Contribution to the statistical thermodynamics of fusion of long heteropolymer chains

I. M. Lifshitz

Institute of Physics Problems, USSR Academy of Sciences

(Submitted March 7, 1973)
Zh. Eksp. Teor. Fiz. **65**, 1100–1110 (September 1973)

A general solution is obtained for the problem of the thermodynamics of fusion of a heteropolymer chain in the case of strong cooperativity. The free energy on the fusion section can be expressed in terms of definite correlation characteristics of the monomer sequence. The fusion curves assume a universal shape for a small correlation radius, and from its scale the correlation radius can be determined.

The statistical-thermodynamic analysis of the melting of hetero-polymer chains (DNA, polypeptides) has been the subject of many articles. This problem was first treated correctly and completely by Vedenov, Dykhne, and Frank-Kamenetskiĭ[1], who succeeded, for instance, in obtaining a closed analytical solution for a random (uncorrelated) sequence of links of two kinds in the most important limiting case, that of strong cooperativity1). This problem was recently re-examined by Azbel'[2], who posed the question of determining the statistical characteristics of the sequence from the thermodynamics of melting of the chain.

In the following we give a general solution for the problem of the thermodynamics of melting of a hetero-polymer chain having a sequence of links with given correlation characteristics, in the same limiting case of strong cooperativity.

1. FORMULATION OF THE MODEL AND DERIVATION OF THE GENERAL FORMULAS

Let the polymer chain be represented by a fixed sequence of links of two kinds (A and B). Each section of the chain can exist in either of two states ("phases"): helical (I) or "melted" (II). If the free energy of each link in a given phase does not depend on the kind of neighbors it has, then the free energy of any section of the chain depends on the average concentrations of the components in that section. If s is the coordinate of a point on the chain, and x(s) is the smoothed-out concentration near this point, then we have for each of the phases in an element of length Δs

$$\Delta F = \Delta s[f_A + x(s)(f_B - f_A)],$$
$$\Delta F = \Delta F^{\mathrm{I}}, \quad \Delta F^{\mathrm{II}}, \quad f = f^{\mathrm{I}},\ f^{\mathrm{II}}.$$

If the free energy depends on the interaction with the nearest neighbors, then ΔF contains the "concentration" y(s) of pairs of unlike neighbors, AB or BA:

$$\Delta F = \Delta s[\lambda_0 + \lambda_1 x + \lambda_2 y],$$
$$\lambda_0 = f_{AA}, \quad \lambda_1 = f_{BB} - f_{AA}, \quad \lambda_2 = f_{AB} - {}^1/_2(f_{AA} + f_{BB}).$$

Here f_{AA}, f_{BB}, and f_{AB} are the respective free energies per link in the chains ....AAAA...., ...BBBB...., and ....ABAB.... .

Longer-range direct interactions are usually insignificant (although they can be included in the calculation by using the same scheme). We can therefore express in this approximation the free energy of a complete chain of length L, in phase I or phase II, in the form

$$F^{\mathrm{I}} = \int_0^L \lambda^{\mathrm{I}}(s)\,ds, \quad F^{\mathrm{II}} = \int_0^L \lambda^{\mathrm{II}}(s)\,ds; \tag{1.1}$$

$$\lambda^{\mathrm{I,II}}(s) = \lambda_0^{\mathrm{I,II}} + \lambda_1^{\mathrm{I,II}} x(s) + \lambda_2^{\mathrm{I,II}} y(s), \quad 0 < x,\ y < 1. \tag{1.2}$$

The functions x(s) and y(s) are specified by the fixed primary structure of the chain. Their statistical (correlation) characteristics will be assumed to be spatially homogeneous over the entire length of the chain. The coefficients λ_i are known functions of the temperature2).

