THE JOURNAL OF PHYSICAL CHEMISTRY B

ARTICLE

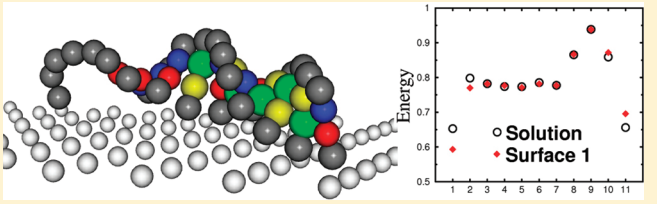
pubs.acs.org/JPCB

Effect of Surface Binding on Heterogeneous DNA Melting Equilibria: A Monte Carlo Simulation Study

John H. Allen, Emily T. Schoch, and John M. Stubbs*

Department of Chemistry and Physics, The University of New England, 11 Hills Beach Road, Biddeford, Maine 04005, United States

ABSTRACT: Deoxyribonucleic acid (DNA) microarrays are constructed with a surface-immobilized single-stranded probe sequence that hydrogen bonds with its complementary target strand from solution and is subsequently detected, making their hybridization equilibrium of central importance. Unexpectedly, the effect of surface immobilization is that if the sequences of probe and target are exchanged, the hybridization equilibrium shifts. Here, configurational-bias Monte Carlo simulations using a coarse-grained model for DNA were carried out for an undecamer double helix both in solution and bound to a surface to determine dissociation equilibria. Four possible surface binding orientations were independently investigated. Analysis shows that the effect of surface binding is to destabilize hydrogen-bonding interactions of bases proximal to the binding site and enhance those of distal bases due to the double helix lying flat on the surface. Results have implications for predicting surface-bound DNA hybridization equilibria.



INTRODUCTION

One of the key properties of deoxyribonucleic acid (DNA) is the transition between a double-stranded helix and two single strands. These transitions, known as hybridization and denaturation or "melting" transitions, depend on strand length and sequence as well as the environment, and are often characterized by thermodynamic properties. The strength of association between single strands in a double helix can also be characterized in terms of the melting temperature, T_m, the temperature at which a double strand is half-dissociated. Measurements of T_m can be made calorimetrically or spectroscopically via hyperchromicity[1] and are useful in determining the thermal stability of a given DNA duplex. Hybridization and melting transitions are at the heart of DNA sensor microarrays, a technology in which a surface-immobilized probe sequence associates with its complementary target sequence in solution and is subsequently detected.[2] Because of the usefulness of being able to test for a particular nucleotide sequence, this technology has a wide variety of applications.[3−6]

The accuracy of DNA microarrays depends upon knowledge of heterogeneous hybridization equilibria. Although solution-based or homogeneous transitions are generally well described by empirical nearest-neighbor models,[7,8] the increased complexity of transitions when one strand is bound to a surface has made this approach difficult to apply to heterogeneous equilibria. One approach to this problem implemented by Zhang et al.[9,10] is to introduce weighting factors for interaction energies that differentiate base−base hydrogen bonding depending upon a base's position within a strand. The result of this methodology is that internal bases have enhanced contributions to the overall stabilization energy and terminal bases have decreased contributions, thereby improving signal prediction. However, outlier sequences remain which are attributed mostly to probes with G-rich regions. Interestingly, they also explore the effect of surface binding for a given double-stranded DNA sequence and find that the amount of signal, which is directly related to the strength of interaction, often dramatically changes depending on which strand is surface bound, i.e., which is the probe and which is the target, even when equal concentrations are considered. This surface-binding effect is then correlated to the probe sequence with the number of G bases as well as the difference between number of A and T bases enhancing differences, though no physical basis is discussed. The analysis consisted of pairs of complementary probe signals from the Affymetrix Mapping50K_Xba240 array and were regressed in terms of

$$Y = \ln\left(\frac{S_{PM}}{S_{cPM}}\right) \quad (1)$$

where PM and cPM are a probe and its perfect complement and S_{PM} and S_{cPM} are the signals due to each, respectively.[10] Since the experimental data contained equal concentrations of each probe's target, the difference in binding interactions of the PM and cPM probes is reflected in the extent to which Y deviates from zero.

The explanation of this result is one of the remaining challenges in modeling DNA microarrays which principally include how the generally robust descriptions of homogeneous DNA hybridization can be extended to include the effect of surface immobilization.[11] Molecular simulation can provide insight into

Received: November 29, 2010
Revised: January 4, 2011
Published: January 31, 2011

the physical basis behind surface binding effects which ideally can then be used to improve predictive models and microarray technology. Currently, however, due to computational demands, simulations are somewhat limited with only a few systems having been modeled.[12−19] Here, we apply Monte Carlo molecular simulation to solution and surface-bound DNA hybridization transitions to investigate the surface-binding effect (which strand is bound as well as its orientation) described by Zhang et al.[10] in terms of structural and thermodynamic properties using an approach we have previously successfully applied to mutation effects and surface−probe spacers.[12]

METHODS

Molecular Model. A modified version of the coarse-grained model of Drukker et al.[20] is used for each DNA single strand. This allows for a much larger region of phase space to be sampled than a completely atomistic model would, and has been previously shown to give good agreement with empirical nearest-neighbor models for melting temperatures.[12] The model represents each nucleic acid as a single interaction site as well as the deoxyribose and phosphate groups (the "backbone") as a single interaction site. Thus, each DNA monomer unit consists of a backbone and base site. The geometric parameters are taken to represent B-DNA with a full helical turn made over 10 base units and a length of 34 Å.

