier and a quantitative investigation of the rate constants requires analysis of the time evolution of ϕ deduced from a improved model (free from those assumptions). This investigation is now being made and the results will be reported in our future work.

V. CONCLUSION

In this study, we utilized fluorescent microscopy to directly observe the structures of probed DNA chains in unprobed DNA/PEG/monovalent salt (NaCl) aqueous solutions. Specifically, we investigated the effect of competition between the depletion and electrostatic interactions on the coil-globule transition and the aggregate formation by varying PEG molecular weight M_{PEG} , DNA concentration C_{DNA} , and salt concentration C_{Na^+} . It was found that the critical C_{Na^+} value for coil-globule and coil-aggregate transitions decreases with increasing M_{PEG} and the globules coexist with aggregates when C_{DNA} is higher than the overlapping concentration of DNA chains.

For understanding these experimental results, we formulated the simple free energy and kinetic models. We modeled the free energy per single chain as the sum of the depletion energy, the electrostatic energy, and the bending energy. This model suggested that the transition from a random coil to a globule or to an elongated chain in the aggregate is governed by the competition of the depletion interaction and electrostatic interactions (the transition occurs when the depletion interaction overwhelms the electrostatic interaction). Experimental results such as the shift in the transition point were described well by the model. However, the above thermodynamic model cannot explain the experimental fact that globules exist with the aggregates stably at high C_{Na^+} . Thus, we also formulated a nonequilibrium kinetic model for the globule/aggregate formation. This model described the observed population ratio of the globules and aggregates, which suggests that the DNA solutions were in a metastable, non-equilibrium state thereby rationalizing the coexistence of the globules and aggregates.

APPENDIX A: THE EFFECT OF YOYO-1

We examine whether the YOYO-1 dye molecules have some effects on the stability of the globule or the aggregate. These molecules are positively charged and thus strongly interact with the negatively charged DNA chains. Consequently, the effective charge of DNA is reduced by the YOYO-1 molecules to a level of $\sim 60\%$ of the bare charge,^{26,51} which decreases the electrostatic repulsion and tend to stabilize the globules and aggregates. At the same time, YOYO-1 forms a complex with DNA to harden the DNA chain and destabilizes the globule when the density of YOYO-1 is high.⁵² However, the YOYO-1 density in this study was rather low (1/5 in the molar ratio of DNA base pair to the dye) and this hardening effect seems to be negligible compared to the electrostatic and depletion effects explained in the main text. Thus, in our modeling, we consider that YOYO-1 molecules affect only the charge density of the DNA chains.

APPENDIX B: DEPLETION INTERACTION IN TWO-DIMENSIONAL SYSTEMS

In this appendix, we derive the depletion interaction energy between two DNA segments based on the Asakura-Oosawa theory.²²⁻²⁴ We model the DNA and PEG chains as a rigid cylinder and a rigid sphere having the diameters D and d , respectively. The depletion interaction energy between DNA segments can be estimated as that between two infinitely long parallel cylinders. The energy per unit length of the cylinder, u_{dep} , can be expressed as

$$u_{\text{dep}}(R) = -\Pi v_{\text{dep}}(R), \quad (\text{B1})$$

where R ($>D$) is the distance between the centers of the cylinders, Π is the osmotic pressure, and $v_{\text{dep}}(R)$ is the volume of the depletion region per unit length of the cylinder. This $v_{\text{dep}}(R)$ can be calculated as an overlapping area of two disks (\sim overlapping volume reduced to unit length of the cylinder) as shown in Fig. 8. For $R > D+d$, there is no overlapped area and $v_{\text{dep}}(R)=0$. For $R \leq D+d$, we find

$$v_{\text{dep}}(R) = 4 \int_{R/2}^{(D+d)/2} \sqrt{\left(\frac{D+d}{2}\right)^2 - r^2} dr = \frac{(D+d)^2}{2} \left[\frac{\pi}{2} - \frac{R}{D+d} \sqrt{1 - \left(\frac{R}{D+d}\right)^2} - \sin^{-1}\left(\frac{R}{D+d}\right) \right]. \quad (\text{B2})$$

Substituting Eq. (B2) into Eq. (B1), we have the following expression for the depletion interaction energy per unit length:

$$u_{\text{dep}}(R) = \begin{cases} -\frac{\Pi(D+d)^2}{2} \left[\frac{\pi}{2} - \frac{R}{D+d} \sqrt{1 - \left(\frac{R}{D+d}\right)^2} - \sin^{-1}\left(\frac{R}{D+d}\right) \right] & (D < R \leq D+d) \\ 0 & (D+d < R) \end{cases}. \quad (\text{B3})$$

Assuming the van't Hoff's relation $\Pi = \rho k_B T$ (which corresponds to our treatment of the PEG chains as hard spheres), we obtain Eq. (2).

APPENDIX C: STEADY STATE SOLUTION FOR KINETIC MODEL

In this appendix, we calculate the steady state solution of Eqs. (13)–(16) for $\alpha=1$. In Eqs. (13)–(16) we have three

independent dynamic variables in general. However, for $\alpha=1$, the number of independent dynamic variables can be reduced to two, and the analytic solution can be obtained as shown below.