If $\bar{x}$ and $\bar{y}$ are the concentrations x and y averaged over the entire chain, then F^{I} and F^{II} take the simple form

$$F^{\mathrm{I}}/L = \lambda_0{}^{\mathrm{I}} + \lambda_1{}^{\mathrm{I}}\bar{x} + \lambda_2{}^{\mathrm{I}}\bar{y}, \quad F^{\mathrm{II}}/L = \lambda_0{}^{\mathrm{II}} + \lambda_1{}^{\mathrm{II}}\bar{x} + \lambda_2{}^{\mathrm{II}}\bar{y}.$$

In the melting process, sections of the melted phase II appear in the chain. If s_i and t_i are the coordinates of the beginning and end of the i-th section, then we have for the free energy $F = F(\{s\}, \{t\})$ of the heterophase chain

$$F(\{s\},\{t\}) - F^{\mathrm{I}} = \sum_i [\Lambda(s_i, t_i) + V], \tag{1.3}$$

$$\Lambda(s,t) = \int_s^t \lambda\,ds = \Lambda(t) - \Lambda(s), \tag{1.4a}$$

$$\lambda = \lambda(s) = \lambda_0 + \lambda_1 x + \lambda_2 y, \quad \lambda_i = \lambda_i{}^{\mathrm{II}} - \lambda_i{}^{\mathrm{I}}. \tag{1.4b}$$

Here V is double the surface energy at the junction of phases I and II, which is assumed to be a large quantity ($V \gg T, \lambda_i$). The average value of λ over the entire length of the chain is

$$\bar{\lambda} = \lambda_0 + \lambda_1\bar{x} + \lambda_2 \qquad \bar{\lambda} = \bar{\lambda}(T). \tag{1.5}$$

The "melting temperature" of the chain is the name usually given to the temperature T_0 at which $F^{\mathrm{I}} = F^{\mathrm{II}}$, i.e., $\bar{\lambda}(T_0) \equiv \lambda(T_0, \bar{x}, \bar{y}) = 0$. It represents the true melting temperature of the chain as a whole as $V \to \infty$. If the concentrations x and y are reckoned from their mean values $\bar{x}$ and $\bar{y}$, we have

$$\begin{gathered} \lambda = \lambda(T, x, y) = \theta + \lambda_1\xi + \lambda_2\eta, \\ \xi = x - \bar{x}, \quad \eta = y - \bar{y}, \quad \theta = \theta(T) = \lambda(T, \bar{x}, \bar{y}). \end{gathered} \tag{1.6}$$

Near the temperature T_0, where the entire melting process evolves, θ plays the part of a temperature shift3):

$$\theta = \varkappa(T - T_0), \quad \varkappa = d\bar{\lambda}/dT, \tag{1.7}$$

and λ_1 and λ_2 are constants:

$$\lambda_1 = \lambda_1(T_0), \quad \lambda_2 = \lambda_2(T_0).$$

The coordinates s_i and t_i of the melted-section boundaries can be regarded as internal dynamic variables of the possible states of the system, and $F(\{s\}, \{t\}) - F^{\mathrm{I}}$ as the energy levels of these states. Then the desired free energy F is given by

$$F - F^{\mathrm{I}} = -T \ln Z,$$
$$Z = \sum_n \int \dots \int \exp\left\{-\frac{1}{T}\sum_{i=1}^{n}[\Lambda(s_i, t_i) + V]\right\}\prod_i ds_i\, dt_i. \tag{1.8}$$

Our original assumption that V is very large ($V \gg \lambda_1, \lambda_2, \theta, T$) means that the only contributions to Z are from sections that are quite long and have large negative values of the integral $\Lambda(s_n t_i)$, so that we can use the method of steepest descent and separate the principal part of the argument of the exponential in Eq. (1.8) as the "ground state" of the system. Since, according to Eq. (1.6), λ depends on the temperature θ, the extremal points s_i^0 and t_i^0, their number n^0, and the "ground state" energy U_0 are also temperature-dependent:

$$U_0 = U_0(\theta) = \min \sum_{i=0}^{n}[\Lambda(s_i, t_i) + V]$$
$$= \sum_{i=1}^{n^0}[\Lambda(s_i^0, t_i^0) + V] = n^0(\theta)[\langle\Lambda(s^0, t^0)\rangle + V] \tag{1.9}$$

(the angle brackets denote averaging over all sections (s_i^0, t_i^0)).

