

Solitons hiding in DNA and their possible significance in RNA transcription

E. W. Prohofsky

Department of Physics, Purdue University, West Lafayette, Indiana 47907

(Received 16 February 1988)

We find that the hydrogen-bond–stretch bands of the double helix appear to be nonlinear enough to support solitary-wave energy concentration. Coupling this fact to predictions of our self-consistent theory of helix melting gives rise to speculations of a mechanism for base pair melting in RNA transcription which is consistent with the known energy needs of that process.

I. INTRODUCTION

Several recent papers have discussed the possibility of solitary waves in double-helical DNA associated with the helical acoustic modes.¹ The work was stimulated by observations in DNA of resonant microwave absorption.² Papers have also appeared which have disputed the theoretical predictions of the likelihood of such solitary waves,³ and papers have also appeared which challenge the experimental observations.⁴ Solitary-wave phenomena require nonlinear effects, and both experimental observations⁵ and theoretical analysis⁶ indicate that helix acoustic modes are quite linear in their dispersion and amplitude dependence. In our theoretical studies of helix lattice dynamics of DNA we have found that a different region of the DNA spectrum is far more nonlinear⁷ and seems to satisfy the quantitative requirements for solitary wave effects.

The highly nonlinear modes are the hydrogen-bond–stretch modes at $10\text{--}120\text{ cm}^{-1}$. These modes have been observed in both low-frequency Raman scattering⁸ and Fourier-transform infrared-absorption⁹ experiments. In both cases the modes have been seen to soften in frequency with increasing temperature. We have predicted theoretically that these modes are temperature dependent in agreement with observations.¹⁰

The same vibrational modes have played a central role in the theory of helix melting that we have developed based on helical lattice dynamics. The temperature dependence discussed above, coupled with thermal expansion does lead naturally to hydrogen-bond breakup and helix melting. These results for poly(dG)·poly(dC) predict melting at a temperature in fair agreement with observations. In that work all the parameters were fitted to spectral data at room temperature and no parameters were fitted to melting.¹¹

Helix melting is an essential element in many biological processes. For example, it is essential in the process of RNA transcription. Recent reviews of this process reaffirm the belief that energy must flow into the RNA polymerase complex from its surroundings.² The ability of solitary-wave effects to concentrate excitations in these bands would help in understanding the problems with energy flow in these processes. This biological role for nonlinear effects is in line with the original suggestions of solitary-wave significance in biological processes.¹³

In this paper we discuss our helical lattice dynamics results for the nonlinear modes and show how these results can be used to extract parameters needed to evaluate the possibilities of nonlinear concentration. We then use the theory of melting developed to give order of magnitude estimates of the amplitudes needed to induce base pair melting, and discuss the significance of these results for RNA transcription.

II. HYDROGEN-BOND–STRETCH MODES IN THE DOUBLE HELIX

In our study of the vibrational modes of the double helix, we have found several bands which can be characterized as vibrations that compress the hydrogen bonds between base pairs. There are overlapping bands for all combinations of base pairs, and several pass bands in this region should exist regardless of base sequence.¹⁴ Dispersion curves, i.e., the change in frequency with wave vector, are shown for several such bands in Fig. 1.¹⁵ Theta is the phase shift of the displacements from one unit cell to the next. These bands have local minimum frequencies. In one case the local minimum is in the middle of a band; in other cases it is at a zone edge. The curves could

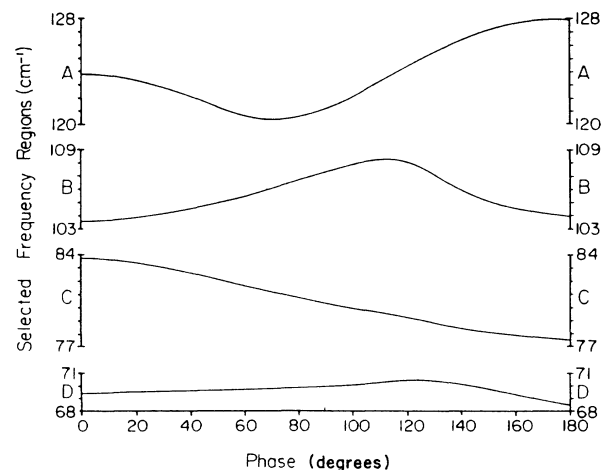


FIG. 1. Dispersion relations for several optical bands in poly(dG)·poly(dC). The frequency is in cm^{-1} , and the phase is the relative phase between neighbor unit cells. In terms of wave vector q , this phase is qa where a is the lattice constant.

