# Theory of pressure-dependent melting of the DNA double helix: Role of strained hydrogen bonds

Y. Z. Chen and E. W. Prohofsky
*Department of Physics, Purdue University, West Lafayette, Indiana 47907*
(Received 22 July 1992)

In this paper we study the effect of hydrostatic pressure on the thermal stability of DNA base pairs. The thermal stability of a base pair is predominantly determined by the fluctuational motion of the interbase hydrogen bonds in the base pair. We postulate that the pressure exerted on the surface of DNA creates a compressive stress in the interbase hydrogen bonds. This stress in turn induces a strain in these hydrogen bonds. The effect of this compressive stress on the thermal fluctuational motion of the strained interbase hydrogen bonds in DNA can be analyzed by the modified self-consistent phonon approximation theory. We calculate individual hydrogen bond disruption probabilities and base pair opening probabilities of two *B*-conformation DNA polymers, an adenine-thymine copolymer and a guanine-cytosine homopolymer, at various pressures and at temperatures both in the premelting region and in the helix-coil transition region. Our calculated pressure dependence of the melting temperature for both polymers is in fair agreement with experimental observations. Our values for the melting temperatures are also in agreement with observation and our theory is successful as a theory of cooperative melting.

PACS number(s): 87.10.+e, 87.15.By, 63.20.Dj, 63.70.+h

## I. INTRODUCTION

A DNA molecule is a double helix of two strands of nucleotides (bases and backbone sugar-phosphate groups). These two strands are linked together through hydrogen bonds between the complementary bases [1]. The dynamic stability of a base pair is largely determined by the stability of these interbase H bonds. This feature has been exploited in our earlier studies of DNA, and it has led to the development of a modified self-consistent phonon approximation theory (MSPA theory) of DNA thermal denaturation. The calculated interbase H-bond disruption probabilities and base pair thermal fluctuational opening probabilities for a number of DNA polymers turned out to be in fair agreement with various experiments both at premelting temperatures and in the helix-coil transition region [2–7]. MSPA is successful at predicting actual melting temperatures without using parameters fitted to melting [3,6,7]. The parameters used in MSPA are fitted to room-temperature data. Our studies exhibited the relationship between individual interbase H-bond thermal motion and the observed DNA proton exchange phenomenon. We showed that the breakdown of an amino interbase H bond can be associated with the "open state" needed to facilitate that particular amino proton exchange [2,4]. This association held for the two amino interbase H bonds in a guanine-cytosine (GC) pair [2] as well as for the one amino interbase H bond in an adenine-thymine (AT) pair [4]. We also showed that a base pair open configuration, i.e., that for which all interbase H bonds are disrupted in a single base pair, could be associated with the open state needed to facilitate imino proton exchange [2,4].

The interbase H bonds lie directly between the strands. As a result, static forces, such as the intramolecular force from the nonbonded phosphate-phosphate Coulomb repulsion and the exterior force from the hydrostatic pressure of the environment, that act between the two strands of a DNA double helix can affect the equilibrium configuration of these H bonds. We postulate that the stress generated from a cross-strand static force induces a strain in the intervening interbase H bonds. These strained H bonds, in turn, generate a compensating stress that keeps the system in equilibrium. This in turn affects the vibrational normal modes of the system as well as the dynamic stability of the base pair. It is this stress-strain relationship that dictates the effect of a cross-strand static force on the dynamic stability of a base pair.

In our earlier studies [8,9], we have used this stress-strain relationship to study the salt effect on the thermal stability of the interbase H bonds in a DNA homopolymer, poly(dG)·poly(dC), which indicates a single strand of all G bases paired with one of all C bases. In that case, the cross-strand static force originates from the cross-strand phosphate-phosphate Coulomb repulsion. This Coulomb force is partially screened by the counterions in the solvent. Therefore, its strength is dependent on the bulk salt concentration of the solvent. Our work led to a theory of the melting temperature of poly(dG)·poly(dC) as a function of salt concentration, in fair agreement with observations [9]. We have also found [7,10] that the concept of stressed H bonds is essential in predicting observed premelting transitions in another DNA homopolymer, poly(dA)·poly(dT). In poly(dA)·poly(dT), there exists a well-organized hydration network—a spine of hydration in its minor groove [11,12]. This hydration spine is connected to the base atoms on the floor of the minor groove through water–base-atom H bonds. The disruption of these water–base-atom H bonds is responsible for the observed premelting transition in this DNA polymer [13]. The spine water molecules are inside the minor groove and therefore are subject to additional pocket van der Waals attraction associated with the minor groove. The compression arising from this pocket attraction

 

should also induce a strain in the water–base-atom H bonds. Our calculations showed that only this stressed H-bond model can give rise to a broadened spine dissociation transition, in agreement with observations.

In this paper, we explore a much simpler case of the stress-strain relationship in the interbase H bonds to study the pressure-temperature stability of the interbase H bonds in DNA polymers. We calculate interbase H-bond disruption probabilities and base pair opening probabilities of a DNA polymer as a function of hydrostatic pressure and temperature. If strained bonds are of general importance in the stability of macromolecules, they should also be important here. There are two different types of base pairs in DNA. One is an adenine-thymine base pair, which has two interbase H bonds. The other is a guanine-cytosine base pair, which has three interbase H bonds. Figure 1 shows the structure of these two types of base pairs. To study the effect of pressure on both types of base pairs, we consider in this work two DNA polymers, each containing only one type of base pair. For the GC pair model, we choose a homopolymer, poly(dG)·poly(dC). For the AT pair model, we choose a copolymer, poly[d(A—T)]. Unlike the AT homopolymer, there is no hydration spine in the minor groove of poly[d(A—T)]. This copolymer is selected so that we do not need to deal with extra interactions between DNA and a hydration spine. We will show that the thermal stability of the interbase H bonds of both types of base pairs is enhanced at elevated pressures such that the melting temperatures increase with pressure linearly, in agreement with observations.