It is easy to prove that keeping the term of the next order of magnitude in F is equivalent to taking into account, in the statistical interval of Z, the motion of the variable boundaries s_i and t_i in the neighborhood of their extremal positions s_i^0 and t_i^0, but does not include the generation of new melted sections or the vanishing of old ones. This gives

$$F - F^{\mathrm{I}} = U_0(\theta) + U_1;$$
$$U_1 = -T \ln \int \dots \int \exp\left\{-\frac{1}{T}\sum_{i=1}^{n_0}[\Lambda(t_i^0, t_i) + \Lambda(s_i, s_i^0)]\right\}\prod_i ds_i\, dt_i \tag{1.10}$$
$$= -n^0 T\left\langle \ln \int_{-\infty}^{\infty} \exp\left\{-\frac{\Lambda(t^0, t)}{T}\right\}dt + \ln \int_{-\infty}^{\infty} \exp\left\{-\frac{\Lambda(s, s^0)}{T}\right\}ds\right\rangle.$$

The method by which the averages in Eqs. (1.9) and (1.10) are calculated is as follows: we formulate inequalities that determine uniquely the extremal points s_i^0 and t_i^0 for melted sections with a given increase of free energy

$$\Lambda(s_n^0, t_i^0) + V = -z \quad (z > 0).$$

We then determine the probability that these inequalities are satisfied. In particular, we find the probability (per unit length) $\mathscr{P}(z)dz$ that an arbitrary point s is the beginning of an extremal melted section with the given z. We can express the free energy in terms of these probabilities. For example, the principal term $U_0(\theta)$ is of the form

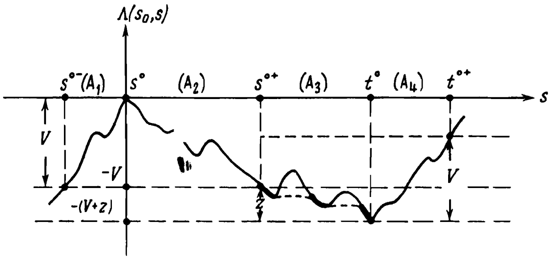
$$U_0(\theta) = -L\int_0^\infty z\mathscr{P}(z)\,dz, \quad \mathscr{P}(z) = \mathscr{P}(z, \theta). \tag{1.11}$$

The total probability (per unit length) that a melted section begins at the given point s is

$$\mathscr{P}_0 = \int_0^\infty \mathscr{P}(z)\,dz \quad (\mathscr{P}_0 = \mathscr{P}_0(\theta)),$$

and the number of melted sections is $n^0 = L\mathscr{P}_0$.

Since the coefficients λ_1, λ_2, and θ in (1.6) are constant parameters at a given temperature, the desired probabilities are themselves correlation characteristics of the link sequence, i.e., of the functions ξ(s) and η(s). However, as we shall see, they can be expressed in terms of simpler and more usual correlation functions. Let us now proceed to determine all these quantities.



Schematic behavior of the quantity $\Lambda(s_0, s) = \int^s \lambda ds = \Lambda(s) - \Lambda(s_0)$ in regions (A_2) and (A_3) of a melted section ($s^0 < s < t^0$) and outside its boundaries ((A_1) and (A_4)). Inequalities (A_1) and (A_2) are necessary and sufficient to determine the extremal position of the point s^0. (A_3) and (A_4) determine the extremal position of the end t^0 of a melted section with a given gain $\Delta F = -z$ in the free energy. To interpret the probability p(z)dz, the thicker line sections represent the pieces of the "trajectory" Λ(s) in region (A_3) where a "particle" moving from left to right first reaches the intermediate values z' ($0 < z' < z$).

2. DETERMINATION OF THE EXTREMAL POINTS AND CONSTRUCTION OF THE SYSTEM OF PROBABILITIES

To formulate the conditions that determine the points s^0 and t^0 uniquely, let us turn to Fig. 1, which shows a characteristic portion of the function $\Lambda(s^0, s)$ near any one of the extremal points $s_i^0 = s^0$. Besides s^0 and t^0, a significant part is played in Fig. 1 by the auxiliary points lying closest to them, s^{0-}, s^{0+}, and t^{0+}, given by the conditions

$$\Lambda(s^0, s^{0-}) = -V, \quad \Lambda(s^0, s^{0+}) = -V, \quad \Lambda(t^{0+}, t^0) = -V.$$

The necessary and sufficient conditions for the extremal points s^0 and t^0 can more easily be formulated and verified than derived:

$$\begin{aligned}
&(A_1) \quad s^{0-} < s < s^0, \quad -V < \Lambda(s, s^0) < 0;\\
&(A_2) \quad s^0 < s < s^{0+}, \quad -V < \Lambda(s^0, s) < 0;\\
&(A_3) \quad s^{0+} < s < s' < t^0, \quad \Lambda(s, s') < V,\\
&\qquad\quad \Lambda(s^{0+}, t^0) = -z, \quad -z < \Lambda(s, t^0) < 0;\\
&(A_4) \quad t^0 < s < t^{0+}, \quad -V < \Lambda(t^0, s) < 0.
\end{aligned} \tag{2.1}$$

Each of the inequalities (A_1)—(A_4) imposes requirements on the behavior of λ(s) (i.e., ξ(s) and η(s)) within only one of the four regions (A_1) to (A_4) (see Fig. 1). Their meaning is quite obvious; it is easy to show that violation of even one of these requirements will increase the energy of the system.