The bonded interactions are modeled with the harmonic potential for sites separated by one bond

$$u_{\mathrm{vib}} = \frac{k_{ij}}{2}(r_{ij} - r_{ij,\mathrm{eq}})^2 \tag{2}$$

where k_{ij}, r_{ij}, and $r_{ij,\mathrm{eq}}$ are the force constant, distance between i and j, and equilibrium distance between i and j, respectively. Sites separated by two bonds interact via the angle between the bonds if one site is a base (eq 3) or via the distance between them (eq 4) if they are both backbone sites

$$u_{\mathrm{bend}} = \frac{k_{ijk}}{2}(\cos\theta_{ijk} - \cos\theta_{ijk,\mathrm{eq}})^2 \tag{3}$$

$$u_{\mathrm{bend}} = \frac{k_{ik}}{2}(r_{ik} - r_{ik,\mathrm{eq}})^2 \tag{4}$$

where k_{ijk}, θ_{ijk}, and $\theta_{ijk,\mathrm{eq}}$ are the force constant, bond angle, and equilibrium bond angle, respectively, and k_{ik}, r_{ik}, and $r_{ik,\mathrm{eq}}$ are the same as in eq 2. Sites separated by three bonds interact via a torsional potential (eq 5) if both end sites are bases

$$u_{\mathrm{tors}} = k_{ijkl}(1 - \cos(\phi_{ijkl} - \phi_{ijkl,\mathrm{eq}})) \tag{5}$$

where k_{ijkl}, ϕ_{ijkl}, and $\phi_{ijkl,\mathrm{eq}}$ are the force constant, torsional angle, and equilibrium torsional angle, respectively. Torsions where one or both of the end sites are backbones have zero interaction. All bonded parameters are given in Table 1.

Nonbonded interactions between sites separated by four or more bonds or located on different molecules are modeled with the Lennard-Jones potential for dispersion−repulsion interactions

$$u_{\mathrm{LJ}} = 4\varepsilon_{ij}\left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right] \tag{6}$$

Table 1. Interaction Parameters for Bonded Interactions

bond vibration	r_{eq} (Å)	k_{ij}/k_{B} (K Å^{-2})
A-BB	3.05	50 322
T-BB	2.70	50 322
C-BB	2.70	50 322
G-BB	3.05	50 322
BB-BB	4.35	50 322
bond bending	θ_{eq} (deg)	k_{ijk}/k_{B} (K)
base-BB-BB	79.0	40 257
	r_{eq} (Å)	k_{ik}/k_{B} (K Å^{-2})
BB-BB-BB	8.60	25 161
torsional rotation	ϕ_{eq} (deg)	k_{ijkl}/k_{B} (K)
base-BB-BB-base	28.5	40 257

where r_{ij}, ε_{ij}, and σ_{ij} are the distance between sites i and j, the Lennard-Jones well depth, and diameter for sites i and j, respectively. Unlike Lennard-Jones parameters were computed with the Lorentz−Berthelot combining rules:[21,22]

$$\sigma_{ij} = \frac{1}{2}(\sigma_{ii} + \sigma_{jj})\ \varepsilon_{ij} = \sqrt{\varepsilon_{ii}\varepsilon_{jj}} \tag{7}$$

Finally, hydrogen bonding between complementary bases are modeled using a combination of potentials (eq 8) that represent not only direct interaction between bases via a Morse potential (eq 9) but also a term that captures implicit solvent effects (eq 10):

$$u_{\mathrm{HB}} = (u_{\mathrm{HB},1} - u_{\mathrm{HB},2})\, f(\theta^{\mathrm{HB}}) \tag{8}$$

$$u_{\mathrm{HB},1} = D_{ij}((\mathrm{e}^{-a_{ij}(r_{ij} - r_{\mathrm{eq}}^{\mathrm{HB}})} - 1)^2 - 1) \tag{9}$$

where D_{ij}, a_{ij}, r_{ij}, and $r_{\mathrm{eq}}^{\mathrm{HB}}$ are the Morse well depth, Morse well shape, distance between sites i and j, and equilibrium hydrogen-bonding distance, respectively;

$$u_{\mathrm{HB},2} = \frac{D_{ij}}{4}(\tanh[b(r_{ij} - r_{\mathrm{eq}}^{*})] - 1) \tag{10}$$

where b and r_{eq}^{*} are the solvent interaction factor and equilibrium solvent distance, respectively;

$$f(\theta^{\mathrm{HB}}) = \begin{cases} \frac{1}{2}(\cos(\gamma\theta^{\mathrm{HB}}) + 1) & \text{if } \theta_{\mathrm{min}}^{\mathrm{HB}} \le \theta^{\mathrm{HB}} \le \theta_{\mathrm{max}}^{\mathrm{HB}} \\ 0 & \text{otherwise} \end{cases} \tag{11}$$

where θ^{HB} is the angle between the hydrogen bond acceptor, donor, and the backbone bonded to the acceptor, $\theta_{\mathrm{min}}^{\mathrm{HB}}$ and $\theta_{\mathrm{max}}^{\mathrm{HB}}$ are the minimum and maximum hydrogen-bonding angles, and γ ensures that $f(\theta^{\mathrm{HB}})$ is continuous at the boundaries. Noncomplementary bases are not considered to hydrogen bond, and the values of all nonbonded interaction parameters are given in Table 2. An alternate set of parameters, set 1, with $\phi_{\mathrm{eq}} = 36°$ and all D values 45.5% higher was employed in our previous work[12] as well as initially here. The differences between these two parameter sets will be discussed below.

