Open-to-Closed Transition of a Hard-Sphere Chain with Attractive Ends

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

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

Received July 24, 2002
Revised Manuscript Received November 5, 2002

The hairpin structure is frequently observed in both single-strand DNA (ssDNA) and RNA and participates in many biological functions.[1−4] A typical hairpin loop consists of a linear polymer segment connected to binding monomers. It fluctuates thermodynamically between different conformations, which, in a simplified description, are divided into two main states: the open state where the binding monomers are separated and the closed one where the binding monomers form a bond. Characterization of the nature of the open-to-closed transition is essential to our understanding of biopolymer dynamics.

Recently, particularly simple DNA hairpin-loop structures were purposefully designed to investigate their conformational fluctuations.[5,6] The stem is formed by a few bases, complementary to each other, at the two ends of ssDNA. The loop is made of a homogeneous sequence such as poly(T) (polydeoxythymidines). By attaching donor and acceptor fluorophores to both ends of a ssDNA hairpin, the open-to-closed conformational dynamics have been investigated.[5,6] The fraction of the open state is displayed in terms of the melting curve, which depicts the variation of the static fluorescence intensity with temperature. The melting temperature T_m of the structure is defined as the temperature where the probabilities of closed and open states are equal.

The all-or-none two-state model is commonly adopted to describe the thermal equilibrium between closed (c) and open (o) conformations,[1−3,5−7]

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

where $k_{i\to j}$ is the rate coefficient jumping from i to j state. The fraction of open state is $f = 1/(1 + K)$. However, other physical mechanisms may cause the deviation from the two-state model. For example, the open state may actually be further classified into an open, random-coil and a mismatched-loop state. On the other hand, it is also difficult to tell whether the closed state corresponds to a hairpin structure with some base pairs intact or whether it is merely a compact, collapsed structure, where the donor and acceptor are quite close.[7] Moreover, factors such as transition state[7] and base stacking[5,7] also complicate the thermodynamics of the open-to-closed transition. In this Communication, we rule out those complications and concentrate on investigating the open-to-closed transition by considering a hard-sphere chain with attractive ends.

We perform off-lattice Monte Carlo (MC) simulations to obtain the melting curves and transition rates. The biopolymer is modeled as a freely jointed, hard-sphere chain with N beads of diameter σ. Bonded beads i and $i + 1$ interact via an infinitely deep square-well potential[8]

$$U_{i,i+1} = \begin{cases} \infty, & r < \sigma \\ 0, & \sigma \le r \le 1.2\sigma \\ \infty, & r > 1.2\sigma \end{cases} \tag{1}$$

The terminal beads 1 and N interact via a square-well potential

$$U_{1,N} = \begin{cases} \infty, & r < \sigma \\ -\epsilon, & \sigma \le r \le 1.2\sigma \\ 0, & r > 1.2\sigma \end{cases} \tag{2}$$

At each MC step, a randomly selected bead on the chain was allowed to move around its previous position with a restriction of the bond fluctuation between σ and 1.2σ. The number of MC steps per bead is more than 10^7. The conformation is clearly identified as the closed state when $U_{1N} < 0$, i.e., $\sigma \le |\mathbf{r}_1 - \mathbf{r}_N| \le 1.2\sigma$, and as the open state otherwise. Without the loss of generality, we assume the binding energy $\epsilon = 15$. Figure 1 depicts the melting curve ($f-T$) for different chain lengths. The result is qualitatively consistent with experimental observations. As the temperature rises, the probable structure shifts from a stable closed state to a stable open one. Moreover, the melting temperature of the structure decreases with increasing chain length. The degree of the melting temperature decline becomes much less substantial when the chain length is long enough.

