The Theory of Melting in Heteropolymers. I.
Random Chains[1]

A. Vilenkin[2,3]

A method is developed for the calculation of the ground-state melting curves (η_0 vs. T) for random, infinite heteropolymers. Here η_0 is the ground-state value of the fraction of melted links η ($\eta_0 \approx \eta$ in the strong cooperativity approximation). It is shown that the differential melting curves ($d\eta/dT$ vs. T) can have a fine structure in the form of several peaks on the bell-shaped main curve. Positions, magnitudes, and widths of these peaks are estimated. The "accidental" fine structure of melting curves, which is caused by a finite length of the polymer, is briefly discussed.

1. INTRODUCTION

The melting of polynucleotides has been a subject of interest over the last decade. A detailed analysis of various models describing this process can be found in review articles.[1,2] In the simplest case of circular polymers, the melting is described by a one-dimensional, two-component Ising model with the effective Hamiltonian[3]

$$H = H^h + \tfrac{1}{2} \sum_{k=1}^{\mathcal{N}} [-(1 - \sigma_k)U(a_k) + \tfrac{1}{2}V(1 - \sigma_k\sigma_{k+1})] \qquad (1)$$

Here H^h is the polymer energy in the completely helical state; $a_1, a_2, ..., a_{\mathcal{N}}$

is a fixed base pair sequence, which we shall assume to be random; $a_k = 1$ and $a_k = 2$ stand for A–T and G–C pairs, respectively; $\sigma_k = 1$ if the kth link (base pair) is in the helical state and $\sigma_k = -1$ if it is melted; $U(1) \equiv U_1$ and $U(2) \equiv -U_2$ are the (free) energy differences between the helical and melted states for links of type 1 and type 2, respectively; $V/2$ is the "surface" energy arising at the phase boundaries. U_1 and U_2 are functions of temperature. A linear temperature dependence is usually assumed, but we shall not specify the form of $U_k(T)$.

The values of U_1 and U_2 can vary widely with change of environmental conditions.[4] On the other hand, the surface energy V and the melting temperature T_m are relatively independent of environment (in terms of percentage), V being much larger than T_m. The ratio

$$T/V \ll 1 \tag{2}$$

is a small parameter of the problem. This inequality means that the temperature is low and thus the polymer is close to its ground state, i.e., to the state minimizing the Hamiltonian.[5–7] This state, however, changes with temperature because the parameters U_1 and U_2 are temperature dependent.

The polymer microstate is completely characterized by indicating the ends of all melted sections. The necessary and sufficient conditions[6,7] determining the melted and helical sections in the ground state are easily formulated in terms of the function

$$h(M) = \sum_{k=0}^{M} U(a_k) \tag{3}$$

where the origin ($k = 0$) can be chosen arbitrarily. Each polymer segment (μ, ν) can be characterized by the quantity $h(\mu, \nu) \equiv h(\nu) - h(\mu) = KU_1 - LU_2$, where K and L are the numbers of component-1 and component-2 links in the segment. A segment (μ, ν) is a melted section if (1) it has $h(\mu, \nu) \geqslant V$, (2) it has no segments with $h < -V$ inside it, (3) $h(\mu) \leqslant h(M) \leqslant h(\nu)$ for $\mu < M < \nu$. A segment (μ, ν) is a helical section if (1) it has $h(\mu, \nu) < -V$, (2) it has no segments with $h \geqslant V$ inside it, (3) $h(\nu) < h(M) < h(\mu)$ for $\mu < M < \nu$. It is easily verified that any violation of these requirements increases the polymer energy.

A new melted section can appear in the ground state in one of three possible ways: (A) it can appear as a whole (an isolated newborn section), (B) it can form a left-hand or right-hand part of an already existing melted section, (C) it can connect two already existing melted sections. The newborn sections of types (A)–(C) satisfy the following conditions: (A) $KU_1 - LU_2 = V$, (B) $KU_1 - LU_2 = 0$, (C) $KU_1 - LU_2 = -V$, respectively. The fraction of melted links in the ground state of an infinite heteropolymer, $\eta_0(T)$, has jumps at all points satisfying one of the conditions (A)–(C) with arbitrary integers $K, L \geqslant 0$.[8] It is easily seen from the equation for condition (B) that

η_0 has a jump at every rational value of the ratio U_1/U_2. Physically, there is no difference between rational and irrational numbers. If $U_1/U_2 = r/s$, where r and s are integers and r/s is an irreducible fraction, then one expects that the jumps of η_0 become arbitrarily small as r, $s \to \infty$ (with finite r/s).[4] The maximal jumps of η_0 corresponding to sections of type (B) are expected at the minimal values of r and s. One can easily convince oneself that the maximal jumps corresponding to sections of type (A) occur when $KU_1 - LU_2 = V$ and at the same time[5] the ratio U_1/U_2 is equal to the ratio of two small integers[6] [and a similar condition for sections of type (C)].