Inequalities (A_1) and (A_2) differ only in the sign of $s - s^0$, i.e., in the direction of motion along s. Inequality (A_4) coincides with (A_1) if Λ is measured from the points t^{0+}:

$$(A_4) \quad t^0 < s < t^{0+}, \quad -V < \Lambda(s, t^{0+}) < 0.$$

If there is no correlation between the behavior of ξ and η in the regions (A_1) to (A_4), the probabilities that each of the inequalities[4] (A_1)—(A_4) is satisfied are respectively

$$\mathscr{P}(A_1) = \mathscr{P}^-, \quad \mathscr{P}(A_2) = \mathscr{P}^+, \quad \mathscr{P}(A_3) = p(z), \quad \mathscr{P}(A_4) = c\mathscr{P}^-. \tag{2.2}$$

(The normalization constant c arises in the conversion from probabilities per length element ds to probabilities per element dz.) Then the above-mentioned probability $\mathscr{P}(z)$ of a melted section with given z will be

$$\mathscr{P}(z) = c\mathscr{P}^-\mathscr{P}^+p(z)\mathscr{P}^-. \tag{2.3}$$

In the general case we must substitute in Eq. (2.3) the conditional probabilities

$$\mathscr{P}^- \mathscr{P}_{A_1}{}^+ p_{A_1,A_2}(z) \mathscr{P}^-_{A_1,A_2,A_3}. \tag{2.4}$$

The probability that an arbitrary point s is the beginning s^0 of a melted section involves the satisfaction of only the first two inequalities, (A_1) and (A_2), so that

$$\mathscr{P}_0 = \mathscr{P}^- \mathscr{P}_{A_1}{}^+; \tag{2.5}$$

$$\begin{gathered} \langle z\rangle = \int_0^\infty z p_{A_1,A_2}(z)\, c\mathscr{P}^-_{A_1,A_2,A_3}\, dz, \\ \int_0^\infty p_{A_1,A_2}(z)\, c\mathscr{P}^-_{A_1,A_2,A_3}\, dz = 1. \end{gathered} \tag{2.6}$$

3. CALCULATION OF THE PROBABILITIES AND THERMODYNAMIC QUANTITIES

To calculate the probabilities we must make some assumptions about the character of the correlations in the informational structure of the chain, i.e., about the characteristics of the functions $\xi(s)$ and $\eta(s)$, or, what amounts to the same thing, about the correlation characteristics of the function

$$\lambda(s) = \theta + \lambda_1\xi + \lambda_2\eta.$$

In general there should exist various scales of correlations: small-scale correlations involving a small number of neighboring links, and long-range correlations between macroscopic sections of the chain. We begin with a case in which there are only small-scale correlations, whose radius is much smaller than the average size of the sections of each phase. In this case all the formulas (2.4)—(2.6) are greatly simplified. We note first that by choosing the s interval in $\Delta\Lambda = \lambda\Delta s$ to be large compared to the correlation radius, we arrive at the problem of the uncorrelated random walks of a particle with allowance for the limitations set by the inequalities (A_1)—(A_4). The variable s plays the role of "time", and the role of "distance" is played by

$$\Lambda = \int^s \lambda\, ds.$$

The diffusion characteristics of this process are the quantities

$$\alpha = \overline{\Delta\Lambda} / \Delta s = \theta, \quad \beta = \overline{(\Delta\Lambda)^2} / 2\Delta s. \tag{3.1}$$

(the superior bar in Eq. (3.1) denotes the average over all positions of the element Δs along the chain).