Simulation Details. The undecamer 5′TCGCAGACAAA3′ (strand A) together with its complement 3′AGCGTCTGTTT5′ (strand B) were studied under five different conditions, the first

dx.doi.org/10.1021/jp111347p |J. Phys. Chem. B 2011, 115, 1720–1726

Table 2. Interaction Parameters for Nonbonded Interactions

site	σ (Å)	ε/k_B (K)		
A	3.0	225		
T	2.4	225		
C	2.4	225		
G	3.0	225		
BB	2.4	150		
HB pair	**donor**	**acceptor**	D/k_B (K)	a (Å^{-1})
AT	A	T	980	1.961
AT	T	A	905	1.909
GC	C	G	1024	2.884
GC	G	C	734	2.349
GC	G	C	993	2.784
HB parameter		**value**		
r_{eq}^{HB}		3.05 Å		
r_{eq}^{*}		4.00 Å		
b		2.00 Å^{-1}		
θ_{min}^{HB}		−0.25 rad		
θ_{max}^{HB}		0.25 rad		
γ		2π		

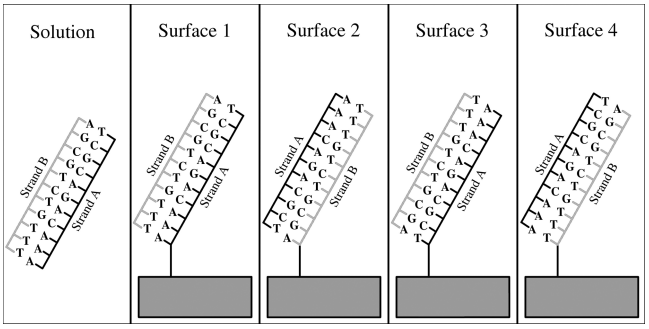


Figure 1. Illustration of the five systems studied.

being in solution where both strands are free. The remaining four systems are with one strand acting as the probe, i.e., bound to a surface, and the other strand as the target, i.e., unbound in solution. The four binding possibilities investigated are (i) strand A at its 3′ end, (ii) strand B at its 3′ end, (iii) strand A at its 5′ end, and (iv) strand B at its 5′ end. These four systems will be referred to as surface 1, surface 2, surface 3, and surface 4, respectively, and Figure 1 illustrates schematically the five different systems. Thus, the same hydrogen-bonding system is studied in all five cases and the effect of surface binding can be studied. The choice of sequence was made based on the work of Zhang et al.[10] We examined the probe sequences in the Affymetrix probe listings for the Mapping50K_Xba240 array[23] in terms of their predicted Y value[10] and chose one with $Y_{predicted}$ among those furthest from zero indicating the most difference between S_{PM} and S_{cPM}. The full sequence, 5′AAAAAAATCGCAGACAAATCTCATA3′, yields $Y_{predicted} = -0.63$ or $S_{PM} = 0.53S_{cPM}$, indicating that the PM signal is only approximately half that of the cPM. Finally, due to computational limits, we reduced the probe size from 25 bases by selecting the central 11 bases to study. This size is large enough to have a substantial number of bases not at or adjacent to the ends of the duplex, and more feasible to model as the calculation of the energy scales as the square of the number of interaction sites. The shorter sequence also requires fewer total moves due to the reduced number of conformational degrees of freedom. Since it is known that a mutation in the middle of a strand has the largest effect on the stability,[24] we similarly reasoned that the central bases would be most important to this investigation.

For each system, Monte Carlo molecular simulations were carried out in the *NVT* ensemble with one strand of each type in a cubic simulation box of length 100 Å for the solution simulations and a rectangular box of dimensions 100 Å × 100 Å × 114 Å for the surface simulations, with a hard surface occupying the additional 100 Å × 100 Å × 14 Å as compared to the solution box. For all systems, the temperature was varied from 300 to 370 K. In addition to normal translation and rotation moves, conformational degrees of freedom were efficiently sampled using both coupled−decoupled configurational-bias Monte Carlo (CBMC)[25−28] and self-adapting fixed-end point (SAFE) CBMC.[29] All intermolecular interactions were truncated at 14 Å. For surface-bound strands, a 5-unit spacer is used between the surface and probe. As described previously,[12] a spacer is necessary to achieve full duplex hybridization and because both 5- and 10-unit spacers allow this and have little difference between their resulting hybridization behaviors, the smaller size was chosen for computational efficiency. At least 4.5×10^7 and 4.8×10^8 Monte Carlo cycles (MC cycles) were used for equilibration and production, respectively, where one cycle consists of $N_{molecule} = 2$ steps. Statistical uncertainties are estimated from block averages of the results with each block 1.5×10^7 MC cycles in length.

