

# Salt dependent premelting base pair opening probabilities of *B* and *Z* DNA Poly [d(G-C)] and significance for the *B*-*Z* transition

Y. Z. Chen and E. W. Prohofsky

Department of Physics, Purdue University West Lafayette, Indiana 47907-1396 USA

**ABSTRACT** We calculate room temperature thermal fluctuational base pair opening probabilities of *B* and *Z* DNA Poly[d(G-C)] at various salt concentrations and discuss the significance of thermal fluctuation in facilitating base pair disruption during *B* to *Z* transition. Our calculated base pair opening probability of the *B* DNA at lower salt concentrations and the probability of the *Z* DNA at high salt concentrations are in agreement with observations. The salt dependence of the probabilities indicates a *B* to *Z* transition at a salt concentration close to the observed concentration.

## INTRODUCTION

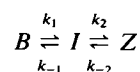
The *B* to *Z* transition involves a complete turnover of every base pair in the DNA double helix. One view of the mechanism of the transition involves base pair separation and unstacking before the separated bases flip over (1, 2, 3). Others have proposed mechanisms which don't require base separation (4). Here we consider the base separation case because we can quantify base opening probabilities, and quantitative calculations relevant to this case can be done. We have recently calculated the base pair separation probabilities in *B* DNA as a function of NaCl concentration and found good agreement with observations of both premelting fluctuational opening rates (5) and the salt dependent melting temperatures (Chen, Y. Z., and E. W. Prohofsky, manuscript submitted for publication). To the extent that base separation is involved in the *B* to *Z* transition, calculations of the base pair opening probabilities can give insight into the mechanism of transition. We find that, with the incorporation of a number of assumptions, our calculations do predict the value of salt at which the transition occurs, as well as predicting the observed fluctuational base pair opening probabilities for both *B* and *Z* DNA far away from the *B*-*Z* transition region and well below the melting temperature. This quantitative agreement strengthens the probability that base opening is involved.

The *B* to *Z* transition has many elements in common with spin reversal transitions in magnetic systems. In the spin case, each spin must reverse its orientation and in the *B* to *Z* transition every base pair has to flip over. Such transitions are characterized by domain formation and domain growth rather than an instantaneous transition involving all of the sample. The instantaneous transition requires an intermediate state of massive disruption of the spin orientation. In the domain growth model, the disruption is confined to a small region at the domain boundaries and this reduction in size greatly reduces the

energy associated with the disrupted region. The transition is then characterized by the formation of a new domain and the advance of the domain boundary into and through the old domain. This domain growth model seems to be supported by the experimental evidence that at most only a small number of base pairs are disrupted at any one time in the *B* to *Z* transition (1, 6). On the other hand, there is considerable evidence that some small number of base fluctuational openings are always present in both *B* and *Z* DNA (7). There is no reason to assume that this process stops at the *B*-*Z* transition although the number of open pairs, including those between domains, is small and not easily detected. In our calculation the open pair probabilities are  $\approx 10^{-6} - 10^{-7}$ . Concepts such as the cooperative unit for *B*-*Z* transition would relate to the size of domains that rapidly convert, rather than the size of the region of disruption at the domain boundary.

The simplest analysis of the dynamics of a *B* to *Z* transition could be carried out by assuming the simultaneous existence of two domains separated by a domain boundary region and to then determine the direction of advance of that boundary region. In DNA one would assume a *B* region and a *Z* region separated by a small region with conformation disruption and likely unpaired bases. We assume that unpaired pairs are necessary for the domain conversion, and we assume that the boundary is most likely to advance into the domain that presents the most fluctuational base pair openings adjoining the boundary. The newly formed domain is further stabilized by favorable free energy. We therefore calculate the base pair opening probability for both *B*-DNA and *Z*-DNA as a function of salt concentration and look for the value of concentration where the probabilities cross.

