Coarse-Grained Simulations of DNA Reveal Angular Dependence of Sticky-End Binding

Published as part of The Journal of Physical Chemistry virtual special issue “Ruth Nussinov Festschrift”.

Nicholas M. Gravina, James C. Gumbart, and Harold D. Kim*

ABSTRACT: Annealing between sticky ends of DNA is an intermediate step in ligation. It can also be utilized to program specific binding sites for DNA tile and origami assembly. This reaction is generally understood as a bimolecular reaction dictated by the local concentration of the sticky ends. Its dependence on the relative orientation between the sticky ends, however, is less understood. Here we report on the interactions between DNA sticky ends using the coarse-grained oxDNA model; specifically, we consider how the orientational alignment of the double-stranded DNA (dsDNA) segments affects the time required for the sticky ends to bind, \( \tau_c \). We specify the orientation of the dsDNA segments with three parameters: \( \theta \), which measures the angle between the helical axes, and \( \phi_1 \) and \( \phi_2 \), which measure rotations of each strand around the helical axis. We find that the binding time depends strongly on both \( \theta \) and \( \phi_2 \): \( \sim 20 \)-fold change with \( \theta \) and 10-fold change with \( \phi_2 \). The binding time is the fastest when the helical axes of duplexes are pointing toward each other and the sticky ends protrude from the farthest two points. Our result is relevant for predicting hybridization efficiency of sticky ends that are rotationally restricted.

INTRODUCTION

Hybridization or annealing between sticky ends is commonly used as an intermediate step to ligation or assembly of DNA nanostructures. The sticky ends are two complementary single strands of DNA that extend from a DNA duplex. In a bimolecular reaction, the sticky ends undergo diffusive encounters at random relative orientations. Hence, the annealing rate would be an average over all relative orientations and depend solely on the local concentration of sticky ends. But in many cases such as in DNA cyclization or in DNA origami platforms or building blocks, the relative orientation of the sticky ends can be restricted. Inside the cell, DNA damage can cause a DNA double-strand break. In a repair pathway called alternative end joining, sticky ends are generated at the broken ends for subsequent ligation. In this process, the sticky ends are held in close proximity and likely to be rotationally restricted as well. Consequently, the annealing rate between rotationally restricted sticky ends will depend on the orientation factor. But how the relative orientation between the sticky ends influences their annealing efficiency is not clear.

In both biological and nanotechnological contexts, theoretical and computational models have been widely exploited to probe the behavior of DNA. Quantum chemistry calculations have been used to study the interactions between nucleotides in great detail. However, the high computational cost of these methods limits their scope to the study of interactions between nearest-neighbor base pairs in a vacuum. Classical all-atom approaches, in which every atom of DNA and surrounding solvent is accounted for as a point particle governed by effective interactions, have been used extensively in the study of short DNA segments and have recently been applied to larger DNA systems. However, because these methods are limited to microsecond time scales, simulating processes such as the breaking and formation of base pairs (so-called “rare-event” processes) remain a challenge. Theoretical approaches, such as the wormlike chain model, have been developed to understand large-scale properties of DNA; however, these approaches are not detailed enough to address processes such as duplex formation.

Coarse-grained DNA models lie between the extremes of analytical and all-atom approaches: the approximations made by these models imply a compromise between computational efficiency and accuracy that allows for simulation of larger time scales. In particular, coarse-grained DNA models have been used to analyze DNA hybridization, the behavior of DNA under torsional and tensile stress, and DNA duplex melting.

One of the most popular coarse-grained models is oxDNA2, which is an improved version of the original oxDNA model.
created by Ouldridge et al. in 2011. Both oxDNA and oxDNA2 were designed with a heuristic, “top-down” approach, focusing on reproducing well-known properties of DNA (such as the helical structure of the B-DNA duplex) and experimental results (such as duplex melting temperatures). The oxDNA model treats each nucleotide as a rigid body with three interaction sites, each of which has mutual, highly anisotropic interactions. This coarse-grained approach is detailed enough to obtain good agreement with measurements of the structural, mechanical, and especially the thermodynamic properties of single-stranded DNA (ssDNA) and double-stranded DNA (dsDNA). Consequently, the model has provided key insights into many different areas of research, including DNA nanotechnology, DNA origami, and DNA biophysics. In this study, we thus employed the coarse-grained oxDNA2 model to investigate the orientation dependence of sticky-end annealing. We found a significant dependence between the orientation of the dsDNA segments and the binding rate of sticky ends, which reflects the finite bend and twist persistence lengths of ssDNA overhangs.