be extended about either end, i.e., the total θ range can run from $-180^\circ < \theta \leq 180^\circ$ or $0 \leq \theta < 360^\circ$. The curves would have symmetry about the extension as the positive and negative phase shifts are connected by time-reversal invariance. Correlating frequency with energy, and phase shift with wave vector, and hence momentum for these excitations, one can relate these curves to particles with energy versus momentum relations approximated by

$$E = E_0 + \frac{(\Delta p)^2}{2m} . \quad (1)$$

The effective mass m is then

$$m = \frac{(\Delta p)^2}{2 \Delta E} ,$$

$$\frac{m}{\hbar} = \frac{(\Delta q)^2}{2 \Delta \omega_d}$$

if a is the lattice spacing

$$\frac{ma^2}{\hbar} = \frac{(\Delta \theta)^2}{2 \Delta \omega_d} , \quad (2)$$

where θ is in radians and $\Delta \omega_d$ is the change in frequency due to dispersion effects for a shift in the phase from cell to cell of $\Delta \theta$. In the solitary-wave analysis by Davydov¹³ the wave equation for such a system is given by the nonlinear Schrödinger equation in one dimension

$$\left[i \hbar \frac{\partial}{\partial t} + \frac{\hbar^2}{2m} \frac{\partial^2}{\partial z^2} + G |\psi|^2 \right] \psi(z, t) = 0 , \quad (3)$$

with the substitutions

$$x = z/a, \quad \tau = \frac{\hbar t}{2ma^2}, \quad \text{and} \quad g = \frac{ma^2 G}{\hbar^2} . \quad (4)$$

This becomes

$$\left[i \frac{\partial}{\partial \tau} + \frac{\partial^2}{\partial x^2} + 2g |\psi|^2 \right] \psi(x, \tau) = 0 , \quad (5)$$

which has the solitary-wave solution

$$\psi = \frac{1}{2} \sqrt{g} e^{i(2kx - \omega\tau)} \operatorname{sech} \left[\frac{g(x - 4k\tau)}{2} \right] , \quad (6)$$

with

$$\omega = 4 \left[k^2 - \frac{g^2}{16} \right], \quad k = \frac{maV}{\hbar} , \quad (7)$$

where V is the velocity of the centroid of the solitary wave. Concentration of the excitation occurs in the square-wave excitation case¹³ for $g > g_{\text{cr}}$ where

$$g_{\text{cr}} = \frac{\pi^2}{4l} \quad (8)$$

and l is the extent of the initial excitation along x .

The term $G |\psi|^2$ is a measure of how the energy is lowered by increased concentration of the probability amplitude. In our self-consistent temperature-dependent lattice calculations we have found that the hydrogen-

bond-stretch modes are anharmonic, i.e., they lower in frequency as the excitation level or temperature increase. We have found this shift in frequency to be in rough agreement with the observed frequency shift from low temperature to room temperature.¹⁰ This decrease in energy with increase in concentration of amplitude provides the self-trapping effect which enhances concentration of the excitation. In Fig. 2 we show a calculation¹⁴ of the frequency of several modes in poly(dG)·poly(dC) as a function of temperature for temperatures above room temperature. For vibrational systems the mean wave amplitude squared is proportional to the temperature. The frequency is seen to drop with increasing temperature and, hence, mean squared amplitude.

