Hot Phonon Induced Bond Breaking: Application to the Advance of a Replicating Fork in DNA

W. Zhuang and E. W. Prohofsky
Department of Physics, Purdue University, West Lafayette, Indiana 47907
(Received 21 September 1992)

We have developed an algorithm for calculating the opening of base pairs in the double helix based on a modification of self-consistent phonon theory. The calculated results are in good agreement with experimental observation for both the fluctuational opening in the premelting regime and the critical behavior at the melting transition. We apply the algorithm to a localized structure, a model of the replicating fork, and find that the addition of two hot phonons added to local modes will lead to the separation of the last closed base pair and the advance of the fork. Several other combinations of hot phonons will also bring about separation and advance of the fork as needed in replication.

PACS numbers: 87.10.+e, 63.20.Dj, 63.70.+h, 87.15.By

In many biological processes directed work is done and one would expect, because of the second law of thermodynamics, that energy flow through the system is required. The system is a complex of molecules and flow of energy into molecules is described in terms of excitations of particular modes of the molecule. The goal then in describing molecular engines is to describe how excitations can bring about the action of the engine. Developing detailed microscopic physical theories of this dynamics is hampered by the fact that these systems are both large and highly nonlinear. Common physics methods can deal with one but not both of these elements simultaneously. The principal microscopic approach to such a combined problem is simulation. Because of the computational demands of simulation it is limited to small time scale events in large systems. The problem discussed in this work involves a large macromolecule involved in the highly nonlinear act of breaking particular bonds where the time scale for the occurrence of the disrupted state is long. We have developed a statistical approach that can predict the bond breaking behavior in terms of either thermal excitations or the absorption of particular quantum excitations, i.e., hot phonons. The method applies to long time scale events. We know of no other practical method for making the calculations described.

The helix is a large molecule and localization of the quantum excitations is necessary for a finite amount of excitations to accomplish anything. The problem requires the solution of the modes of a structure known as the replicating fork. This system is close to disruption by thermal levels of excitation and our calculations indicate that as little as two additional far-infrared phonons into a local mode of frequency ≈ 303 cm^{-1} can lead to rapid dissociation of the bonds at the replicating fork. Dissociation can also be caused by the addition of five 106 or six 89 cm^{-1} phonons to inband modes at the fork. The replicating fork is the point at which a double helix splits into two single strands. The conformation of the double helix is well known but less is known of the conformation in the single strand region which is further complicated by the attachment of the replicating complex of enzymes. Our model is fairly accurate for the double helical region and crudely approximates the single strand separated region; the atoms on the single strand side are taken to be in the same positions as if the helix were double helix. All dynamic interaction between the single strands has been cut and they are essentially invisible to each other. The justification for this approximation is that the important dynamics is principally in the stretch of the interbase H bonds and this dynamics only occurs in the part which is double helix. Changes in conformation of the separated strands may only effect the H-bond dynamics to higher order. To the extent that this is true the calculated frequencies and energies should be a reasonable first approximation to the actual fork dynamics.

As can be seen in Table II the local mode at 303 cm^{-1} is principally a scissoring motion opening and closing the major groove; the largest stretch is on the H bond adjacent to the major groove. The excitation first disrupts the major groove H bond and this then destabilizes the remaining H bonds leading to base pair opening. The 106 cm^{-1} inband mode more evenly effects the major groove and center H bonds as seen in Table III. The 89 cm^{-1} mode is from Table III, centered on the central H bond, and is a breathing mode of the central H bond.

At body temperature, the thermal occupancy of the 303 cm^{-1} modes from Bose-Einstein statistics is ≈ 0.32. These phonons are therefore rare, as are a packet of five or six of the inband phonons. They could arise from an energetic event. The absorption of this energy into the helix can be thought of as the necessary input of energy for directed work described at the molecular level. The energy of the two local mode phonons is ≈ 0.075 eV. That for the required inband phonons is ≈ 0.066 eV. The energy of the likely energetic event, the hydrolysis of ATP, is ≈ 0.3 eV. The dissociation would also occur for some combination of phonons such as one 303 and three 89 cm^{-1} modes, etc. If hot phonon effects are important in bringing about separation of the helix, some biological process may be enhanced by irradiation at critical frequencies. An experimental verification of the role of specific modes in such processes may be possible.