RESULTS AND DISCUSSION

Model Parameters. In preliminary examination of our initial results it became apparent that our first parameter set, set 1, gave rise to a double helix structure that is somewhat overwound with respect to the average B-DNA structure. At 300 K, a temperature at which the system is predominantly in a double-stranded form for the duration of the simulation, those configurations without end fraying were selected and the twist around the helical axis was determined by locating the ideal axis center point between two complementary bases (slightly closer to the purine nucleotides due to their increased size) and the dihedral angle made between base to backbone bonds around the axis center to axis center vector was calculated and averaged over the simulation run, and was determined to be 38.7 ± 1.0°. For comparison to global dihedral rotations (see below) the rotation angle about a best-fit overall linear helix axis was also calculated and determined to be 41.5 ± 2.4°. The rise per base plane as calculated by the center-to-center distance was 3.59 ± 0.03 Å. Finally, the average duplex was somewhat curved from one end to the other with an average deflection from linear of 47.5 ± 4.2°, larger than the amount expected for a short oligomer, estimated at 19° for a dodecamer as measured by X-ray crystallography.[30]

It is known experimentally that, although the average structure of B-DNA is well described with 10.0 ± 1.2 bases per full helical turn, each separated by a distance of 3.32 ± 0.19 to 3.36 ± 0.46 Å along the helix axis, there can be significant variation due not only to thermal but also to base sequence effects.[31,32] The static structure, as determined by single-crystal X-ray diffraction analyses of short oligomers (10 or 12 base pairs) can be characterized by both a local rotation angle, measured only between neighboring base planes, and a global rotation angle, measured with respect to an overall linear helix axis, and are estimated to be 38.0 ± 4.4° and 35.9 ± 4.2°

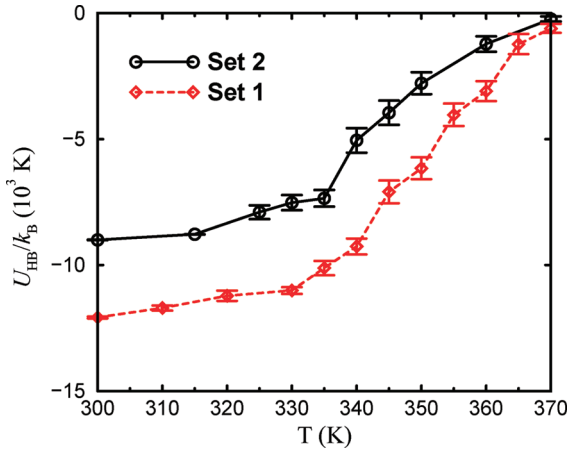


Figure 2. Average hydrogen-bonding energy as a function of temperature for the solution system as modeled with parameter sets 1 and 2. Lines are drawn as guides for the eye.

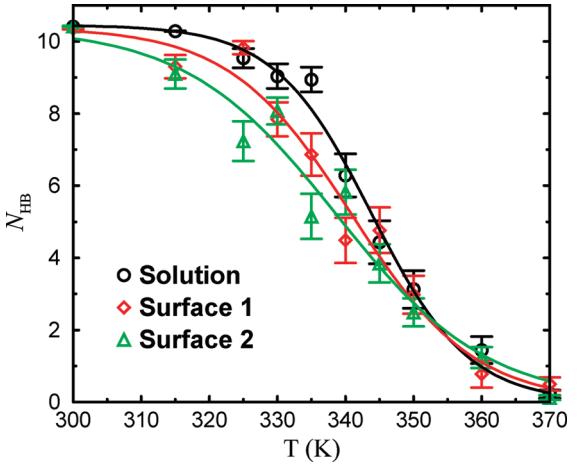


Figure 3. Average number of hydrogen-bonding base pairs as a function of temperature for solution, surface 1, and surface 2 systems. Symbols represent simulation data and lines are the best fit sigmoidal curves.

for local and global rotation, respectively.[31−33] Another measure of the range is given by the root-mean-square fluctuation in helix dihedral angle as measured by spectroscopic investigation of an ethidium−DNA complex in solution[34] and was found to be 5.9°. Nevertheless, although parameter set 1 is close to or within the experimental ranges for base plane rise and local and global rotation angles, it is on the high end and additionally the axis curvature is more than twice as large as the dodecamer value. Because the results of the present study are characterized using the double-stranded geometric properties and how they are affected by the surface, we felt that the use of set 2 with a geometry closer to the average of B-DNA structure was important enough to justify changing to, which when analyzed at 300 K gives average local and global rotation angles of 34.7 ± 0.4° and 35.4 ± 0.7°, respectively, and a base plane rise and axis curvature of 3.53 ± 0.04 Å and 32.1 ± 4.4°, respectively. Finally, for comparison, Figure 2 shows the average hydrogen-bonding energy as a function of temperature for the undecamer in solution and shows that although set 1 gives a lower average value (consistent with larger D values in Table 2) the qualitative melting behavior is identical.

Table 3. Summary of T_m from Simulation Data

system	T_m (K)
solution	344
surface 1	341
surface 2	339
surface 3	340
surface 4	343

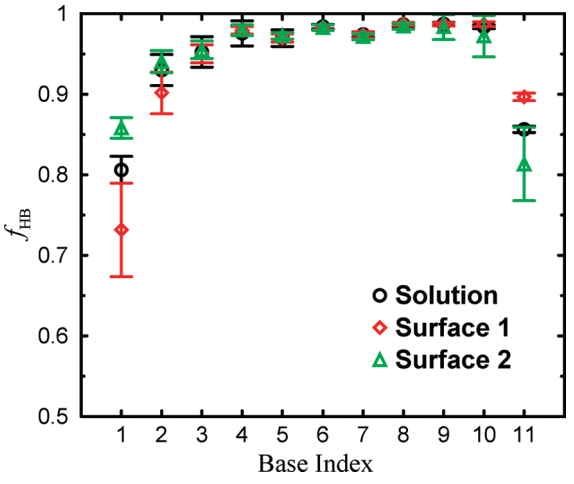


Figure 4. Fraction of configurations that a base is hydrogen bonded at 300 K for solution and surfaces 1 and 2, strand A, from 3′ to 5′ ends.