Since the two-state model is perfectly fulfilled, a thermodynamic theory can be adopted to obtain the melting curve, which is examined by simulation results. At equilibrium, the principle of detailed balance gives

$$k_{o\to c}P_o = k_{c\to o}P_c \tag{3}$$

where P_i is the probability of the state i. The rate coefficient is assumed to follow the Arrhenius kinetics, $k_{i\to j} = k^*_{i\to j}\exp(-\beta F_{i\to j})$, where β is the inverse temperature $\beta = 1/kT$ and $F_{i\to j}$ denotes the free energy barrier associated with jumping from the i to j states. When the conformation changes from the open state to the closed one, the free energy barrier corresponds only to the entropy loss from a random coil to a ring polymer, $\beta F_{o\to c} = -(S_c - S_o)/k$. Hence, $k_{o\to c}$ is temperature-independent. Moreover, the entropy cost varies with chain length and is logarithmic.[9,10] In a Gaussian chain, the closed-conformation probability is proportional to $N^{-3/2}$ and the open-conformation probability is approximately $1 - O(N^{-3/2})$.[10,11] As a result, $\Delta S/k \approx \alpha \ln N$ with $\alpha = 3/2$. However, for a hard-sphere chain, $\alpha \approx 2$, as will be shown later. On the other hand, when the structure fluctuates from the closed state to the open one, the free energy barrier is simply the binding energy, $\beta F_{c\to o} = \beta\epsilon$.

---

[*] To whom correspondence should be addressed. E-mail hktsao@cc.ncu.edu.tw.
[†] National Taiwan University.
[‡] University of Waterloo.
[§] National Central University.

10.1021/ma025604x CCC: $22.00

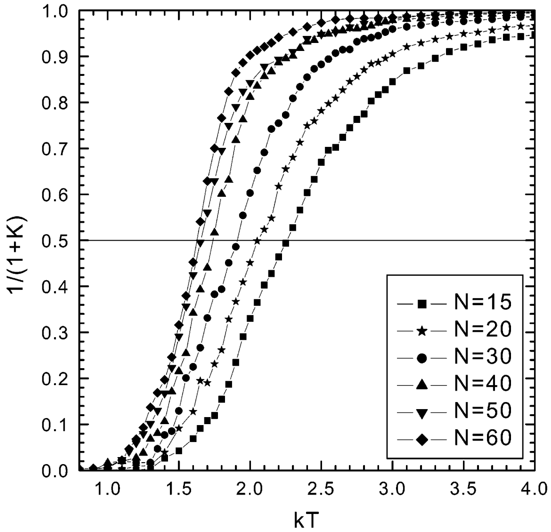

Figure 1. Melting curves for different chain lengths.

Since $f = P_o$ and $P_o + P_c = 1$, using eq 3, one obtains

$$f = \frac{1}{1+K} = \frac{1}{1 + \left[\frac{k^*_{o\to c}}{k^*_{c\to o}} \exp\left(-\frac{\Delta S}{k}\right)\right] \exp(\beta\epsilon)} \qquad (4)$$

ΔS is the entropy change from open to closed state. At $T = T_m$, $f = 1/2$. Consequently, the preexponential factor in eq 4 is related to the melting temperature,

$$\frac{k^*_{o\to c}}{k^*_{c\to o}} \exp\left(-\frac{\Delta S}{k}\right) = \exp(-\beta_m\epsilon) \qquad (5)$$

Combing eqs 4 and 5 yields the melting curve,

$$f(T) = \frac{1}{1 + e^{\epsilon(\beta-\beta_m)}} \qquad (6)$$

Equation 5 also shows that the melting temperature decreases with increasing chain length logarithmically,

$$\frac{\epsilon}{kT_m} = \alpha \ln N + \ln\left(\frac{k^*_{c\to o}}{k^*_{o\to c}}\right) \qquad (7)$$

Thermodynamic quantities commonly analyzed at the transition of polymer conformations are the energy per monomer and the specific heat. Because the melting curve represents the probability of the open state, heat capacity curve can be evaluated directly from eq 6,

$$\frac{C}{k} = \frac{f(1-f)\epsilon^2}{(kT)^2} = (\beta\epsilon)^2 \frac{e^{\epsilon(\beta-\beta_m)}}{[1 + e^{\epsilon(\beta-\beta_m)}]^2} \qquad (8)$$