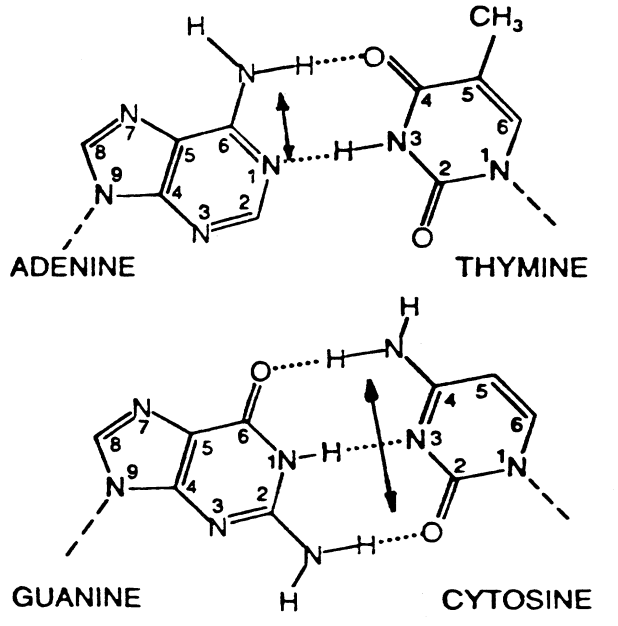


FIG. 1. Structure of an adenine-thymine base pair and the structure of a guanine-cytosine base pair. The sugar-phosphate backbones are not shown in the figure. An arrow in the interbase region represents the width of the cross section for force compressing the imino interbase H bond (see text for details).

This paper makes no attempt to calculate from first principles the ground-state conformation of the helix. Attempts to calculate the conformation are as yet unsuccessful, particularly when long-range Coulomb repulsion is properly included. We do follow the path, proven so successful in solid-state studies, of determining the excitations of a system from a given ground state. The role of salt and water of hydration are important in determining that ground state, but the calculation of such a ground state from first principles is hindered by the fact that there are many conformations very close in energy to the true ground state. The accuracy with which one can calculate the energy of these conformations is less than the differences in energy between them. A way around this is to expand all calculations from an experimentally determined ground-state configuration. In such an expansion, the question of what determines the true ground-state conformation does not explicitly arise. What we do concentrate on are changes in forces like pressure that arise from the surrounding medium. All such factors enter as changes from a previous condition known to be stable.

We do not explicitly put in factors relating to hydrophobic-hydrophilic effects. Our assumption is that, although they are certainly important in helix stability, they do not change appreciably up through the disruption of interbase H bonds. They may change appreciably after. Therefore, changes due to such effects are part of the stabilizing effects implicitly balanced in the force constant expansion about equilibrium positions. It is the strength of an approach like ours that complex interactions such as hydration effects, that do not change appreciably, can be eliminated as problems in the calculations. Only factors that enter into the dynamic factors need be considered in a dynamic analysis, as distinct from an equilibrium analysis.

## II. MSPA FORMULATION OF DNA DYNAMICS

The MSPA theory is a theory that emulates a canonical ensemble system in a self-consistent box [14]. This theory is based on the self-consistent phonon formulation of anharmonic lattice dynamics [15]. In this theory, a system such as a DNA duplex is described at the atomic level of detail. In principle, we conceive of a DNA helix with all interatomic interactions represented by realistic bounded potentials. These potentials are then replaced, at an atomic level of detail, by MSPA self-consistent unbounded interactions, i.e., connected by effective harmonic force constants. These force constants can be evaluated self-consistently at each temperature and in a particular environment.

A DNA molecule can be described by a Hamiltonian similar to the standard Hamiltonian used in simulation studies [16,17]. Following Baird's argument [18], we choose a Morse potential as the potential for the interbase H bonds and other H bonds. The Hamiltonian is given by

$$H=\sum_j \frac{\mathbf{P}_j^2}{2M_j}+\sum_{\text{bonds}} K_r(r-r^{\text{eq}})^2+\sum_{\text{angles}} K_\theta(\theta-\theta^{\text{eq}})^2 + \sum_{\text{dihedrals}} K_\Phi[1+\cos(n\Phi-\delta)] + \sum_{\text{H bonds}} [V^0(1-e^{-a(r-r^0)})^2-V^0] + \sum_{\text{nonbond pairs}} \left[\frac{A}{r^{12}}-\frac{B}{r^6}+\frac{q_1q_2}{\epsilon r}\right] . \quad (1)$$

In this Hamiltonian, the bond stretch and angle bending terms are harmonic. The torsion terms for small dihedral deviation motions can also be regarded as harmonic. For the remaining anharmonic terms, one can expand the thermal average of a potential around its equilibrium value in the following manner:

$$\langle V(r)\rangle = V(r^{\text{eq}})+\frac{1}{2}\frac{d^2V(r^{\text{eq}})}{dr^2}\langle dr^2\rangle+\cdots . \quad (2)$$

This anharmonic potential can be treated in the framework of the self-consistent phonon approximation (SCPA) theory of anharmonic lattice dynamics, explored in a large number of publications. An extensive review article has been written by Werthamer [15], which also contains references to earlier work. In the SCPA approach, the cumulant expansion of the free energy for the Hamiltonian system [Eq. (1)] is given by

$$F=F_0+\sum_{\text{H bonds+nonbonds}} V_i(r^{\text{eq}}) + \tfrac{1}{2}\sum_{\text{H bonds+nonbonds}} \{[\exp(\tfrac{1}{2}D_i\nabla\nabla)-1]V_i(r^{\text{eq}}) - \tfrac{1}{2}D_i\phi_i\}+\cdots , \quad (3)$$

where $F_0$ is the contribution to the free energy from a MSPA effective harmonic system, represented by a phonon Hamiltonian $H_0$:

$$H_0=\sum_j \frac{\mathbf{P}_j^2}{2M_j}+\sum \tfrac{1}{2}\phi_i\mathbf{u}_i^2 . \quad (4)$$

The second sum in Eq. (4) is for all the bonding, angle bending, torsion, H-bonding, and nonbonded interactions. $\mathbf{u}_i$ is the displacement vector of the relevant motion in the internal coordinate system. $\phi_i$ is the related force constant. The force constants are determined self-consistently at different temperatures and in different environments. They are determined such that the effective phonon Hamiltonian [Eq. (4)] is adjusted to the standard Hamiltonian [Eq. (1)] through a minimization procedure on the cumulant expansion of the free energy [Eq. (3)] [14,15]. The force constant related to an anharmonic potential is found in this way to be given by an integration over the second derivative of that potential.