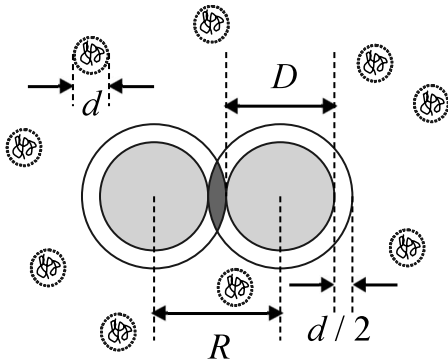


FIG. 8. Illustration of the model for the depletion interaction. Gray circles represent cylinders (DNAs) and the dark gray represents the depletion region.

From Eq. (16), Eqs. (13) and (14) can be reduced to the following set of equations:

$$\frac{d\phi_c(\bar{t})}{d\bar{t}} = -(1 + \tilde{k})\phi_c(\bar{t}) + \tilde{k}\phi_c(\bar{t})\phi_g(\bar{t}), \quad (\text{C1})$$

$$\frac{d\phi_g(\bar{t})}{d\bar{t}} = \phi_c(\bar{t}). \quad (\text{C2})$$

Equations (C1) and (C2) completely specify the time evolution of ϕ_c and ϕ_g . Substituting Eq. (C2) into Eq. (C1), we find

$$\begin{aligned} \frac{d^2\phi_g(\bar{t})}{d\bar{t}^2} &= -(1 + \tilde{k})\frac{d\phi_g(\bar{t})}{d\bar{t}} + \tilde{k}\frac{d\phi_g(\bar{t})}{d\bar{t}}\phi_g(\bar{t}) \\ &= \frac{d}{d\bar{t}} \left[-(1 + \tilde{k})\phi_g(\bar{t}) + \frac{\tilde{k}}{2}\phi_g^2(\bar{t}) \right]. \end{aligned} \quad (\text{C3})$$

Equation (C3) can be integrated to give

$$\frac{d\phi_g(\bar{t})}{d\bar{t}} = -(1 + \tilde{k})\phi_g(\bar{t}) + \frac{1}{2}\tilde{k}\phi_g^2(\bar{t}) + 1. \quad (\text{C4})$$

At the steady state [$d\phi_g(\bar{t})/d\bar{t} = 0$ at $\bar{t} \rightarrow \infty$], Eq. (C4) gives

$$\phi_g^{(\text{st})} = \phi_g(\bar{t} \rightarrow \infty) = \frac{1 + \tilde{k} - \sqrt{1 + \tilde{k}^2}}{\tilde{k}}. \quad (\text{C5})$$

From Eqs. (C2) and (C5), we find $\phi_c^{(\text{st})} = \phi_c(\bar{t} \rightarrow \infty) = 0$. Finally, from Eq. (16) we obtain

$$\phi_a^{(\text{st})} = \phi_a(\bar{t} \rightarrow \infty) = 1 - \phi_c^{(\text{st})} - \phi_g^{(\text{st})} = \frac{-1 + \sqrt{1 + \tilde{k}^2}}{\tilde{k}}. \quad (\text{C6})$$

Thus we have Eqs. (17)–(19) as the steady state solution for $\alpha=1$ at $\bar{t} \rightarrow \infty$. As mentioned in the text, the steady state solutions do not significantly depend on α and thus the above analytical solutions for $\alpha=1$ serve as a good approximation in the entire range of α between 0 and 1.

¹L. S. Lerman, *Proc. Natl. Acad. Sci. U.S.A.* **68**, 1886 (1971).

²S. Forster and M. Schmidt, *Adv. Polym. Sci.* **120**, 51 (1995).

³V. A. Bloomfield, *Curr. Opin. Struct. Biol.* **6**, 334 (1996).

⁴K. Yoshikawa, *Adv. Drug Delivery Rev.* **52**, 235 (2001).

⁵B. M. Baysal and F. E. Karasz, *Macromol. Theory Simul.* **12**, 627 (2003).

⁶K. Minagawa, Y. Matsuzawa, K. Yoshikawa, A. R. Khokhlov, and M. Doi, *Biopolymers* **34**, 555 (1994).

⁷H. Mayama, T. Iwataki, and K. Yoshikawa, *Chem. Phys. Lett.* **318**, 113 (2000).

⁸V. V. Vasilevskaya, A. R. Khokhlov, Y. Matsuzawa, and K. Yoshikawa, *J. Chem. Phys.* **102**, 6595 (1995).

⁹Y. Matsuzawa and K. Yoshikawa, *Koubunshi Ronbunshu* **52**, 753 (1995).

¹⁰A. Y. Grosberg and A. R. Khokhlov, *Statistical Physics of Macromolecules* (AIP, New York, 1994).

¹¹L. C. Gosule and J. A. Schellman, *Nature (London)* **259**, 333 (1976).

¹²V. A. Bloomfield, *Biopolymers* **44**, 269 (1997).

¹³S. M. Mel'nikov, V. G. Sergeev, and K. Yoshikawa, *J. Am. Chem. Soc.* **117**, 2401 (1995).