The rate constants are evaluated from MC simulations and compared to our thermodynamic theory. For the two-state model, the rate constant $k_{i\to j}$ is inversely related to the mean lifetime of the state i, i.e., $k_{i\to j} = \tau_i^{-1}$. As shown in Figure 2a, when $\ln k_{c\to o}^{-1}$ is plotted against β, all the rate constants evaluated from different chain lengths collapse into a single line with a slope of the binding energy ϵ. This consequence indicates that thermal fluctuations provide a probability of $\exp(-\beta\epsilon)$ to unbind the closed conformation regardless of the chain length. On the other hand, the rate constant from the open to closed state $k_{o\to c}$ is independent of the temperature but decreases with increasing chain length. This result clearly shows that thermal fluctuations furnish a constant probability of $\exp[-\Delta S(N)/k]$ to form the closed state. When $\ln(k_{o\to c}^{-1}/N^2)$ is plotted against β, as illustrated in Figure 2b, all the rate constants computed from different chain lengths fall into a constant line with zero slope. In accordance with our simulation results, the entropy loss from a coil to a ring is $\Delta S(N)/k \sim N^2$. In comparison with the Gaussian chain, the scaling exponent of N rises from $3/2$ to 2 due to excluded volume.

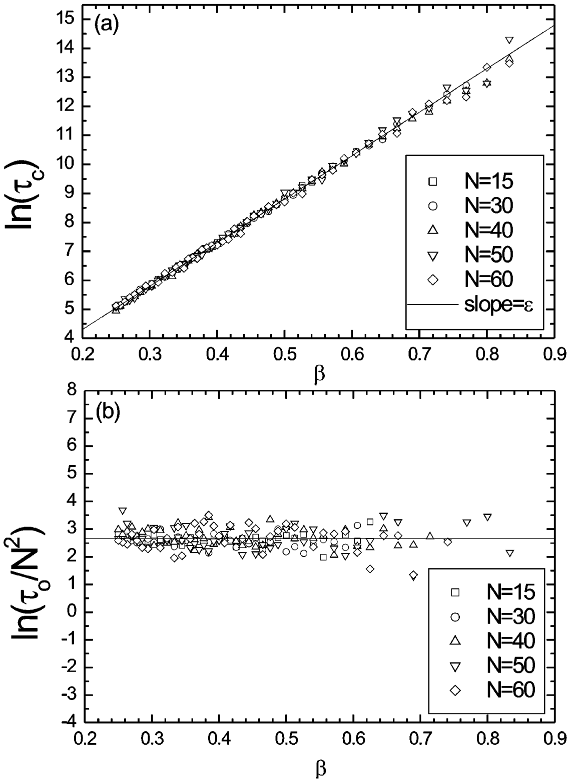

Figure 2. Variation of the rate constants with the temperature β for different chain lengths. (a) $\tau_c = k_{c\to o}^{-1}$; (b) $\tau_o = k_{o\to c}^{-1}$.

The entropy loss from the open to closed state can be further confirmed by considering the distribution function associated with a self-avoiding walk of N steps.[11] The probability for a terminal point adjacent to the origin is given by $p_N \sim N^{-(3+g)\nu}$. For good solvents, one has $\nu \approx 3/5$ and $g \approx 0.28$. Consequently, $\Delta S/k \approx \alpha \ln N$ with an entropy reduction exponent $\alpha = 1.968$, which is quite close to our numerical value 2. For the ideal chain, $\nu = 1/2$ and $g = 0$, one recovers $\alpha = 3/2$. This result is also closely related to the mean first passage time for the close approach of the ends of a polymer.[12]