The jumps of η_0 correspond to the peaks on the differential melting curve ($d\eta/dT$ vs. T).[7] These peaks form what we shall call the intrinsic fine structure of melting curves, as distinguished from the "accidental" fine structure, which is caused by the finite length of the polymer.[9]

To explain the origin of the intrinsic fine structure, let us consider the newborn sections of type (B) in some more detail. If $U_1/U_2 = 1$, then the equation of condition (B) gives $K = L$. The possible values of K and L are $(K, L) = (1, 1), (2, 2),\dots$ The lengths of the corresponding newborn sections range from $K + L = 2$ to infinity. In the case $U_1/U_2 = 99/100$ we have $(K, L) = (100, 99), (200, 198),\dots$, the minimal length being 199. We see that the "density" of possible pairs (K, L) in the first case is much greater than that in the second case. In other words, the probability of a newborn section of type (B) in the case $U_1/U_2 = 1$ is much greater than that in the case $U_1/U_2 = 99/100$. Generally, one expects that this probability has maxima at the values of temperature at which U_1/U_2 is equal to the ratio of two small integers. The existence of such maxima explains the intrinsic fine structure of melting curves. Strictly speaking, our discussion corresponds to the low-temperature limit ($T \ll U_1, U_2, V$). If $T \gtrsim U_1, U_2$, then the function $\eta_0(T)$ with multiple sharp peaks on it gives, so to say, a "bare" melting curve. On the "dressed" melting curve $\eta(T)$, all the peaks acquire a finite width, so that they overlap with each other and only the largest peaks (if any) can survive.

---

[4] It can be shown that $\Delta\eta_0$ is an exponentially decreasing function of $r + s$.[15]

[5] It is not necessary that both conditions be satisfied exactly at the same temperature. The quantity U_1/U_2 has only to be close enough to the ratio of two small integers when $KU_1 - LU_2 = V$ (see Section 5).

[6] Obviously, only few (if any) such points can occur in the melting interval. If there are none, one can adjust the sodium ion concentration (see Ref. 4) so that conditions for principal peaks of types (A) or (C) are satisfied for at least one point.

[7] The physical origin of the fine structure of melting curves in the case of finite sequences was first explained by Azbel' as related to the coiling of the ground-state segments (Ref. 6; see also Refs. 10 and 11). Prof. Azbel' had completed this work and reported it at the Moscow Seminar on Condensed State before the previous paper of the present author[8] was submitted for publication. I am grateful to Prof. Azbel' for bringing this fact to my attention and regret that a reference to his work is missing in Ref. 8.

In this paper the melting of random heteropolymers is studied, using the method developed in Ref. 8.[8] A system of equations is derived for the calculation of the fraction of melted links in the ground state, η_0, and its jumps, $\Delta\eta_0$. To estimate the magnitudes and the widths of the peaks on the differential melting curves, these equations are solved in the simplest case $U_1 = U_2$. In the case of a finite heteropolymer, the fluctuations of η_0 corresponding to the accidental fine structure are estimated. The possibilities of experimental detection of the intrinsic fine structure are briefly discussed.

2. GENERAL FORMULAS

Assuming that U_1/U_2 is a rational number, we have $U_1 = rU$ and $U_2 = sU$, with r and s integers. It is convenient to introduce a dimensionless parameter $N \equiv V/U$ and a dimensionless function $n(M) \equiv U^{-1}h(M)$. This function can assume arbitrary integral values. For a polymer with a random sequence of links of the two types, the graph of $n(M)$ represents a Brownian motion of a "particle," n playing the role of coordinate and M the role of time. The "particle" steps r units up with probability w_1 and s units down with probability $w_2 = 1 - w_1$, where w_1 and w_2 are the concentrations of links of the first and of the second type, respectively. The fraction of melted links in the ground state, η_0, can be expressed in terms of the probabilities of certain Brownian trajectories of the "particle."[7,8]

Let us denote by $\mathscr{P}_m^+$ $(\mathscr{P}^-, \tilde{\mathscr{P}}^-, q_k, q_0)$ the probability that the "particle," starting from the point $n = 0$, will for the first time get to the point $n = N + m$ ($n < -N; n < -N; n = k; n = 0$) without previously leaving the domain $0 \leqslant n \leqslant N - 1$ ($-N \leqslant n \leqslant 0; -N \leqslant n \leqslant 1; -N \leqslant n \leqslant 0; -N \leqslant n \leqslant -1$). The parameters m and k take values from 0 to $r - 1$ and from 1 to r, respectively. For simplicity, the definitions above are given for the case of integral N. If N is not an integer, then the probabilities $\mathscr{P}_m^+,...$ ought to be replaced by

$$\mathscr{P}_m^+(N) \to \mathscr{P}_m^+([N] + 1), \qquad \mathscr{P}^-(N) \to \mathscr{P}^-([N])$$
$$\tilde{\mathscr{P}}^-(N) \to \tilde{\mathscr{P}}^-([N]), \qquad q_k(N) \to q_k([N]), \qquad q_0(N) \to q_0([N]) \tag{4}$$

where $[N]$ is the whole part of N. The probabilities $\mathscr{P}^-$, $\tilde{\mathscr{P}}^-$, q_k, and q_0 are connected by the relations

$$\mathscr{P}^- = 1 - \sum_{l=1}^{r} q_l, \qquad \mathscr{P}^-(1 - q_0) = \tilde{\mathscr{P}}^- \tag{5}$$

which can be proved proceeding from the definitions given above.