The SCPA is a phenomenologically appealing approach, but it has a number of problems that arise due to its inability to incorporate thermal expansion and deal correctly with odd power terms in the potential expanded about the potential minimum. The MSPA theory is a modification of the SCPA, which retains the simple phenomenological features of the SCPA but specifically introduces an effective thermal expansion. In MSPA, thermal expansion of a bond is introduced by simply shifting the origin of the vibration from the potential minimum to a self-consistent thermal mean position [14]. This approach is in keeping with the spirit of an effective oscillatory solution that results from self-consistent calculations.

A simplification of our application of MSPA is that not all the force constants are determined by a SCPA integration. Some of the force constants that can be shown to vary little with temperature and pressure are assumed to be temperature and pressure independent. Other nonbonded force constants are determined by an effective potential. All of these force constants are fitted to experimental observations at a temperature well below the melting temperature. The refined valence force constants can be found in Ref. [19]. The long-range nonbonded force constants are formulated in Ref. [20], and the van der Waals base-stacking force constants are formulated in Ref. [21]. Only the interbase H-bond force constants are calculated by using the SCPA integration. This simplification is a good approximation because, compared to the other bonds, the hydrogen bonds are much weaker, and the disruption of these bonds plays a dominant role in DNA thermal denaturation.

For a system with a regular repeating sequence, it is useful to exploit the structural symmetry inherent in the system to reduce the calculation. Therefore we divide a DNA sequence into unit cells. A unit cell contains a single repeating section, which is composed of one or several base pairs and the associated backbones. For example, the unit cell of a homopolymer contains one base pair, and the unit cell of a copolymer contains two base pairs. The calculation is then reduced to a number of calculations each of the dimensionality of a single unit cell. As a result of the use of the DNA helical symmetry in our calculation, a phase angle $-\pi<\theta\le\pi$ appears in the MSPA formalism. Using this description, the $F_0$ for a repeating system is given by [14,15]

$$F_0=-k_BT\ln \mathrm{Tr}\exp\left[-\frac{H_0}{k_BT}\right] = \frac{Nk_BT}{\pi}\sum_\lambda \int_0^\pi d\theta \ln\left[2\sinh\left[\frac{\omega_\lambda(\theta)}{2k_BT}\right]\right] , \quad (5)$$

where $N$ is the total number of unit cells, $k_B$ is Boltzmann's constant, $\lambda$ is the index of the $\lambda$th band, and $\omega_\lambda(\theta)$ is the eigenfrequency of the secular equation in the degree of freedom of a unit cell derived from the MSPA effective Hamiltonian [Eq. (4)].

The force constant of an interbase H bond is found in MSPA to be given by an integration over the second derivative of a Morse potential weighted by a vibrational distribution function. This force constant corresponds to the one for an intact bond:

$$\phi_i^{\text{int}}=A_i\int_{-h_i}^{\infty} du\, e^{-u^2/2D_i}\frac{d^2}{du^2}V_{\text{Morse}}(R_i+u) , \quad (6)$$

where $D_i$ is the mean vibrational square amplitude:

TABLE I. Key parameters of the H bonds in AT base pair and GC base pair. $a$, $r^0$, and $V^0$ are the Morse parameters, and $L^{\max}$ is the maximum bond stretch length.

| System | Bond | $a$ (Å$^{-1}$) | $r^0$ (Å) | $V^0$ (kcal/mol) | $L^{\max}$ (Å) |
|---|---|---|---|---|---|
| AT base pair | N(6)—H—O(4) | 1.961 | 2.738 | 3.543 | 3.191 |
| | N(1)—H—N(3) | 1.909 | 2.789 | 3.269 | 3.119 |
| GC base pair | N(4)—H—O(6) | 2.884 | 2.693 | 3.701 | 3.124 |
| | N(3)—H—N(1) | 2.349 | 2.805 | 2.650 | 3.048 |
| | O(2)—H—N(2) | 2.784 | 2.706 | 3.586 | 3.049 |

$$D_i=\frac{1}{\pi}\sum_{\lambda}\int_0^{\pi}d\theta\frac{|s_i^{\lambda}(\theta)|^2}{2\omega_{\lambda}(\theta)}\coth\left[\frac{\omega_{\lambda}(\theta)}{2k_BT}\right], \tag{7}$$

$R_i$ is the mean atom-atom distance, which will be given later, $A_i$ is a normalization factor,

$$A_i^{-1}=\int_{-h_i}^{\infty}du\,e^{-u^2/2D_i}, \tag{8}$$

and $-h_i$ is the hard-core inner boundary. The Morse potential is given in Eq. (1) (the fifth term), and the Morse parameters for both AT and GC base pairs are given in Table I.

A problem associated with the mean-field treatment outlines above, i.e., the use of the symmetry related to the repeating sequence, is that the cooperative effects contributed from the open base pairs and disrupted bonds are neglected in the calculation. These effects can be neglected at premelting temperatures because of low open-state probabilities. However, they cannot be ignored at temperatures near and in the helix-coil transition region. In our earlier work [3,6], we have shown that these cooperative effects can be incorporated by use of a probability-weighted linear combination of the intact-state force constant and the open-state force constant. An interbase H-bond force constant $\phi_i^h$ and an interstrand base-stacking force constant $\phi_{ll'}^s$ are therefore given by

$$\begin{aligned}\phi_i^h&=(1-P_i)\phi_i^{\rm int}+P_i\phi_i^{\rm op},\\ \phi_{ll'}^s&=[1-(P_l^{\rm op}+P_{l'}^{\rm op})^{1/2}]\phi_{ll'}^{\rm int}+(P_l^{\rm op}P_{l'}^{\rm op})^{1/2}\phi_{ll'}^{\rm op}.\end{aligned} \tag{9}$$