We can estimate G for particular excitation modes from Fig. 2. In Eq. (3), $|\psi|^2$ is a normalized nondimensional squared wave amplitude. We associate $\Delta |\psi|^2$ with $\Delta T/T$ for regions of curve in Fig. 2. We also associate the appropriate ΔE with $\hbar \Delta \omega_A$ over ΔT . $\Delta \omega_A$ is the change in frequency due to anharmonicity not dispersion. Then

$$\frac{G}{\hbar} = \frac{\Delta \omega_A}{\Delta T/T} . \quad (9)$$

Using this and Eq. (2) in Eq. (4),

$$g = \frac{(\Delta \theta)^2}{2 \Delta \omega_d} \frac{\Delta \omega_A}{\Delta T/T} . \quad (10)$$

For the 85-cm⁻¹ band from Figs. 1 and 2 at $T \approx$ room temperature we estimate $g \approx 10$. For reasons discussed in the next section we estimate l_{min} the minimum initial size of the excitation to be 2.5 in units of a . For this value $g_{\text{cr}} \approx 1$. One can expect nonlinear concentration of excitations in these bands in DNA. In Fig. 2 one can see that G would be even larger at the higher levels of excitation needed for helix melting.

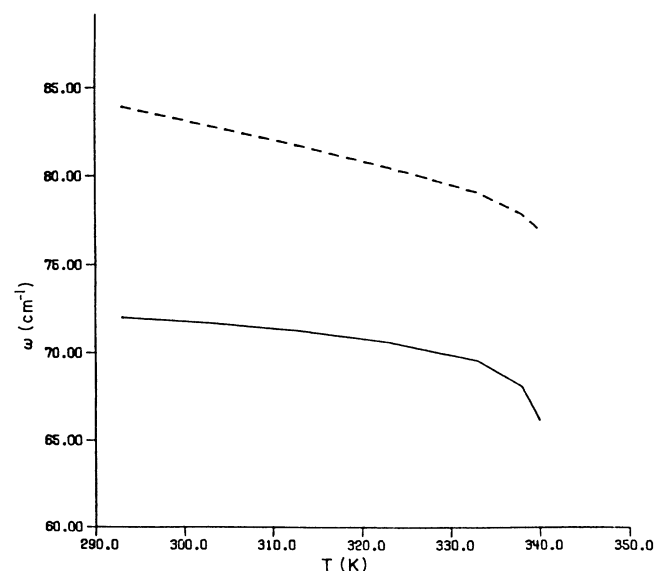


FIG. 2. Calculated change in frequency of selected zone-center modes of poly(dG)·poly(dC) as a function of temperature. For details see Ref. 14.

III. DISCUSSION

From the review of RNA transcription the basic process, a "step," fuses an RNA nucleotide to the growing RNA transcript. For each step a DNA base pair is split on the downstream side, a DNA-RNA hybrid base pair is formed, a DNA-RNA hybrid base pair is split, a DNA base pair upstream is reformed, and a polymerization event in the RNA backbone takes place. As a result of the backbone chemistry a pyrophosphate group is removed from the nucleoside triphosphate which is the source of the added RNA section. The pyrophosphate group is later hydrolyzed but not at a site directly connected to the other operations. The overall energy flow in the transcription complex is such that on the average 0.5 kcal/mole transcribed must be absorbed by the complex from its surroundings.

For the purposes of this discussion we will explore ways in which energy introduced into the complex can bring about base pair melting. The free energy needed to melt a GC base pair is generally accepted to be 3.5 kcal/mole and that for an AT base pair 1 kcal/mole.¹² If inflow of this amount of energy occurred, the net energy requirements of transcription would easily be met. The reason to consider this form of energy transfer to the transcription complex is that we believe it would involve the nonlinear hydrogen-bond stretch modes. In our theory of melting we calculate the vibrational fluctuations D on each hydrogen bond where $D = \langle uu \rangle$ and u is the displacements of atoms from equilibrium due to phonon amplitude. The self-consistent softening of the bond and thermal expansion are calculated based on D . The softening and expansion leads to a melting of the hydrogen bonds. In this sense D , which grows with excitation level, drives the melting. In our model where we looked at melting adjacent to a melted base pair in poly(dG)·poly(dC) we found that when a hydrogen bond has $D \approx 0.040 \text{ \AA}^2$ melting occurred.¹¹

