Effect of Intrachain Mismatch on Loop-to-Coil Transition of an Associating Chain

Heng-Kwong Tsao,[†] Jeff Z. Y. Chen,[‡] and Yu-Jane Sheng*[,§]

Department of Chemical and Materials Engineering, National Central University, Chung-li, Taiwan 320, Republic of China; Department of Physics, University of Waterloo, Waterloo, Ontario, Canada N2L 3G1; and Department of Chemical Engineering, National Taiwan University, Taipei, Taiwan 106, Republic of China

Received January 4, 2003; Revised Manuscript Received May 20, 2003

ABSTRACT: The conformation of a telechelic chain, such as ssDNA with complementary bases at both ends, fluctuates from loop (closed) to coil (open) state. The transition between open and closed states can be successfully described by the two-state model. However, the mismatch between one end with complementary bases in the central region along the chain complicates the transition. We investigate the kinetics of conformational fluctuations for an associating polymer of chain length N with an attractive site at both ends (α and β) and another attractive site in the midpoint (γ) by Monte Carlo simulations. The binding energy of α with β and γ is $-\epsilon_1$ and $-\epsilon_2$, respectively. The probability of the coil state, which varies with temperature and corresponds to the melting curve in experiments, is obtained. A three-state model including open (o), closed (c), and mismatch (m) states is proposed and agrees quite well with the simulation results. It is interesting that direct transition between mismatch and closed states is observed significantly at low enough temperature. The rate constants $k_{i,j}$ from i to j state are determined as well. $k_{o,m}$, $k_{o,c}$, and $k_{m,c}$ are found to be independent of temperature but proportional to N^{-2}. On the contrary, $k_{m,o}$ and $k_{c,o}$ are independent of chain length but proportional to $e^{-\epsilon_2/k_BT}$ and $e^{-\epsilon_1/k_BT}$, respectively. Nevertheless, the transition from closed to mismatch states $k_{c,m}$ follows $N^{-2}e^{-(\epsilon_1-\epsilon_2)/k_BT}$. For a long enough chain with $\epsilon_1 > \epsilon_2$, the two-state model is adequate since the mismatch state becomes insignificant.

I. Introduction

An associating chain contains a small number of active groups, such as hydrophobic or ionic groups, which are attached at some points along the chain backbone. These groups may attract one another if the chain motions happen to bring them into close proximity of one another. One of the simplest examples of associating polymers is a telechelic polymer, which is a water-soluble chain with relatively small associating groups at the ends such as poly(ethylene oxide) with end-capped $C_{16}-C_{18}$ alkanes.[1] Since they can be tailored to display specific rheological properties, associating polymers are ideal systems for fundamental studies in addition to wide industrial applications like thickeners for food.[2−4] Recently, telechelic biopolymers have attracted attention and have been experimentally investigated by fluorescence measurement.[5−9] A typical example is a single-stranded DNA (ssDNA) or RNA, which is purposefully designed to be made of a homogeneous sequence such as poly(T) (polydeoxythymidines) with a few complementary bases at both ends.

A copolymer with self-complementarity in the sequence forms secondary structures. For example, ssDNA and RNA frequently form a hairpin-loop structure due to intramolecular base pairing. A hairpin loop consists of a linear polymer loop connected to the stem, which is constructed by a few bases, complementary to each other, at the two ends of the biopolymer. It is involved in many biological functions such as gene expression and regulation.[10] Thermodynamically, hairpin structures fluctuate among different conformations, including random coil and intact hairpin loop. In a simplified description, the conformations can be classified into two states: the open state where the binding monomers are separated and the closed state where the binding monomers form bonds. Lately, the conformation dynamics of an associating polymer have been investigated by attaching a donor fluorophore at one end of a short ssDNA and an acceptor at the other.[6,8] The fluorescence is quenched when donor and acceptor are within transfer range. Hence, a hairpin (closed) is quenched and a random coil (open) fluoresces. The probability of the open state can then be revealed by the melting curve, which depicts the variation of the static fluorescence intensity with the temperature. The melting temperature T_m of a hairpin loop is defined as the temperature where open and closed states have equal probabilities.

The thermal equilibrium between open (o) and closed (c) conformations of a ssDNA is often described by the all-or-none two-state model,[5−8,9,11]

$$\text{open} \underset{k_{c,o}}{\overset{k_{o,c}}{\rightleftharpoons}} \text{closed}, \quad K(T) = \frac{k_{o,c}}{k_{c,o}}$$

where $k_{i,j}$ denotes the rate constant jumping from i to j state. K is the equilibrium constant. It represents the probability or time scale ratio of the closed to open states. In other words, the fraction of the open state is given by $P_o = 1/(1 + K)$. To access the validity of the two-state model, we have studied the conformational fluctuations of a hard-sphere chain with a reacting bead attached to each end.[12] The binding energy between them is $-\epsilon$. On the basis of the two-state model, the melting curve can be derived, $P_o = [1 + e^{\epsilon(\beta_t-\beta_m)}]^{-1}$, where k_B is the Boltzmann constant and the inverse temperature $\beta_t = 1/k_BT$ and $\beta_m = 1/k_BT_m$. This analytical

† National Central University.
‡ University of Waterloo.
§ National Taiwan University.
* To whom correspondence should be addressed. E-mail yjsheng@ccms.ntu.edu.tw.

10.1021/ma034011z CCC: $25.00
Published on Web 06/27/2003

melting curve agrees excellently with the simulation result. Moreover, the kinetic behavior of such a model telechelic polymer can also be well depicted by the two-state model. The transition rate constant from a closed state to an open state $k_{c,o}$ is found to be independent of chain length N but proportional to $e^{-\beta_t \epsilon}$. This indicates that the free energy barrier from closed to open states is simply the binding energy. In contrast, $k_{o,c}$ is independent of temperature and proportional to N^{-2}. Such results points out that the free energy barrier from open to closed states is associated with entropy reduction of the chain conformation. As a consequence, the melting temperature decreases with increasing the chain length logarithmically, $\beta_m \epsilon \sim 2 \ln N$.

Since the two-state model is simple, it is frequently adopted to explain experimental observations. Nonetheless, there exist a few possible mechanisms,[5] which may cause substantial deviation from the two-state modeling. In fact, the open state may be further classified into the random coil and intrachain mismatch states.[6,13,14] For example, consider a homogeneous sequence poly(C) with TTGGG-3′ at one end and its complement 5′-CCCAA at the other. The bases 3′-GGG may zip with CCC within the sequence and form a mismatched loop conformation. Unlike in the absorbance experiment,[15] this misfolded structure gives signals just like the open state for a typical fluorescence experiment. It is then doubtful to extract thermodynamic as well as kinetic properties from the melting curves according to the two-state model. In this paper, we therefore intend to investigate the effect of intrachain mismatch on the loop-to-coil transition. The model associating polymer is a hard-sphere chain of chain length N with an attractive site at both ends (1 and N) and another attractive site in the middle along the backbone (N/2).

The present paper is organized as follows. In section II, we take into account the intrachain mismatch state and introduce the three-state model, which consists of three equilibriums, open ⇔ closed, open ⇔ mismatch, and mismatch ⇔ closed. On the basis of this model, the thermodynamic theory can be developed to obtain the melting curve and the transient rate constants. Section III briefly describes the model associating chain and the Monte Carlo simulation details. Finally, in section IV we compare the simulation results with those obtained from the three-state model. The effect of the misfolded state is discussed.

