PREMELTING OF SUPERHELICAL DNA: AN EXPRESSION FOR SUPERHELICAL ENERGY

A. V. VOLOGODSKII and M. D. FRANK-KAMENETSKII
Institute of Molecular Genetics, USSR Academy of Sciences, Moscow 123182, USSR

Received 26 June 1981

1. Introduction

Numerous experimental data show that negatively supercoiled closed circular DNA (ccDNA) has a weakened secondary structure [1–8]. However, these data per se cannot offer a detailed quantitative picture of the base-pair opening probability at particular sites of a DNA molecule. As a result attempts have been made to obtain theoretical estimations of this probability [9–13]. The theory is based on the ordinary helix–coil model and the expression of the superhelix energy G as a function of two variables: superhelix density σ and the fraction of open base-pairs ϑ. For the case of $\vartheta = 0$ the function $G(\sigma,\vartheta)$ was determined directly from experiment [14,15]:

$$G(\sigma,0) = 10RTN\sigma^2 \tag{1}$$

Starting from this equation we proposed the following expression [10,11]:

$$G(\sigma,\vartheta) = 10RTN(\sigma + \vartheta)^2 \tag{2}$$

However this equation as well as the considerably different equation used in [12,13] have no reliable rationale. Here, we show that a very general consideration leads to a theoretically sound equation for the $G(\sigma,\vartheta)$ function. Moreover, using experimental data on the early melting of ccDNA one can obtain a definite expression for the superhelix energy as a function of σ and ϑ.

2. Theory

The total energy of a given state of ccDNA may be regarded as consisting of two terms. The first one is the ordinary term which is used in the helix–coil theory of linear DNA. The second term, $G(\sigma,\vartheta)$, allows for the fact that we consider a closed circular molecule and is called the superhelix energy. Our objective is to obtain an expression for the superhelix energy $G(\sigma,\vartheta)$ applicable to the region where both variables are much smaller than 1 ($\sigma \ll 1$, $\vartheta \ll 1$). Indeed, the physiological values of superhelix density are always smaller than 0.1 (see [17]), and for these values of σ the fraction of open base-pairs ϑ is also very small. So we can expand the $G(\sigma,\vartheta)$ function in a series and confine ourselves to linear and quadratic terms:

$$G(\sigma,\vartheta) = A\sigma + B\vartheta + C\sigma^2 + D\vartheta^2 + E\sigma\vartheta \tag{3}$$

It is possible to determine all but one of the coefficients A,B,C,D and E from the reliable considerations that follow:

(1) It follows from eq; (1) that $A = 0$ and $C = 10RTN$;
(2) For given σ the G function has a minimum value at $\vartheta = -\frac{\gamma_0}{10}\sigma \approx -\sigma$, where γ_0 is the number of base pairs per turn of the double helix under given external conditions.

This follows from the fact that at a given σ the minimum energy corresponds to a state where the helical part of ccDNA is completely relaxed and the open regions consist of two non-interwound strands.

From condition $\partial G/\partial\vartheta|_{\vartheta = -\sigma} = 0$ one obtains:

$$B - 2D\sigma + E\sigma = 0 \tag{4}$$

Because eq. (4) has to be valid for any σ, we obtain: $B = 0$ and $E = 2D$. The final equation may be written as:

Published by Elsevier/North-Holland Biomedical Press
00145793/81/0000–0000/$02.50

$$G(\sigma,\vartheta) = 10RTN[(1-b)(\sigma + \vartheta)^2 + b\sigma^2] \qquad (5)$$

where b is the only undetermined parameter. Our early eq. (2) is a special case of eq. (5) for $b = 0$. So eq. (2) assumes that at $\vartheta = -\sigma$ the superhelix energy has to be not only minimal but zero, which is incorrect. Indeed, the entropy of the open state in ccDNA is always lower than in linear DNA due to topological constraints. That is $G(\sigma,\vartheta = -\sigma) > 0$ and $b > 0$. To obtain a quantitative estimation of the b value we need additional experimental data.