Substituting in

$$\overline{(\Delta\Lambda)^2} = \int_{\Delta s}\int_{\Delta s} \overline{\lambda\lambda'}\, ds\, ds'$$

the possible discrete values of λ in each link, we obtain

$$\beta = \frac{1}{2}\sum_{n=-\infty}^{\infty} \{\lambda_1{}^2(\overline{x_0x_n} - \bar{x}^2) + 2\lambda_1\lambda_2(\overline{x_0y_n} - \bar{x}\bar{y}) + \lambda_2{}^2(\overline{y_0y_n} - \bar{y}^2)\} \quad (x_n, y_n = 0, 1). \tag{3.2}$$

Thus, for example, if $\lambda_2 = 0$ and $\bar{x}W_n$ is the probability that two sites separated by a distance n are both occupied by links of type B, then

$$\beta = {}^1/_2\lambda_1{}^2\bar{x}(1-\bar{x}) + {}^1/_2\lambda_1{}^2\bar{x}\sum_{n\neq 0}(W_n - \bar{x}). \tag{3.3}$$

If the placement of the links is completely random, we have $W_n = \bar{x}$ and

$$\beta = {}^1/_2\lambda_1{}^2\bar{x}(1-\bar{x}).$$

In the solution of the random-walk problem with the given limitations (A_1)—(A_4), a change in the direction of motion along s is equivalent to the replacement of θ by $-\theta$. We therefore obtain $P^+(\theta) = P^-(-\theta)$, and accordingly,

$$\mathscr{P}(z) = c\mathscr{P}^-(\theta)\mathscr{P}^-(-\theta)p(z)\mathscr{P}^-(\theta). \tag{3.4}$$

On the other hand, according to Eq. (2.6),

$$c\mathscr{P}^-P_0 = 1, \quad P_0 = P_0(\theta) = \int_0^\infty p(z,\theta)\, dz.$$

Therefore from Eq. (2.5)

$$\mathscr{P}_0 = 1 / c^2P_0(\theta)P_0(-\theta) \tag{3.5}$$

and the problem reduces to the calculation of the probability $p(z) \equiv p(z, \theta)$ that a particle wandering randomly with "drift velocity" θ and "diffusion coefficient" β will first reach the point $-z$ by moving in such a way that never, during the motion process, does it return to a distance V from any point $-z'$ that was already attained ($\Lambda(s'') < \Lambda(s') + V$, if $s'' > s'$).

Clearly, with this definition of p(z), we have for an infinitesimal Δ

$$p(z+\Delta) = p(z)(1 - q\Delta), \quad dp(z) / dz = -qp(z), \tag{3.6}$$

where $q\Delta$ is the probability that a particle leaving the point $-z$ will return at least once to the point $-z + V$ during the time it takes to reach $-z - \Delta$. The probability $q\Delta$ obviously does not depend on z, and thus q has the meaning of an inverse "mean free path" of the particle with respect to return over a distance V:

$$p(z) = e^{-qz}, \quad P_0 = \int_0^\infty e^{-qz}\, dz = 1/q, \tag{3.7}$$

$$\begin{gathered} \bar{z} = q\int_0^\infty z e^{-qz}\, dz = 1/q, \quad q = q(\theta); \\ \mathscr{P}_0 = q(\theta)q(-\theta) / c^2. \end{gathered} \tag{3.8}$$

For the principal part of the free energy $\langle U_0\rangle$ we have, from Eq. (1.11),

$$\langle U_0\rangle / L = \mathscr{P}_0\langle z\rangle = q(-\theta) / c^2.$$

The value of $q(\theta)$ can be obtained from its definition.

Let $\varphi(s, \Lambda)$ be the probability that a particle leaving the origin at s = 0 will be found at the point Λ at the "instant" s. Then

$$\frac{\partial\varphi}{\partial s} = -\theta\frac{\partial\varphi}{\partial\Lambda} + \beta\frac{\partial^2\varphi}{\partial\Lambda^2}, \quad \varphi(0,\Lambda) = \delta(\Lambda). \tag{3.9}$$

Since we are interested only in particles that reach the point $\Lambda = V$ first before reaching $\Lambda = -\Delta$, the boundary conditions for φ will be

$$\varphi(s, V) = \varphi(s, -\Delta) = 0.$$

Then the sought quantity $q\Delta$ is the total flux of particles through the point $\Lambda = V$:

$$q\Delta = \int_0^\infty \beta\left.\frac{\partial\varphi}{\partial\Lambda}\right|_V ds = \beta\left.\frac{d\Phi}{d\Lambda}\right|_V, \tag{3.10}$$

where

$$\Phi = \int_0^\infty \varphi\, ds.$$

For Φ, from Eq. (3.8), we have the equation

$$\beta\frac{\partial^2\Phi}{\partial\Lambda^2} - \theta\frac{\partial\Phi}{\partial\Lambda} = -\delta(\Lambda), \quad \Phi(V) = \Phi(-\Delta) = 0.$$