We also investigated the contribution of individual vibrational bands $D(\omega)$ to the total fluctuation D as seen in Fig. 3. The hydrogen-stretch band at $95\text{--}100 \text{ cm}^{-1}$ at 340 K has $D(\omega) \approx 5.5 \times 10^{-3}$ from the entire band. Ignoring the contributions of other bands one would only have to increase the excitation of this band by a factor of 7.27 to bring on melting from the energy in this band alone. Since the thermal occupation is about two phonons at this temperature, one would require $\approx 14\text{--}15$ phonons to bring about melting. The energy of ≈ 14 phonons is approximately the 3.5 kcal/mole generally associated with the binding energy of GC base pair. The localization to the hydrogen bond of these excitations is in the localized amplitudes calculated and plotted in Fig. 3.

If we consider melting a GC base pair in terms of the needed energy in one mode while all others are in thermal equilibrium at room temperature a different estimation is appropriate. One still needs a total of $D \approx 0.04 \text{ \AA}^2$. The thermal activation of all modes at room temperature is $D_R \approx 0.028 \text{ \AA}^2$.⁷ The needed additional fluctuation is $\Delta D = 0.012 \text{ \AA}^2$. For $D(\omega)$ of the $(95\text{--}100)\text{-cm}^{-1}$ band $D(\omega) = 0.0055$. To make up this difference the excitation level would have to be 2.8 times that at 343 K. This re-

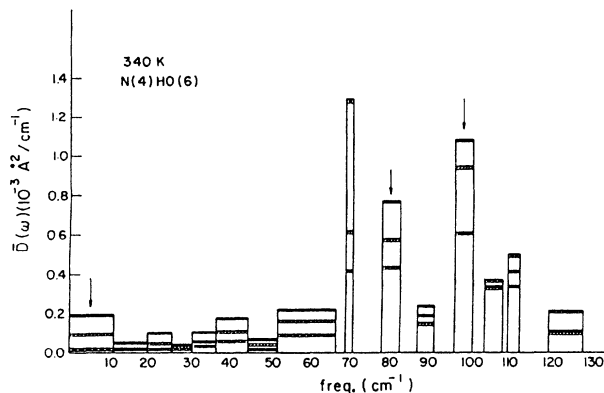


FIG. 3. The contribution to the fluctuation amplitude by individual vibrational bands for the hydrogen bond adjacent to the major groove in poly(dG)·poly(dC). The three divisions are for the two base pairs adjoining a melted base pair on either side and for a base pair far away from the melted base pair. For the $95\text{--}100\text{-cm}^{-1}$ band the highest division is for the (+1) base pair, i.e., $3' \rightarrow 5'$ direction in the G chain. The next highest is for the (-1) base pair, i.e., on the other side of the melt. The lowest division is for the (∞) base pair, i.e., far away from the melt. See Ref. 11 for details.

quires approximately six phonons in a packet arriving at the base pair adjacent to the melted region. The thermal activation level is ≈ 2 . The mean large fluctuation value from Bose-Einstein statistics is $\approx 4\text{--}5$. Such a concentration of energy may be achieved by capturing energy from the surroundings as is apparently needed to explain the energetics of DNA transcription. Such an excitation could travel in unmelted DNA without disrupting the DNA as the $D(\omega)$ for all bands is less for base pairs far from a melted section. $D(\omega)$ for the $(95\text{--}100)\text{-cm}^{-1}$ band is about half the value at a far away cell (∞) as for cell (+1) next to the melt, as seen in Fig. 3. When the energy propagates into the cell adjacent to the melt the change in localized mode behavior greatly increases fluctuation amplitudes and can disrupt the base pair.