II. Three-State Model

Consider a chain made of N hard spheres. We assume that the end beads attract each other and are designated as α and β. The binding energy between α and β is $-\epsilon_1$. To investigate the effect of intrachain mismatch, the bead in the middle part along the chain, named as γ, is also attracted to the α bead. The binding energy of α with γ is $-\epsilon_2$. The conformations of such an associating chain can then be classified into three possible states, open (o), closed (c), and mismatch (m), as shown in Figure 1. It is anticipated that the probability of the open state $P_o \to 1$ at high temperature and the probability of the closed state $P_c \to 1$ at low temperature if $\epsilon_1 > \epsilon_2$. At thermodynamic equilibrium, one would expect that open ⇔ closed and open ⇔ mismatch. Intuitively, the direct transition between the closed and mismatch states is quite rare and therefore can be ignored. However, to avoid overcoming the free energy barrier ϵ_2 associated with unbinding the mismatch loop

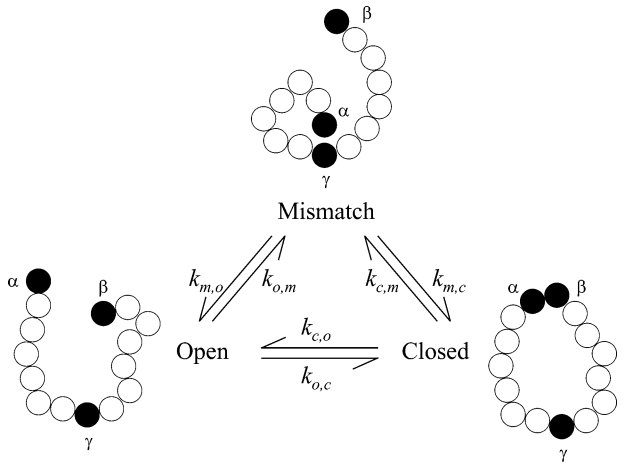

Figure 1. Schematic representations of the open, mismatch, and closed states of a model associating chain. The equilibrium relationships among these three states are present.

at low enough temperature, the transition mismatch ⇔ closed does take place substantially. The mechanism will be discussed in detail later.

At equilibrium, a thermodynamic three-state model ($o \Leftrightarrow c$, $o \Leftrightarrow m$, $m \Leftrightarrow c$) is proposed to describe the conformational kinetics of the model associating chain, as shown in Figure 1. The principle of detailed balance gives

$$k_{o,c} P_o = k_{c,o} P_c \tag{1}$$

$$k_{o,m} P_o = k_{m,o} P_m \tag{2}$$

and

$$k_{m,c} P_m = k_{c,m} P_c \tag{3}$$

where P_i is the probability of the state i and $k_{i,j}$ the rate coefficient jumping from i to j state. In terms of the probability of the closed state, one has

$$P_o = \frac{k_{c,o}}{k_{o,c}} P_c$$

$$P_m = \frac{k_{o,m}}{k_{m,o}} P_o = \frac{k_{o,m}}{k_{m,o}} \frac{k_{c,o}}{k_{o,c}} P_c \tag{4}$$

Because $P_o + P_m + P_c = 1$, the probabilities of the open and closed states are obtained respectively as

$$P_c = \frac{1}{1 + \frac{k_{c,o}}{k_{o,c}} + \frac{k_{c,o}}{k_{o,c}} \frac{k_{o,m}}{k_{m,o}}} \tag{5}$$

and

$$P_o = \frac{\frac{k_{c,o}}{k_{o,c}}}{1 + \frac{k_{c,o}}{k_{o,c}} + \frac{k_{c,o}}{k_{o,c}} \frac{k_{o,m}}{k_{m,o}}} \tag{6}$$

The rate coefficients are assumed to follow the Arrhenius kinetics

$$k_{i,j} = k^*_{i,j} \exp(-\beta_t F_{i,j}) \tag{7}$$

where β_t is the inverse temperature $\beta_t = 1/k_BT$ and $F_{i,j}$ represents the free energy barrier when jumping from i to j state. $k^*_{i,j}$ is a constant and depends on the chemical nature of the chain. There are six rate coefficients associated with three reversible equilibriums in this triangular, three-state model. As a result, the main task is to decide the free energy barrier $F_{i,j}$, which can be examined by the Monte Carlo simulation. The study of coil-to-loop transition of a telechelic chain[12] indicates that $F_{o,c}$ for the conformational change from the random coil state to the loop state is associated with the entropy reduction only. For the model associating chain, we also expect that the free energy barrier from open to closed state ($o \rightarrow c$) can be attributed to the entropy loss from coil to ring conformation,

$$\beta_t F_{o,c} = (S_o - S_c)/k_B \tag{8}$$

As a result

$$k_{o,c} = k^*_{o,c} \exp\left(-\frac{\Delta S_{o,c}}{k_B}\right) \tag{9}$$

This relation indicates that $k_{o,c}$ is temperature-independent. Similarly, for jumping from open to mismatch state ($o \rightarrow m$), one has

$$k_{o,m} = k^*_{o,m} \exp\left(-\frac{\Delta S_{o,m}}{k_B}\right) \tag{10}$$

The entropy loss due to conformational change is related to the chain length N and can be evaluated by considering the distribution function associated with a self-avoiding walk of N steps.[15] The probability for a terminal point adjacent to the origin is given by

$$P_N \sim N^{-(3+g)\nu}$$

where the exponent g and ν are $g \approx 0.28$ and $\nu \approx {}^3/_5$ for good solvents.[16] As a consequence, the entropy cost of a hard-sphere chain is increased with chain length and is estimated as $\Delta S_{o,c}/k_B \approx \alpha \ln N$ with $\alpha \approx 1.968$. One recovers $\alpha = {}^3/_2$ for a Gaussian chain with $g = 0$ and $\nu = {}^1/_2$. For $o \rightarrow m$, we can regard the mismatch conformation as the closed state associated with a hard-sphere chain with chain length $N/2$. Therefore, we also have the entropy reduction, $\Delta S_{o,m}/k_B \sim 2 \ln N$. Some experimental works have similar results.[9,13] In the actual case of ssDNA, other factors such as stacking effect are involved, and the exponent of N dependence may be quite different. For example, in the work of Kuznetsov et al.,[15] they pointed out that the free energy of the hairpin scales with N^α with $\alpha \geq 7$. They proposed that the stability of hairpins depends on both the entropic effect and stabilizing interactions within the loops, and these interactions may be related to the hydrophobic interactions of the bases.

On the other hand, for jumping from closed to open state, previous studies by Monte Carlo simulations confirm that the free energy barrier $F_{c,o}$ is simply the binding energy between the ends for loop-to-coil transition of a telechelic chain.[12] Following the same reasoning, we anticipate that the free energy barriers are $\beta_t F_{c,o} = \beta_t \epsilon_1$ for jumping from closed to open state and $\beta_t F_{m,o} = \beta_t \epsilon_2$ for jumping from mismatch to open state. Thus, one has

$$k_{c,o} = k^*_{c,o} \exp(-\beta_t \epsilon_1) \tag{11}$$

and

$$k_{m,o} = k^*_{m,o} \exp(-\beta_t \epsilon_2) \tag{12}$$

The direct transition between the mismatch configuration and the closed state is somewhat unexpected. For such an event to happen, the beads α, β, and γ have to be in the neighborhood and hence the associating polymer has to form an "8"-like configuration. As a result, there is an additional entropy cost of free energy barrier as the chain fluctuates from mismatch to closed state ($m \rightarrow c$),

$$k_{m,c} = k^*_{m,c} \exp\left(-\frac{\Delta S_{m,c}}{k_B}\right) \tag{13}$$

This entropy reduction is associated with the change from open to ringlike state for the chain segment (β–γ) of length $N/2$, though there already exists a ringlike conformation of the chain segment (α–γ) in the mismatch state. Note that there is no attraction between β and γ beads. Consequently, one can estimate the free energy barrier as $\Delta S_{m,c}/k_B \sim 2 \ln N$. On the other hand, for $c \rightarrow m$, in addition to the entropy reduction for the formation of figure "8"-like configuration, one also has to overcome the binding energy difference ($\epsilon_1 - \epsilon_2$). The free energy barrier $F_{c,m}$ then consists of both binding energy and entropy cost and is expressed as

$$k_{c,m} = k^*_{c,m} \exp[-\beta_t(\epsilon_1 - \epsilon_2)] \exp\left(-\frac{\Delta S_{c,m}}{k_B}\right) \tag{14}$$

with $\Delta S_{c,m}/k_B \sim 2 \ln N$.