3. Fitting theory to experiment

To determine the value of b in eq. (5) we used the experimental data in [8]. The points in fig.1 show their very accurate data concerning the dependence of the fraction of open pairs (the degree of denaturation) on the superhelix density at different temperatures. In the same figure we show theoretical curves calculated for different values of b. The calculations were performed using the algorithm in [10]. We used the thermodynamic parameters of linear DNA corresponding to the conditions in [7,8]: melting temperature of AT-pairs $T_{AT} = 16°C$; melting temperature of GC-pairs $T_{GC} = 58°C$; the enthalpy of melting of AT pairs $U_{AT} = 7.1$ kcal/mol, corresponding to the melting temperature observed; cooperatively factor $\hat{\sigma} =$

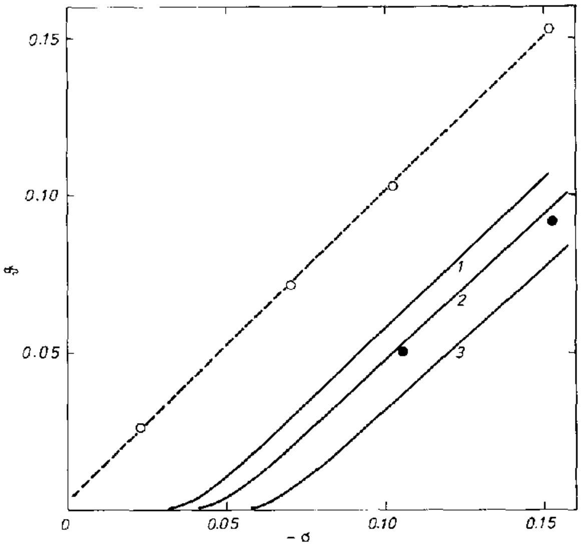

Fig.1. The dependence of the function of open base-pairs ϑ on superhelix density σ for conditions used in [8]. Points are taken from [8]: (○) 35°C; (●) 15°C. Curves are calculated using eq. (5): (– – –) 35°C; (——) 15°C. Curves 1–3 correspond to different values of parameter b: 0(1); 0.2(2) and 0.4(3). At 35°C all three calculated curves practically coincide because this is the melting temperature of linear DNA under these conditions. Calculations were performed for the sequence of ϕX174 DNA.

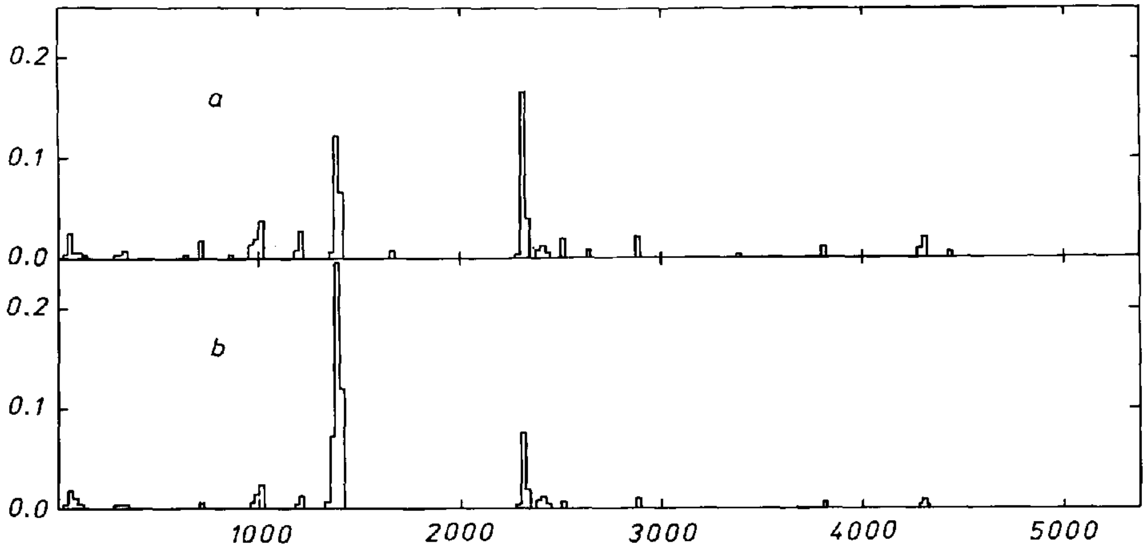

Fig.2. Theoretical dependence of base-pair opening probability on the base-pair order number along ϕX174 sequence at $T_m = 20°C$, for the superhelix density value $\sigma = -0.05$ and the external condition [8]: (a) $\alpha = 0$; $b = 0.4$; (b) $\alpha = 1.5$ and $b = 0.2$.

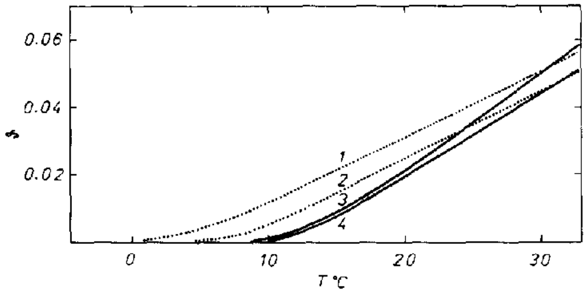

Fig.3. Theoretical dependences of the degree of denaturation of ϕX174 ccDNA on temperature within the premelting zone. The parameters correspond to the conditions in [8] (see text). Calculations were performed for different theoretical models: (1) $\alpha = 0, b = 0$; (2) $\alpha = 1.5, b = 0$; (3) $\alpha = 0, b = 0.4$; (4) $\alpha = 1.5; b = 0.2$.