In Sec. II we estimated l the initial size of the excitation as being at least ≈ 2.5 base units in length. This is based on the idea of a planar longitudinal wave or shock wave from pyrophosphate hydrolysis impinging on a section of DNA helix from the side. A helical structure is just what is needed to convert such an external excitation into a short-wavelength excitation in the internal coordinates of the DNA because of k -conservation selection rules. For some forms of excitation a tenfold helix would select excitations of 10 base pair wavelength or $\theta = 36^\circ$ on dispersion curves such as Fig. 1. To generate a hydrogen-bond stretch wave where compression of either strand leads to compression of the hydrogen bonds the wavelength would be five base pairs or $\theta = 72^\circ$. We have associated half this wavelength with the minimum extent of the initial square-wave excitation and set $l = 2.5$. Short wavelengths can lead to more localized wave packets and localization or concentration is needed to efficiently melt base pairs adjacent to the already melted area. It is in-

teresting to speculate on how the existence of a small integer-fold helical structure for DNA interacts with the need for energy transfer in process such as transcription in early evolution of the system.

This discussion is couched in terms of linear concepts such as wave packets and thermal occupation of phonons. A proper theory would involve spontaneous development of the highly nonlinear entities which is not simply described in terms of linear excitations. The nonlinear Schrödinger equation used here to determine a criterion for nonlinear effects is far too simple for a case in which the nonlinearities are cooperative leading to the breakdown of mean-field effects and ultimately the melting of bonds. Particular problems arise in such nonlinear

systems.¹⁶ The better nonlinear models of wave propagation more appropriate to the case of dispersive optical modes would allow a more quantitative estimate of many of the energies which have been crudely estimated in this phenomenological discussion. The role of DNA excitations in biological processes may be important, and seems likely to be based on these simple estimates.

ACKNOWLEDGMENTS

This work was supported in part by U.S. Office of Naval Research (ONR) Grant No. N00014-86-K-0252 and U.S. National Institutes of Health (NIH) Grant No. GM-24443.

¹A. C. Scott and J. H. Jensen, *Phys. Lett.* **109A**, 243 (1985).

²G. S. Edwards, C. C. Davis, J. D. Saffer, and M. L. Swicord, *Phys. Rev. Lett.* **53**, 1284 (1984).

³L. L. Van Zandt (private communication).

⁴C. Gabriel, E. H. Grant, R. Tata, P. R. Brown, B. Gestblom, and E. Noreland, *Nature* **328**, 145 (1987).

⁵S. M. Lindsay and J. Powell, in *Structure and Dynamics—Nucleic Acides and Proteins*, edited by E. Clementi and R. H. Sarma (Adenine, New York, 1983), pp. 241–260.

⁶M. N. Mei, M. Kohli, E. W. Prohofsky, and L. L. Van Zandt, *Biopolymers* **20**, 833 (1981).

⁷Y. Kim, K. V. Devi-Prasad, and E. W. Prohofsky, *Phys. Rev. B* **32**, 5185 (1985).

⁸H. Urabe and Y. Tominaga, *J. Phys. Soc. Jpn.* **50**, 3543 (1981).

⁹J. W. Powell, G. S. Edwards, L. Genzel, and A. Wittlin, *Phys.*

Rev. A **35**, 3929 (1987).

¹⁰Y. Kim and E. W. Prohofsky, *Phys. Rev. B* **36**, 3449 (1987).

¹¹Y. Kim and E. W. Prohofsky, *Phys. Rev. B* **35**, 2399 (1987).

¹²T. D. Yager and P. H. von Hippel, in *Escherichiacoli and Salmonella Typhimurium: Cellular and Molecular Biology*, edited by F. Neidhart *et al.* (American Society of Microbiology, New York, in press).

¹³A. S. Dovydyov, *Solitons in Molecular Systems* (Reidel, Dordrecht, 1985).

¹⁴Y. Gao, K. V. Devi-Prasad, and E. W. Prohofsky, *J. Chem. Phys.* **80**, 6291 (1984).

¹⁵E. W. Prohofsky, *Comments Mol. Cell. Biophys.* **2**, 65 (1983).

¹⁶D. W. Brown, K. Lindenberg, and B. J. West, *Phys. Rev. A* **37**, 642 (1988).