Taking into account the infinitesimal smallness of Δ, we then obtain

$$q(\theta)=\frac{\chi(-\tau)}{V},\quad \chi(\tau)=\frac{\tau}{1-e^{-\tau}},$$
$$\tau=\frac{\theta V}{\beta}=\frac{T-T_0}{\delta},\quad \delta=\beta/V\frac{d\lambda_0}{dT}. \tag{3.11}$$

δ has the meaning of the width of the transition temperature range.

From Eqs. (3.7), (3.8), and (3.11) we obtain

$$\mathscr{P}^-=\chi(-\tau)\,/\,cV,\quad \mathscr{P}^+=\chi(\tau)\,/\,cV,$$
$$\mathscr{P}_0=\chi(\tau)\chi(-\tau)\,/\,c^2V^2,\quad \langle z\rangle=V\,/\,\chi(-\tau),\quad \langle U_0\rangle\,/\,L=\chi(\tau)\,/\,c^2V. \tag{3.12}$$

The normalization constant c can be found from the requirement that as $V \to \infty$ and at $\theta > 0$ the entire chain passes into the melted state, i.e., $\langle U_0\rangle/L = \lambda = \theta$. Then $c = \beta^{-1/2}$ and, finally,

$$\frac{\langle U_0\rangle}{L}=\frac{\beta}{V}\frac{\tau}{1-e^{-\tau}},\quad \tau=\frac{T-T_0}{\delta}. \tag{3.13}$$

This expression has exactly the same structure as that obtained in[1], and at $\lambda_2 = 0$ and $\beta = \frac{1}{2}\lambda_1^2\bar{x}(1-\bar{x})$ it transforms into the corresponding formula of[1]. However, by explaining the nature of this formula and its connection to the probabilities $\mathscr{P}^+$, $\mathscr{P}^-$, and p(z) we can investigate more complex correlation structures.

The simplest case of "structured probabilities" is the case in which, in addition to the small-scale correlations in the informational structure of the chain, there exists a second large correlation radius. If this radius is large compared to the average dimensions of the sections in each phase⁵⁾, then the values of τ and β in Eqs. (3.13) and (3.14) must be regarded as local values in a large heterophase section, and we must carry out additional averaging over the values of τ and β (i.e., T_0, δ, and β):

$$\frac{\langle U_0\rangle}{L}=\left\langle\frac{\beta}{V}\frac{\tau}{1-e^{-\tau}}\right\rangle,\quad \tau=\frac{T-T_0}{\delta}.$$

When the dimensions of sections I or II become comparable with the larger correlation radius, the averaging is carried out by a somewhat more complex method that follows directly from the definitions given above for the probabilities. In the most general case, the meaning established in Sec. 2 for the probabilities $\mathscr{P}(A_1)$, $\mathscr{P}(A_2)$, $\mathscr{P}(A_3)$, and $\mathscr{P}(A_4)$ is itself the answer to the posed question.

The minimum correlation scales in DNA are evidently related to the "three-letter words" that make up its code. To the contrary, the maximum scales for the spatially inhomogeneous "block" DNA correspond to the dimensions of the separate blocks, as has been shown by Lazurkin and co-workers[3,4]. In the two-scale problem for block DNA, the standard behavior should occur within the individual blocks[3,4], and this makes it possible to determine the radius of the close correlations from the value of β. The quite natural assumption that an intermediate correlation scale exists for a statistically homogeneous chain should alter this behavior. However, the accuracy of the available experimental data is not yet sufficient to permit us to answer this question.