Hydrogen Bonding. The hydrogen-bonding pattern for each system was studied by calculating the hydrogen-bonding interaction (via eq 8) for each complementary pair which was then considered to be hydrogen bonded if the energy was less than the minimum threshold value of −200 K. This value, which we have also used previously,[12] was chosen so that most closely separated interacting pairs are considered hydrogen bonded while further apart next-nearest-neighbor pairs are excluded. Figure 3 shows the average number of hydrogen bonding pairs for the solution, surface 1, and surface 2 systems as a function of temperature together with sigmoidal line shapes from which T_m values in Table 3 are derived. Surfaces 3 and 4 are generally similar to 2 and 1, respectively, and are not shown for clarity. Although all surface systems are somewhat less stable (as indicated by lower T_m values) than solution, surfaces 1 and 4 are closer to the solution result and surfaces 2 and 3 are more destabilized. Since surfaces 2 and 3 have more GC base pairs near the surface, this indicates that their presence at the surface-bound end of the duplex is more detrimental than that of AT base pairs. This is consistent with the surface having a destabilizing effect on the bases closest to the bound end,[12] likely due to the bound end being unable to match its complement through its conformational fluctuations. Since GC pairs, due to their extra hydrogen bond, have a stronger interaction, the absolute value of the destabilization is larger even though an AT pair may be destabilized to the same extent, relatively, with respect to solution.

Surface-Binding Effects on Duplex Structure and Energetics. In order to investigate the differences in duplex stability, energetic and geometric properties were calculated for each base in each strand. The fraction of configurations in which a base is hydrogen bonded is shown at 300 K in Figure 4 for strand A aligned from 3′ to 5′. For the solution system the reduced values

dx.doi.org/10.1021/jp111347p |J. Phys. Chem. B 2011, 115, 1720–1726

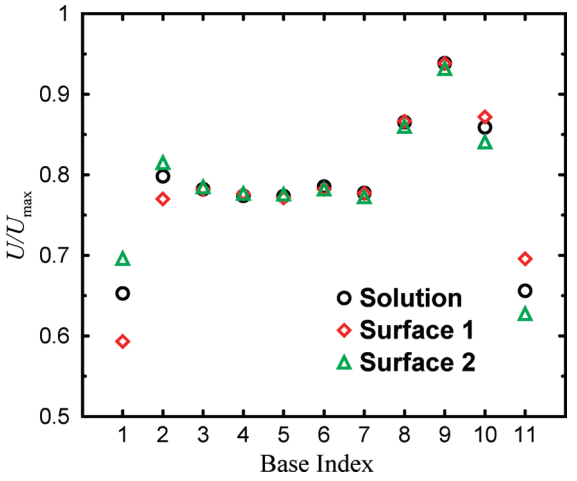


Figure 5. Average intermolecular energy of a base at 300 K relative to its maximum value for solution and surfaces 1 and 2, strand A from 3′ to 5′ ends. Error bars are smaller than symbol size.

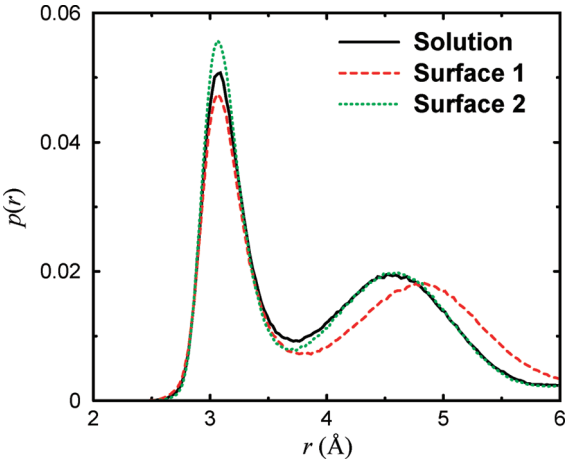


Figure 6. Probability to find a complementary base as a function of distance at 300 K for the terminal A base on strand A for solution, surface 1, and surface 2.

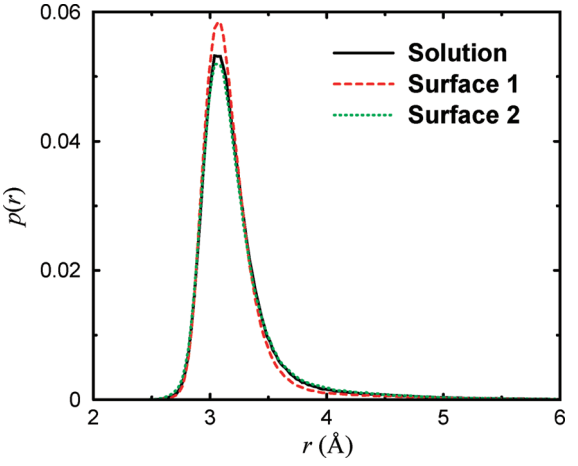


Figure 7. Probability to find a complementary base as a function of distance at 300 K for the terminal T base on strand A for solution, surface 1, and surface 2.

at the ends correspond to the phenomenon of end fraying where the terminal bases, lacking a neighbor on one side, have a destabilized interaction due to the lack of cooperativity from both sides. For both surfaces 1 and 2, the end proximal to the surface is hydrogen bonded less frequently with respect to solution (bases 1 and 2 for surface 1 and bases 11 and 10 for surface 2). This is consistent with the above discussion of the surface having a detrimental effect on the closest bonded base's interactions. Interestingly, the surface distal end is enhanced compared to solution, a result that will be discussed below.