To be specific, we consider a system capable of the following localized transitions:



where *B* is a base pair in *B*-conformation attached to the

Address correspondence to Dr. Earl W. Prohofsky, Department of Physics, Purdue University, 1396 Physics Building, West Lafayette, IN 47907-1396, USA.

end of a *B* domain. *I* is the base pair with the unspecified conformation and unspecified base orientation in the boundary between a *B* and *Z* domain, and *Z* is the base pair with *Z*-conformation attached to the end of a *Z* domain. In the base separation model *I* becomes the common open state of *B*- and *Z*-form base pair in the *B*-*Z* transition. Therefore we can associate the probability of a spontaneous base pair opening for *B* and *Z* conformation as  $P_B^{\text{op}} = k_1/k_{-1}$  and  $P_Z^{\text{op}} = k_{-2}/k_2$ , respectively. Our calculations determine open state probabilities, we do not determine the  $k_i$  rate parameters directly.

We describe a DNA at the atomic level of detail. We assume that the hydrogen atoms are bound to their parent atoms and their masses are added to the masses of these parent atoms. The coordinates are from fiber and crystal data. The system can be represented by a Hamiltonian similar to the standard form used in simulations:

$$H = \sum_{\alpha} \frac{P_{\alpha}^2}{2M_{\alpha}} + \sum_{\text{bonds}} K_r(r - r^{\text{eq}})^2 + \sum_{\text{angles}} K_{\theta}(\theta - \theta^{\text{eq}})^2 + \sum_{\text{dihedrals}} K_{\phi}[1 + \cos(n\phi - \delta)] + \sum_{\text{H-bonds}} [V^0(1 - e^{-a(r-r^0)})^2 - V^0] + \sum_{\text{non-bond-pairs}} \left[ \frac{A}{r^{12}} - \frac{B}{r^6} + \frac{q_1 q_2}{\epsilon r} \right]. \quad (1)$$

We use a Morse potential as the potential for the H-bonds. This potential is an effective potential for the H-bond end atoms and the Morse parameters for a GC base pair can be found in (5). The valence force fields for the bases and the backbones are from references 8 and 9, respectively. This Hamiltonian can be replaced by a MSPA effective harmonic or phonon Hamiltonian in the sense that the cumulant expansion of the free energy  $F$  of the Hamiltonian system (Eq. 1):

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

is minimized (per reference 10). Here  $F_0$  is the free energy of the MSPA effective harmonic system and  $D_i$  is the mean square vibrational amplitude.

We employ Soumpasis et al.'s potential of mean force (SPMF) approach (11) in treating the effect of the contribution to the free energy arising from the coupling of the hydrated DNA phosphate charge distribution to bulk water and the diffuse cloud of mobile hydrated ions. We replace the equilibrium electrostatic interactions between phosphate groups by the effective SPMF interactions obtained from statistical mechanics. The contribution of these effective interactions to the free energy is given by (per reference 11):

$$F_1 = \sum_{j>j'}^{N_{\text{ph}}} W_{11}(r_{jj'}^{\text{eq}}) \quad (3)$$

$W_{11}(r_{jj'}^{\text{eq}})$  is the SPMF of two anions in a homogeneous 1:1 electrolyte and its detailed formalism can be found in (reference 11). The static dielectric constants that we use to calculate the SPMF can be found in (reference 5).

In addition to the valence force field, we also consider the contribution of the nonbonded interactions to the dynamical motions of DNA. The long-range nonbonded force constants are calculated by a simplified approach in which these force constants are given by second derivatives of two effective potentials (12). The temperature dependence of the cross strand base stacking force constants is assumed to be the same as that of the interbase H-bonds (13). While the equations of motions are solved for all degrees of freedom, the selfconsistent loop calculation can be reduced to the degree of freedom of the interbase H-bonds (10). For a system like Poly[d(G-C)] it is useful to divide it into unit cells according to the helical symmetry inherent in the system. A unit cell contains two base pairs and the associated backbones. The calculation can then be reduced to a number of calculations each of the dimensionality of a single unit cell. The use of the DNA helical symmetry in this manner introduces vibrational bands characterized by a phase angle  $\theta$  which resides in the first Brillouin zone ( $-\pi < \theta \leq \pi$ ) for eigenfrequencies and eigenvectors of the Secular equation of the system.