## METHODS

For all simulations, we employed the same sequence to create our dsDNA segments and sticky ends. Each dsDNA segment was composed of a dsDNA segment of seven base pairs in length and a sticky end of seven nucleotides in length. The dsDNA segments were both composed of the sequence TCCCGAT (5′ to 3′)’s complementary sequence. The sticky end for strand 1 has sequence AATGTCC, and the sticky end for strand 2 has sequence GGACATT, both given here from 5′ to 3′. Although the oxDNA2 model is capable of accounting for sequence-dependent effects, we chose to forego this feature and used the hard-coded average parameters for base pairing and stacking interaction strengths.

To interrogate the relationship between the orientation of the sticky ends and the time required to reach the bound state, \( \tau_{s,0} \), we imposed restraints on the DNA strands to fix the values of the orientation variables (see Figure 1a). These restraints prevented the DNA strands from freely diffusing and thereby altering their relative orientation. We restrained four phosphate backbone particles on each dsDNA segment during the simulations: for the shorter ssDNA strand that did not contain the sticky end, we restrained the phosphate backbone particles at the 3′ and 5′ ends; for the longer ssDNA that contained the sticky end, we restrained the 5′ end phosphate backbone particle and the phosphate backbone particle forming the terminal base pair before the start of sticky end (see Figure 1a). We utilized the external force feature of oxDNA to apply harmonic potentials to each of these points, placing the equilibrium position at each particle’s starting position. Each harmonic potential used a force constant of 57.09 pN/nm. The other particles comprising the DNA segments in oxDNA were free to move during the simulation. In addition to considering the effect of orientation on the time required for the sticky ends to bond in our simulations, we also considered its effect on the pathway of sticky-end binding, namely, the order in which base pairs formed between the annealing sticky ends. To do this, we assigned a number 0–6 to each sticky-end base pair and recorded the identity of the base pair(s) present in the first bound configuration (see Figure 5).

To initialize the simulation, we chose to select sticky-end configurations from previous runs and attach them to the initialized duplexes to form different sticky-end initial configurations (SEICs). The default initial condition for the sticky ends chosen by oxDNA is a single helix, in which the sticky end spirals out from the dsDNA segment as if it were bound to a complementary strand. In addition to being a thermodynamically unlikely configuration, this configuration results in steric clash when \( \theta \) is close to 180°. The SEICs served to replace the improbable default initial configuration and to minimize steric clash by allowing the sticky ends to leave the plane containing the helical axes of the dsDNA.
To conduct the first round of simulations, we selected a total of five SEICs (SEIC 0−4). Each of these SEICs was chosen from previous runs that used the same DNA segments. For the second round of simulations varying temperature, we sampled SEICs from a simulation with each sticky end exchanged for polyT segments. This yielded a...
The four trials yielded the mean binding time between all complementary sticky-end nucleotides. We present corresponding rotations around the helical axes; therefore, the sticky ends are directions at \( \theta \) point in the same direction at \( \theta = 0^\circ \), \( \theta = 180^\circ \), \( \theta = 270^\circ \), and \( \theta = 360^\circ \). These configurations were chosen because they correspond to low, medium, and high \( \tau_b \) as determined by the first round of simulations. Hence, we will refer to them as the low \( \tau_b \), medium \( \tau_b \), and high \( \tau_b \) configurations (see Figure 6). For each of the nine temperature and configuration pairs, we also calculated the free energy as a function of the number of sticky-end base pairs (see Figure 7). The free energies were computed by using biased virtual-move Monte Carlo (VMMC) simulations in combination with the dynamic histogram analysis method. Every simulation was conducted with a salt concentration of 0.5 M and a \( D \) of 2.66 nm, a value obtained by using oxDNA to measure the length of a double helix identical in sequence to the sticky ends used in simulation (see Figure 1). During our simulations, we did not make any special considerations for transient states with mismatched base pairs.