Another way to characterize the effect of surface binding is to calculate the average intermolecular energy for each base. As the magnitude of GC interactions is greater than that of AT pairs, Figure 5 shows the average values scaled by their maximum value, U_{max}, to give a uniform plot for solution and surfaces 1 and 2. As in Figure 4, this shows strand A in the same alignment for all three systems and has several interesting features. First, there is a peak at base 9 which is a G with two C neighbors, which achieves nearly 95% of the ideal interaction as compared to bases 2 through 7 that are closer to 78%. This is another indication of cooperativity where having the more strongly bonded C neighbors reduces the available conformational space of base 9, i.e., holding it in place, so that it hydrogen bonds on average more strongly in contrast to the G or C bases with A neighbors at sites 6 and 4 that are more weakly bonded. This feature is present for all double-stranded structures at all temperatures in all systems studied here. Also shown in Figure 5 is that the end proximal to the surface shows the largest decrease in average energy at the terminal base, less so at its neighbor, and little change beyond that for surface 1 and a slight drop compared to solution for surface 2, consistent with Figure 4. Finally, there is an increase in magnitude with respect to solution in the average energy for the surface distal end, most evident in the last two sites (bases 11 and 10 for strand 1 and bases 1 and 2 for strand 2). This unexpected result, also evident in Figure 4, coincides with an increased primary peak in the complementary pair distance distributions shown in Figures 6 and 7 which give the probability to find a complementary base as a function of distance at 300 K for both ends of the solution and surface 1 and 2 systems. This enhancement is consistent across all surface simulations' distal ends, less so for its neighbor, and absent for remaining bases. The destabilization of proximal sites is also evident in Figures 6 and 7 where the decreased primary peak height is likely due to the inability of the bound strand to match the conformational flexibility of its complement. Figure 6 also shows the secondary peak for surface 1 shifted to larger distances. As this peak corresponds to the penultimate T base on strand B, this indicates a structural change from one base plane to the next at the surface proximal end. Analysis of the helix geometrical parameters shows an increase in base-plane rise from 3.45 to 3.59 Å and decrease in local dihedral angle from 35.3 to 33.4° for the proximal end of the surface 1 system as compared to solution. The changes for base planes 2 to 3 are similar in nature but smaller in size, after which there is little difference with the solution result. Thus, the helical conformation at the surface-bound end is somewhat expanded and less wound than solution and likely again attributable to lack of conformational flexibility of strand A.

Examination of configurations sampled throughout the simulations indicates that the enhanced stability for the surface distal end is likely due to the subset of configurations in which both the probe and target strands are aligned flat on the surface. As can be seen in Figure 8, a snapshot taken from a simulation of surface 1,

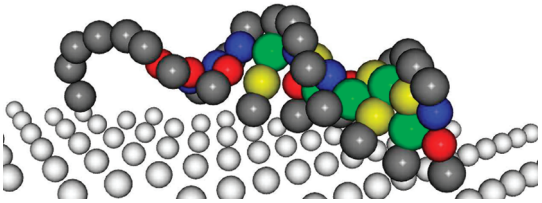


Figure 8. Snapshot from surface 1 system at 300 K, with the surface indicated by white spheres.

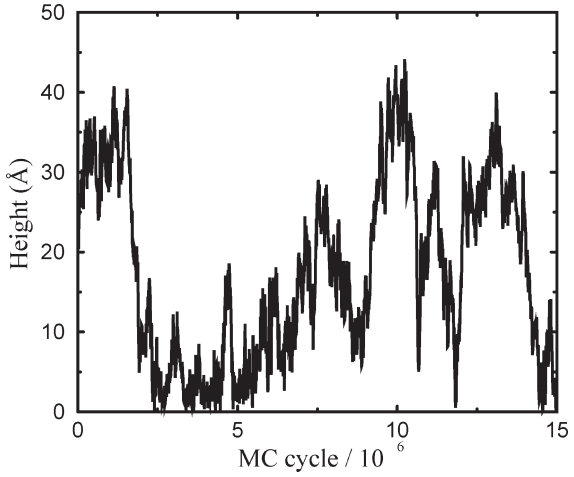


Figure 9. Height from the surface of the terminal distal backbone site of the target strand, strand B, for one simulation block of surface 1 at 300 K.

the probe strand is pinning down the target at the surface distal end. This orientation with respect to the surface does not allow as much end fraying of the distal end because the target is blocked by the surface from moving away from the probe, thus enhancing hydrogen-bonding configurations. Figure 9 shows the height off the surface of the target strand's distal end backbone site for one of the 1.5×10^7 MC cycle simulation blocks at 300 K. As can be seen, the height fluctuates between ≈2 and 35 Å, indicating that the duplex repeatedly samples configurations both flat on the surface and elevated from it, meaning that this result is unlikely to be due simply to nonergodicity within the simulation. Taken together, this indicates that configurations where the double helix is flat on the surface are significant and important to how immobilization on a surface affects hybridization energetics and equilibrium. Recent experiments have confirmed the tendency of a double-stranded probe−target pair to lie more along the surface.[35] Thus, based on the results of this work, the effect of surface binding is a destabilization of hydrogen-bonding interactions for bases closest to the binding site and increased stabilization of hydrogen-bonding interactions for bases furthest from the binding site.