The inverse rate constant $k_{o\to c}^{-1}$ is the mean lifetime of the open state τ_o, which is irrelevant to the binding energy and expected to be closely related to the mean first passage time of cyclization τ_R of a diffusion-limited intrachain reaction. Previous studies point out that τ_R is proportional to N^α with $3/2 \leq \alpha \leq 2$ for a Rouse chain of bond length b.[12,13] When the contact distance a is small compared to $b/N^{1/2}$, τ_R is much greater than the longest relaxation time τ_m and a local equilibrium assumption is valid. Thus, one has $\tau_R \propto N^{3/2}$. On the other hand, in the limit of large N at fixed a, τ_R is believed to be comparable to τ_m.[13,14] Since $\tau_m \propto N^{1+2\nu}$,[10] τ_R is found to be N^2 for a Gaussian chain ($\nu = 1/2$). For a hard-sphere chain adopted in our simulation, one anticipates that, similar to the Gaussian chain, the

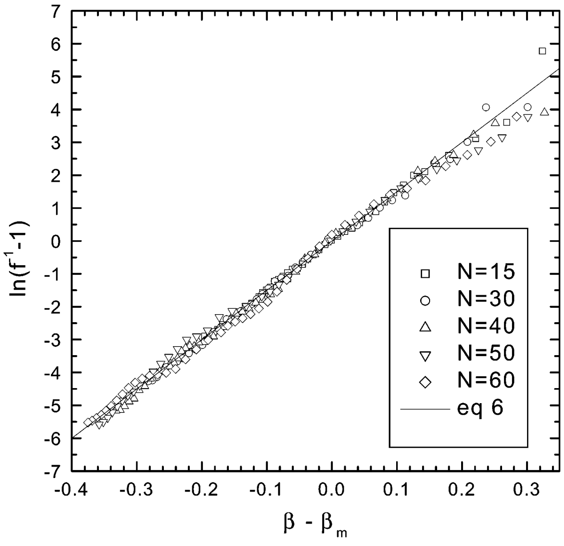

Figure 3. Melting curves for different chain lengths are replotted with $\ln(f^{-1} - 1)$ as a function of $(\beta - \beta_m)$ and compared to eq 6.

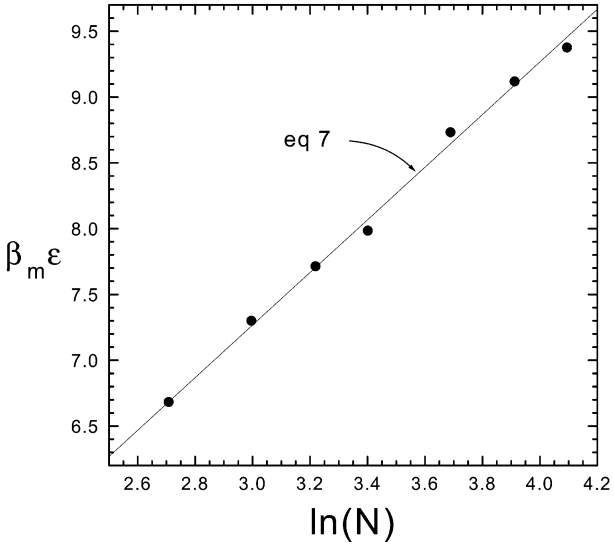

Figure 4. Inverse melting temperature is plotted against the chain length. The straight line denotes eq 7 with $\alpha = 2$.

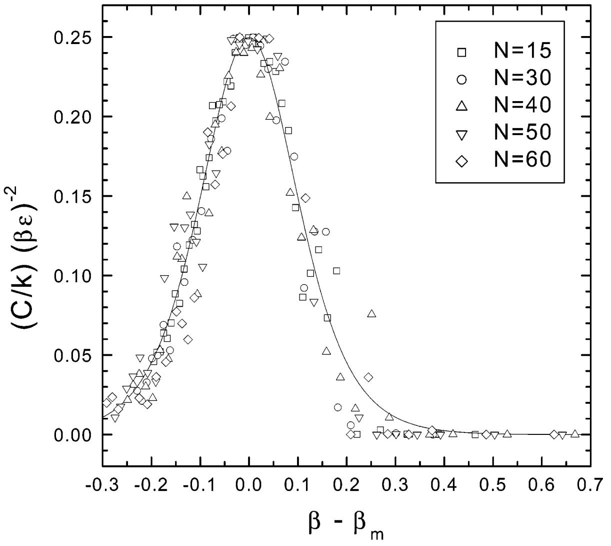

Figure 5. Heat capacity curves for different chain lengths are plotted with $(C/k)(\beta\epsilon)^{-2}$ as a function of $(\beta - \beta_m)$ and compared to eq 8.