In conclusion, we must write out the previously neglected term U_1 in Eq. (1.10) for the free energy:

$$\frac{U_1}{L}=-\mathscr{P}_0(\tau)\cdot 2T\,\langle\ln J\rangle,\quad J=\int_{-\infty}^{\infty}\exp\left[\frac{\Lambda(s^0,s)}{T}\right]ds. \tag{3.14}$$

The calculation given in the Appendix shows that in the melting region $\tau \lesssim 1$ ($\theta \sim \beta/V \ll 1$) we have $\langle \ln J\rangle = \ln(2T^2/\beta)$, whence

$$\frac{U_1}{L}=-\frac{2\beta T}{V^2}\chi(\tau)\chi(-\tau)\ln\frac{2T^2}{\beta}.$$

We thus have ultimately

$$\frac{F-F^{\mathrm{I}}}{L}=-\frac{\beta}{V}\chi(\tau)\left[1+\frac{2T}{V}\chi(-\tau)\ln\frac{2T^2}{\beta}\right]. \tag{3.15}$$

The second term can be neglected if

$$\frac{2T}{V}\ln\frac{T^2}{\beta}\ll 1.$$

Finally, when we consider the melting of a double helix (DNA), the surface energy V acquires, as we know, an entropy increment $\delta V = \frac{3}{2}T\ln l_{\mathrm{II}}$, that depends on the length l_{II} of the melted section. This adds to the free energy of the heterophase system an extra term $U_2 = \frac{3}{2}L\mathscr{P}_0(\tau)T\langle\ln l_{\mathrm{II}}\rangle$, which inside the melting region ($\tau \lesssim 1$) takes on the value

$$\frac{U_2}{L}=\frac{3}{2}\frac{\beta T}{V^2}\chi(\tau)\chi(-\tau)\ln\frac{V^2}{\beta}.$$

The relative contribution from this increment is

$$\frac{U_2}{U_0}=\frac{3T}{2V}\chi(-\tau)\ln\frac{V^2}{\beta}.$$

This term can be neglected when $(T/V)\ln(V^2/\beta) \ll 1$.

APPENDIX

Calculation of ⟨ln J⟩

By choosing the appropriate scale for $\Lambda(s)$ and s,

$$\Lambda(s)=-TX(t),\quad ds=\frac{T^2}{\beta}dt,\quad \Lambda(s)=\int_0^s\lambda\,ds, \tag{A.1}$$

we can, as is apparent from Fig. 1 and inequalities (A_1) and (A_2), represent J as the sum of two independent random quantities of very simple structure:

$$J=\frac{T^2}{\beta}(H_++H_-);$$

$$H=\int_0^{\infty}e^{-X(t)}\,dt,\quad 0<X(t)<\frac{V}{T}, \tag{A.2}$$

$$\left\langle\frac{(\Delta X_\pm)}{\Delta t}\right\rangle=\pm\frac{\theta T}{\beta},\quad \left\langle\frac{(\Delta X)^2}{\Delta t}\right\rangle=1 \tag{A.3}$$

The distribution laws for the $X_\pm(t)$ and the corresponding $H_\pm$ differ only in the sign of $\pm\theta$ in Eq. (A.3).

Let $\varphi(t, X_0, X)$ be the solution of the equation

$$\frac{\partial\varphi}{\partial t}=\frac{\partial^2\varphi}{\partial X^2}-\frac{\theta T}{\beta}\frac{\partial\varphi}{\partial X}$$
$$\varphi|_{t=0}=\delta(X-X_0),\quad \varphi|_{X=0}=\varphi_{X=V}=0. \tag{A.4}$$

Let us further introduce the quantity

$$\Phi(X_0,X)=\int_0^{\infty}\varphi(t,X_0,X)\,dt,$$

which, according to Eq. (A.4), satisfies the equation

$$\frac{\partial^2\Phi}{\partial X^2}-\frac{\theta T}{\beta^2}\frac{\partial\Phi}{\partial X}=-\delta(X-X_0),\quad \Phi|_{X=0}=\Phi_{X=V}=0, \tag{A.5}$$

and also the "particle flux" $Q(X_0)$ through the boundary $X = V/T$:

$$Q(X_0)=\beta\left.\frac{\partial\Phi(X_0,X)}{\partial X}\right|_{X=V/T}. \tag{A.6}$$