Substituting eqs 9–14 into eqs 5 and 6, we obtain the probabilities of the three states and the melting curves,

$$P_o = \frac{K^*_1 \exp(-\beta_t \epsilon_1)}{1 + K^*_1 \exp(-\beta_t \epsilon_1) + K^*_1 K^*_2 \exp[-\beta_t(\epsilon_1 - \epsilon_2)]} \tag{15}$$

$$P_c = \frac{1}{1 + K^*_1 \exp(-\beta_t \epsilon_1) + K^*_1 K^*_2 \exp[-\beta_t(\epsilon_1 - \epsilon_2)]} \tag{16}$$

and

$$P_m = \frac{K^*_1 K^*_2 \exp[-\beta_t(\epsilon_1 - \epsilon_2)]}{1 + K^*_1 \exp(-\beta_t \epsilon_1) + K^*_1 K^*_2 \exp[-\beta_t(\epsilon_1 - \epsilon_2)]} \tag{17}$$

where the equilibrium constants K^*_1 and K^*_2 are independent of the temperature,

$$K^*_1 = \frac{k_{c,o}}{k_{o,c}} e^{\beta_t \epsilon_1} = \frac{k^*_{c,o}}{k^*_{o,c}} N^2 \tag{18}$$

and

$$K^*_2 = \frac{k_{o,m}}{k_{m,o}} e^{-\beta_t \epsilon_2} = \frac{k^*_{o,m}}{k^*_{m,o}} N^{-2} \tag{19}$$

The two equilibrium constants vary with the nature of the system, such as bond length and interaction

distance. In the two-state model, only one equilibrium constant is involved, and hence one can define the melting temperature as $P_o = P_c = {}^1/_2$. To describe the melting curve interfered by the mismatch state, however, we have to introduce two characteristic temperatures, which correspond to K^*_1 and K^*_2 according to eqs 15−17. The choice is quite arbitrary. We define the two characteristic temperatures, T_{m1} and T_{m2}, corresponding to the melting temperatures in the melting curves $P_o(T)$ and $P_c(T)$, as

$$T = T_{m1}, \quad \text{at } P_o = {}^1/_2$$
$$T = T_{m2}, \quad \text{at } P_c = {}^1/_2 \tag{20}$$

One can also define the third equilibrium constant K^*_3 as

$$K^*_3 = \frac{k_{m,c}}{k_{c,m}} e^{-\beta_t(\epsilon_1-\epsilon_2)} = \frac{k^*_{m,c}}{k^*_{c,m}} \tag{21}$$

According to the detailed balance, eqs 1−3, these equilibrium constants are not independent and must satisfy the condition

$$K^*_1 K^*_2 K^*_3 = 1 \tag{22}$$

In other words,

$$\frac{k^*_{c,o}}{k^*_{o,c}} \frac{k^*_{o,m}}{k^*_{m,o}} \frac{k^*_{m,c}}{k^*_{c,m}} = 1$$

The above equation provides an assessment of the three-state model and can be examined by Monte Carlo simulations.

III. Model and Simulation Details

The model associating chain is made of N hard spheres of diameter σ with attractive beads located at both ends (α and β) and in the middle (γ). This polymer is a freely jointed chain in continuous space. The interactions between the bonded beads are through the infinite deep square-well potentials,[17]

$$U_{i,i+1} = \begin{cases} \infty & r < \sigma \\ 0 & \sigma \le r < \xi\sigma \\ \infty & r \ge \xi\sigma \end{cases} \tag{23}$$

where $\xi = 1.2$. Bond crossing (phantom chain) can be prevented by such a choice. For this model associating chain, the interaction between attractive end beads α (1) and β (N) is represented by a standard square-well potential,

$$U_{\alpha,\beta} = \begin{cases} \infty & r < \sigma \\ -3\epsilon & \sigma \le r < \lambda\sigma \\ 0 & r \ge \lambda\sigma \end{cases} \tag{24}$$

where $\lambda = 1.2$. Without the loss of generality, we assume the binding energy $\epsilon = 5$. The other pair of interacting beads (α−γ) interact through a similar potential as

$$U_{\alpha,\gamma} = \begin{cases} \infty & r < \sigma \\ -2\epsilon & \sigma \le r < \lambda\sigma \\ 0 & r \ge \lambda\sigma \end{cases} \tag{25}$$

Here γ denotes the bead in the middle of the chain, and we choose the bead at $N/2 + 1$ if N is even. Note that the end bead β does not interact with the bead γ. In eqs 24 and 25, the binding energies of α with β and γ are taken to be $\epsilon_1 = 3\epsilon$ and $\epsilon_2 = 2\epsilon$, respectively.

In this work, Monte Carlo (MC) simulations were performed to study the loop-to-coil transition kinetics of a short ssDNA. Noted that in general MC simulation is not the method of choice for studying kinetics. However, since the actual dynamics has reached equilibrium, the MC method could be used to estimate the reaction constants of the equilibrium dynamics. The open, closed, and mismatch states of this model associating chain are shown in Figure 1. The chain is identified as in the "open" state when the attractive ends of the chain are not within square-well interaction regimes. That is, $|\mathbf{r}_\alpha - \mathbf{r}_\beta| > \lambda\sigma$. On the other hand, the "closed" state is defined as the formation of an attractive pair from two ends. That is, $|\mathbf{r}_\alpha - \mathbf{r}_\beta| < \lambda\sigma$. Similarly, the "mismatch" state is defined as $|\mathbf{r}_\alpha - \mathbf{r}_\gamma| < \lambda\sigma$. It is noted that we allow α bead to be within $\lambda\sigma$ distance of β and γ beads at the same time, but α bead can only "bind" with one of them. α bead chooses its binding partner according to the probability of $\exp(-\Delta E/k_BT)$. The rate constant $k_{i,j}$ is evaluated from the inverse of the mean time $\tau_{i,j}$. It is defined as the mean period of staying at the state i, which jumps to the state j later. The probability of the state i is calculated by the total Monte Carlo steps (MCs) staying at the state i to the total MCs of simulation. The detailed balance is fulfilled in our simulations.

The systems simulated contain a single polymer chain with chain length N ranging from 15 to 60. The simulations are performed under the conditions of constant temperature and total number of beads. In the present study, the reduced temperature T^* is varied to obtain the melting curves, which are represented by P_o and P_c. The trial moves employed for chains of the equilibration and production process are bead displacement motions. They involve randomly picking a bead and displacing it to a new position in the vicinity of the old position. The distance away from the original position is chosen with a probability, which satisfies the condition of equal sampling of all points in the spherical shell surrounding the initial position. The new configurations resulting from this move are accepted according to the standard Metropolis acceptance criterion.[18] Runs for the same chain length at different temperatures are performed starting with the final configuration from a previous temperature and are equilibrated for 200 million steps. Measurements for static properties such as the probabilities of open or closed states are taken over a period of 5−10 million MCs per bead.