The mean vibrational stretch of an interbase H-bond can be given by

$$D_{li} = \frac{1}{\pi} \sum_{\lambda} \int_0^{\pi} d\theta \frac{|s_{li}^{\lambda}(\theta)|^2}{2\omega_{\lambda}(\theta)} \coth \left[ \frac{\omega_{\lambda}(\theta)}{2k_B T} \right] \quad (4)$$

where  $k_B$  is the Boltzmann's constant,  $T$  is the absolute temperature,  $l$  and  $i$  are the index of the base pairs in a unit cell and the index of the interbase H-bonds in base pair, respectively,  $\lambda$  is the phonon band number,  $\omega_{\lambda}(\theta)$  is the eigenfrequency and  $s_{li}^{\lambda}(\theta)$  is the projection of the bond end atom eigenvectors onto the bond orientation. With this  $D_{li}$  new effective MSPA force constants for intact H-bonds are calculated from an integration over the second derivative of a Morse potential weighted by a vibrational distribution function:

$$\phi_{li}^{\text{int}} = A_{li} \int_{r_{\text{min}}}^{\infty} dr e^{-(r-R_{li})^2/2D_{li}} \frac{d^2}{dr^2} V_{\text{Morse}}(r) \quad (5)$$

where  $A_{li}$  is a normalization constant and  $R_{li}$  is the mean bond length which will be given below. To study base pair separation we define the bond disruption probability  $P_{li}$  and base pair opening probability  $P^{\text{op}}$  as (per reference 10):

$$P_l^{\text{op}} = \prod_i P_{li} \equiv \prod_i A_{li} \int_{L_{\text{min}}}^{\infty} dr e^{-(r-R_{li})^2/2D_{li}}. \quad (6)$$

$L_{ii}^{\max}$  is the separation of atoms that cause a disruption of the H-bond, which has been determined by an analysis of the stability of the H-bonds with increasing temperature. The values of  $L_{ii}^{\max}$  for a GC pair can be found in (reference 5). With this definition of an open bond we introduce cooperativity into the calculation by redefining the effective force constant as (per reference 13):

$$\phi_{ii} = (1 - P_{ii})\phi_{ii}^{\text{int}} + P_{ii}\phi_{ii}^{\text{op}}. \quad (7)$$

$\phi_{ii}^{\text{op}}$  is the open bond effective force constant and is set to zero.

Similarly, the mean bond length  $R_{ii}$  can also be given by a probability weighted combination of intact bond length  $R_{ii}^{\text{int}}$  and open bond end atom distance  $R_{ii}^{\text{op}}$  (per reference 13):

$$R_{ii} = (1 - P_{ii}^{\text{op}})R_{ii}^{\text{int}} + P_{ii}^{\text{op}}(L_{ii}^{\max} + 2P_{ii}\sqrt{2D_{ii}}). \quad (8)$$

$R_{ii}^{\text{int}}$  is determined by several factors involving cross strand phosphate-phosphate repulsion and thermal expansion. In general  $R_{ii}^{\text{int}}$  can be given by

$$R_{ii}^{\text{int}} = r_{ii}^0 + dr_{ii}^C + dr_{ii}^T \quad (9)$$

where  $r_{ii}^0$  is the Morse parameter,  $dr_{ii}^C$  is the change caused by the change of cross strand static forces at different salt concentrations, and  $dr_{ii}^T$  is the change caused by bond thermal expansion.  $dr_{ii}^T$  was found to be (as in reference 5)

$$dr_{ii}^T = \frac{1}{a_{ii}} \ln [\cosh (2a_{ii} \sqrt{2D_{ii}} \ln 2)] \quad (10)$$

where  $a_{ii}$  is the Morse parameter.