**RESULTS**

For each initial configuration of the DNA strands specified by \((\theta, \phi_1, \phi_2)\) and an SEIC, we conducted four trials. In each trial, we recorded the time at which the hydrogen bonding energy first reached \(-517.79\) pN-nm, i.e., the energy for full binding between all complementary sticky-end nucleotides. We present a typical time trajectory of hydrogen bonding energy in Figure 2. The four trials yielded the mean binding time \( \tau_b \) which we then averaged again across runs with different SEIC to obtain a \( \bar{\tau}_b \) for a given \( \theta, \phi_1 \), and \( \phi_2 \). These runs were completed in two groups: one varying \( \theta \) from \( 0^\circ \) to \( 180^\circ \) in steps of \( 5^\circ \) with \( \phi_1 = 0^\circ \) and \( \phi_2 = 180^\circ \), and another varying \( \phi_1 \) from \( 0^\circ \) to \( 180^\circ \) in steps of \( 5^\circ \) at various \( \theta \) values. We employed a temperature of 310 K and a salt concentration of 0.5 M for these simulations. As \( \theta \) represents the angle between the two helical axes, the helices point in the same direction at \( \theta = 0^\circ \) (parallel) and in opposite directions at \( \theta = 180^\circ \) (antiparallel). \( \phi \) angles represent rotations around the helical axes; therefore, the sticky ends are on the same side of the helix when \( \phi_1 = 0^\circ \), \( \phi_2 = 180^\circ \) (the “cis” configuration) and on opposite sides when \( \phi_1 = 0^\circ \), \( \phi_2 = 180^\circ \) (the “trans” configuration). These two extreme configurations are presented in Figure 1. In general, \( \Delta \phi \equiv |\phi_1 - \phi_2| < 180^\circ \).

In Figure 3 we present \( \tau_b \) as a function of \( \theta \), the angle between the two helical axes. Figure 3a shows \( \tau_b \) as a function of \( \theta \) with \( \phi_1 = 0^\circ \) for three different SEICs. Figure 3b shows \( \tau_b \) as a function of \( \theta \) with \( \phi_1 = 180^\circ \) for three different SEICs. We also present the average across these SEICs \( \bar{\tau}_b \) (shown in blue). The standard deviation for these measurements is \( \pm 20\% \) of the corresponding \( \bar{\tau}_b \). In these graphs, \( \theta \) increases in steps of \( 5^\circ \) from \( 0^\circ \), where the two dsDNA segments are parallel, to \( 180^\circ \), where the two dsDNA segments are antiparallel. In both cases (\( \phi_1 = 0^\circ \) and \( \phi_1 = 180^\circ \)), the binding time decreases as \( \theta \) increases: sticky-end annealing is slowest when the helices are parallel and fastest when the helices are antiparallel. In addition to the \( \theta \) dependence, \( \phi_1 \) also modulates \( \bar{\tau}_b \). By comparing Figures 3a and 3b, one can see that \( \bar{\tau}_b \) with \( \phi_1 = 180^\circ \) is globally shorter than \( \bar{\tau}_b \) with \( \phi_1 = 0^\circ \), indicating a preference for the trans configuration as opposed to the cis configuration.

To further explore the relationship between \( \bar{\tau}_b \) and the \( \phi \) angles, we measured \( \tau_b \) when varying \( \phi_2 \). In Figure 3, we present \( \tau_b \) as a function of \( \phi_2 \) at various \( \theta \) with \( \phi_1 = 180^\circ \). For \( \theta = 0^\circ \) there is little to no increase in \( \tau_b \) as \( \phi_2 \) increases from \( 0^\circ \) to \( 180^\circ \). For all other \( \theta \), there is an increase in \( \tau_b \) as \( \phi_2 \) increases, indicating that bonding occurs more rapidly when the strands are closer to the trans configuration than the cis configuration. As shown in Figure 3, there is a decrease in \( \bar{\tau}_b \) as \( \theta \) increases; the \( \bar{\tau}_b \) curve for a particular \( \theta \) is typically less than the \( \bar{\tau}_b \) curve for a smaller \( \theta \) at any given \( \phi_2 \).