In practice, an open bond force constant can be regarded as zero: $\phi^{\rm op}=0$. In Eq. (9), $P_i$ is the individual bond disruption probability and $P_l^{\rm op}$ is the base pair open probability. They are given respectively by

$$\begin{aligned}P_i&=A_i\int_{L_i^{\max}}^{\infty}du\,\exp[-(u-R_i)^2/2D_i],\\ P_l^{\rm op}&=\prod_i^{n_l}P_i,\end{aligned} \tag{10}$$

where $n_l$ is the number of interbase H bonds in the $l$th base pair, and $L_i^{\max}$ is the maximum stretch length of the H bond before disruption. $L_i^{\max}$ is the point at which the self-consistent solution becomes unphysical, such that the mean potential energy of the H bonds begins to decrease [2]. It is calculated in a model in which the cooperativity of the broken bonds and open base pairs is completely ignored. The parameters of this trial model, however, are the same as the model with the cooperative effects. The calculated $L_i^{\max}$ is therefore independent of the status of other bonds and base pairs. Table I gives the value of $L_i^{\max}$ for the interbase H bonds in both AT and GC base pairs.

In MSPA, the mean bond length $R_i^{\rm int}$ of an intact H bond is not determined by using the potential minimum. Instead, it is determined by taking into consideration thermal expansion of the bond. The thermal expansion of an interbase H bond can be given by simply shifting the origin of the H-bond vibration displacement from the potential minimum to a self-consistent mean position. This mean position is determined by the classical condition that, at the classical turnaround point, all oscillator energy is in potential energy. The oscillator would then be described by harmonic displacements from a point midway between the classical turnaround points. The value of $R_i^{\rm int}$ can therefore be determined by the condition $V_{\rm Morse}(R_i^{\rm int}+\mu_i)=V_{\rm Morse}(R_i^{\rm int}-\mu_i)$, where $\mu_i=2\sqrt{2D_i\ln 2}$. We find

$$R_i^{\rm int}=r_i^0+\frac{1}{a_i}\ln[\cosh(\mu_i a_i)]. \tag{11}$$

When the cooperative effect from the open base pairs is taken into account, the mean H-bond length $R_i$ that enters into MSPA formalism should also be given by a probability-weighted linear combination of intact bond length $R_i^{\rm int}$ and open bond end-atom distance $R_i^{\rm op}$:

$$R_i=(1-P_l^{\rm op})R_i^{\rm int}+P_l^{\rm op}R_i^{\rm op}. \tag{12}$$

$R_i^{\rm op}$ can be determined [3,6] by two theoretical considerations. The first is the condition that $R_i^{\rm op}\to L_i^{\max}$ when $P_l^{\rm op}\to 0$. The second consideration is the requirement that, when $P_l^{\rm op}\to 1$, the calculated $R_i^{\rm op}$ should be such that they give a consistent value of $P_l^{\rm op}$. $R_i^{\rm op}$ is found in this way to be

$$R_i^{\rm op}=L_i^{\max}+2P_i\sqrt{2D_i}. \tag{13}$$

Given initial force constants, one can obtain the normal-mode eigenfrequencies and eigenvectors of the system by solving the secular equation obtained from the MSPA effective phonon Hamiltonian [Eq. (4)]. These normal-mode eigenfrequencies and eigenvectors can then

be used to calculate vibrational amplitudes, bond lengths, bond disruption probabilities, etc. These in turn are used to redefine the effective force constants. This process continues until the system reaches self-consistency.

## III. STRESS-INDUCED STRAIN IN THE INTERBASE H BONDS

The formulation outlined in Sec. II is based on the condition that all the static forces are balanced such that the DNA double helix is stabilized at the x-ray observed configuration. However, there are situations in which some of the static forces change when the environment is changed. This would subsequently induce a change in the system so that the reaction forces adjust to the change and a new equilibrium is achieved, i.e., a new balance of forces is reached. In DNA, the most significant structural changes occur in the interbase H bonds. This is because these bonds are the weakest in the system. One can see from MSPA formulation that a change in the interbase H-bond equilibrium configuration or an undetectable change of the mean bond length will significantly affect the H-bond force constants. As a result, the vibrational normal modes as well as the dynamic stability of these interbase H bonds are changed. In this section, we discuss how this effect can be explicitly incorporated into MSPA formalism.

There are several cross-strand forces that are dependent on environment. In this study, we focus on the static force that arises from the hydrostatic pressure in the environment. The hydrostatic pressure in the solvent creates a compressive stress that is expressed across all the intervening interbase H bonds in a base pair. This cross-strand stress $f^P(P)\equiv\sum_i f_i^P(P)$ [$f_i^P(P)$ is the force on an individual H bond] in turn generates a compensating stress that arises from an induced strain or change in the mean H-bond length. This compensating stress $f^h(P)$ can be calculated from the Morse potential $V_{\rm Morse}$:

$$f^h(P)\equiv\sum_i^{n_l} f_i^h(P)=-\sum_i^{n_l}\frac{dV_{\rm Morse}(R_i)}{dR}\ . \tag{14}$$

In addition to the compressive stress from the pressure and the compensating stress from the strained H bonds, there is another cross-strand force that arises from nearest-neighbor interstrand van der Waals base-stacking interactions. This force can be calculated using a van der Waals potential $V_{\rm VW}$. The calculated force is then projected onto the interbase H-bond orientation within the base plane, and the resulting effective static force is given by

$$f^s(P)=-\tfrac{1}{2}\sum_{\alpha,\beta}\frac{dV_{\rm VW}(r_{\alpha\beta}^{\rm eq})}{dr}\mathbf{r}_{\alpha\beta}^0\cdot\mathbf{r}_h^0\ . \tag{15}$$

