I. INTRODUCTION

Theoretical studies on protein folding have often been performed with simple models that incorporate information on the native structure [1–13]. The Wako-Saitô-Muñoz-Eaton (WSME) model is one such example [1–4], described by Ising-like binary variables with long-range interactions on a one-dimensional lattice. The transfer matrix formalism was developed so that the exact partition function can be computed for any given temperature [9]. Since the partition function contains all the information on thermodynamics, various quantities relevant for the conformational transition of a protein can be calculated.

In this work I study the partition function zeros (PFZs) of the WSME models of β hairpins. Partition function zeros are much more sensitive indicators of phase transitions than real-valued quantities such as specific heat and the PFZ method has been used not only in equilibrium [14–27] and nonequilibrium statistical physics [28,29], but also in other fields as diverse as nuclear [30–32] and particle physics [33–35]. Although the WSME model is an exactly solvable model for studying protein folding [9], analysis of its PFZs is lacking; an exception is the recent work by the present author [27], where the features of zeros distinguishing two-state and barrierless folding transitions have been identified.

Although a β hairpin is a very simple structure, it captures nontrivial aspects of protein folding because contacts are formed between residues far away in the sequence. Therefore, β hairpins have been the subject of extensive experimental and computational research [2–4,36–43], including the study using the WSME model [2,3,9]. Exact PFZs for β hairpins were studied also in Ref. [27], under the restriction that the number of peptide bonds is even, the native contacts are antiparallel, and their strengths are uniform [Fig. 1(a)]. The native contacts with uniform strength can be considered as a backbone-hydrogen bond, so the model describes the structure without additional hydrophobic interactions of the side chains.

II. EXACT PARTICITION FUNCTION ZEROS OF THE WSME MODEL OF SIMPLE β HAIRPINS

The WSME model describes a peptide or protein of length \(N+1\) by an Ising-type variable \(m_i\), \((i = 1, \ldots, N)\), which denotes the state of the \(i\)th peptide bond connecting \(i\)th and \((i+1)\)th residues. The variable \(m_i\) takes the value 1 or 0 depending on whether the bond is in the ordered or disordered state. If the entropy of the ordered bond relative to the disordered one is denoted as \(\Delta s_i < 0\), then \(\lambda_i = \exp(-\Delta s_i) > 0\).
can be considered as the effective number of microstates of a disordered bond. We assume that the local entropy cost for ordering a bond is the same throughout the protein chain, which is a common assumption for a β hairpin [2,3,9], and write the effective number of disordered bond states as \( \lambda = \exp(-\Delta s) \).

The number of configurations of an ordered bond is 1 by definition. Note that \( \lambda \) does not have to be an integer in general.

The Hamiltonian of the WSME model is

\[
H(m_k) = \sum_{i=1}^{N+1} \sum_{j=i+1}^{N} \epsilon_{ij} \Delta_{ij} \prod_{k=i}^{j} m_k, \tag{1}
\]

where \( \epsilon_{ij} \) is the contact energy of the \( i \)-th and \( j \)-th bonds and \( \Delta_{ij} \) is 1 only if the bonds are in contact in the native structure and 0 otherwise. Thus the contact energy is assigned if and only if the corresponding pair of bonds is in contact in the native structure and the stretch of sequence between them are all in the ordered state. The contact energy \( \epsilon_{ij} \) can represent either the backbone hydrogen bond or hydrophobic interactions between the side chains. The density of the states \( \Omega(E) \) for given energy value \( E \) is computed using this Hamiltonian and the partition function zeros are then obtained by solving a polynomial equation using these densities as coefficients.

We first concentrate on simplified classes of β hairpins in which the \( i \)-th bond forms native contacts only with the \( (N - i + 1) \)-th bond at the opposite side of the hairpin. The model with an even value of \( N \) [Fig. 1(a)] has an advantage that an analytic formula can be obtained for the partition function zeros in the limit of large \( \lambda \) [27]. However, the structure is rather unrealistic in that all the peptide bonds participate in the antiparallel hydrogen or hydrophobic bonds. All the known experimental structures of β hairpins have extra bonds in the turn region. Therefore, I compute the PFZs for the structure with an odd value of \( N \) [Fig. 1(b)] that has an additional peptide bond at the turn. The analytic formula for the density of states can be obtained as a simple extension of that for the even number of peptide bonds, but the polynomial equation for the zeros can be solved only numerically.