---

[8] It should be mentioned that this method can be used for studying disordered magnetic chains as well (see Ref. 12).

Each melted section has a segment of type $\mathscr{P}_m{}^+$ at the beginning, then a number of segments of types q_k, and then a number of segments of type q_0.[8] Such a section can be characterized by the numbers $m, n_1, n_2,..., n_r, n_0$, where n_k and n_0 are the numbers of segments of types q_k and q_0, respectively. The probability that an arbitrarily chosen link of the chain is the beginning of such a section is given by

$$\mathscr{P}(m, n_k, n_0) = \mathscr{P}_m{}^+ (\tilde{\mathscr{P}}^-)^2 q_0^{n_0} \prod_{k=1}^{r} q_k^{n_k} \left(\sum_{l=1}^{r} n_l \right)! \left(\prod_{i=1}^{r} n_i! \right)^{-1} \tag{6}$$

The average length of such a section is

$$L(m, n_k, n_0) = L_m{}^+ + \sum_{k=1}^{r} n_k l_k + n_0 l_0 \tag{7}$$

where $L_m{}^+$, l_k, and l_0 are the average lengths of the segments corresponding to the probabilities $\mathscr{P}_m{}^+$, q_k, and q_0, respectively. The factor $(\sum_{l=1}^{r} n_l)! / \prod_{i=1}^{r} n_i!$ in Eq. (6) takes account of all possible transpositions of segments with different values of k.

The fraction of melted links in an infinite polymer is given by

$$\eta_0 = \sum_{m=0}^{r-1} \sum_{n_k=0}^{\infty} \sum_{n_0=0}^{\infty} \mathscr{P}(m, n_k, n_0) L(m, n_k, n_0) \tag{8}$$

Using Eqs. (6) and (7), we can rewrite Eq. (8) in the form

$$\eta_0 = (\tilde{\mathscr{P}}^-)^2 \sum_{m=0}^{r-1} \mathscr{P}_m{}^+ \left(L_m + \sum_{k=1}^{r} l_k q_k \frac{\partial}{\partial q_k} + l_0 q_0 \frac{\partial}{\partial q_0} \right) \Phi(q) \Phi_0(q_0) \tag{9}$$

where

$$\Phi(q) = \sum_{n_1=0}^{\infty} \cdots \sum_{n_r=0}^{\infty} \left\{ \left(\sum_{k=1}^{r} n_k \right)! \prod_{i=1}^{r} \frac{q_i^{n_i}}{n_i!} \right\} \tag{10}$$

$$\Phi_0(q_0) = \sum_{n_0=0}^{\infty} q_0^{n_0} = (1 - q_0)^{-1} \tag{11}$$

We shall prove that[8]

$$\Phi(q) = \left(1 - \sum_{k=1}^{r} q_k \right)^{-1} \tag{12}$$

Indeed

$$\left(1 - \sum_{k=1}^{r} q_k \right)^{-1} = \sum_{n=0}^{\infty} \left(\sum_{k=1}^{r} q_k \right)^n$$

$$= \sum_{n=0}^{\infty} \sum_{\{\Sigma n_k = n\}} \frac{n!}{n_1! \, n_2! \cdots n_r!} q_1^{n_1} q_2^{n_2} \cdots q_r^{n_r}$$

$$= \sum_{n_1=0}^{\infty} \cdots \sum_{n_r=0}^{\infty} \left\{ \left(\sum_{k=1}^{r} n_k \right)! \prod_{k=1}^{r} \frac{q_k^{n_k}}{n_k!} \right\}$$

Substituting Eqs. (11) and (12) in Eq. (9) and using the identities (5), we obtain the final expression for η_0:

$$\eta_0 = \sum_{m=0}^{r-1} \mathscr{P}_m^+ \left[\tilde{\mathscr{P}}^- L_m + (1 - q_0) \sum_{k=1}^{r} l_k q_k + \mathscr{P}^- q_0 l_0 \right] \tag{13}$$

Now let us calculate the number of links that change their states at certain values of r, s, N. Suppose that N is an integer (N can be made an integer with whatever accuracy by a suitable choice of r and s). Then the number of links in the newborn sections of types (A)–(C) is given by, respectively,

$$\Delta\eta_A = (\tilde{\mathscr{P}}^-)^2 \mathscr{P}_0^+ \sum_{n=0}^{\infty} q_0^n (L_0^+ + nl_0) = (\mathscr{P}^-)^2 \mathscr{P}_0^+ [(1 - q_0) L_0^+ + q_0 l_0] \tag{14}$$

$$\Delta\eta_B = (\tilde{\mathscr{P}}^-)^2 \tilde{\mathscr{P}}^+ \sum_{m,n=0}^{\infty} q_0^{m+n} (m+n) l_0 = 2(\mathscr{P}^-)^2 \mathscr{P}^+ q_0 l_0 \tag{15}$$

$$\Delta\eta_C = (\tilde{\mathscr{P}}^+)^2 \mathscr{P}_0^- \sum_{n=0}^{\infty} q_0^n (L_0^- + nl_0) = (\mathscr{P}^+)^2 \mathscr{P}_0^- [(1 - q_0) L_0^- + l_0 q_0] \tag{16}$$