As suggested by Zhang et al.,[9,10] base−base interaction parameters in an empirical model could be scaled with weighting factors to improve predicted thermodynamic properties. Based on the results in Figures 4 and 5, these could be separated into solution weighting factors that would have both strand ends lowered, and surface weighting factors which, as compared to solution, would have the surface proximal end again lowered but the surface distal end raised with respect to solution. Finally, although the prediction for surface interaction differences

Table 4. Summary of Thermodynamic Parameters

system	$\Delta U°$ (kJ mol^{-1})	$\Delta S°$ (J K^{-1} mol^{-1})
solution	-209 ± 15	-545 ± 45
surface 1	-197 ± 33	-511 ± 95
surface 2	-150 ± 18	-382 ± 52
surface 3	-158 ± 19	-403 ± 56
surface 4	-188 ± 26	-484 ± 77

(see Simulation Details) indicates surface 1 should be less stable than surface 2 for the full 25-base oligomer, our results for the shorter undecamer indicate it is more stable. This apparent contradiction is actually in line with surface interaction results given above and is due to the identity of the excluded bases. Whereas the full sequence has seven A bases distal to the surface in the analogous surface 1 (PM) setup and two C bases within the first seven proximal bases, our abbreviated sequence reverses this trend with three A bases proximal to the surface and two out of the first three distal bases being C or G. Because our proposed effect of surface binding destabilizes the surface proximal interactions and enhances surface distal ones, the result is that, for a given sequence, proximal GC base pairs lead to a more destabilized system. Thus, both relative stabilities are consistent with this hypothesis. Another result in apparent contradiction with Zhang et al. is that because their $Y_{predicted}$ is based solely on probe nucleotide identity but not base ordering, one would predict that the relative stability of surface 3 with respect to surface 4 should be identical as for surface 1 to surface 2 since surfaces 1 and 3 have the same bases but in reverse order. However, simulation results indicate the opposite stability trend. Although consistent with our proposed effect of surface immobilization, the disagreement with $Y_{predicted}$ is possibly due to the sequences in the results $Y_{predicted}$ was fit to (i.e., the four binding modes we investigated here are not simultaneously present) that allowed a good correlation between experiment and probe identity. Our results indicate that this fit could be improved if the location of a base within a probe is also taken into account.

Thermodynamic Quantities. The effect of surface binding on the thermodynamics of association can be determined by considering the equilibrium

$$A + B \leftrightarrow AB \tag{12}$$

with equilibrium constant

$$K_c = \frac{(c_{AB}/c°)}{(c_A/c°)(c_B/c°)} \tag{13}$$

where c_A, c_B, and c_{AB} are the concentrations of strand A, strand B, and the duplex AB and $c°$ is the standard-state concentration and is taken to be 1 mol L^{-1}. For a given configuration, the strands were considered to form a duplex if more than half of the bases were hydrogen bonded, i.e. if $N_{HB} \geq 6$. An alternate duplex definition based on the total intermolecular energy gave essentially identical results. The definition of the Helmholtz energy and its relation to K_c (appropriate due to constant volume conditions) gives

$$\ln K_c = \frac{-\Delta U°}{R}\frac{1}{T} + \frac{\Delta S°}{R} \tag{14}$$

which under the assumption that $\Delta U°$ and $\Delta S°$ are independent of temperature allow their determination via how ln K_c varies with $1/T$. A least-squares analysis yields the values of $\Delta U°$ and

$\Delta S°$ in Table 4. The trend in values for $\Delta U°$ is consistent with what was seen in the melting curves, i.e., that surfaces 2 and 3 have reduced interaction energies as compared to surfaces 1 and 4, which are closer to solution. In terms of the entropy of association, surfaces 2 and 3 have values ≈15−25% smaller in magnitude than surfaces 1 and 4. This lessening of the entropic penalty of hybridization is consistent with the hydrogen-bond cooperativity of the CGC sequence of strand A as shown in Figure 5 and discussed above, where in surfaces 2 and 3 proximity of this sequence to the surface-bound end of the duplex destabilizes hydrogen-bonding interactions that would otherwise form a more rigid, lower conformational entropy, structure.

CONCLUSIONS

Monte Carlo simulations using a coarse-grained model of a DNA duplex in solution and bound to a hard surface in four alternate orientations were carried out over a range of temperatures to determine the effect of surface binding on the structure and energetics of double-stranded DNA and its equilibrium with single-stranded DNA. The results indicate that, for a given sequence, binding orientations with more GC base pairs distal to the surface-bound end are more stable than those with GC pairs proximal to the binding site. This corresponded to destabilization of surface proximal sites with respect to solution and increased stabilization of distal sites due to less end fraying for configurations in which the double-stranded form lies flat along the surface. These results could be used to improve the prediction of heterogeneous hybridization thermodynamics using empirical models.

AUTHOR INFORMATION

Corresponding Author

*E-mail: jstubbs@une.edu.