The share of this force on an individual H bond can be given by $f_i^s(P)=f^s(P)/n_h$, where $n_h$ is the number of interbase H bonds per base pair. These forces have to be balanced at atmospheric pressure $P_0$ to give the x-ray observed stable double-helix configuration, from which our coordinates of DNA are obtained. This implies that

$$f^h(P_0)+f^p(P_0)+f^s(P_0)=0\ . \tag{16}$$

At an elevated pressure $P$, the strength of cross-strand compressive stress would change, and this results in the change of interbase H-bond length. Let us assume the change of the $i$th H bond to be $dr_i^P$. The mean bond length is now $R_i(P_0,T_0)+dr_i^P$, $R_i(P_0,T_0)$ is the x-ray observed bond length at room temperature $T_0$, from which we can calculate a new static force $f^h(P)$ for interbase H bonding from the Morse potential and a new static force for cross-strand base stacking from the van der Waals potential. These new forces should also be balanced to give rise to a stable double helical structure, and they satisfy the same equation as Eq. (16). We obtain

$$f^h(P)-f^h(P_0)=-df\equiv\sum_i df_i\ , \tag{17}$$

where $df=f^P(P)-f^P(P_0)+f^s(P)-f^s(P_0)$ is the change of the total cross-strand force, and $df_i=f_i^P(P)-f_i^P(P_0)+f_i^s(P)-f_i^s(P_0)$ is the share of the total force on the $i$th interbase H bond. By substituting the Morse potential and Eq. (14) into Eq. (17), we obtain

$$dr_i^P=\frac{1}{a_i}\ln\frac{\eta_i}{1+[(\eta_i-1)^2-2df_i/a_iV_i^0]^{1/2}}\ , \tag{18}$$

where $\eta_i=2\exp\{-a_i[R_i(P_0,T_0)]-r_i^0\}$, and $a_i$, $r_i^0$, and $V_i^0$ are the Morse parameters for the $i$th H bond. The Morse potential can be found in Eq. (1), and the Morse parameters for both AT and GC paris are given in Table I. The mean H-bond length $R_i$ that enters into MSPA calculation now becomes

$$R_i=(1-P_l^{\rm op})(R_i^{\rm int}+dr_i^P)+P_l^{\rm op}R_i^{\rm op}\ . \tag{19}$$

## IV. THE EFFECT OF PRESSURE ON THE STABILITY OF INTERBASE H BONDS

The cross-strand static force $f_i^P(P)$ (per bond) can be calculated by projecting the forces normal to the surface of a section of DNA onto the interbase H-bond orientation and apportioning this force among the H bonds. This section covers one base and its associated sugar-phosphate backbone. This force is equivalent to the product of pressure $P$ and the area of an effective cross section that is perpendicular to the interbase H-bond orientation. We take this effective cross section to be a parallelepiped. The height of this is just the helical pitch of DNA. That height is relatively straightforward, as neighbor base pairs of the helix exclude water above and below a given base pair. The width of the parallelepiped is complicated by the fact that solvent can penetrate the grooves right up to the amino H bonds, i.e., those involving N—H—O bonds, and partially surround the atoms at the ends of these bonds. We find that we have to treat the application of pressure to the various interbase H bonds differently for additional reasons as well, as the amino interbase H bonds are connected to the rigid base-ring structure by an additional flexible element. The bond between an amino group and the base ring can bend relative to the ring structure, reducing the compression of

the amino interbase H bond, and thus shift more of the compressive force to the imino interbase H bond. The imino or N(1)—H—N(3) bond does not have this additional degree of flexibility and must directly absorb the compressive force between the bases. For this reason, we first derive a cross-section area for force compressing the imino interbase H bond. The width of this area for the AT pair is taken to be the distance from the adenine imino N(1) to the C(6) [or equivalently from the thymine N(3) to C(4)] plus half the distance to the amino group (see Fig. 1). For the GC pair, the width is from the guanine C(2) to the C(6) plus half the length of the C(2)—N(2) bond plus half the length of the C(6)—O(6) bond. The same width can also be drawn from the cytosine base (see Fig. 1).

As pointed out above, the outer H bonds (amino interbase H bonds) are connected to the more rigid ring structures by additional CN and CO bonds that have an additional degree of flexibility, the bond-bending degree of freedom, between the H bond and the base ring. Therefore, the force exerted on the bulk of the rigid structure of the base-sugar-phosphate group is unlikely to be fully transmitted to these outer H bonds. Only the force exerted on the bond end atoms can be transmitted directly to these outer H bonds. The pressure in the solvent can partially surround these atoms and exert a force perpendicular to the outer H bonds, which can act to pry the H bond apart. This is because the solvent water molecules can wedge into the grooves and collide with the hydrogen atoms of the outer H bonds. This force is perpendicular to the H bonds and therefore it can deform the H bonds. The extent of the deformation of these H bonds should depend on the pressure. Several three-dimensional hydrogen-bond calculations [22,23] have shown that the force constant of a deformed hydrogen bond becomes significantly weaker as the degree of deformation increases. This weakening of outer H bonds by pressure-induced bond bending should also be taken into consideration. To the authors' best knowledge, no theoretical calculation has been carried out for such deformed H bonds. Because of the lack of theoretical estimates, we are unable to formulate the pressure effect on these bonds. Therefore, we simply assume that the pressure-induced bond compression in the amino interbase H bonds is canceled by the pressure-induced bond deformation. This is equivalent to taking $f_i^P(P)=0$ for these amino interbase H bonds. The imino interbase H bond [N(1)—H—N(3) bond] of an AT pair, which resides on the floor of the minor groove, is shielded by base atoms. Therefore, the deformation of this H bond by the pressure should be much smaller than that in an amino interbase H bond, and its effect can be neglected.

We find that this apportionment of the force among the H bonds is the one that gives quantitative agreement with experiments, as shown below. A different analysis of the areas and force distribution will not alter the basic linear pressure dependence of the melting temperature. It will only alter the slope. The slope, however, is not too sensitive to the limited range of alternative formulations of this force distribution. In calculations with different approximations, the slope changes at most by a factor of 2.

By the above discussion, the effective static force $f_i^P(P)$ across the central interbase H bond can be given by the following equation:

$$f_i^P(P)=PA_i^{\text{eff}}, \tag{20}$$

where $A_i^{\text{eff}}$ is the area of the effective base cross section: $A_i^{\text{eff}}=hd$. $h$ is the helical pitch (for a $B$-DNA double helix, $h=3.4$ Å). The value of $d$ depends on the base pair type. For a GC pair, it is the distance between the center point of the O(6)—C(6) bond and the center point of the N(2)—C(2) bond in the guanine base. For an AT pair, it is the distance between the center point of the N(6)—C(6) bond and the N(1) atom in the adenine base. The length of $d$ for both AT and GC pairs is illustrated in Fig. 1.