Here $\tilde{\mathscr{P}}^+$ ($\mathscr{P}_0^-$) is the probability that the "particle," starting from the point $n = 0$, will for the first time get to the point with $n > N$ ($n = -N$) without previously leaving the domain $1 \leqslant n \leqslant N$ ($-N + 1 \leqslant n \leqslant 0$),

$$\mathscr{P}^+ = \tilde{\mathscr{P}}^+ (1 - q_0)^{-1} \tag{17}$$

L_0^- is the average length of the segment corresponding to the probability $\mathscr{P}_0^-$. The $\mathscr{P}_0^-$ and L_0^- can be obtained from $\mathscr{P}_0^+$ and L_0^+ by interchanging w_1, w_2 and r, s:

$$\begin{aligned} \mathscr{P}_0^-(w_1, w_2, r, s) &= \mathscr{P}_0^+(w_2, w_1, s, r) \\ L_0^-(w_1, w_2, r, s) &= L_0^+(w_2, w_1, s, r) \end{aligned} \tag{18}$$

The total jump of η_0 is given by

$$\Delta\eta_0 = \Delta\eta_A + \Delta\eta_B + \Delta\eta_C \tag{19}$$

A system of equations for the probabilities $\mathscr{P}_m^+, \mathscr{P}^-, \tilde{\mathscr{P}}^-$, and q_k has been determined in Ref. 8. In fact, there is a simpler system of equations that is equivalent to the initial system. These equations, together with equations for the average lengths, are given in the next section.

3. EQUATIONS FOR PROBABILITIES AND FOR AVERAGE LENGTHS

The probabilities $\mathscr{P}_m{}^+$ and the average lengths $L_m{}^+$ can be found from the equations

$$\phi_{M,n} = w_1\phi_{M-1,n-r} + w_2\phi_{M-1,n+s} \qquad (0 \leqslant n \leqslant N-1) \qquad (20a)$$

$$\phi_{M,n} = 0 \qquad (n \leqslant -1; \quad n \geqslant N) \qquad (20b)$$

$$\phi_{0,n} = \delta_{0,n} \qquad (20c)$$

$$\mathscr{P}_m{}^+ = w_1 \sum_{M=0}^{\infty} \phi_{M,N+m-r} \qquad (0 \leqslant m \leqslant r-1) \qquad (21)$$

$$L_m{}^+ = (w_1/\mathscr{P}_m{}^+) \sum_{M=0}^{\infty} (M+1)\phi_{M,N+m-r} \qquad (22)$$

where $\delta_{m,n} = 1$ if $m = n$ and $\delta_{m,n} = 0$ if $m \neq n$; $\phi_{M,n}$ is the probability that the "particle" has coordinate n at time M; and M and n are integers. Equations (20b) and (20c) play the role of boundary and initial conditions for Eq. (20a). It is easily verified that Eqs. (20) are equivalent to the more complicated system used in Ref. 8.

Equations for q_k, l_k:

$$\phi_{M,n} = w_1\phi_{M-1,n-r} + w_2\phi_{M-1,n+s} \qquad (-N \leqslant n \leqslant 0) \qquad (23a)$$

$$\phi_{M,n} = 0 \qquad (n \geqslant 1; \quad n \leqslant -N-1) \qquad (23b)$$

$$\phi_{0,n} = \delta_{0,n} \qquad (23c)$$

$$q_k = w_1 \sum_{M=0}^{\infty} \phi_{M,k-r} \qquad (k = 1,...,r) \qquad (24)$$

$$l_k = (w_1/q_k) \sum_{M=0}^{\infty} (M+1)\phi_{M,k-r} \qquad (25)$$

Equations for q_0, l_0:

$$\phi_{M,n} = w_1\phi_{M-1,n-r} + w_2\phi_{M-1,n+s} \qquad (-N \leqslant n \leqslant -1) \qquad (26a)$$

$$\phi_{M,n} = 0 \qquad (n \geqslant 0, \quad n \leqslant -N-1) \qquad (26b)$$

$$\phi_{0,n} = w_2\delta_{n,-s} \qquad (26c)$$

$$q_0 = w_1 \sum_{M=0}^{\infty} \phi_{M,-r} \qquad (27)$$

$$l_0 = (w_1/q_0) \sum_{M=0}^{\infty} (M+2)\phi_{M,-r} \qquad (28)$$

Similar equations can be written for $\mathscr{P}^-$ and $\tilde{\mathscr{P}}^-$. However, this is not neces-sary: $\mathscr{P}^-$ and $\tilde{\mathscr{P}}^-$ can be found from the identities (5).