IV. Results and Discussion

The kinetics of conformational fluctuations of an associating chain with attractive beads located at both ends and in the middle is investigated by Monte Carlo simulations. The two-state model may fail to depict the loop-to-coil transition due to the existence of the mismatch state. A three-state model is therefore proposed, and the corresponding thermodynamic theory is developed. Different from the two-state model, two characteristic temperatures, corresponding to the two equilibrium constants K^*_1 and K^*_2, are required to characterize the melting curves. By the comparison between the three-state model and the simulation results, the kinetic

rate constants are determined, and the effect of the mismatch state is realized in this section.

The melting curve, which depicts the variation of the probability of the open state with temperature, is shown in Figure 2a for different chain lengths. The general trends agree qualitatively with experimental observations.[6−8] The probable structure changes from a stable random coil to a mismatch or loop state as the temperature declines. That is, the conformational entropy dominates over the binding energy at high temperature. As the chain length is increased, the melting curves shift toward the lower temperature side. Since more entropy loss associated with a longer chain is needed for the transition from the coil to loop state, the melting temperature T_{m1}, defined at $P_o = {}^1/_2$, decreases with increasing N. It is very common to adopt the two-state model to describe the melting curve. If such a simple model is valid in the present case, then the melting curve ought to be well predicted by[12]

$$P_o(T) = \frac{1}{1 + e^{\epsilon_1(\beta_t - \beta_{m1})}} \tag{26}$$

Here β_{m1} can be readily determined from the melting curve. Unfortunately, eq 26 is unable to depict those data points even by regarding β_{m1} as a fitting parameter. This outcome is not surprising at all because of the intrachain mismatch.

The failure of the two-state model is able to be seen from another melting curve $P_c(T)$. Figure 2b illustrates the variation of the probability of the closed state with the temperature for different chain lengths. As the temperature is decreased, the binding energy gradually overcomes the thermal energy. Therefore, P_c grows with decreasing temperature, and its behavior is opposite to that of P_o. It is natural to define another melting temperature T_{m2} where $P_c = {}^1/_2$. If the two-state model is applicable, one has $P_o = P_c = {}^1/_2$ at $T_{m1} = T_{m2}$. Comparing parts a and b of Figure 2, one finds that both melting temperatures are not equal and $T_{m1} > T_{m2}$. This results indicate again that the two-state model fails to depict the coil-to-loop transition for an associating chain with intrachain mismatch. On the other hand, the simulation results in parts a and b of Figure 2 can be excellently described by the three-state model, eqs 15 and 16, respectively.

A clear evidence of the existence of the third state, i.e., the mismatch state, is displayed through Figure 2c for P_m. The result agrees quite well with that calculated by the three-state model, eq 17. The probability of the mismatch state varies with temperature for a given chain length and shows a peak value at $T = T_p(N)$. On the basis of eq 17, one is able to obtain T_p for $\epsilon_1 > \epsilon_2$,

$$\beta_p = \frac{1}{\epsilon_1} \ln K^*_1 \frac{\epsilon_2}{\epsilon_1 - \epsilon_2} \tag{27}$$

This result indicates that the larger the chain length, the smaller the value of T_p. That is, $\beta_p\epsilon_1 \sim 2 \ln N$ as $N \gg 1$ by using eq 18. This behavior is consistent with that of T_{m1} and T_{m2}, which will be discussed later. Moreover, T_p is less than T_{m2}. Comparing P_m in eq 17 to P_c in eq 16 gives that $P_m/P_c = K^*_1 K^*_2 e^{-\beta_t(\epsilon_1 - \epsilon_2)}$. It points out that in general $P_m < P_c$ if $\epsilon_1 > \epsilon_2$. Nevertheless, P_m might be greater than P_c at high enough temperature, $\beta(\epsilon_2 - \epsilon_1) \ll 1$, if $K^*_1 K^*_2 > 1$. The present study belongs to this case because the entropy reduction is higher for the closed state than the mismatch state.

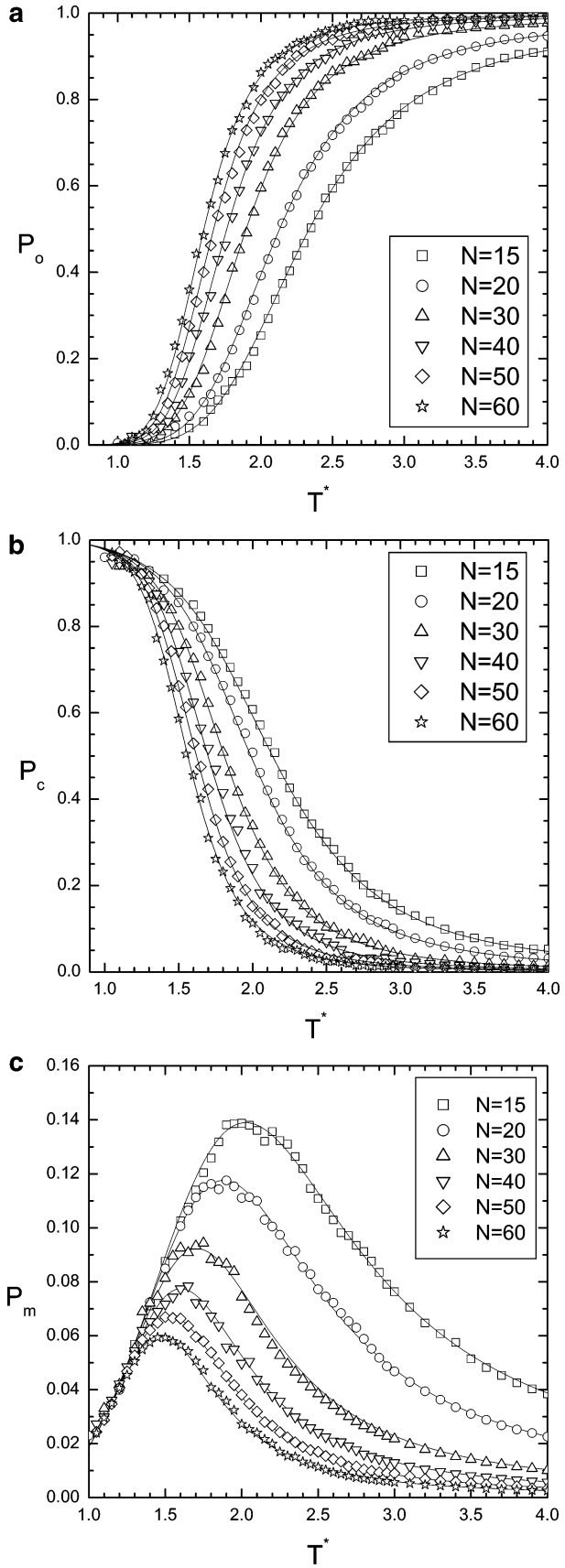

Figure 2. Variation of the probabilities of the three states with temperature for different chain lengths N: (a) open state P_o; (b) closed state P_c; (c) mismatch state P_m. The solid curves correspond to analytical predictions according to the three-state model.

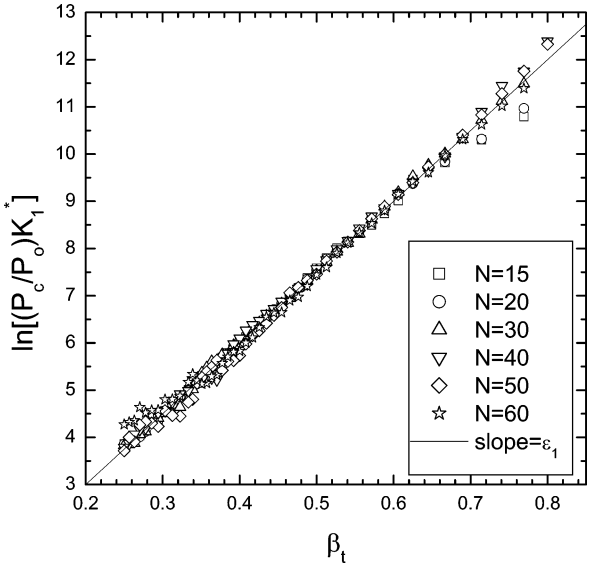

Figure 3. $\ln[(P_c/P_o)K^*_1]$ is plotted against β_t for various chain lengths. The solid line is a straight line with slope ϵ_1 and zero intercept based on the three-state model.