When the lines of native contacts cross each other as in Fig. 1(c), the analytic formula for the density of states becomes more complicated and consequently less useful. However, one can easily compute the exact density of states and partition function zeros for a given set of parameters, using a transfer matrix [9].

Let us call the contact between the \( i \)-th and \( (N - i + 1) \)-th bonds the \( i \)-th contact and rewrite the corresponding energy as \( \epsilon_i \equiv \epsilon_{i,N-i+1} \) \( [i = 1, \ldots, (N - 1)/2] \) for simplicity of notation. The broken native contacts can appear only as a sequential stretch in the tip region due to the restriction that the native contacts can form only when all the intervening bonds are ordered. Suppose that the \( i \)-th native contacts with \( i > j \) are all formed and those with \( i \leq j \) are all broken. The corresponding energy value is

\[
E_j = \sum_{i=j+1}^{n} \epsilon_i = E_F - \sum_{i=1}^{j} \epsilon_i \quad (0 \leq j \leq n), \tag{2}
\]

where \( n \equiv (N - 1)/2 \) and \( E_F \equiv \sum_{i=1}^{n} \epsilon_i \) is the energy value of the fully folded conformation. If \( j < n \), then at least one of the bonds forming the \( j \)-th contact has to be disordered, so they cannot both be in the ordered states. Therefore, the total number of states that these pairs can be in is

\[
\omega_j = (\lambda + 1)^2 - 1. \tag{3}
\]

The \( n \)-th native contact is special in that it can be broken due to the disorder in the \( (n + 1) \)-th peptide bond at the turn. When the \( (n + 1) \)-th bond is in the native conformation, at least one of the residues in the \( n \)-th pair should be in a non-native state, leading to \((\lambda + 1)^2 - 1\) states. In contrast, if the \( (n + 1) \)-th bond is in one of \( \lambda \) unfolded states, then the residues forming the \( n \)-th pair can be any of the \((\lambda + 1)^2\) conformations. Therefore, the total number of conformations for the three bonds at the turn regions is

\[
\omega_n = (\lambda + 1)^2 - 1 + \lambda(\lambda + 1)^2 = (\lambda + 1)^3 - 1. \tag{4}
\]

All the other bonds with broken native contacts can be in any of \( \lambda + 1 \) states, whereas those forming the native contact are in the ordered state, whose number is 1 by definition. By multiplying these numbers by \( \omega_j \) for \( j > 0 \), the total number of conformations for a given value of \( j \) is obtained as

\[
\Omega(E_j; \lambda) = \begin{cases} 
1 & (j = 0) \\
[\lambda(\lambda + 1)^2 - 1](\lambda + 1)^{2j/2} & (1 \leq j < n) \\
[\lambda(\lambda + 1)^3 - 1](\lambda + 1)^{2n-2} & (j = n),
\end{cases} \tag{5}
\]

where \( j = 0 \) corresponds to the fully folded conformation. The last line of (5) is the only difference from the density of states obtained for even value of \( N \).

If all the native contacts are due to hydrogen bonds, we may assign an equal energy value \( \epsilon_i = \epsilon < 0 \) to each contact and the partition function zeros are obtained by solving the polynomial equation

\[
Z(z) = \sum_{j} \Omega(E_j; \lambda) z^j = 0, \tag{6}
\]

where \( z \equiv e^{\beta \epsilon} \). The solution for the even value of \( N \) was obtained analytically in Ref. [27] as

\[
z_j = \frac{1}{(\lambda + 1)^{2j}} \exp \left( \frac{2\pi i j}{N/2 + 1} \right) \quad (j = 1, \ldots, N/2) \tag{7}
\]

under the approximation

\[
\frac{\lambda^2 + 2\lambda + 1}{\lambda^2 + 2\lambda} \simeq 1. \tag{8}
\]