Equations (13) and (20)–(28) can be used for the direct calculation of melting curves on computers. Of course, in this case the summation over M in Eqs. (22), (24), etc., has to be cut off at some value M_0 which is much larger than the corresponding average length.

4. CALCULATION OF THE AVERAGE LENGTHS

To estimate the jumps $\Delta\eta_0$, we shall determine the average lengths L_0, l_0, and l_1 in the simplest case, $U_1 = U_2 = U$, or $r = s = 1$. In this case Eqs. (20)–(22) for $\mathscr{P}_m^+$ and L_m^+ take the form

$$\phi_{M,n} = w_1\phi_{M-1,n-1} + w_2\phi_{M-1,n+1} \qquad (0 \leqslant n \leqslant N-1) \qquad (29\text{a})$$

$$\phi_{M,n} = 0 \qquad\qquad\qquad\qquad (n \geqslant N, \quad n \leqslant -1) \qquad (29\text{b})$$

$$\phi_{0,n} = \delta_{0,n} \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad (29\text{c})$$

$$\mathscr{P}_0^+ = \sum_{M=1}^{\infty} w_1\phi_{M,N-1} \qquad\qquad\qquad\qquad\qquad (30)$$

$$L_0^+ = (w_1/\mathscr{P}_0^+) \sum_{M=1}^{\infty} (M+1)\phi_{M,N-1} \qquad\qquad (31)$$

Let us introduce the notations

$$\phi_n = \sum_{M=1}^{\infty} \phi_{M,n} \qquad\qquad\qquad\qquad\qquad (32)$$

$$\psi_n = \sum_{M=1}^{\infty} (M+1)\phi_{M,n} \qquad\qquad\qquad (33)$$

The equations for ϕ_n can be obtained by summing the left- and right-hand parts of Eqs. (29a)–(29b) over all M from $M = 1$ to infinity:

$$\phi_n = w_1\phi_{n-1} + w_2\phi_{n+1} \qquad (2 \leqslant n \leqslant N-1) \qquad (34\text{a})$$

$$\phi_1 = w_1\phi_0 + w_2\phi_2 + w_1 \qquad\qquad\qquad\qquad (34\text{b})$$

$$\phi_0 = w_2\phi_1 \qquad\qquad\qquad\qquad\qquad\qquad\qquad (34\text{c})$$

$$\phi_n = 0 \qquad\qquad\qquad\qquad (n \geqslant N) \qquad\qquad (34\text{d})$$

The general solution of Eq. (34a) is

$$\phi_n = A_1 + A_2X^n \qquad (1 \leqslant n \leqslant N) \qquad\qquad (35)$$

where A_1 and A_2 are arbitrary constants, $X = w_1/w_2$. Substituting Eq. (35)

into the boundary conditions (34b) and (34d) and using Eq. (34c) to eliminate ϕ_0, we obtain

$$A_1 + A_2 X^{-1} = 1/\omega_1, \qquad A_1 + A_2 X^N = 0 \tag{36}$$

The solution of this system is

$$A_1^{-1} = \omega_1(1 - X^{-N-1}); \qquad A_2^{-1} = -\omega_2(X^{N+1} - 1) \tag{37}$$

and thus

$$\mathscr{P}_0^+ = \omega_1 \phi_{N-1} = (1 - X^{-1})/(1 - X^{-N-1}) \tag{38}$$

To find the equations for ψ_n, let us multiply Eqs. (29a)–(29c) by $M + 1$ and then sum over M from $M = 1$ to infinity. We get

$$\psi_n = \omega_1 \psi_{n-1} + \omega_2 \psi_{n+1} + \phi_n \qquad (2 \leqslant n \leqslant N - 1) \tag{39a}$$

$$\psi_1 = \omega_1 \psi_0 + \omega_2 \psi_2 + \phi_1 + \omega_1 \tag{39b}$$

$$\psi_0 = \omega_2 \psi_1 + \phi_0 \tag{39c}$$

$$\psi_n = 0 \qquad\qquad (n \geqslant N) \tag{39d}$$

The general solution of Eq. (39a) is

$$\psi_n = C_1 + C_2 X^n + n(D_1 + D_2 X^n) \qquad (1 \leqslant n \leqslant N) \tag{40}$$

where C_1 and C_2 are arbitrary constants and

$$D_1 = A_1(\omega_1 - \omega_2)^{-1}; \qquad D_2 = -A_2(\omega_1 - \omega_2)^{-1} \tag{41}$$

Substituting Eq. (40) into the boundary conditions Eqs. (39b)–(39d), we find the equations for C_1 and C_2:

$$C_1 + C_2 X^N + N(D_1 + D_2 X^N) = 0$$

$$C_1 X + C_2 = A_1 + A_2 X + D_1 + D_2 X + \omega_2^{-1} \tag{42}$$

and after an elementary calculation we get

$$L_0^+ = \frac{\omega_1}{\mathscr{P}_0^+} \psi_{N-1} = \frac{X+1}{(X-1)(1-X)^{-N-1}} \left[N(1 + X^{-N-1}) - 2\frac{1 - X^{-N}}{X - 1} \right] \tag{43}$$