One can define a nominal salt concentration  $C_0$  as that concentration where the shielded Coulomb forces are canceled by other nonbonded forces. That is, it is the concentration where the interbase H-bonds can be regarded as unstrained. We found that for the geometry of B-DNA  $C_0 = 0.05$  NaCl (reference 13) and for Z-DNA  $C_0 = 0.25$  NaCl. The stress that crosses the interbase H-bonds at other concentrations is dependent on the bulk salt concentration, and this stress in our formulation is expanded about the equilibrated nominal salt condition and is

$$df = f_p(C) - f_p(C_0) + f_s(C) - f_s(C_0) \quad (11)$$

where  $f_p(C)$  is the force from the cross strand phosphate-phosphate repulsion projected onto the interbase H-bond orientation and it is given by the derivative of the relevant SPMF.  $f_s(C)$  is force from the interstrand stacking interactions projected onto the interbase H-bond orientation, and it is dominated by near neighbor Van der Waals forces. Using the condition of balance of forces, one obtains

TABLE 1 Dependence of the premelting base pair opening probability  $P^{\text{op}}$  of B and Z DNA Poly[d(G-C)] on NaCl concentration

C (M)	$P^{\text{op}}$		
	B-form fiber	Z-form fiber	Z-form crystal
0.05	$4.42 \times 10^{-6}$	$1.16 \times 10^{-3}$	$3.36 \times 10^{-2}$
0.1	$1.97 \times 10^{-6}$	$9.20 \times 10^{-5}$	$5.79 \times 10^{-5}$
0.5	$9.47 \times 10^{-7}$	$2.85 \times 10^{-6}$	$2.18 \times 10^{-6}$
1.0	$1.01 \times 10^{-6}$	$1.62 \times 10^{-6}$	$1.23 \times 10^{-6}$
2.0	$1.08 \times 10^{-6}$	$1.17 \times 10^{-6}$	$9.21 \times 10^{-7}$
2.5	$1.06 \times 10^{-6}$	$1.06 \times 10^{-6}$	$7.30 \times 10^{-7}$
3.0	$1.01 \times 10^{-6}$	$9.68 \times 10^{-7}$	$6.31 \times 10^{-7}$
4.0	$8.60 \times 10^{-7}$	$7.24 \times 10^{-7}$	$3.61 \times 10^{-7}$
5.0	$9.30 \times 10^{-7}$	$4.63 \times 10^{-7}$	$1.85 \times 10^{-7}$
6.0	$2.37 \times 10^{-6}$	$3.41 \times 10^{-7}$	$2.34 \times 10^{-7}$

$$dr_{ii}^C = \frac{1}{a_{ii}} \ln \left[ \frac{\eta_{ii}}{1 + \sqrt{(\eta_{ii} - 1)^2 - 2df_{ii}/a_{ii}V_{ii}^0}} \right] \quad (12)$$

where  $\eta_{ii} = 2 \exp[-a_{ii}(R_{ii}(T_0) - r_{ii}^0)]$ ,  $R_{ii}(T_0)$  is the bond length at room temperature and  $df_{ii}$  is the stress across the  $i$ th H-bond in the  $l$ th base pair.  $df_{ii}$  is determined such that the cross strand stress is assumed to concentrate on the imino interbase H-bond unless that bond is broken. This is done because the amino CN bonds have an additional rotational degrees of freedom and the more rigid imino bond is expected to directly absorb the cross strand stress. The entire system, i.e., the variable parameters  $\phi_{ii}$ ,  $R_{ii}$ ,  $D_{ii}$ ,  $P_{ii}$ , and the eigenvalue equation in full dimensionality are solved iteratively until self-consistency is reached.