exponent is in the range $2.0 \le \alpha \le 2.2$ for a swollen chain ($\nu = {}^3/_5$).[15] The current result implies that our contact distance in eq 2 leads to a local equilibrium condition, and hence $\alpha = 2$. Following this line of reasoning, one also expects that the contact distance would shift the melting curve and hence alter the melting temperature, as will be discussed later.

Using $\epsilon = 15$, eq 6 gives the melting curves in excellent agreement with those in Figure 1. The melting temperature obtained by the fitting procedure is consistent with that determined directly from the simulation data at $f = {}^1/_2$. Using the predetermined T_m, all the melting curves associated with different chain lengths collapse into a straight line with a slope ϵ as shown in Figure 3, when we plot $\ln(f^{-1} - 1)$ against $(\beta - \beta_m)$. The data become slightly scattered at low temperature because of the metastable structure of the closed state. The variation of the melting temperature $\beta_m\epsilon$ with chain length $\ln N$ is depicted in Figure 4. The data points can be well represented by a straight line with a slope α equal to 2. This result is predicted in eq 7 and consistent with that obtained from the chain length dependence of $k_{o\to c}{}^{-1}$ in Figure 2b.

It is natural to identify the peak temperature of heat capacity curve T_c as the transition temperature of open-to-closed transition. When we plot $(C/k)(\beta\epsilon)^{-2}$ against $(\beta - \beta_m)$ in Figure 5, all data points of different chain lengths collapse into a single curve as indicated by eq 8. Solving eq 8 yields $T_c \approx T_m/[1 + 4(\beta_m\epsilon)^{-2}]$ as $(\beta_c - \beta_m)\epsilon < 1$. Consequently, the peak temperature is always less than the melting temperature. Figure 1 and eq 6 show that open-to-closed transition is continuous for a chain of finite length. As the chain length N increases, the entropy loss of forming a closed loop also increases, and the formation of a closed conformation becomes more difficult. To maintain a finite T_m, the binding energy ϵ must also increase as well, according to eq 7. In reality, this is usually achieved by increasing the number of binding monomers at the stem. In the current model, we consider two single attractive ends with an effective binding energy ϵ which should increase logarithmically in N. In the thermodynamic limit of $N \to \infty$, the derivative of the melting curve yields $(T\, \mathrm{d}f/\mathrm{d}T)_{T=T_m} = (1/4)(\epsilon/kT_m) \to \infty$ based on eq 6. The smooth open-to-closed transition of a finite chain becomes a step function. In the meantime, T_c approaches T_m, and the peak position in heat capacity curve represents the transition temperature. One can further show that the free energy undergoes a finite jump across the closed-to-open transition temperature in the thermodynamic limit, indicating a first-order phase transition.