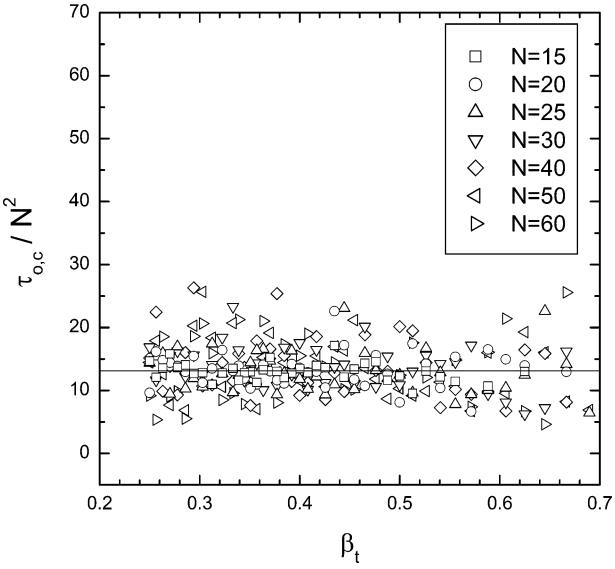

Figure 4. Mean lifetime of the open state jumping to the closed state later, $\tau_{o,c}/N^2$, is plotted against β_t for different chain lengths.

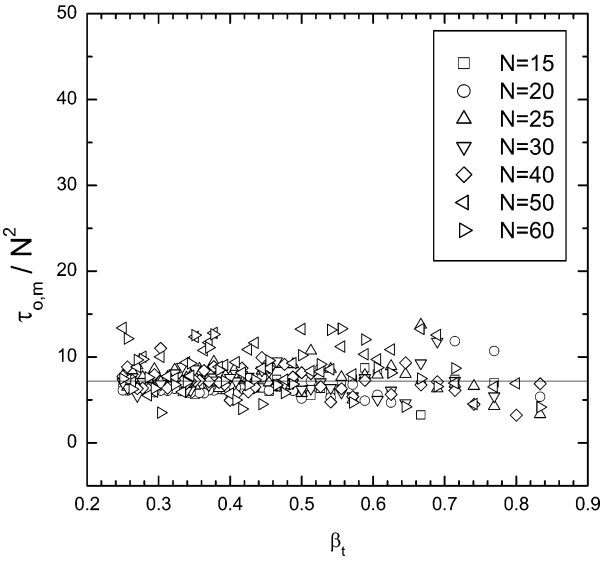

Figure 5. Mean lifetime of the open state jumping to the mismatch state later, $\tau_{o,m}/N^2$, is plotted against β_t for various chain lengths.

To confirm the validity of the three-state model furthermore, we anticipate that an expression derived from eqs 15 and 16 can represent all results from different chain lengths. Dividing eq 16 by eq 15 gives

$$K^*_1(P_c/P_o) = \exp(\beta_t\epsilon_1) \tag{28}$$

The above equation is a function of temperature only. Consequently, when $\ln[K^*_1(P_c/P_o)]$ is plotted against β_t, all data points of different chain lengths should collapse into a straight line, which passes through the origin with a slope ϵ_1. Figure 3 shows that the simulation results behave as expected.

The excellent agreement between the simulation results and the expressions based on the three-state model confirms the success of the proposed theory. The values of the equilibrium constants $K^*_1(N)$ and $K^*_2(N)$ used in eqs 15−17 can be determined by two independent approaches. The first one is simply in accordance with the melting curves P_o, P_c, and P_m. Following eqs 18 and 19 for their dependence on chain length, one obtains

$$K^*_1 = 3.64N^2, \quad K^*_2 = 0.688N^{-2} \tag{29}$$

The second approach is to decide the equilibrium constants from the kinetic rate constants, i.e., the ratios of $k_{c,o}$ to $k_{o,c}$ and of $k_{o,m}$ to $k_{m,o}$. It will be shown later that these two approaches yield consistent results. Equation 29 indicates that $K^*_1 K^*_2$ is independent of chain length, and hence one has $P_c \to 0.275N^{-2}$, $P_m \to 0.688N^{-2}$, and $P_o \to 1 - 0.963N^{-2}$ as $T \to \infty$ from eqs 15−17. In other words, at high temperature, the probabilities of the closed and mismatch states are insignificant for a long enough chain.

The equilibrium constants are related to the melting temperatures. According to eq 20, one has

$$e^{-\epsilon_1\beta_{mo}} = \frac{1 \pm K^*_1 K^*_2 e^{-\beta_{mo}(\epsilon_1-\epsilon_2)}}{K^*_1} \tag{30}$$

where + and − denote 1 and 2, respectively. Figure 8 shows the good agreement of the melting temperatures between the simulation and the three-state model. Although the analytical solution of the melting temperatures $\beta_{m\pm}$ is not available, the asymptotic analysis for $N \gg 1$ ($\beta_{m\pm} \gg 1$) yields

$$\beta_{m\pm} \cong \frac{1}{\epsilon_1} \ln \frac{K^*_1}{1 \pm (K^*_1)^{\epsilon_2/\epsilon_1} K^*_2} \tag{31}$$

This analytical result agrees quite well that obtained from simulations as shown in Table 1. For $N \to \infty$, eq 31 gives

$$\beta_{m\pm} \approx \frac{1}{\epsilon_1}\{2 \ln N + \ln 3.64 \mp O[N^{-2(1-\epsilon_2/\epsilon_1)}]\} \tag{32}$$

The above result indicates that the two melting temperatures approach each other and turn to be $\beta_m\epsilon_1 = 2 \ln N$ essentially. This outcome is the same as that of the two-state model and hence concludes that the two-

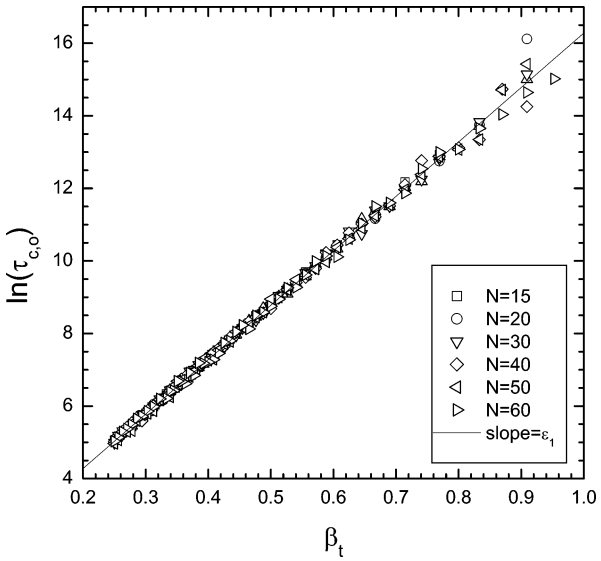

Figure 6. Mean lifetime of the closed state jumping to the open state later, ln $\tau_{c,o}$, is plotted against β_t for different chain lengths. The solid line is a straight line with slope ϵ_1 based on the three-state model.