The normalized zeros \((\lambda + 1)^2 z_j\) for \( N = 14 \) are displayed in Fig. 2 as intersections of the straight lines with the unit circle. The exact numerical solution for \( N = 14 \) and 15 are also plotted as closed and open symbols; those for \( N = 14 \)
EXACT PARTITION FUNCTION ZEROS OF THE WAKO-

FIG. 2. Partition function zeros of the WSME model of β hairpins with seven native contacts and uniform interaction strengths in the complex plane of \((\lambda + 1)^2z\). The analytic solution (7) lies on the intersection of the unit circle and the straight lines, which fits quite well with the numerical solutions for \(N = 14\). The zeros for \(N = 15\) are located inside the circle due to the entropy cost for disordering the extra bond at the turn region. The introduction of the entropic barrier also makes the angular distribution closer to the uniform distribution.

agree quite well with the analytic solution even for the extreme case of \(\lambda = 1\) \((\Delta s = 0)\).

In contrast, the zeros for \(N = 15\), corresponding to the same number of native contacts as \(N = 14\), are located inside the circle because now the fully unfolded structure is more favorable compared to that with an even \(N\) due to the additional entropic contribution from the extra bond at the turn. The deviation from \(N = 14\) increases as \(\lambda\) increases, as expected. The gap of the distribution of zeros, defined as the separation between the first zeros relative to the average separation between the zeros, is a signature of the low free energy barrier at the transition temperature, indicating a first-order-like barrierless (or weak barrier) transition [27].

For the same number of native contacts we see that the gap for the odd value of \(N\) is smaller than that for the even \(N\), which is due to the introduction of the entropic barrier between the fully unfolded state and the rest of the states.

III. PARTITION FUNCTION ZEROS OF β HAIRPINS WITH HYDROPHOBIC CORES

By introducing hydrophobic interactions in addition to the hydrogen bond, we can observe the collapse transition to

\[ Z = z^{-p_{\text{HB}}} \sum_{j=0}^{h-1} (\lambda + 1)^2 j z^{pj} \]

where the approximation (8) is used. The first and the second factors give two concentric circles for zeros:

\[ z_j = \frac{1}{(\lambda + 1)^2/p} \exp \left( \frac{2\pi i (j + hk)}{hp} \right) \quad (j = 1, \ldots, h-1), (k = 0, \ldots, p - 1), \]  \(\tag{10}\)

The angular distribution of the zeros on the inner circle has the same form as the one without the hydrophobic core, corresponding to the first-order-like barrierless transition [27], whereas that on the outer circle is a uniform distribution corresponding to the two-state transition, due to the energy cost of breaking the hydrophobic core during the transition from the intermediate to the fully unfolded state. It is easy to see, from the analytic solution of the loci (10), that the folding and collapse transition occurs at

\[ T_f = -\epsilon/p/[2k_B \ln(\lambda + 1)] \]

and

\[ T_c = -\epsilon/(hp + q)/[2k_B \ln(\lambda + 1)] \].

As expected, a larger value of \(|\Delta G_{\text{SC}}|\) corresponds to a higher value of \(T_c\) and a denser distribution of zeros on the outer locus, signifying a sharper collapse transition. Also, for \(q = 0\), the two circles

\[ \exp \left( \frac{2i(j+ik)}{hp+q} \right) \quad (j = 0, \ldots, hp + q - 1). \]
collapse into one circle corresponding to the folding transition (7) by setting \( p = 1 \) without loss of generality. The exact partition function zeros for \( n = 9, p = 1, q = 2, \) and \( \lambda = 2, \) with an extra hydrophobic interaction at the \( k \)th contact, are plotted on a complex \( z \) plane, in Fig. 4, along with the circles at radii \( 1/(\lambda + 1)^{2/p} = 1/9 \simeq 0.111 \) and \( 1/(\lambda + 1)^{2/(\lambda p + q)} = 1/9^{5/7} \simeq 0.208. \) We see that the zeros for \( k = 5 \) are extremely well described by the analytic solution (10), being distributed on the inner and outer circles at an angular interval of \( 2\pi /hp = 2\pi /5 \) (4\( \pi /hp = 4\pi /5 \) between the first zeros) and \( 2\pi /hp + q = 2\pi /7, \) respectively. We see that as the position of the hydrophobic core is moved toward the tip, the radius of the outer locus decreases because the intermediate becomes unfavorable entropically. In addition, the density of zeros at the inner locus decreases because the intermediate and the fully folded conformation become less distinguishable [Fig. 3(a)]. Eventually, at \( k = 1, \) the zeros form one locus corresponding to the folding transition. In contrast, as the hydrophobic core moves toward the turn, the radius of the outer locus increases and its density decreases because the entropy of the intermediate increases and it becomes less distinguishable from the unfolded state [Fig. 3(b)].