We carry out MSPA calculations on both B and Z conformation Poly[d(G-C)] at room temperature (293°K). The coordinates for the B-conformation are from a fiber study (14). For the Z-conformation both the fiber coordinates (14) and the  $Z_l$  coordinates from crystal study (15) are used. In all three calculations the same Morse parameters and  $L_s^{\max}$  as well as the same valence force fields are used. The results are given in Table 1. One finds from Table 1 that the calculated  $P_s^{\text{op}}$  of both the B conformation at lower salt concentration and the Z-conformation at high salt concentrations are in fair agreement with available experimental estimates. For example, the  $P^{\text{op}}$  for the B-Poly[d(G-C)] at 0.1 M NaCl is  $1.96 \times 10^{-6}$ . This is compared to the estimate of  $2.5 \times 10^{-6}$  for the base pairs in a B-DNA oligomer d(CG)<sub>12</sub> from imino proton exchange measurement (7). At 4- M NaCl the calculated  $P^{\text{op}}$  of the Z-Poly[d(G-C)] is  $7.24 \times 10^{-7}$  for the fiber coordinates and  $3.60 \times 10^{-7}$  for the crystal coordinates. These are compared to the estimate of  $5 \times 10^{-7}$  for the base pairs in the Z-form d(CG)<sub>12</sub> from imino proton exchange measurement (7).

We expect Z form when the  $P^{\text{op}}$  of the B form is larger, i.e., the domain wall advances through the B domain and vice versa. The newly formed domain is further sta-

Received for publication 4 November 1992 and in final form 12 January 1993.

## REFERENCES

1. Pohl, F. M. 1982. Salt-induced transition between two double-helical forms of oligo(dC-dG). *Cold Spring Harbor Symp. Quant. Biol.* 47:113-117.
2. Manzini, G., L. E. Xodo, and F. Quadrioglio. 1987. dC-dG alternating oligonucleotides: thermodynamic and kinetic aspects of the B-Z transformation. *J. Biomol. Struct. Dyn.* 4:651-662.
3. Haworth, I. S., A. Rodger, and W. G. Richards. 1992. A molecular dynamics simulation of a polyamine-induced conformation change of DNA. A possible mechanism for the B to Z transition. *J. Biomol. Struct. Dyn.* 10:195-211.
4. Harvey, S. C. 1983. DNA structural dynamics: longitudinal breathing as a possible mechanism for B-Z transition. *Nucleic Acids Res.* 11:4867-4878.
5. Chen, Y. Z., W. Zhuang, and E. W. Prohofsky. 1992. Salt-dependent stability of Poly(dG)·Poly(dC) with potential of mean force Coulomb interactions. *Biopolymers.* 32:1123-1127.
6. Sarma, M. H., G. Gupta, M. M. Dhingra, and R. H. Sarma. 1983. During B-Z transition there is no large scale breakage of Watson-Crick base pairs. A direct demonstration using 500 MHz  $^1\text{H}$  NMR spectroscopy. *J. Biomol. Struct. Dyn.* 1:59-81.
7. Kochoyan, M., J. L. Leroy, and M. Gueron. 1990. Processes of base-pair opening and proton exchange in Z-DNA. *Biochemistry.* 29:4799-4805.
8. Tsuboi, M., S. Takahashi, and I. Harada. 1973. Infrared and Raman spectra of nucleic acids—vibrations in the base-residues. In *Physico-chemical properties of nucleic acids*. Vol. 2. J. Duchesne, editor. Academic Press. New York. 91-145.
9. Lu, K. C., E. W. Prohofsky, and L. L. Van Zandt. 1977. Vibrational modes of A-DNA, B-DNA and A-RNA backbones: An application of a Green-function refinement procedure. *Biopolymers.* 16:2491-2506.
10. Chen, Y. Z., Y. Feng, and E. W. Prohofsky. 1991. Premelting thermal fluctuational base pair opening probability of Poly(dA)·Poly(dT) as predicted by the modified self-consistent phonon theory. *Biopolymers.* 31:139-148.
11. Klement, R., D. M. Soumpasis, E. Kitzing, and T. M. Jovin. 1990. Inclusion of ionic interactions in force field calculations of charged biomolecules-DNA structural transitions. *Biopolymers.* 29:1089-1103.
12. Prabhu, V. V., L. Young, and E. W. Prohofsky. 1989. Hydrogen-bond melting in B-DNA copolymers in a mean-field self-consistent phonon approach. *Phys. Rev. B* 39:5436-5443.
13. Chen, Y. Z., and E. W. Prohofsky. 1992. A cooperative self-consistent microscopic theory of thermally induced melting of a repeat sequence DNA polymer. *Biopolymers.* In press.
14. Chandrasekaran, R., and S. Arnott. 1989. The structures of DNA and RNA helices in oriented fibers. In *Landolt-Börnstein Numerical Data and Functional Relationships in Science and Technology*, Vol. VII/1b. W. Saenger, editor. Springer Verlag. 31-170.
15. Wang, A. H.-J., G. J. Quigley, F. J. Kolpak, G. V. D. Marel, and J. H. van Boom. 1980. Left-handed double helical DNA: Variations in the backbone conformation. *Science (Wash. DC)*. 211:171-176.