In this Communication, we have established a simple thermodynamic theory for the open-to-closed transition based on a two-state picture. It describes excellently the melting curves of a hard-sphere chains with attractive ends, which are obtained by MC simulations. For a chain with $N = 15$, we obtain $kT_m/\epsilon = 0.15$ for the contact distance $a = 1.2\sigma$. In a fluorescence experiment, Wallace et al.[6] attached a single strand of poly-T to stems consisting of five-base sequence TTGGG at one end and its complementary AACCC at the other end. They evaluated the apparent enthalpy change $\Delta H \approx -40kT$ at $T = 298$ K. If we adopt the binding energy $\epsilon = -\Delta H$ in our model, the estimation of the melting temperature is evidently too high. To obtain a reasonable melting temperature between the freezing and boiling temperatures of water, one must have $0.023 < kT_m/\epsilon < 0.031$. Since the melting temperature declines with decreasing the contact distance, we are able to obtain T_m within this range as $a \to \sigma$. For example, $kT_m/\epsilon \approx 0.10$ for $a = 1.01\sigma$ and $kT_m/\epsilon \approx 0.08$ for $a = 1.001\sigma$. This small value of the contact distance $(a - \sigma)$ seems unrealistic. It is because we have employed a spherical excluded volume associated with attractive interactions. In the ssDNA experiment, two stems can come much closer to each other. We have performed simulations for a hard-sphere chain of $N = 40$ with five attractive spheres at both ends. The melting temperature for $a = 1.2\sigma$ is approximately $kT_m/\epsilon \approx 0.037$, which is smaller but still not within reasonable range. Note that the effect of misfolded configuration, the AT bonding of the AA segment in the stem with any TT segment in poly-T loop, is not considered in the present model. Our MC studies show that such mismatched loop conformations can significantly reduce the melting temperature. The base-stacking interactions, which are also not accounted for here, lower the melting temperature furthermore.[5]

Finally, it is worth mentioning that our work could have relevance beyond the world of DNA. For example, it may be helpful in understanding the behavior of associating systems such as telechelic polymers.[16] Earlier work focus mainly on interactions between many chains below the melting temperature.[17] That is, the binding energy is large compared to the thermal energy. Our present study provides the effect of temperature kT/ϵ on conformational kinetics of a telechelic polymer.

Acknowledgment. Y.-J.S. and H.-K.T. thank NSC of Taiwan for financial support, and J.Z.Y.C. gratefully acknowledges financial support from National Center of Theoretical Science of Taiwan. We also thank B.-Y. Ha for bringing ref 13 to our attention.

References and Notes

(1) Wennmalm, S.; Edman, L.; Rigler, R. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 10641.

(2) Bonnet, G.; Krichevsky, O.; Libchaber, A. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 8602.

(3) Ansari, A.; Kuznetsov, S. V.; Shen, Y. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 7771.

(4) Zhang, W.; Chen, S.-J. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 1931.

(5) Goddard, N. L.; Bonnet, G.; Krichevsky, O.; Libchaber, A. *Phys. Rev. Lett.* **2000**, *85*, 2400.

(6) Wallace, M. I.; Ying, L.; Balasubramanian, S.; Klenerman, D. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 5584.

(7) Ying, L.; Wallace, M. I.; Klenerman, D. *Chem. Phys. Lett.* **2001**, *334*, 145.

(8) Zhou, Y.; Hall, C. K.; Karplus, M. *Phys. Rev. Lett.* **1996**, *77*, 2822.

(9) Bundschuh, R.; Hwa, T. *Phys. Rev. Lett.* **1999**, *83*, 1479.

(10) Doi, M.; Edwards, S. F. *The Theory of Polymer Dynamics*; Oxford University Press: New York, 1988.

(11) 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.

(12) Wilemski, G.; Fixman, M. *J. Chem. Phys.* **1974**, *60*, 866. Wilemski, G.; Fixman, M. *J. Chem. Phys.* **1974**, *60*, 878. Szabo, A.; Schulten, K.; Schulten, Z. *J. Chem. Phys.* **1980**, *72*, 4350.

(13) Pastor, R. W.; Zwanzig, R.; Szabo, A. *J. Chem. Phys.* **1996**, *105*, 3878.

(14) Doi, M. *Chem. Phys.* **1975**, *9*, 455.

(15) A. Dua; B. Cherayil, J. *Chem. Phys.* **2002**, *116*, 399.

(16) Winnik, M. A.; Yekta, A. *Curr. Opin. Colloid Interface Sci.* **1997**, *2*, 424.

(17) Baljon-Haakman, A. R. C.; Witten, T. A. *Macromolecules* **1992**, *25*, 2969.

MA025604X