In Figure 4, we present the effects of \( \theta \) and \( \phi \) on the binding time \( \tau_b \) using a heatmap. The vertical axis represents \( \theta \) while the horizontal axis represents \( \phi_2 \), \( \phi_1 \) is held constant at \( 180^\circ \). Within a given \( \phi_2 \) value, the binding time \( \tau_b \) decreases as \( \theta \) increases in the same manner as shown in Figure 3, although the difference between the maximum \( \tau_b \), which occurs around \( \theta = 0^\circ \), and the minimum \( \tau_b \) which occurs around \( \theta = 180^\circ \), changes depending on the \( \phi_2 \) value.
In addition to recording $\tau_B$, we also recorded the system configuration at intervals of $9.09 \times 10^{-10}$ s. We used this data to determine which of the sticky-end base pairs had formed in the first time step with any sticky-end binding. In the majority of cases, multiple sticky-end base pairs formed before the first time step where sticky-end binding could be observed. This meant that the first observable instance of binding had multiple base pairs, making it impossible to determine precisely which base pair formed first. To display the dominant binding pathway for various $\phi_2 - \theta$ pairs, we considered the average initial base pair position $B_0$ which we present in Figure 5.

Figure 5. (a) Schematic showing the labeling scheme for the sticky-end base pairs. Base pair 0 is the base pair closest to strand 2, and base pair 6 is the base pair closest to strand 1. (b) Average initial base pair for various $\phi_2 - \theta$ pairs shown with a linear scale. A value of $\sim 3$ indicates no preference for initial hydrogen bond position. For each $\phi_2 - \theta$ pair, base pair data were collected from a total of eight runs with two SEICs, four runs from each SEIC. $B_0 > 3$ are shown in shades of red; $B_0 < 3$ are shown in shades of blue. Thus, red squares indicate a preference for forming initial base pairs closer to strand 1; blue squares indicate a preference for forming initial base pairs closer to strand 2. When $B_0 = 3$ (white squares), initial base pairs form between all sticky-end nucleotides with equal probability.

When $\phi_2 > 60^\circ$, the average initial base pair position $B_0$ shifts from being closer to the attachment point on strand 2 ($B_0 < 3$) to being closer to the attachment point on strand 1 ($B_0 < 3$).

Figure 3 shows the binding time averages for each SEIC and the overall average for both $\phi_1 = 0^\circ$ and $\phi_1 = 180^\circ$. In both the $\phi_1 = 0^\circ$ and $\phi_1 = 180^\circ$ cases, plotting $\tau_B$ as a function of $\theta$ on a semilog plot reveals a sinusoidal curve with a minimum at $\theta \approx 140^\circ$. For $\theta > 140^\circ$, both curves are relatively flat. In this regime where the DNA strands are nearly antiparallel, further increasing the angle between the helical axes ($\theta$) does not greatly impact the binding time. Note also that the $\tau_B$ for $\phi_1 = 180^\circ$ is globally less than $\tau_B$ when $\phi_1 = 0^\circ$. This indicates that binding between the sticky ends occurs more rapidly when the junction between the sticky ends and their corresponding dsDNA segments are on opposite sides of the plane containing the helical axes of both DNA strands (i.e., when the strands are in the trans configuration).

To further investigate the angular dependence of sticky-end binding, we characterize the binding interaction as a function of $\phi_2$ (see Figure 1). In Figure 4 we plot the binding time, $\tau_B$, as a function of $\theta$ and $\phi_2$ with $\phi_1$ held constant at $180^\circ$. The difference between $\phi_1$ and $\phi_2$ represents the relative azimuthal angle between the two strands. In Figure 4, for all $\theta$ values there is a clear increase in $\tau_B$ as $\phi_2$ increases from $0^\circ$ to $180^\circ$. When $\phi_2 < 90^\circ$, the sticky ends are in the trans configuration, and when $\phi_2 > 90^\circ$, the sticky ends are in the cis configuration. Therefore, this increase in $\tau_B$ indicates that binding between the sticky ends occurs more rapidly in the trans configuration than in the cis configuration, regardless of $\theta$.