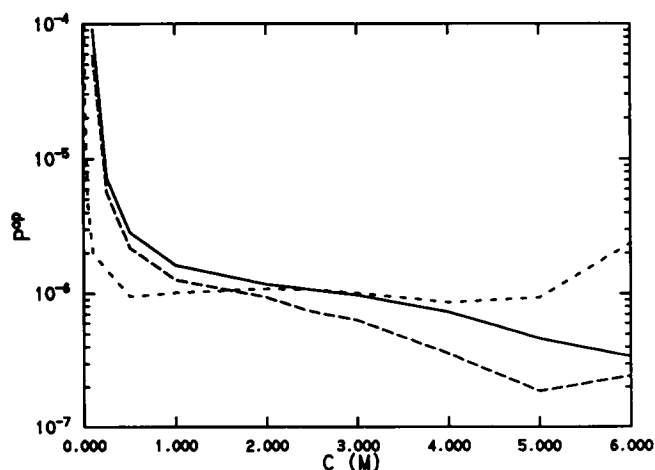


FIGURE 1 Calculated base pair opening probabilities  $P^{\text{op}}$  of B and Z DNA Poly[d(G-C)] as functions of NaCl concentration. The solid line is for the Z-form fiber structure, the long dashed line is for the Z-form crystal structure, and the short dashed line is for the B-form.

bilized by favorable free energy. From Fig. 1 one finds that at low salt concentrations the  $P^{\text{op}}$  of the two Z structures are higher than that of the B-structure. On the other hand, at high salt concentrations the  $P^{\text{op}}$  of both Z structures are lower than that of the B structure. The crossover salt concentrations for the curves are also close to the observed 2.5 M NaCl B-Z transition concentration for Poly[d(G-C)]. The crossover concentration for the fiber structure is 2.5 M NaCl and that for the crystal structure is 1.6 M NaCl. We point out that although the  $P^{\text{op}}$  seems to be small, the measured base pair life time (7) is comparable to the observed B-Z transition time, which indicates the relevance of thermal fluctuation to the transition. No free parameters fitted to the crossover concentration or to opening probability were used in this calculation. All parameters are fitted to spectral data at room temperature and  $C_0$  is a phenomenological parameter associated with the specific Morse parameters we used. We have shown that the salt dependence of  $P^{\text{op}}$  is determined by the particular cross strand potential of mean force (Chen, Y. Z. and Prohofsky, E. W. manuscript submitted for publication). A change of the hydrogen bond parameters will not change the slope of the  $P^{\text{op}}$  curve. The crossover concentration however depends on  $C_0$ . Since the  $P^{\text{op}}$  curves are flat near the crossover concentration, small change in  $C_0$  only leads to small change in the crossover concentration. The crossover concentration is therefore relatively insensitive to changes in parameters. Changes in conformation, such as that associated with the difference between fiber and crystal Z conformation, can affect the calculated SPMF and consequently effect the crossover somewhat as can be seen from Fig. 1.