The model we calculate is shown at the bottom of Fig. 1 which has an intact semi-infinite double helix to the right and two semi-infinite single strands to the left. The long solid lines are the DNA backbones and the crossing lines represent bases. The angle of the backbones at the fork is drawn only to allow space between the open base pairs of the single strands. In our calculations all segments have proper helical geometry. In replication, the single strands are doubled. Our calculations ignore the enzymes and newly added DNA and deal only with the DNA strands shown. The number of dynamic elements per base pair double helical unit cell is 41 and the dimension of the dynamic matrix in the helix calculations is 123×123. Force constants are fitted to Raman and ir observations. The calculations then give good agreement with the observed spectra of double helical DNA and the model can be considered a fairly realistic dynamic model for the double helix. Base pairs are either adenine and thymine (AT) pairs or guanine cytosine (GC) pairs. The GC pairs are more strongly bound and the disruption of these pairs is the rate limiting step in advance of the replicating fork. We therefore do our calculation on poly(dG)·poly(dC), the simplest system of GC pairs.

The bases in the double helical part are connected by H bonds which at physiological temperatures are highly nonlinear. The nonlinear effects lead to effective force constants derived by a modification of self-consistent phonon approximation theory (MSPA) [1]. This method leads to a theory of bond disruption and even cooperative melting [2] which is in good agreement with observation.

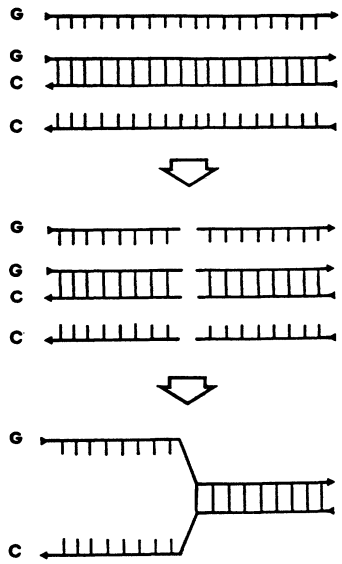

FIG. 1. The top section represents a poly(dG) single strand, a poly(dG)·poly(dC) double helix, and a poly(dC) single strand. The helix is unwound in the figure but all elements are in proper helical geometry in the calculations. Helical lattice methods are used for the dispersion calculations. The middle section is the same strands after cutting in half. The bottom section is after splicing some elements to create a model of a replicating fork. The bend in backbone at the bottom is not related to the geometry used in the calculation.

The fork dynamics is first solved in the harmonic approximation. We start by solving by lattice dynamics an infinite double helix and two infinite single strands that have the same bases as in the double helix. The model is shown at the top of Fig. 1. We then cut each element in half by introducing the negative of all force constants across the cut in a defect matrix generating the system in the middle of Fig. 1. We then splice two single strands to the opposite half double strand by adding appropriate force constants between the proper degrees of freedom to the defect matrix as at the bottom of Fig. 1. We then use a Green function to solve for the modes of the new structure.

From the imaginary part of the Green function we can calculate the mean square displacement, or fluctuation, $D_\alpha=\langle u_\alpha u_\alpha\rangle$. The principal contribution to the nonlinearities of the H bonds is motion that separates the end atoms of the bond. The largest contribution is from breathing modes where motion is transverse to the helix axis and directly stretches the H bonds. The total square stretch, however, has contributions from all modes, some with quite different character. New effective MSPA force constants for an intact H bond are calculated from

$$\phi_\alpha^{\rm int}=C_\alpha\int_{r_\alpha^{\rm min}}^{\infty}dr\,e^{-(r-R_\alpha)^2/2D_\alpha}\frac{d^2}{dr_\alpha^2}V_\alpha(r)\,, \tag{1}$$

where

$$C_\alpha^{-1}=\int_{r_\alpha^{\rm min}}^{\infty}dr\exp[-(r-R_\alpha)^2/2D_\alpha] \tag{2}$$

and

$$V_\alpha(r)=V_\alpha^0(1-e^{-a_\alpha(r-r_\alpha^0)})^2-V_\alpha^0\,, \tag{3}$$

where V_α^0, a_α, and r_α^0 are the parameters of the Morse potential listed in Table I. $r_\alpha^{\rm min}$ is the near position where the potential energy equals the ionization energy of the potential. We define the open pair probability of a single interbase H bond as [2]

$$P_\alpha^{\rm op}=C_\alpha\int_{L_\alpha^{\rm max}}^{\infty}dr\exp[-(r-R_\alpha)^2/2D_\alpha]\,. \tag{4}$$

$L_\alpha^{\rm max}$ is the separation of atoms that cause a disrupting of the H bond and has been determined by an analysis of the stability of H bonds with increasing temperature [3].

TABLE I. Parameters used for the interbase H bonds of poly(dG)·poly(dC).

Bond	V^0 (mdyn Å)	a (Å^{-1})	r^0 (Å)	$L^{\rm max}$ (Å)
N(4)-H-O(6)	0.0257	2.885	2.693	3.132
N(1)-H-N(3)	0.0184	2.351	2.804	3.057
N(2)-H-O(2)	0.0251	2.818	2.702	3.051

TABLE II. Localized modes and their contribution to mean-square H-bond stretch.