Then, as we can easily show, the probability distribution $\rho_t(X)$ for a random value of the "path" $X(t) = X$ at the "instant" t on the trajectories $X(t)$ that reach the point $X = V/T$ (i.e., on the melted sections) is

$$\rho_t(X) = \lim \varphi(t, X_0, X)\frac{Q(X)}{Q(X_0)}, \quad X_0 \to 0.$$

For the averages ⟨H⟩, $\langle H^2\rangle$, ... we obtain

$$\langle H\rangle = \int_0^\infty \langle e^{-X(t)}\rangle\, dt = \int_0^\infty\int_0^\infty e^{-X}\rho_t(X)\, dt\, dX = \int_0^\infty e^{-X}\rho(X)\, dX,$$

$$\langle H^2\rangle = 2\int_0^\infty\int_0^\infty \langle \exp[-X(t_1) - X(t_1+t_2)]\rangle\, dt_1\, dt_2$$

$$= 2\int_0^\infty\int_0^\infty \exp\{-X_1 - X_2\}\, \rho(X_1, X_2)\, dX_1\, dX_2,$$

$$\rho(X) = \lim_{X_0\to 0} [\Phi(X_0, X) Q(X)/Q(X_0)],$$

$$\rho(X_1 X_2) = \lim_{X_0\to 0} [\Phi(X_0, X_1)\Phi(X_1, X_2) Q(X_2)/Q(X_1)].$$

Because of the conditions $V/T \gg 1$, $\theta T/\beta \ll 1$ we can set $V/T \to \infty$, $\theta T/\beta \to 0$ in the calculation of ⟨H⟩, $\langle H^2\rangle$, ... According to Eqs. (A.5) and (A.6), this gives

$$Q(X)/Q(X_0) = X/X_0,$$

$$\Phi(X, X') = \begin{cases} X, & X < X' \\ X', & X > X' \end{cases}$$

and accordingly

$$\rho(X) = X, \quad \rho(X, X') = \Phi(X, X')X'.$$

Therefore

$$\langle H_+ + H_-\rangle = 2\langle H\rangle = 2, \quad \langle h^2\rangle = \left\langle\left(\frac{H - \langle H\rangle}{\langle H\rangle}\right)^2\right\rangle = \frac{1}{2}$$

and finally

$$\langle \ln J\rangle = \ln\langle J\rangle + \left\langle \ln\left(1 + \frac{h_+ + h_-}{2}\right)\right\rangle = \ln\frac{2T^2}{\beta} - \varepsilon^2,$$

$$\varepsilon^2 = {}^1/_4\langle h^2\rangle - {}^1/_2\langle h^3\rangle + \ldots \approx 0{,}1.$$

---

[1] I.e., when there is a large "surface energy" at the junction between the helical and melted sections. The term "melting" refers to the loss of rigidity in the chain when its secondary helical structure breaks down.

[2] In the case of an arbitrary number of components (kinds of links), the function λ(s) in Eqs. (1.1) to (1.4) is replaced by $\Sigma\lambda_i x_i(s)$.

[3] The melting temperatures T_{01} and T_{02} of the pure components A and B are usually close to each other ($|T_{01} - T_{02}| \ll T_0$); then

$$\varkappa = \left.\frac{d\lambda_0}{dT}\right|_{T_0}, \quad \lambda_1 = \varkappa(T_{02} - T_{01}) \ll T_0\frac{d\lambda_0}{dT}.$$

[4] We recall that the probability of an inequality is the relative frequency with which this inequality is satisfied when the reference point s runs through all values in the chain length L.

[5] The average length $\langle l\rangle$ of section I or II is $1/\mathscr{P}_0(\tau)$, i.e.,

$$\langle l\rangle = \frac{V^2}{\beta\tau^2}\cdot 4\,\mathrm{sh}^2\frac{\tau}{2}.$$

[1] A. A. Vedenov and A. M. Dykhne, Zh. Eksp. Teor. Fiz. **55**, 357 (1968) [Sov. Phys.-JETP **28**, 187 (1969)]; A. A. Vedenov, A. M. Dykhne, and M. D. Frank-Kamenetskiĭ, Usp. Fiz. Nauk **105**, 479 (1971) [Sov. Phys.-Uspekhi **14**, 715 (1972)].

[2] M. Ya. Azbel', ZhETF Pis. Red. **16**, 183 (1972) [JETP Lett. **16**, 128 (1972)].

[3] Yu. S. Lazurkin, M. D. Frank-Kamenetskiĭ, and E. N. Trifonov, Biopolymers **9**, 1253 (1970).

[4] A. V. Shugalin, M. D. Frank-Kamenetskiĭ, and Yu. S. Lazurkin, Molek. Biol. **4**, 275 (1970); **5**, 766 (1971). [Molecular Biology **4**, 221 (1970); Vol. 5 (in press)].

Translated by R. Rutherford
112