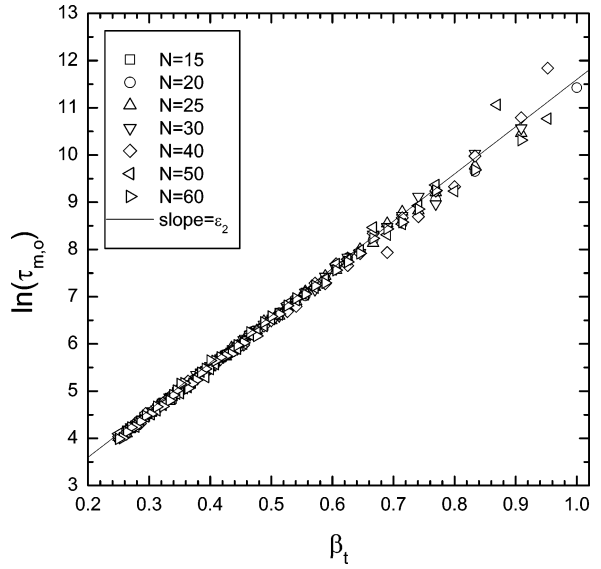

Figure 7. Mean lifetime of the mismatch state jumping to the open state later, ln $\tau_{m,o}$, is plotted against β_t for different chain lengths. The solid line is a straight line with slope ϵ_2 based on the three-state model.

state model is adequate if the chain length is long enough and $\epsilon_1 > \epsilon_2$. In other words, the entropy reduction for a very long chain, growing about ln N, requires the compensation from the larger binding energy ϵ_1 instead of ϵ_2. Hence, the mismatch state is getting insignificant with increasing N as shown in Figure 2c. However, the story is totally different if $\epsilon_1 \leq \epsilon_2$ though the three-state model, eqs 15−17, is still applicable.

We have confirmed that the proposed three-state model has to be employed to describe the thermodynamic equilibrium of the model associating chain. The kinetics of conformational fluctuations is also developed in eqs 9−14. To assess our theory, the rate constants are evaluated from Monte Carlo simulations and compared to the proposed model. On the basis of the principle of detailed balance, the rate constant $k_{i,j}$ is inversely proportional to the mean lifetime of the state i, i.e., $k_{i,j} = \tau_i^{-1}$. In accord with eqs 9 and 10, the rate

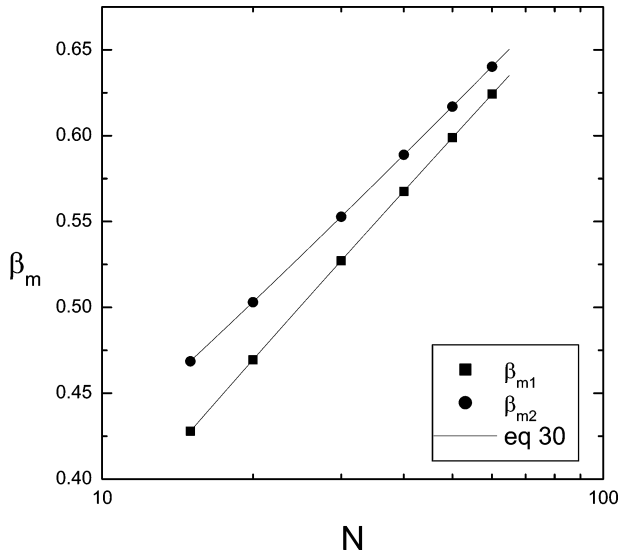

Figure 8. Variation of the two melting temperatures as a function of the chain length. The lines are calculated according to eq 30.

Table 1. Characteristic Temperatures Estimated from $C(T)$, $P_o(T)$, and $P_c(T)$ Curves for Various Chain Lengths

N	$T_c(C)$	$T_{m2}(P_c)$	$T_{m1}(P_o)$
15	2.10	2.13	2.34
20	1.95	1.99	2.13
30	1.78	1.81	1.90
40	1.67	1.70	1.76
50	1.59	1.62	1.67
60	1.54	1.56	1.60

constants from open to closed or mismatch state, i.e., $k_{o,c}$ or $k_{o,m}$, are proportional to N^{-2} and independent of temperature. As demonstrated in Figures 4 and 5, when $k_{o,c}^{-1}/N^2$ and $k_{o,m}^{-1}/N^2$ are plotted against β_t, all the rate constants for different chain lengths fall together. The straight line with a zero slope points out that the transition from open to closed or mismatch state is irrelevant to internal energy change and therefore is related to the entropy reduction. In other words, the formation of a loop requires to overcome thermal fluctuations associated with the chain. This consequence verifies our argument that the entropy loss from a coil to a ring $-\Delta S(N)$ is proportional to α ln N with $\alpha \approx 2$ very close to $(3 + g)\nu$.

In contrast to $k_{o,c}$ and $k_{o,m}$, eqs 11 and 12 give that $k_{c,o}$ and $k_{m,o}$ are proportional to $\exp(-\beta\epsilon_1)$ and $\exp(-\beta\epsilon_2)$, respectively. They do not vary with chain length. As illustrated in Figures 6 and 7, ln $k_{c,o}^{-1}$ and ln $k_{m,o}^{-1}$ are plotted against β_t for different chain lengths. All the rate constants evaluated for different chain lengths collapse into a straight line with a slope of the binding energy ϵ_1 for $k_{c,o}$ and ϵ_2 for $k_{m,o}$. This result indicates that the thermal energy provides a probability of $\exp(-\beta\epsilon_1)$ to unbind the intact loop to the random coil conformation. Similarly, the mismatch loop is unbounded with a probability $\exp(-\beta\epsilon_2)$ by thermal fluctuations. The chain length does not play any role in these processes.

The direct transition between closed and mismatch states means that the formation of the $\alpha-\gamma$ pair is right after the breakage of the $\alpha-\beta$ pair, or vice versa. As a result, this process requires all attractive beads α, β, and γ nearby. Otherwise, the transition from mismatch to closed state is indirect and ought to go through the open state, i.e., $m \Leftrightarrow o \Leftrightarrow c$. Direct transition $m \Leftrightarrow c$ involves a chain conformation of two loops, i.e., "8"-like

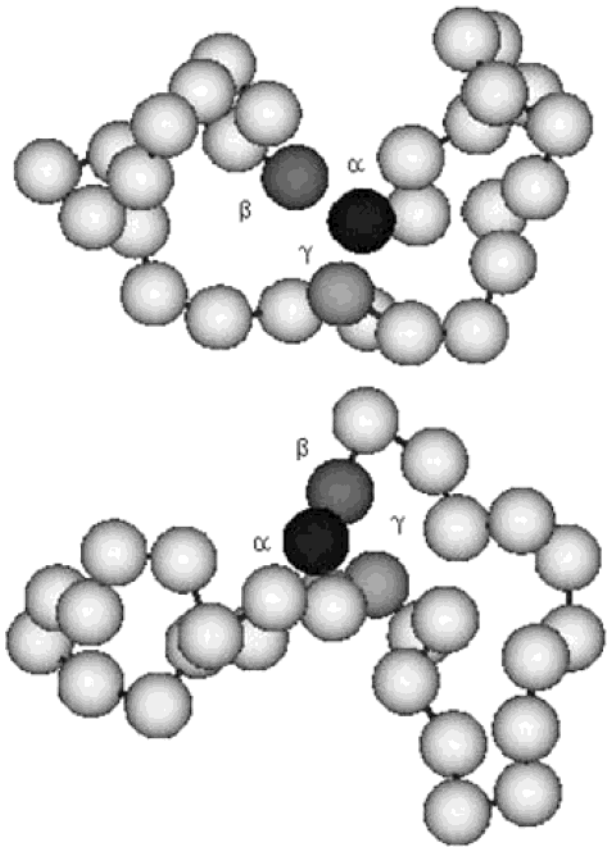

Figure 9. Two typical snapshots of the associating chain conformations correspond to the direct transition between closed and mismatch states.