The average lengths l_0 and l_1 can be determined in a similar way. Omitting the calculations, we show only the final results:

$$l_0 = 2\frac{X^{N+2} - X^{-N} - (N + 1)(X^2 - 1)}{(X - 1)(1 - X^{-N})(X^{N+1} - 1)} \tag{44}$$

$$l_1 = 2\frac{X^{N+2} - X^{-N} - (N + 1)(X^2 - 1)}{(X - 1)(1 - X^{-N-1})(X^{N+2} - 1)} + 1 \tag{45}$$

The quantities l_0 and l_1 satisfy the identities

$$l_1 = \omega_1^{-1}q_0q_1l_0 + 1; \qquad l_0(N + 1) = l_1(N) + 1 \tag{46}$$

which follow from their definitions.

The expressions for the average lengths, Eqs. (43)–(45), are rather cumbersome. To have an idea of the order of magnitude of L_0^+, l_0, and l_1, let us find their values in the case $\omega_1 = \omega_2 = 0.5$, $X = 1$:

$$L_0^+ = N(N + 2)/3; \qquad l_0 = \tfrac{2}{3}(N + 2); \qquad l_1 = \tfrac{2}{3}N + 1 \tag{47}$$

Finally, we can find the average lengths of the three types of newborn sections L_A, L_B, and L_C:

$$L_A = \Delta\eta_A/\mathscr{P}_A = L_0^+ + q_0(1 - q_0)^{-1}l_0 \tag{48}$$

where $\mathscr{P}_A$ is the probability of a newborn section of type (A):

$$\mathscr{P}_A = (\tilde{\mathscr{P}}^-)^2\mathscr{P}_0^+ \sum_{n=0}^{\infty} q_0^n = \tilde{\mathscr{P}}^-\mathscr{P}^-\mathscr{P}_0^+ \tag{49}$$

Similarly,

$$L_B = l_0; \qquad L_C = L_0^- + q_0(1 - q_0)^{-1}l_0 \tag{50}$$

In the case $r = s = 1$, $\omega_1 = \omega_2$ we get

$$L_A = L_C = (N + 2)^2/3; \qquad L_B = 2(N + 2)/3 \tag{51}$$

5. CALCULATION OF η_0

Now we can calculate the fraction of melted links in the ground state in the case $r = s = 1$:

$$\eta_0 = \mathscr{P}_0^+[\tilde{\mathscr{P}}^-L_0^+ + (1 - q_0)q_1l_1 + \mathscr{P}^-q_0l_0] \tag{52}$$

Using the identities

$$(1 - q_0)q_1 = \omega_1; \qquad \omega_1l_1 - q_0q_1l_0 = \omega_1 \tag{53}$$

we find

$$\eta_0 = \mathscr{P}_0^+(\tilde{\mathscr{P}}^-L_0^+ + q_0l_0 + \omega_1) \tag{54}$$

The probabilities $\mathscr{P}_0^+$, etc., have been found in Ref. 8:

$$\mathscr{P}_0^+ = \frac{1 - X^{-1}}{1 - X^{-N-1}}; \qquad \mathscr{P}^- = \frac{X - 1}{X^{N+2} - 1}; \qquad q_0 = \omega_2\frac{1 - X^{-N}}{1 - X^{-N-1}}$$

$$q_1 = \frac{1 - X^{-N-1}}{1 - X^{-N-2}}; \qquad \tilde{\mathscr{P}}^- = \frac{\omega_2(X - 1)}{X^{N+1} - 1} \tag{55}$$

Substituting Eqs. (43)–(45) and (55) in Eq. (54), we get

$$\eta_0 = \frac{w_2(1 - X^{-1})}{1 - X^{-N-1}} \left[\frac{2X^2(X^N - 1) - N(X^2 - 1)}{(X^{N+1} - 1)(X - 1)} + X \right] \qquad (56)$$

An approximate expression for η_0 near the melting temperature for arbitrary values of U_1 and U_2 can be found by substituting $X = e^{\theta U/\beta}$, where $\theta = w_1 U_1 - w_2 U_2$, $\beta = U_1 U_2/2$ (see Ref. 8 for justification). Near the melting temperature $|\theta| \ll U_1, U_2$; $|X - 1| \ll 1$; and

$$\eta_0 \approx \frac{1 - X^{-N} - N(X - 1)X^{-N}}{(1 - X^{-N})^2} \qquad (|X - 1| \ll 1) \qquad (57)$$

or

$$\eta_0 = \frac{e^{\theta V/\beta} - 1 - \theta V/\beta}{4 \, \text{sh}^2(\theta V/2\beta)} \qquad (58)$$

which coincides with the expression obtained by Vedenov and Dykhne.[13]