In Figure 6, we present the binding time $\tau_B$ for low, medium, and high temperatures (283, 310, and 363 K, respectively) for three angular configurations. The angular configurations chosen correspond to high, medium, and low values for $\tau_B$ as determined from previous simulations. With the exception of the high-$\tau_B$ configuration measured at high temperature, each point corresponds to a measurement of $\tau_B$ using $\sim 20$ SEICs each with four replicas. This specific number of SEICs (20) was chosen because it allowed for measurements of $\tau_B$ to be reasonably converged while remaining computationally tractable. For the high-$\tau_B$ configuration at high temperature, only a small number of runs yielded a binding time (see Figure 7). For the low- and medium-$\tau_B$ configurations there is a slight increase in binding time when moving from 283 to 310 K, although this increase is within the margin of error. For the high-$\tau_B$ configuration this trend reverses, and $\tau_B$ decreases when moving from 283 to 310 K. But for all configurations, the longest binding time was observed at 363 K.

Figure 7 shows the free energy profiles for three temperatures and angular configurations. Changing the angular configuration has a clear impact on the free energy curves, which indicates that the differences in the binding times observed at these three configurations is due in part to thermodynamic changes. Of particular note is the free energy profiles at high temperature (see Figure 7c). Changing the angular configuration of the strands at this temperature shifts the minimum of the free energy curve and changes the
The equilibrium number of sticky-end base pairs: in the low-$r_b$ configuration full sticky-end binding (seven sticky-end base pairs) is thermodynamically favorable, but in the high-$r_b$ configuration, three base pairs are thermodynamically favorable.

**Trends in Binding Time $r_b$.** Next, we attempted to discern whether $\theta$ or $\phi_2$ exerts the most control over $r_b$. To do this, we will consider the ratio between the maximum and minimum $r_b$ within a given $\theta$ and $\phi_2$:

$$R(\theta') = \frac{\max\{r_0(\theta = \theta', \phi_1 = 180^\circ, \phi_2)\}}{\min\{r_0(\theta = \theta', \phi_1 = 180^\circ, \phi_2)\}}$$

and

$$S(\phi_2') = \frac{\max\{r_0(\theta, \phi_1 = 180^\circ, \phi_2 = \phi_2')\}}{\min\{r_0(\theta, \phi_1 = 180^\circ, \phi_2 = \phi_2')\}}$$

$R(\theta')$ is the ratio between the maximum and minimum $r_b$ values that occur with $\theta = \theta'$; it measures the degree to which changing $\phi_2$ influences $r_b$. $S(\phi_2')$ is the ratio between the maximum and minimum $r_b$ values that occur with $\phi_2 = \phi_2'$; it measures the degree to which changing $\theta$ influences $r_b$. We find that averaged over all $\theta$ values, $R \approx 10$, and when averaged over all $\phi_2$ values, $S \approx 20$. This indicates that $\theta$ has a greater effect on $r_b$ than $\phi_2$. The dependence on $\phi_2$ is roughly constant across each theta value, with $R$ never deviating far from the average of $9.8 \pm 2.5$. This is not the case for $r_b$’s dependence on $\theta$. When $\Delta \phi < 90^\circ$, the maximum and minimum $r_b$ differ a factor of $\sim 10$; when $\Delta \phi > 90^\circ$, the $\theta$ dependence is stronger, with maximum and minimum $r_b$ separated by a factor of $\sim 30$.

The transition between these two regimes for $S$ occurs at $\phi_2 = 90^\circ$, reflecting the transition between the cis and trans configurations. In other words, when the strands are in the cis configuration, $\theta$ has less influence on the binding time than it does when the strands are in the trans configuration.

On the basis of our simulation results, we present a simple equation that phenomenologically describes the angular dependence of the sticky-end annealing time:

$$r_b(\phi_2, \theta) = \exp[C + A \cos(\phi_2 - \delta) + B \cos(\theta - \epsilon) + E \sin(\phi_2 - \delta) \cos(\theta - \epsilon)]$$

The parameter values that produce the best fit are $A = -0.4453$, $B = 1.554$, $C = -16.5$, $E = -0.3446$, $\delta = 38.68^\circ$, and $\epsilon = -25.26^\circ$. The primary trends of increase in $r_b$ as $\phi_2$ increases and decrease in $r_b$ as $\theta$ increases are included in the model through the $A$ and $B$ terms, respectively. $B$ is larger than $A$, indicating that the $\theta$-dependence is stronger than the $\phi_2$-dependence. These basic trends are altered by the cross-term associated with $E$ and the two phase shift parameters $\delta$ and $\epsilon$. One effect of these terms is to alter the position of the minimum $r_b$ for $\phi_2 < 90^\circ$ to $\theta = 150^\circ$. This shifted minimum matches the data well, as $r_b$ is slightly less at $\theta = 150^\circ$ compared to $\theta = 180^\circ$ when $\phi_2 < 90^\circ$. In addition, the model also contains intermediate minima between $\phi_2 = 0^\circ$ and $\phi_2 = 180^\circ$ within a given $\theta$. These features are also present in the data as well, although at slightly different $\phi_2$. In both cases, we speculate that the minima are the result of the changing distance between the attachment points as a function of $\theta$ and $\phi_2$ and, therefore, that the annealing reaction between the