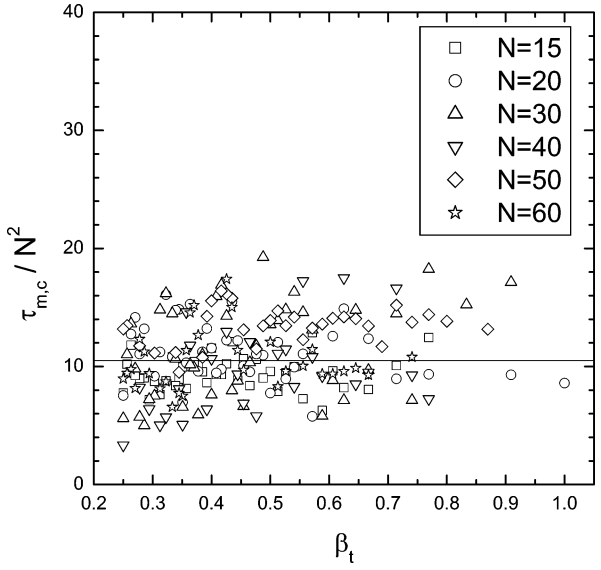

Figure 10. Mean lifetime of the mismatch state jumping to the closed state later, $\tau_{m,c}/N^2$, is plotted against β_t for different chain lengths.

configuration, and therefore more entropy loss than that associated with one loop is required. Intuitively, this event would be rare. Nevertheless, the indirect transition consists of overcoming the binding energy and entropy reduction in series. Therefore, the free energy barrier associated with the first process may be relatively low. We are able to observe this particular conformation at low enough temperature, e.g., T_p. Two typical snapshots are demonstrated in Figure 9. To

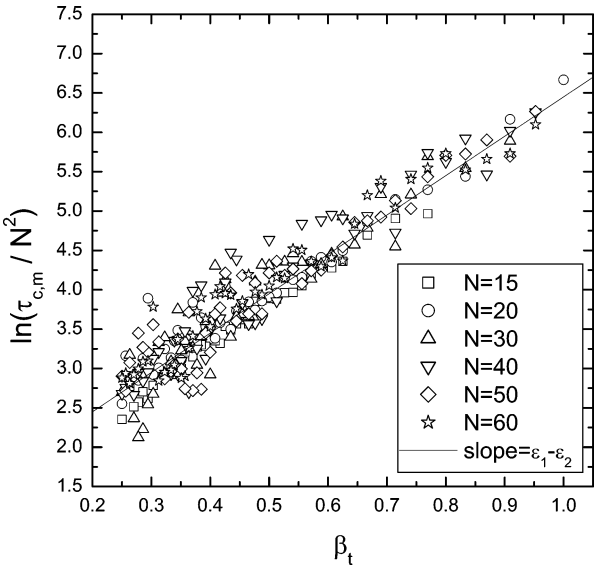

Figure 11. Mean lifetime of the closed state jumping to the mismatch state later, $\ln(\tau_{c,m}/N^2)$, is plotted against β_t for different chain lengths. The solid line is a straight line with slope $\epsilon_1 - \epsilon_2$ based on the three-state model.

further confirm the direct $m \leftrightarrows c$ transition, we evaluate the rate constants $k_{m,c}$ and $k_{c,m}$ from simulations. As pointed out in eq 13, the free energy from one mismatched loop to two loops needs an entropy cost, $kT \ln N^2$. Thus, $k_{m,c}$ should be independent of temperature for $m \rightarrow c$. The plot of $k_{m,c}^{-1}/N^2$ for different chain lengths against β_t yields a straight line with zero slope as shown in Figure 10. The transition $c \rightarrow m$, however, involves both entropy reduction and binding energy decrease. Equation 9 gives that $k_{c,m} \propto N^{-2}\exp[-\beta_t(\epsilon_1 - \epsilon_2)]$. Consequently, as shown in Figure 11, the plot of $\ln k_{c,m}^{-1}/N^2$ against β_t collapses all data points from different chain lengths into a single, straight line with a slope $\epsilon_1 - \epsilon_2$. The agreement illustrated in Figures 10 and 11 verifies the direct transition $m \leftrightarrows c$ and the three-state model.

In accordance with eqs 18 and 19, the equilibrium constants can be determined by those kinetic rate constants as well. The above analysis gives

$$k^*_{o,c} = 0.0763, \quad k^*_{c,o} = 0.278$$

and

$$k^*_{o,m} = 0.139, \quad k^*_{m,o} = 0.202$$

The ratio of $k^*_{c,o}$ to $k^*_{o,c}$ yields 3.64 for K^*_1 and $k^*_{o,m}/k^*_{m,o} = 0.688$ for K^*_2. They are identical to those determined from the melting curves, eq 29. From the direct transition between closed and mismatch states, we have

$$k^*_{m,c} = 0.0952, \quad k^*_{c,m} = 0.235$$

The equilibrium constant K^*_3 can then be evaluated from $k^*_{m,c}/k^*_{c,m} = 0.406$. The product $K^*_1 K^*_2 K^*_3 = 1.016$ is very close to 1 and hence satisfies the constraint inherent in the three-state model, eq 22.

The transition temperature of a system is often characterized by the heat capacity curve $C(T)$ because it is closely related to the variation of the internal

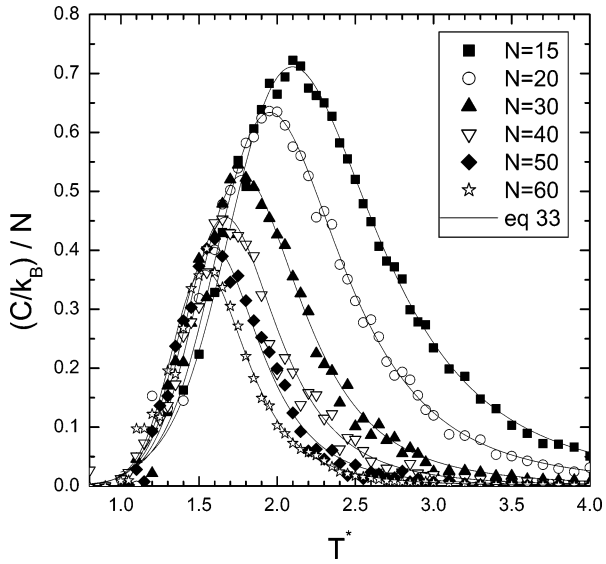

Figure 12. Heat capacity curves for different chain lengths are plotted against temperature and compared to eq 33 based on the three-state model.

energy U with the change of the configuration. The heat capacity is defined as

$$\frac{C}{k_B} = \frac{\langle U^2\rangle - \langle U\rangle^2}{(k_B T)^2}$$

We can define the peak temperature of the heat capacity curve, T_c, as the transition temperature and compare it to the melting temperatures obtained from the melting curves, P_o and P_c. In addition, one can also make a comparison of $C(T)$ computed by the three-state model and the simulation. Since $\langle U\rangle = \epsilon_1 P_o + \epsilon_2 P_m$ and $\langle U^2\rangle = \epsilon_1^2 P_o + \epsilon_2^2 P_m$ for our model associating chain, the heat capacity is evaluated as

$$\frac{C}{k_B} = \beta_t^2[(\epsilon_1^2 P_c + \epsilon_2^2 P_m) - (\epsilon_1 P_c + \epsilon_2 P_m)^2] \quad (33)$$

Substituting eqs 16 and 17 into eq 33, one is able to obtain the heat capacity curve based on the three-state model. As illustrated in Figure 12, eq 33 can depict the simulation results both qualitatively and quantitatively. Once more, this proves the validity of the three-state model in representing the behavior of our model associating chain. The peak temperature can be decided from eq 33 by $dC(T)/dT|_{T=T_c} = 0$ or simply by visual inspection. Both ways yield practically identical results. Figure 12 shows that T_c declines with increasing chain length. This tendency is similar to that associated with the melting temperatures. The three characteristic temperatures are listed in Table 1 for different chain lengths. Clearly, it indicates that $T_{m1} > T_{m2} > T_c$ for all chain lengths. It is not difficult to verify this order from the proposed three-state model. Table 1 also reveals that their differences decline with increasing the chain length. As $N \gg 1$, $P_c \gg P_m$ and the peak temperature is approximately $\beta_c\epsilon_1 \approx \ln K^*_1 \propto \ln N^2$. This result coincides with those of the melting temperatures, i.e., eq 32. In fact, in the thermodynamical limit of $N \to \infty$, the smooth loop-to-coil transition for an associating chain of finite chain length will become a step function, and all three characteristic temperatures converge together.