5×10^{-5}, loop weighting factor $\alpha = 3/2$, as usual (see, e.g., [18]). As the sequence of PM2 DNA is unknown, we used in our calculation the sequence of ϕX174 DNA [19] which has nearly the same GC-content. Such a rough characteristic as the degree of helicity ϑ is known to be quite insensitive to the particular sequence at the given GC-content.

One can see that at 15°C the theoretical curves are highly sensitive to the b value. Hence the b value can be determined with a high precision: $b = 0.2$.

The above results correspond to the most realistic helix–coil model for linear DNA, allowing for loop formation in open regions. Unfortunately this model is inconvenient for calculations of fluctuational violations in ccDNA because it has great difficulty in allowing for cruciform structures. Anyway the parameter b serves to adjust the chosen helix-coil model of linear DNA to the case of ccDNA. So we have attempted to use the simplest helix–coil model of linear DNA, i.e., the one formally corresponding to the case of $\alpha = 0$. The results of calculations within the framework of this model proved to meet the experimental data in [8] for $b = 0.4$.

Of course this model can be used only if the resulting value of the opening probability is consistent with that obtained by the model allowing for loop formation in the open state ($\alpha = 1.5$). Figures 2 and 3 show that both models yield quite similar results, considering the accuracy that can generally be expected from such theoretical calculations.

4. Conclusion

We have obtained a reliable expression of the superhelix energy as a function of two variables, superhelix density σ and the fraction of open base-pairs ϑ (eq. (5)). The parameter was shown to depend on the choice of the particular helix–coil model and its value was determined by comparing the calculated premelting behaviour of ccDNA with the experimental data of Burke and Bauer [8]. The final results are shown to be reasonably insensitive to the choice of the model.

Our results form a basis for an estimation of the probability of open and cruciform states in ccDNA. Preliminary data show that our early results based on eq. (2) remain essentially valid if one uses the more sound expression of the superhelix energy obtained here.

References

[1] Vinograd, J., Lebowitz, J. and Watson, R. (1968) J. Mol. Biol. 33, 173–197.
[2] Wang, J. C. (1974) J. Mol. Biol. 89, 783–801.
[3] Hsieh, T. and Wang, J. C. (1975) Biochemistry 14, 527–535.
[4] Wang, J. C. (1974) J. Mol. Biol. 87, 797–816.
[5] Dean, W. W. and Lebowitz, J. (1971) Nature New Biol. 231, 5–8.
[6] Lau, P. and Gray, H. (1979) Nucleic Acids Res. 6, 331–357.
[7] Burke, R. L. and Bauer, W. R. (1978) Nucleic Acids Res. 5, 4819–4836.
[8] Burke, R. L. and Bauer, W. R. (1980) Nucleic Acids Res. 8, 1145–1165.
[9] Laiken, N. (1973) Biopolymers 12, 11–26.
[10] Anshelevich, V. V., Vologodskii, A. V., Lukashin, A. V. and Frank-Kamenetskii, M. D. (1979) Biopolymers 18, 2733–2744.
[11] Vologodskii, A. V., Lukashin, A. V., Anshelevich, V. V. and Frank-Kamenetskii, M. D. (1979) Nucleic Acids Res. 6, 967–982.
[12] Benham, C. J. (1979) Proc. Natl. Acad. Sci. USA 76, 3870–3874.
[13] Banham, C. J. (1980) J. Chem. Phys. 72, 3633–3639.
[14] Depew, R. E. and Wang, J. C. (1975) Proc. Natl. Acad. Sci. USA 72, 4275–4279.
[15] Pulleyblank, D. E., Shure, M. T., Tang, D., Vinograd, J. and Vosberg, H. P. (1975) Proc. Natl. Acad. Sci. USA 72, 4280–4284.
[16] Lilley, D. M. (1980) Proc. Natl. Acad. Sci. USA 77, 6468–6472.
[17] Bauer, W. R. (1978) Annu. Rev. Biophys. Bioeng. 7, 287–313.
[18] Lyubchenko, Yu. L., Vologodskii, A. V. and Frank-Kamenetskii, M. D. (1978) Nature 271, 28–31.