![Figure 7](https://doi.org/10.1021/acs.jpcb.1c00432)
sticky ends is a function of this distance as well as the orientation between the dsDNA strands.

**DISCUSSION**

**Factors Impacting Sticky-End Rigidity.** We suspect that the observed angular dependence present in sticky-end binding primarily results from stacking interactions between adjacent bases of sticky ends. Compared to dsDNA, the persistence length of ssDNA is small. Accordingly, when considering a system composed of ssDNA and dsDNA segments, the ssDNA segments are often assumed to be very flexible.\(^{42}\) Such an assumption would be valid for sticky ends markedly longer than the persistence length of ssDNA. But in the case of sticky ends as short as 10 nucleotides used in this study, the semiflexible nature of ssDNA must be considered.\(^{43\text{--}47}\)

Stacking between the first base of ssDNA and the end base pair of dsDNA also further restricts the orientational freedom of the sticky ends. A model that ignores this intrinsic rigidity would lead one to assume that the contortions required to achieve sticky-end binding are not thermodynamically unfavorable and, therefore, that sticky-end binding is largely independent of the angular orientation of the dsDNA segments. An individual segment of ssDNA is, however, known to have a nonzero persistence length, typically on the order of 6–12 Å.\(^{45,48}\) Therefore, sticky ends, being ssDNA segments, possess intrinsic rigidity, which causes the binding reaction to be dependent on the mutual orientation between the sticky ends.

Adding to the rigidity of the sticky ends is the nearest-neighbor stacking interaction, which occurs between the first nucleotide of the sticky ends and the final base pair of the dsDNA segment. Stacking interactions are well-known to occur between adjacent base pairs in dsDNA segments.\(^{49\text{--}51}\) These interactions also increase the thermodynamic stability of overhanging single-stranded segments via stacking between the terminal base pair and the first nucleotide of the overhang.\(^{52}\) This interaction is a differentiating factor between ssDNA strands freely diffusing in solution and sticky ends. Therefore, ssDNA as an overhang to dsDNA is effectively stiffer than ssDNA with free ends. This difference in the boundary condition is borne out by previous measurements of the persistence length of ssDNA. The persistence length of an overhang ssDNA\(^{53}\) is significantly larger than that of a freely diffusing ssDNA.\(^{22,48}\)

In this study, we used a generic coarse-grained model and did not account for sequence-dependent stacking interactions. However, experimental studies\(^{34}\) and sequence-dependent models\(^{32}\) both show that poly dA has a longer persistence length than poly dT due to stronger helical stacking. In addition, computational studies have elucidated sequence-dependent effects in the formation of DNA minicircles.\(^{55}\) However, it is not obvious how sequence dependence would affect the annealing rate because of the complementarity of sticky ends; increasing the rigidity of one sticky end will necessarily decrease the rigidity of the other sticky end. Therefore, we expect the binding time to change nonmonotonically as one sticky-end sequence is varied from low to high purine content.

As shown in Figure 6, \(\tau_b\) depends on temperature in a nontrivial manner. For the high-\(\tau_b\) configuration, increasing the temperature from 283 to 310 K reduces the average binding time. Conversely, for the low-\(\tau_b\) configuration, increasing temperature from 283 to 310 K increases the average binding time. We attribute these dependencies to the decrease in the persistence lengths of the ssDNA overhangs. For angular configurations that are unfavorable to binding (i.e., the high-\(\tau_b\) configuration), this reduction in persistence length enables the sticky ends to twist and bend more easily, allowing them to bind with greater efficiency. For angular configurations that are favorable to binding (i.e., the low-\(\tau_b\) configuration), however, a decrease in the persistence length allows the sticky ends to more freely deviate away from the favorable binding configuration established by the geometry, increasing the binding time. As the temperature is further increased from 310 to 363 K, the binding rate is no longer limited by the conformational fluctuations of the sticky ends, but instead governed by the enthalpic barrier of the transition state (\(\Delta H^\ddagger\)). Since \(\Delta H^\ddagger\) is negative for base-pair formation, the binding rate would decrease with temperature (\(\sim e^{-\Delta H^\ddagger/k_B T}\)). This heuristic explanation is consistent with our result that \(\tau_b\) for all configurations increases and converges when the temperature is changed from moderate (310 K) to high (363 K).