Freq. (cm^{-1})	Phonons in equilibrium	Mean-square H-bond stretch per phonon ($Å^2$) N(4)-H-O(6)	N(1)-H-N(3)	N(2)-H-O(2)	Threshold phonons	Threshold energy (eV)
302.60	0.32	7.345×10^{-4}	5.019×10^{-5}	1.489×10^{-6}	2	0.075
311.80	0.31	4.773×10^{-4}	5.437×10^{-5}	7.889×10^{-5}	3	0.116
381.72	0.20	2.621×10^{-4}	1.148×10^{-4}	1.064×10^{-3}	3	0.142
134.06	1.16	4.317×10^{-5}	1.675×10^{-4}	2.650×10^{-4}	10	0.166

We conceive of a particular H bond as disrupted when the instantaneous separation of the end atoms is greater than $L_\alpha^{\max}$. The time scale for electron transitions as well as H-atom motion is fast compared to the time scale of H-bond end atom breathing motion which we calculate at 85–303 cm^{-1} frequency. The lasting disruption of the H bonds is therefore associated with this slowest motion and the extension to $L_\alpha^{\max}$.

In MSPA bounded true potentials are approximated by unbounded effective harmonic potentials. The harmonic interactions, no matter how weak, cannot display bond disruption and must be further modified to study disruption. The ability to display disruption can be incorporated in a self-consistent manner using factors already developed in this study. We set

$$\phi_\alpha=(1-P_\alpha^{\rm op})\phi_\alpha^{\rm int}, \tag{5}$$

where $P_\alpha^{\rm op}$ is from Eq. (4) and $\phi_\alpha^{\rm int}$ is the intact effective force constant from Eq. (1). The incorporation of such a bond breaking operator into MSPA calculations has produced melting curves in very good agreement with experimental observation [2].

The difference between each iteration of ϕ_α and the earlier ϕ_α's are treated as an additional local defect. This defect matrix is then used in a further Green function calculation in which we recalculate D_α. The entire process is iterated until self-consistency. In the final solution we have the structural and anharmonic factors both contributing synergistically to a self-consistent final solution. The solution is now an anharmonic solution which depends on the level of excitation. In our thermal melting papers this procedure was performed at each temperature to find $P_\alpha^{\rm op}(T)$. Poly(dG)·poly(dC) has three interbase H bonds at each base pair. Then we define the open pair probability as

$$P^{\rm op}=\prod_{\alpha=1}^{3}P_\alpha^{\rm op}. \tag{6}$$

The fluctuation D_α for a particular H bond is

$$D_\alpha=\frac{\hbar}{\pi}\int d\omega\,{\rm Im}[G_{\alpha\alpha}(\omega^2)]\coth[\hbar\omega/2kT] + \frac{\hbar}{2}\sum_k\frac{1}{\omega_k}q_{\alpha k}q_{\alpha k}^*\coth[\hbar\omega/2kT], \tag{7}$$

where the first term is the contribution from inband modes and the second from local modes. The $q_{\alpha k}$ is the displacement component on the α coordinate of the eigenvector of the kth local mode. In both terms the coth term is the thermal weighting, i.e., $\bar{n}_k+\frac{1}{2}$. We can simulate an athermal distribution of phonons by putting in $\bar{n}_k+n_k$ for $\bar{n}_k$ where n_k is the extra hot phonons added to fork mode k. The iteration to self-consistency is run as before and $P^{\rm op}$ is monitored. This then is the anharmonic solution to the problem at the higher level of excitation. We consider the base pair to have separated when $P^{\rm op}$ approaches 1. Our calculated values for the open base probability are typically <0.1 at an excitation of one phonon less than threshold. On reaching self-consistency at threshold the value jumps typically to 0.999. This is likely due to the sharp cooperativity of our model which includes the $1-P_\alpha^{\rm op}$ factor. Even without this factor the open probability would be $>\frac{1}{2}$ and would indicate a base pair dissociation. Table II shows the data for when this condition is satisfied for several local modes. Table III shows the same information for inband modes where the threshold number of phonons is calculated assuming the energy is all in the virtual local mode, i.e., in a wave packet initially localized to the base pair at the fork.

A wave packet in an inband mode will in time propagate away from the fork base pair with an appropriate group velocity. The calculated mean group velocity for the two inband modes discussed is such that the packet would undergo a minimum of seven oscillations during the time it traveled the distance to the next base pair. Since the period of oscillation is the characteristic time

TABLE III. Inband modes' contribution to mean-square H-bond stretch.