ACKNOWLEDGMENT

We thank Nick Tito for work on analysis programs, J. I. Siepmann for helpful discussion, the University of New England's Junior Faculty Leave Program (J.M.S.), UNE CAS Dean's office for a student summer research stipend (E.T.S.), and finally the Minnesota Supercomputing Institute where the calculations were partially carried out.

REFERENCES

(1) Saenger, W. *Principles of Nucleic Acid Structure*; Springer-Verlag: New York, 1984.
(2) Southern, E.; Mir, K.; Shchepinov, M. *Nat. Genet. (Suppl.)* **1999**, *21*, 5–9.
(3) Brown, P. O.; Botstein, D. *Nat. Genet.* **1999**, *21*, 33–37.
(4) Lipshutz, R. J.; Fodor, S. P. A.; Gingeras, T. R.; Lockhart, D. J. *Nat. Genet.* **1999**, *21*, 20–24.
(5) Lockhart, D. J.; Winzeler, E. A. *Nature* **2000**, *405*, 827–836.
(6) Pozhitkov, A.; Noble, P. A.; Domazet-Lošo, T.; Nolte, A. W.; Sonnenberg, R.; Staehler, P.; Beier, M.; Tautz, D. *Nucleic Acids Res.* **2006**, *34*, e66.
(7) SantaLucia, J., Jr. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 1460–1465.
(8) SantaLucia, J., Jr.; Hicks, D. *Annu. Rev. Biophys. Biomol. Struct.* **2004**, *33*, 415–440.
(9) Zhang, L.; Miles, M. F.; Aldape, K. D. *Nat. Biotechnol.* **2003**, *21*, 818–821.
(10) Zhang, L.; Wu, C.; Carta, R.; Zhao, H. *Nucleic Acids Res.* **2007**, *35*, e18.
(11) Levicky, R.; Horgan, A. *Trends Biotechnol.* **2005**, *23*, 143–149.
(12) Tito, N. B.; Stubbs, J. M. *Chem. Phys. Lett.* **2010**, *485*, 354–359.
(13) Jayaraman, A.; Hall, C. K.; Grenzer, J. *Biophys. J.* **2006**, *91*, 2227–2236.
(14) Yao, L.; Sullivan, J.; Hower, J.; He, Y.; Jiang, S. *J. Chem. Phys.* **2007**, *127*, 195101.
(15) Qamhieh, K.; Wong, K.-Y.; Lynch, G. C.; Pettitt, B. M. *Int. J. Numer. Anal. Model.* **2009**, *6*, 474–488.
(16) Lee, O.-S.; Schatz, G. C. *J. Phys. Chem. C* **2009**, *113*, 15941–15947.
(17) Irving, D.; Gong, P.; Levicky, R. *J. Phys. Chem. B* **2010**, *114*, 7631–7640.
(18) Monti, S.; Cacelli, I.; Ferretti, A.; Prampolini, G.; Barone, V. *J. Phys. Chem. B* **2010**, *114*, 8341–8349.
(19) Ambia-Garrido, J.; Vainrub, A.; Pettitt, B. M. *Comput. Phys. Commun.* **2010**, *181*, 2001–2007.
(20) Drukker, K.; Wu, G.; Schatz, G. C. *J. Chem. Phys.* **2001**, *114*, 579–590.
(21) Lorentz, H. A. *Ann. Phys.* **1881**, *248*, 127–136.
(22) Berthelot, D. *C. R. Hebd. Séances Acad. Sci., Paris* **1898**, *126*, 1703–1706.
(23) Affymetrix, Inc., http://www.affymetrix.com (accessed June 2009).
(24) Wick, L. M.; Rouillard, J. M.; Whittam, T. S.; Gulari, E.; Tiedje, J. M.; Hashsham, S. A. *Nucleic Acids Res.* **2006**, *34*, e26.
(25) Siepmann, J. I. *Mol. Phys.* **1990**, *70*, 1145–1158.
(26) Siepmann, J. I.; Frenkel, D. *Mol. Phys.* **1992**, *75*, 59–70.
(27) Frenkel, D.; Mooij, G. C. A. M.; Smit, B. *J. Phys.: Condens. Matter* **1992**, *4*, 3053–3076.
(28) de Pablo, J. J.; Laso, M.; Suter, U. W. *J. Chem. Phys.* **1992**, *96*, 2395–2403.
(29) Wick, C. D.; Siepmann, J. I. *Macromolecules* **2000**, *33*, 7207–7218.
(30) Dickerson, R. E.; Drew, H. R. *J. Mol. Biol.* **1981**, *149*, 761–786.
(31) Dickerson, R. E.; Drew, H. R.; Conner, B. N.; Wing, R. M.; Fratini, A. V.; Kopka, M. L. *Science* **1982**, *216*, 475–485.
(32) Dickerson, R. E. *Methods Enzymol.* **1992**, *211*, 67–111.
(33) Drew, H. R.; Samson, S.; Dickerson, R. E. *Proc. Natl. Acad. Sci. U.S.A.* **1982**, *79*, 4040–4044.
(34) Millar, D. P.; Robbins, R. J.; Zewail, A. H. *Proc. Natl. Acad. Sci. U.S.A.* **1980**, *77*, 5593–5597.
(35) Barhoumi, A.; Zhang, D.; Halas, N. J. *J. Am. Chem. Soc.* **2008**, *130*, 14040–14041.