## V. RESULTS AND DISCUSSION

We employ the formalism outlined in Secs. II–IV to specifically calculate individual interbase H-bond disruption probabilities and base pair opening probabilities for both the AT copolymer poly[d(A—T)] and the GC homopolymer poly(dG)·poly(dC) as a function of both pressure and temperature. In this work, we only consider the effect associated with the pressure on the stability of DNA. Other effects, such as that associated with the solvent salt, are not considered. Experiments [24,25] indicated that, at lower salt concentrations, the melting temperature $T_m$ of DNA becomes less dependent on the pressure. There seems to exist a critical salt concentration at which the $T_m$ becomes independent of pressure, and there are suggestions that the pressure may even destabilize the DNA below that salt concentration. Our MSPA calculations of DNA correspond to a nominal salt concentration. This nominal salt concentration is related to a number of physiological salt concentrations corre-

TABLE II. Dependence of the static force $f^P$, across the imino interbase H bond [N(1)—H—N(3) bond], amino interbase H-bond [N(6)—H—O(4) bond] disruption probability $P_{\text{aam}}$, base pair opening probability $P^{\text{op}}$, and the melting temperature $T_m$, of poly[d(A—T)] on hydrostatic pressure $P$. The abbreviation bp denotes base pair.

| | | $P_{\text{aam}}$ | | | $P^{\text{op}}$ (units of $10^{-3}$) | | | |
|---|---|---|---|---|---|---|---|---|
| $P$ (bar) | $f^P$ (mdyn/bp) | $T=293$ K | $T=313$ K | $T=323$ K | $T=293$ K | $T=313$ K | $T=323$ K | $T_m$ (K) |
| 1 | $3.44\times10^{-7}$ | 0.045 | 0.075 | 0.104 | 3.76 | 8.29 | 14.10 | 330 |
| 1000 | $3.44\times10^{-4}$ | 0.044 | 0.071 | 0.096 | 3.27 | 7.01 | 11.15 | 333 |
| 4000 | $1.38\times10^{-3}$ | 0.041 | 0.064 | 0.083 | 2.11 | 4.39 | 6.59 | 342 |
| 8000 | $2.75\times10^{-3}$ | 0.039 | 0.058 | 0.073 | 1.21 | 2.47 | 3.63 | 354 |

sponding to various fiber x-ray studies and the Raman and ir studies from which our parameters are refined. We assume that this nominal salt concentration is higher than the critical salt concentration.

The calculated force $f_i^P(P)$ across the central interbase H bond, together with the individual amino interbase H-bond disruption probabilities and base pair opening probabilities for both polymers, are given in Tables II and III, respectively. We carry out the calculation in a pressure range from atmospheric pressure ($\sim 1$ bar) to 8000 bar (1 bar $=10^5$ N/m$^2$). We find that at premelting temperatures the individual amino interbase H-bond disruption probabilities decrease with increasing pressure. Because the static force $f_i^P(P)$ across the central interbase H bond is relatively small, the dependence of base pair opening probability $P^{op}$ on pressure is not strong. For example, the $P^{op}$ of the AT copolymer is 0.0037 at atmospheric pressure. It only reduces by a factor of 3, to 0.0012, at 8000 bar. The $P^{op}$ for the GC homopolymer has a slightly higher pressure dependence. At atmospheric pressure, the $P^{op}$ is $5.42\times10^{-6}$. It is reduced by an order of magnitude, to $4.38\times10^{-7}$, at 8000 bar.

In Fig. 2, we plot the temperature-dependent profile of $P^{op}$ for both the AT copolymer and the GC homopolymer at several pressures. We find from Fig. 2 that the interbase H bonds of the intact double helix are all disrupted ($P^{op}\rightarrow 1$) when the temperature increases to a critical point. Near the critical temperature, the disruption is highly cooperative, in agreement with many experimental observations. We have shown in our earlier studies [3,6,7] that the MSPA approach is successful as a theory of cooperative melting of DNA, and the calculated critical temperature can be regarded as the melting temperature of DNA. In Tables II and III we give the calculated melting temperature $T_m$ for both the AT copolymer and the GC homopolymer as a function of pressure. The $T_m$ versus $P$ curves for the two polymers are also presented in Fig. 3. Our calculated melting temperatures for both polymers at atmospheric pressure are the same as those given in an earlier calculation [7]. We have shown in that work that these calculated melting temperatures are in agreement with experimental measurements carried out at the MSPA nominal salt concentration ($\sim 0.05M$ NaCl).

A number of experiments have been carried out that measure the effect of pressure on the melting temperature of DNA [24–30]. These experiments show that the melting temperature of DNA is a monotonically increasing

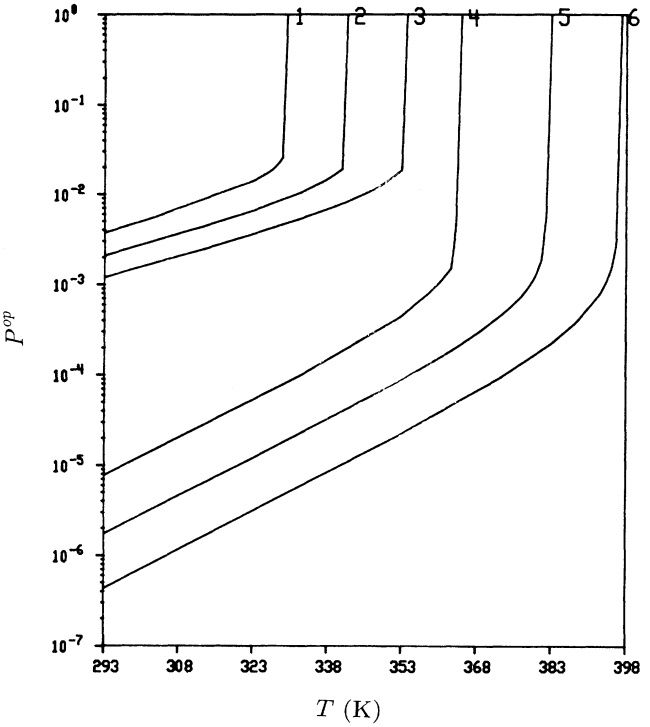


FIG. 2. Calculated base pair opening probability of poly[d(A—T)] and poly(dG)·poly(dC) as a function of temperature at several pressures. The assignments of the lines are (1) AT polymer at 1 bar, (2) AT polymer at 4000 bar, (3) AT polymer at 8000 bar, (4) GC polymer at 1 bar, (5) GC polymer at 4000 bar, and (6) GC polymer at 8000 bar.