Frequency (cm^{-1})	Mean-square H-bond stretch (10^{-4} $Å^2$) N(4)-H-O(6)	N(1)-H-N(3)	N(2)-H-O(2)	Threshold Phonon No.	Energy (eV)	Group velocity (m/s)
72.06–76.19	11.69	4.553	1.077	11	0.104	83.8
82.84–88.79	6.634	12.64	9.842	6	0.066	121.0
91.44–94.48	3.300	12.94	17.65	12	0.141	61.6
99.67–106.48	6.663	2.758	0.460	5	0.066	138.1
107.61–111.81	13.54	7.562	6.762	8	0.111	85.2

for dissociation, the diffusion of energy is not a limitation to the action of these hot phonons. The 303 cm^{-1} local mode is specific to GC pairs. AT helices have local modes at forks but at different frequencies. The inband frequencies discussed here are part of the broad ≈ 85 cm^{-1} band common to all DNA helices which is observed experimentally [4]. Hot phonon base opening effects at these frequencies due to inband absorption are likely to operate for all DNA sequences.

In DNA replication cyclase enzymes are attached to one single strand fairly close to the replicating fork. These enzymes have been shown to burn fuel in the form of adenine triphosphate (ATP) molecules when advancing the fork. Since the enzymes are unchanged, the bulk of this energy must be dissipated in the generation of excitations which can be used to bring about advance of the fork in two ways. In the more direct scheme some of the excitations give rise to hot phonons that can be directly absorbed at the fork bringing about hot phonon advance of the fork. In the second scheme the excitations created by hydrolysis become thermalized before absorption in the fork. They then maintain the temperature of the system. Thermal fluctuations are absorbed by the fork to bring about melting. In this second scenario the melting calculation should use thermal values for all fork excitations. It is the first scenario that conforms to the idea of a Carnot-like engine running off energy input. The second scenario runs on thermal fluctuations and its direction is driven by a chemical potential rather than energy flow. There is no reason why both scenarios cannot be operative.

The second or thermal fluctuation scenario seems more likely in DNA transcription. There is no cyclase enzyme and the associated energetic event, the hydrolysis of pyrophosphate, takes place far from the transcription fork [5]. Earlier thermal calculations of ours on bulk DNA [6] indicate fluctuational base opening times to be milliseconds. They are likely faster at a transcription fork [7]. The observed transcription rate is ≈ 50 base pairs per second [5] and the less efficient thermal opening rate is at least 20 times faster than needed for observed transcription rates. Replication is, however, a much faster process. It runs at ≈ 1000 base pairs per second matching the bulk fluctuation opening rate and the source of energetic excitations is quite close to the fork. The speed of the process makes it unlikely that it is dependent on thermal fluctuational base opening for advancing the fork [8]. An additional advantage to the efficiency of the hot phonon mechanism is that the advance of the fork is correlated in time to an event on cyclase. Cyclase is assumed to wind the single strand, advancing one base pair per hydrolyzed ATP. Energy released by the cyclase will propagate in both directions along the one-dimensional strand of single helix with small leakage into the loosely coupled solvent. This would lead to approximately one-half of the energy dissipated propagating down one single strand to the fork. The nonlinearity introduced by this energy can break down the orthogonalities of our modes, leading to absorption into fork modes.

It should be pointed out that in slow replicating systems, under certain circumstances, the advance of the fork has been observed to occur when the burning of the fuel molecule has been disabled [9]. The mechanism of base opening in the advancing of the fork seems to be a dual mechanism, one that can use energy flow to achieve rapid processivity but is capable of running at some slower rate on thermal fluctuations consistent with our analysis.

This work was supported in part by ONR Contract No. N00014-92-K-1232.

[1] Y. Kim and K. V. Devi-Prasad, and E. W. Prohofsky, Phys. Rev. B **32**, 5185–5189 (1985).
[2] Y. Z. Chen and E. W. Prohofsky, J. Chem. Phys. **94**, 4665–4667 (1991).
[3] Y. Z. Chen, Y. Feng, and E. W. Prohofsky, Phys. Rev. B **42**, 11335–11338 (1990).
[4] T. Weidlich and S. M. Lindsay, Phys. Chem. **92**, 6497 (1488).
[5] W. R. McClure, Annu. Rev. Biochem. **54**, 171–204 (1985).
[6] M. Gueron, M. Kochoyan, and J. L. Leroy, Nature (London) **328**, 89–92 (1987).
[7] Y. Z. Chen, W. Zhuang, and E. W. Prohofsky, J. Biomolec. Struct. Dyn. **10**, 415–427 (1992).
[8] R. H. Das, G. T. Yarranton, and M. L. Gefter, J. Biol. Chem. **264**, 8069–8073 (1980).
[9] T. C. Jarvis, J. W. Newport, and P. H. von Hippel, J. Biol. Chem. **266**, 1830–1840 (1991).