6. THE FINE STRUCTURE OF MELTING CURVES

From Eqs. (14)–(16) and (43)–(45) we can find the magnitudes of the jumps $\Delta\eta_A, \Delta\eta_B, \Delta\eta_C$ in the case $r = s = 1$, $N = [N]$. The expressions one obtains are rather involved, but they are essentially simplified if $w_1 = w_2 = 0.5$, $X = 1$:

$$\Delta\eta_A = \Delta\eta_C = \frac{N(N + 4)}{6(N + 1)^2(N + 2)}; \qquad \Delta\eta_B = \frac{2}{3} \frac{N}{(N + 1)(N + 2)^2} \qquad (59)$$

According to the discussion in the introduction, Eq. (59) gives the order of magnitude of the maximal possible jumps of η_0. If $N \gg 1$, which is usually the case, we can write

$$\Delta\eta_A = \Delta\eta_C \approx U/6V; \qquad \Delta\eta_B \approx 2U^2/3V^2 \qquad (60)$$

It is expected that $\Delta\eta$ at $U_1 = U_2$ gives the order of magnitude of the largest peaks on the differential melting curve.

New sections in the ground state appear "instantly." Actually, however, the melting of a section of length L is spread over an interval[1] $\delta\theta \sim 4T/L$. If we assume a linear dependence of U_1 and U_2 on temperature,

$$U_1 = a(T - T_1), \qquad U_2 = -a(T - T_2) \qquad (61)$$

where T_1 and T_2 are the melting temperatures of pure component-1 and component-2 polymers, then[9] $\theta = a(T - T_m)$ and

$$\delta T = 4T/aL \qquad (62)$$

[9] $T_m = w_1 T_1 + w_2 T_2$ is the melting temperature.

The widths of the peaks corresponding to the three types of sections are of order

$$\delta T_A = \delta T_c \approx 12TU^2/aV^2 = 4T\Delta T/V \tag{63}$$

$$\delta T_B \approx 2T\Delta T/U \tag{64}$$

where $\Delta T = 3U^2/aV$ is the total width of the melting curve.[13,2] For a peak to be observable, it is necessary that $\delta T < \Delta T$.

For the standard values of the parameters U, V, and T, $U/T \sim 2/3$, $V/T \sim 10$, we get $\delta T_A \sim \delta T_C \sim 0.4\Delta T$, $\delta T_B \sim 3\Delta T$. Thus, only the peaks corresponding to A and C sections can be observed and even these peaks are rather wide. The situation can be different in so-called block heteropolymers consisting of long regions with different mean concentrations of components. In this case the melting interval can be much wider than in the case of random heteropolymers, while the widths of the peaks are approximately the same. Another possibility is to consider the melting of open DNA molecules, in which case the loop entropy has to be taken into account.[1,2] The loop entropy term renormalizes the surface energy V,[14] so that V effectively increases by a factor 1.7.[9] This results in $\delta T_A \sim \delta T_C \sim 0.23\Delta T$.

In our calculations we assumed that N is an integer. Physically, it is evident that the results will not change if N is not an integer exactly, but is close enough to some integral value. In this case the sections that contribute, say, to $\Delta\eta_A$ will melt at different temperatures, depending on their lengths. If T changes by δT, the energy difference between the helical and coiled states for a section of length L changes by a $L\,\delta T$. The shortest section of type (A) has length $\sim V/U_1$, and thus $\delta T_{\max} = U_1 V^{-1}\,\delta V$ is the maximal width a peak can gain due to a nonintegral N {here $\delta V = (N - [N])U$}. The peak will remain essentially unchanged if $\delta T_{\max} \ll \delta T_A$. In the case $U_1 = U_2$, $w_1 = w_2 = 0.5$ this implies $\delta V \ll 12TU/V$.

Finally, let us discuss the accidental fine structure of melting curves that is caused by the finite length of a polymer. If $\mathscr{P}$ is the probability of a newborn melted section and $\mathscr{N}$ is the number of links in the chain, then $\mathscr{N}\mathscr{P}$ is the average number of newborn sections and $(\mathscr{N}\mathscr{P})^{1/2}$ is the fluctuation of this number. The fluctuation of η is of order

$$\delta\eta \sim L_A(\mathscr{N}\mathscr{P})^{1/2}/\mathscr{N} \tag{65}$$

The maximal fluctuations of η are expected at values of r and s corresponding to the peaks of the intrinsic fine structure (i.e., to maxima of $\mathscr{P}$) or to relatively rare but very long sections. In the latter case Eq. (65) may not be applicable since it is valid only if $\mathscr{N}\mathscr{P} \gg 1$.

In the case $U_1 = U_2 = U$, $w_1 = w_2 = 0.5$ Eq. (65) gives

$$\delta\eta \sim (V/18U\mathscr{N})^{1/2} \tag{66}$$

Comparing Eqs. (60) and (66), we see that the accidental fine structure grows with the increase of V/U, while the intrinsic fine structure smoothes down and vice versa. The condition that the fluctuation $\delta\eta$ is negligible compared to the intrinsic variation $\Delta\eta$ is given by

$$(\mathcal{N}U^3/V^3)^{1/2} \gg 1 \tag{67}$$

At the standard conditions, $V/U \sim 15$, and Eq. (67) requires rather big polymer lengths ($\mathcal{N} \sim 3 \times 10^5$). However, at low sodium-ion concentrations, the values of U_1 and U_2 increase significantly and the ratio V/U can become as small as 6 (or even smaller).[4] Under such conditions the intrinsic fine structure may be observable even for relatively short sequences ($\mathcal{N} \sim 2 \times 10^4$).