function of pressure at least up to 4000 bar. The slope of $T_m$ versus $P$ fits to a linear curve in this pressure range. The derivative of the slope $dT_m/dP$ depends on the GC content of the DNA [30], although the variation is small. A DNA with a higher GC content was found to have a larger $dT_m/dP$ value. These features are all seen in Fig. 3. Both curves in Fig. 3 are roughly linear below 4000 bar. As can be seen in Table II, in this pressure range our calculated $dT_m/dP$ for the AT copolymer poly[d(A—T)] is $\sim 3$ K/kbar (K is kelvin). This is compared to the observed value of 2.43 K/kbar for poly[d(A—T)] [30]. As can be seen in Table III, our calculated $dT_m/dP$ for the GC homopolymer poly(dG)·poly(dC) is $\sim 4.5$ K/kbar. This is compared to the experimentally extrapolated

TABLE III. Dependence of the static force $f^P$, across the central interbase H bond (N(1)—H—N(3) bond), base pair opening probability $P^{op}$, and the melting temperature $T_m$, of poly(dG)·poly(dC) on hydrostatic pressure $P$.

| | | $P^{op}$ | | | |
|---|---|---|---|---|---|
| $P$ (bar) | $f^P$ (mdyn/bp) | $T=293$ K | $T=323$ K | $T=353$ K | $T_m$ (K) |
| 1 | $1.15\times10^{-6}$ | $7.91\times10^{-6}$ | $5.49\times10^{-5}$ | $4.47\times10^{-4}$ | 365 |
| 1000 | $1.15\times10^{-3}$ | $5.41\times10^{-6}$ | $3.67\times10^{-5}$ | $3.08\times10^{-4}$ | 369 |
| 2000 | $2.31\times10^{-3}$ | $3.75\times10^{-6}$ | $2.52\times10^{-5}$ | $1.99\times10^{-4}$ | 374 |
| 4000 | $4.62\times10^{-3}$ | $1.76\times10^{-6}$ | $1.21\times10^{-5}$ | $8.83\times10^{-5}$ | 383 |
| 6000 | $6.93\times10^{-3}$ | $8.72\times10^{-7}$ | $6.09\times10^{-6}$ | $4.37\times10^{-5}$ | 389 |
| 8000 | $9.25\times10^{-3}$ | $4.39\times10^{-7}$ | $3.13\times10^{-6}$ | $2.26\times10^{-5}$ | 396 |

TABLE IV. Dependence of the cytosine amino interbase H-bond (N(4)—H—O(6) bond) disruption probability $P_{cam}$, and guanine amino interbase H-bond (O(2)—H—N(2) bond) disruption probability $P_{gam}$, of poly(dG)·poly(dC) on hydrostatic pressure $P$.

| | $P_{cam}$ (units of $10^{-3}$) | | | $P_{gam}$ (units of $10^{-2}$) | | |
|---|---|---|---|---|---|---|
| $P$ (bar) | $T=293$ K | $T=323$ K | $T=353$ K | $T=293$ K | $T=323$ K | $T=353$ K |
| 1 | 5.75 | 14.20 | 37.20 | 3.23 | 6.20 | 12.59 |
| 1000 | 5.37 | 13.08 | 33.72 | 3.10 | 5.79 | 11.92 |
| 2000 | 5.07 | 12.15 | 30.18 | 2.98 | 5.49 | 10.89 |
| 4000 | 4.57 | 10.79 | 25.50 | 2.77 | 5.03 | 9.51 |
| 6000 | 4.22 | 9.86 | 22.72 | 2.61 | 4.71 | 8.67 |
| 8000 | 3.93 | 9.12 | 20.74 | 2.49 | 4.44 | 8.06 |

value of 4.75 K/kbar for a DNA with 100% GC content [30].

Our calculations at pressures higher than 4000 bar show that the AT curve follows the same linear curve as that in the lower-pressure range. However, the GC curve begins to depart from the lower-pressure linear curve above 4000 bar. (See Table IV.) The calculated rate of $dT_m/dP$ for poly(dG)·poly(dC) becomes slightly lower.

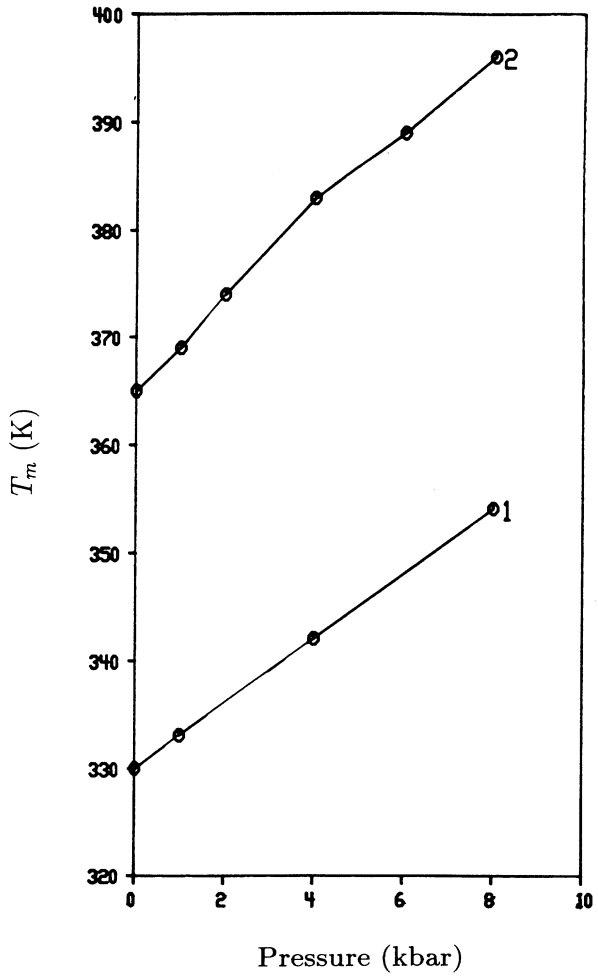


FIG. 3. Calculated melting temperature $T_m$ of poly[d(A—T)] and poly(dG)·poly(dC) as a function of pressure. Line 1 is for the AT polymer and line 2 is for the GC polymer.