It should be noted that a dead-end three-state model may also be used for describing the transition dynamics which include the open, close, and mismatch states. However, the dead-end mechanism is without the direct transition between close and mismatch states as

$$\text{mismatch} \leftrightharpoons \text{open} \leftrightharpoons \text{close}$$

The dead-end mechanism and the triangular mechanism are equivalent. As can be seen from eq 22, the detailed balance in the triangular nechanism pose a constraint on the three equilibrium constants, i.e., $K^*_1 K^*_2 K^*_3 = 1$, which indicate that only two of transition processes (mismatch ⇔ open, open ⇔ close, and mismatch ⇔ close) are independent. The system can be represented by any two of the transition processes. As can be seen in this work, any information about transition between mismatch and closed (mismatch ⇔ close) states gathered from simulations is absolutely not needed in describing the essential features, such as P_o, P_c, P_m, and C_v of the system (eqs 15, 16, 17, and 33). If the dead-end mechanism is used instead of the triangular mechanism in this work, all the main conclusions would be just the same. In other words, $k_{m,c}$ and $k_{c,m}$ are not used in describing the loop-to-coil transition. However, substantial direct transitions between mismatch and closed states were observed in our simulation, and thus the triangular mechanism was used to model the system.

As to whether or not direct transitions between mismatch and closed states occur in reality, it is believed to depend on the "time interval" of observations and flexibility of the chain. In the actual case of ssDNA, it is possible that due to the directionality of the chain, direct transition mismatch ⇔ closed rarely happens. In that case, the dead-end model is a better choice. However, many works[19,20] show that the persistence length of many ssDNA chains is about 1 nm, which is approximately the size of a base. It indicates that, unlike a double-stranded DNA chain, ssDNA can usually be taken as a fully flexible chain. In our work, direct transition between mismatch and closed states is a natural cause of our model. For a fully flexible chain (without any stacking effect), when temperature is low, the energy factor dominates over the entropy effect. Thus, α, β, and γ beads tend to be together (see Figure 9) to lower the system's free energy. Since they are so close to each other, direct transitions between mismatch and closed states can then take place. In our simulation, we can indeed independently collect kinetic constants for open ⇔ closed, mismatch ⇔ open, and mismatch ⇔ closed and check whether they satisfy the detailed balance set as eq 22. This is an important consistency check of our work.

In conclusion, we have proposed a three-state model to describe the loop-to-coil transition of an associating polymer which exists as open, closed, or mismatch state. Our model agrees well with the MC computational results for the same one-chain system. The characteristic times $\tau_{i,j}$ for i to j state are determined. $\tau_{o,m}$, $\tau_{o,c}$, and $\tau_{m,c}$ are found to be independent of temperature but proportional to N^2. $\tau_{m,o}$ and $\tau_{c,o}$ are independent of N but proportional to $e^{\epsilon_2/k_B T}$ and $e^{\epsilon_1/k_B T}$. $\tau_{c,m}$ follows $N^2 e^{(\epsilon_1-\epsilon_2)/k_B T}$.

The success of the three-state model on the loop-to-coil transition of an associating chain with three attractive sites clearly identifies the inadequacy of a simple two-state model which is used to extract kinetic and

thermodynamic information from experiments. The existence of the mismatch state may significantly influence the outcome. Therefore, to be able to quantitatively model the equilibrium kinetics of a ssDNA, we should know more about the mismatch (or intermediate) states. In this work, the model chain used was essentially hard-sphere chain with full flexibility. In the actual case of ssDNA and RNA, other factors such as stacking effect (hydrophobic interactions) may complicate the hairpin-loop transition. These factors should be taken into consideration for further investigation of this issue. The study of a more realistic chain is currently in process.

Acknowledgment. H.-K.T. and Y.-J.S. thank NSC of Taiwan for financial support, and J.Z.Y.C. gratefully acknowledges financial support from National Center of Theoretical Science of Taiwan. Computing time, provided by the Nation Center for High-Performance Computing of Taiwan, is gratefully acknowledged.

References and Notes

(1) Pham, Q. T.; Russel, W. B.; Thibeault, J. C.; Lau, W. *Macromolecules* **1999**, *32*, 2996.
(2) Ma, S. X.; Cooper, S. L. *Macromolecules* **2001**, *34*, 3294.
(3) Eiser, E.; Klein, J.; Witten, T. A.; Fetter, L. J. *Phys. Rev. Lett.* **1999**, *82*, 5076.
(4) Khalatur, P. G.; Khokhlov, A. R.; Kovalenko, J. N.; Mologin, D. A. *J. Chem. Phys.* **1999**, *110*, 6039.
(5) Ying, L.; Wallace, M. I.; Klenerman, D. *Chem. Phys. Lett.* **2001**, *334*, 145.
(6) Wallace, M. I.; Ying, L.; Balasubramanian, S.; Klenerman, D. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 5584.
(7) Ansari, A.; Kuznetsov, S. V.; Shen, Y. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 7771.
(8) Goddard, N. L.; Bonnet, G.; Krichevsky, O.; Libchaber, A. *Phys. Rev. Lett.* **2000**, *85*, 2400.
(9) Bonnet, G.; Krichevsky, O.; Libchaber, A. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95* 8602.
(10) Zazopoulos, E.; Lalli, E.; Stocco, D. M.; Sassone-Corsi, P. *Nature (London)* **1997**, *390*, 311.
(11) Wennmalm, S.; Edman, L.; Rigler, R. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 10641.
(12) Sheng, Y.-J.; Chen, J. Z. Y.; Tsao, H.-K. *Macromolecules* **2002**, *35*, 9624.
(13) Ansari, A.; Shen, Y.; Kuznetsov, S. V. *Phys. Rev. Lett.* **2002**, *88*, 069801.
(14) Goddard, N.; Bonnet, G.; Krichevsky, O.; Libchaber, A. *Phys. Rev. Lett.* **2002**, *88*, 069802.
(15) Kuznetsov, S. V.; Shen, Y.; Benight, A. S.; Ansari, A. *Biophys. J.* **2001**, *81*, 2864.
(16) De Gennes, P.-G. *Scaling Concepts in Polymer Physics*; Cornell University Press: Ithaca, NY, 1993. Des Cloizeaux, J.; Jannink, G. *Polymers in Solution: Their Modelling & Structure*; Oxford University Press: Oxford, 1990.
(17) Zhou, Y.; Hall, C. K.; Karplus, M. *Phys. Rev. Lett.* **1996**, *77*, 2822. Sheng, Y.-J.; Jiang, S.; Tsao, H.-K. *Macromolecules* **2002**, *35*, 7865.
(18) Allen, M. P.; Tildesley, D. J. *Computer Simulations of Liquids*; Oxford University Press: New York, 1987.
(19) Ansari, A.; Kuznetsov, S. V.; Shen, Y. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 7771.
(20) Smith, S. B.; Cui, Y.; Bustamante, C. *Science* **1996**, *271*, 795.

MA034011Z