**IV. PARTITION FUNCTION ZEROS OF A REAL \( \beta \) HAIRPIN**

So far we have concentrated on a general class of simplified hairpin models. A real \( \beta \) hairpin, the 16 C-terminal residues of streptococcal protein G B1, was also studied with the WSME model [2,3,9], which includes crossed lines of native contacts [Fig. 1(c)]. We compute the density of states using the transfer matrix formalism [9], where the native contacts are given in Refs. [2,3]. The hydrogen bond and hydrophobic interaction energies are \( \Delta H_{HB} = -1.1 \) kcal/mol and \( \Delta G_{SC} = -2.0 \) kcal/mol and the local entropic cost of folding is \( \Delta s = -3.12 \) cal/Kmol, which corresponds to \( p = 11 \) and \( q = 20 \) with \( \epsilon = -0.1 \) kcal/mol and \( \lambda = 4.80. \) The zeros are obtained as the solutions to a 137th-order polynomial equation, which are plotted in the plane of \( e^{i\epsilon} \) in Fig. 5. We see that the loci of zeros also form concentric circles with radii 0.727 and 0.845, corresponding to the temperatures \( T = 158 \) and 299 K.

Note that since a real \( \beta \) hairpin is an intrinsically finite system, there is no phase in a rigorous sense. The PFZs may provide a generalized definition of a conformational transition in a finite-size system [17,18]. Here, since the loci are circles, we simply define conformational transitions to exist at the intersections of these circles with the positive real axis.

The transition temperature \( T = 299 \) K is quite consistent with the value of 297 K, which was defined in terms of kinetic rates and unfolding curves in Ref. [2]. A two-state transition behavior at this temperature was reported experimentally [2] and also confirmed by theoretical study using the WSME model [9]. In contrast, the distribution of the zeros on the inner locus is not only nonuniform but also very sparse, especially near the positive real axis. Therefore, any possible transition from the hydrophobically collapsed intermediate to the fully folded state near \( T \sim 158 \) K might be quite smooth, making it hard to be connected with experimental data on real-value quantities such as a peak in the specific heat. Moreover, the WSME model is based on native structure that is obtained by experiments performed at temperature far above \( T \sim 158 \) K, so the validity of extrapolating the WSME model to such a low temperature may need to be put under further scrutiny.
V. CONCLUSION

I computed the partition function zeros of β hairpins in the framework of the Wako-Saitō-Muñoz-Eaton model, using both analytic and numerical methods. The zeros for the β hairpin with an odd number of peptide bonds were computed, which is much more realistic than the one where all the peptide bonds participate in the hydrogen bonds. The distribution of the zeros exhibits features that correspond to an additional entropic barrier between the fully unfolded conformation and the rest of the states. By introducing a hydrophobic core, the zeros for a hairpin that undergoes multiple transitions could be obtained, consisting of concentric circles. By moving the position of the hydrophobic core toward either tip of the turn region, one of the circles is seen to dissolve, which has a clear interpretation: the intermediate structure becomes indistinguishable from either the fully folded or fully unfolded states. The zeros of a real β hairpin, the 16 C-terminal residues of streptococcal protein G B1, were also numerically computed, where the structure of the hydrophobic core is more complex due to native contacts that cross each other. I found that these zeros also lie on concentric circles, the difference from the simpler cases being that the distribution of the zeros on the inner circle is not uniform. Also the zeros on the inner circle are rather sparse, indicating that the corresponding transition has an extremely weak cooperativity.

ACKNOWLEDGMENTS

The author thanks Seung-Yeon Kim for useful discussions. This study was supported by a Grant No. A120757 of the Korea Healthcare Technology R&D, Ministry of Health and Welfare, Republic of Korea.


022710-5