We have found no experimental measurements of the DNA melting temperature above 4000 bar. Therefore it is difficult to quantitatively access our calculated melting temperatures at these high pressures. It is worth noting that, for some proteins, their melting temperatures do not increase indefinitely as the pressure is raised [31,32]. At a sufficiently high pressure ($\sim$3000 bar), the melting temperatures of these proteins actually decrease rapidly with further pressure increase. One is naturally interested in the question of whether a similar phenomenon can be found in DNA. Experiments on DNA have found no low-temperature denaturation of DNA at pressures up to 10 000 bar [24]. However, since these experiments did not reach a high enough temperature to actually observe melting of DNA at these very high pressures, it does not rule out the possibility that the increasing rate of the melting temperature of some DNA may become smaller at this pressure range. Our calculation for the GC polymer may, in fact, indicate a possible slowdown of the increasing rate of $T_m$ at these elevated pressures. It is also not surprising that our calculation shows such a slowdown only for the GC polymer, because both our calculation and experiments show that DNA polymers with higher GC contents are more sensitive to pressure. This higher pressure sensitivity is related to the fact that a GC pair has a larger effective cross section than an AT pair. Another factor is that a GC pair has an additional amino interbase H bond.

We point out that our analysis is based on the assumption that several effects on the outer interbase H bonds caused by pressure completely cancel. This assumption seems to work well both for an AT polymer and a GC polymer at pressures at least up to 4000 bar, as shown by the agreement between our calculation and observations. This assumption is therefore principally justified by agreement with experiment when it is applied. This assumption can be dropped without altering the predicted linear dependence of melting temperature on pressure. The slope is only affected by a factor of 2 for several other reasonable alternative assumptions. This particular choice brings about the best fit for both AT and GC base pairs.

## ACKNOWLEDGMENT

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

[1] R. E. Dickerson, Sci. Am. **249**, 94 (1983).
[2] Y. Z. Chen, Y. Feng, and E. W. Prohofsky, Biopolymers **31**, 139 (1991).
[3] Y. Z. Chen and E. W. Prohofsky, J. Chem. Phys. **94**, 4665 (1991).
[4] Y. Z. Chen, W. Zhuang, and E. W. Prohofsky, Biopolymers **31**, 1273 (1991).
[5] Y. Z. Chen and E. W. Prohofsky, Nucleic, Acids Res. **20**, 415 (1992).
[6] Y. Z. Chen and E. W. Prohofsky, Biopolymers (to be published).
[7] Y. Z. Chen and E. W. Prohofsky (unpublished).
[8] Y. Z. Chen, W. Zhuang, and E. W. Prohofsky, Biopolymers **32**, 1123 (1992).
[9] Y. Z. Chen and E. W. Prohofsky (unpublished).
[10] Y. Z. Chen and E. W. Prohofsky (unpublished).
[11] V. P. Chuprina, Nucleic Acids Res. **15**, 293 (1987).
[12] A. V. Teplukhin, V. I. Poltev, and V. P. Chuprina, Biopolymers **32**, 1445 (1992).
[13] J. E. Herrera and J. B. Chaires, Biochemistry **28**, 1993 (1989).
[14] Y. Gao, K. V. Devi-Prasad, and E. W. Prohofsky, J. Chem. Phys. **80**, 6291 (1984).
[15] N. R. Werthamer, Phys. Rev. B **1**, 572 (1970).
[16] B. Tidor, K. K. Irikura, B. R. Brooks, and M. Karplus, J. Biomol. Struct. Dyn. **1**, 231 (1984).
[17] S. J. Weiner, P. A. Kollman, D. T. Nguyen, and D. A. Case, J. Comput. Chem. **7**, 230 (1986).
[18] N. C. Baird, Int. J. Quantum Chem. Symp. **1**, 49 (1974).
[19] K. C. Lu, E. W. Prohofsky, and L. L. Van Zandt, Biopolymers **16**, 2491 (1977).
[20] K. V. Devi-Prasad and E. W. Prohofsky, Biopolymers **23**, 1795 (1984).
[21] V. V. Prabhu, L. Young, and E. W. Prohofsky, Phys. Rev. B **39**, 5436 (1989).
[22] R. Schroeder and E. R. Lippincott, J. Phys. Chem. **61**, 921 (1957).
[23] L. L. Van Zandt and W. K. Schroll, J. Biomol. Struct. Dyn. **8**, 431 (1990).
[24] S. A. Hawley and R. M. Macleod, Biopolymers **13**, 1417 (1974).
[25] E. Nordmeier, J. Phys. Chem. **96**, 1494 (1992).
[26] C. G. Heden, T. Lindahl, and I. Toplin, Acta Chem. Scand. **18**, 1150 (1964).
[27] B. Weida and S. J. Gill, Biochim. Biophys. Acta **112**, 179 (1966).
[28] F. Hughes and R. F. Steiner, Biopolymers **4**, 1081 (1966).
[29] T. E. Gunter and K. K. Gunter, Biopolymers **11**, 667 (1972).
[30] S. A. Hawley and R. M. Macleod, Biopolymers **16**, 1833 (1977).
[31] S. A. Hawley, Biochemistry **10**, 2436 (1971).
[32] A. Zipp and W. Kauzmann, Biochemistry **12**, 4217 (1973).