¹⁴R. W. Wilson and V. A. Bloomfield, *Biochemistry* **18**, 2192 (1979).

¹⁵T. Iwataki, S. Kidoaki, T. Sakaue, K. Yoshikawa, and S. S. Abramchuk, *J. Chem. Phys.* **120**, 4004 (2004).

¹⁶M. Kojima, K. Kubo, and K. Yoshikawa, *J. Chem. Phys.* **124**, 024902 (2006).

¹⁷R. L. Rill, *Proc. Natl. Acad. Sci. U.S.A.* **83**, 342 (1986).

¹⁸F. Livolant and A. Leforestier, *Prog. Polym. Sci.* **21**, 1115 (1996).

¹⁹H. H. Strey, R. Podgornik, D. C. Rau, and V. A. Parsegian, *Curr. Opin. Struct. Biol.* **8**, 309 (1998).

²⁰J. Ubbink and T. Odijk, *Biophys. J.* **68**, 54 (1995).

²¹Z. Ou and M. Muthukumar, *J. Chem. Phys.* **123**, 074905 (2005).

²²S. Asakura and F. Oosawa, *J. Chem. Phys.* **33**, 183 (1954).

²³S. Asakura and F. Oosawa, *J. Polym. Sci.* **33**, 183 (1958).

²⁴A. Vrij, *Pure Appl. Chem.* **48**, 471 (1976).

²⁵R. R. Netz and D. Andelman, *Phys. Rep.* **380**, 1 (2003).

²⁶A. N. Glazer and H. S. Rye, *Nature (London)* **359**, 859 (1992).

²⁷H. S. Rye, S. Yue, D. E. Wemmer, M. A. Quesada, R. P. Haugland, R. A. Mathies, and A. N. Glazer, *Nucleic Acids Res.* **20**, 2803 (1992).

²⁸A. Kishino and T. Yanagida, *Nature (London)* **334**, 74 (1988).

²⁹T. T. Perkins, D. E. Smith, and S. Chu, *Science* **276**, 2016 (1997).

³⁰H. Sato, Y. Masubuchi, and H. Watanabe, *J. Polym. Sci., Part B: Polym. Phys.* **47**, 1103 (2009).

³¹K. Yoshikawa and Y. Matsuzawa, *J. Am. Chem. Soc.* **118**, 929 (1996).

³²N. Makita, M. Ullner, and K. Yoshikawa, *Macromolecules* **39**, 6200 (2006).

³³R. Musti, J. L. Sikorav, D. Lairez, G. Jannink, and M. Adam, *C. R. Seances Acad. Sci., Ser.* **320**, 599 (1995).

³⁴M. P. Taylor and J. E. G. Lipson, *J. Chem. Phys.* **104**, 4835 (1996).

³⁵H. H. Gan and B. C. Eu, *J. Chem. Phys.* **109**, 2011 (1998).

³⁶N. V. Hud, M. J. Allen, K. H. Downing, J. Lee, and R. Balhorn, *Biochem. Biophys. Res. Commun.* **193**, 1347 (1993).

³⁷U. K. Laemmli, *Proc. Natl. Acad. Sci. U.S.A.* **72**, 4288 (1975).

³⁸S. L. Brenner and V. A. Parsegian, *Biophys. J.* **14**, 327 (1974).

³⁹J. Ray and G. S. Manning, *Langmuir* **10**, 2450 (1994).

⁴⁰P. Wissenburg, T. Odijk, P. Cirkel, and M. Mandel, *Macromolecules* **28**, 2315 (1995).

⁴¹V. V. Vasilevskaya, A. R. Khokhlov, S. Kidoaki, and K. Yoshikawa, *Biopolymers* **41**, 51 (1997).

⁴²S. Kawaguchi, G. Imai, J. Suzuki, A. Miyahara, T. Kitano, and K. Ito, *Polymer* **38**, 2885 (1997).

⁴³N. Yoshinaga and K. Yoshikawa, *J. Chem. Phys.* **116**, 9926 (2002).

⁴⁴Y. Yamasaki, Y. Teramoto, and K. Yoshikawa, *Biophys. J.* **80**, 2823 (2001).

⁴⁵M. Morton, *Anionic Polymerization: Principles and Practice* (Academic, New York, 1983).

⁴⁶K. Yoshikawa, *Macromol. Symp.* **106**, 367 (1996).

⁴⁷Yu. A. Kuznetsov and E. G. Timoshenko, *J. Chem. Phys.* **111**, 3744 (1999).

⁴⁸H. Noguchi and K. Yoshikawa, *J. Chem. Phys.* **113**, 854 (2000).

⁴⁹T. Sakaue and K. Yoshikawa, *J. Chem. Phys.* **117**, 6323 (2002).

⁵⁰B. Chu, Q. C. Ying, and A. Y. Grosberg, *Macromolecules* **28**, 180 (1995).

⁵¹H. Schiessel, *J. Phys.: Condens. Matter* **15**, R699 (2003).

⁵²N. Yoshinaga, T. Akitaya, and K. Yoshikawa, *Biochem. Biophys. Res. Commun.* **286**, 264 (2001).