7. COMPUTER CALCULATION OF MELTING CURVES

A study of $\Delta\eta_{A,B,C}$ and $\delta T_{A,B,C}$ as functions of r, s, N, and ω_1 requires computer calculations, which are now in progress. Results will be published
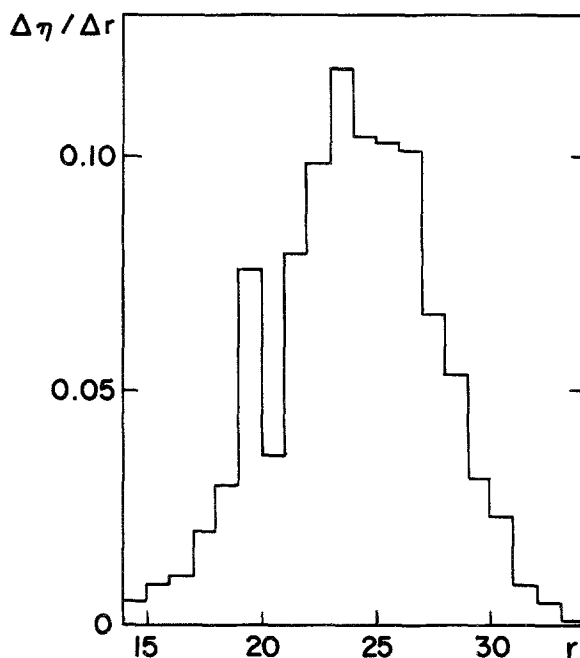

Fig. 1. Ground-state differential melting curve for an infinite heteropolymer with a random sequence of components. Calculation was performed for the following parameter values: $N = 119$, $r + s = 40$, $\omega_1 = 0.4$, $\omega_2 = 0.6$. The peak at $r = 20$ corresponds to $U_1 = U_2$, $V/U_1 \approx 6$.

elsewhere. In this work I would like to show only one ground-state differential melting curve, which was calculated by I. Simon using Eqs. (13) and (20)–(28). The following values of parameters were chosen: $\omega_1 = 0.4$, $\omega_2 = 0.6$, $N = 119$, $r + s = 40$, $r = 15, 16,..., 35$. The summation over M was cut off at $M = 200$. The whole calculation took 37 sec of computer time. The results are shown in Fig. 1. The peak at $r = 20$ corresponds to $U_1 = U_2$ and $V/U_1 \approx 6$ ($V/U_1 = 5.95$).[10]

REFERENCES

1. R. W. Wartell and E. W. Montroll, *Adv. Chem. Phys.* **22**:129 (1972).
2. A. A. Vedenov, A. M. Dykhne, and M. D. Frank-Kamenetskii, *Usp. Fiz. Nauk* **105**:479 (1971) [*Sov. Phys.—Usp.* **14**:715 (1972)].
3. E. W. Montroll and N. S. Goel, *Biopolymers* **4**:855 (1966).
4. M. D. Frank-Kamenetskii and Yu. S. Lazurkin, *Ann. Rev. Biophys. Bioeng.* **3**:127 (1974).
5. M. Ya. Azbel', *Zh. Eksp. Teor. Fiz. Pis'ma Red.* **16**:183 (1972) [*Sov. Phys.—JETP Lett.* **16**:128 (1972)].
6. M. Ya. Azbel', *Phys. Rev. Lett.* **31**:589 (1973).
7. I. M. Lifshitz, *Zh. Eksp. Teor. Fiz.* **65**:1100 (1973) [*Sov. Phys.—JETP* **38**:545 (1974)].
8. A. Vilenkin, *Biopolymers* **16**:1657 (1977).
9. Yu. L. Lyubchenko, M. D. Frank-Kamenetskii, A. V. Vologodskii, Yu. S. Lazurkin, and G. G. Gasue, Jr., *Biopolymers* **15**:1019 (1976).
10. M. Ya. Azbel', Invited paper at STATPHYS 13, *Ann. Isr. Phys. Soc.*, 2 (1977).
11. M. Ya. Azbel', Thermodynamic Decoding of DNA and Double-thread Replicative RNA, to be published.
12. A. Vilenkin, *Phys. Rev. B* **18**, No. 3.
13. A. A. Vedenov and A. M. Dykhne, *Zh. Eksp. Teor. Fiz.* **55**:375 (1968) [*Sov. Phys.—JETP* **28**:187 (1969)].
14. M. Ya. Azbel', *J. Chem. Phys.* **62**:3635 (1975).
15. I. Simon and A. Vilenkin, to be published.

---

[10] $V/U \approx 6$ corresponds to $T_{GC} - T_{AT} \approx 100$ K. Such values of $T_{GC} - T_{AT}$ have been observed at low [Na$^+$]. (See Ref. 4.)