**Dominant Pathway Analysis.** Figure 5 shows the change in the average initial base pair \(\bar{B}_0\) as a function of \(\phi_2\) and \(\theta\) with \(\phi_1 = 180^\circ\). We made an interesting observation that the dominant binding pathway as determined by the average initial base pair changes as a function of the strand orientation. Specifically, for \(\phi_2 > 60^\circ\), \(\bar{B}_0\) increases as \(\theta\) increases from 0° to 180°. Because we did not employ the sequence-dependent capabilities of oxDNA2, this change cannot be a result of any sequence-dependent differences in persistence length between the two sticky ends. We speculate that this change in binding site preference is due to stacking interactions between the sticky-end nucleotides, which serve to bias the nucleotides toward facing inward toward the helical axis.\(^{43\text{--}47}\) As \(\phi_2\) increases, the attachment point for the second sticky end moves counterclockwise. For \(\phi_2 < 110^\circ\), this decreases the distance between the attachment points and, in combination with stacking interactions, changes the preferred direction of the nucleotides. When \(\phi_2 \approx 110^\circ\), stacking interactions bias the nucleotides toward facing away from sticky end 1. For \(\phi_2 > 110^\circ\) the distance between the attachment points begins to increase, but the preferred direction of the sticky end 2 nucleotides becomes closer to horizontal. Therefore, \(\phi_2 \approx 110^\circ\) decreases the probability of sticky end 2 assuming a configuration that allows binding with sticky end 1. Meanwhile, the nucleotides of sticky end 1, pointing “horizontally” with \(\phi_1 = 180^\circ\), can bend toward sticky end 2 and bind with comparative ease and bind to their complementary nucleotide in sticky end 2. Thus, the sticky ends will prefer to bind closer to strand 2, and \(\bar{B}_0\) is more likely to be < 3. When 120° < \(\theta\) < 180°, this effect is minimal, as much less bending is required to achieve binding since the strands are almost antiparallel.

Therefore, one would expect no preference in the initial binding site in this regime, which is what the data show: in this regime, \(\bar{B}_0\) randomly fluctuates around 3.

**CONCLUSION**

The oxDNA2 model is capable of simulating a system of 42 nucleotides on a microsecond time scale, which is sufficient to simulate rare events such as the breaking and forming of base pairs that occurs in sticky-end binding. In this study, we measured the time required for the sticky ends to transition from their initial configuration to the bound state by noting the time step at which the hydrogen binding energy dropped
below the critical value of $-517.79$ pN-nm, which indicated full binding between all the sticky-end nucleotides. The binding time $\tau_b$ was measured as a function of the angle between the helical axes of the dsDNA segments ($\theta$), rotation around the helical axes for DNA segment 1 ($\phi_1$), and rotation around the helical axis for DNA segment 2 ($\phi_2$). Our simulations show that the binding rate between sticky ends is influenced by the orientation of the dsDNA segments they are attached to. Changing $\theta$ has a stronger effect on the binding rate than changing $\Delta \phi_i$ for DNA segment $i$, where $\Delta \phi_i = \phi_i - \phi_{i-1}$. However, the degree to which $\theta$ influences the binding rate is modulated by $\Delta \phi$: when $\Delta \phi > 5$ the flexural rigidity of sticky ends is reduced. The stand orientation energy, significant fluctuations in the binding rate is modulated by $\Delta \phi$: when the strands occupy the trans conformation, the $\theta$ dependence is stronger, with maximum and minimum $\tau_b$ separated by a factor of $\sim 30$. Changing temperature, in addition to influencing the free energy, significantly impacts the binding rate possibly by reducing the rigidity of the sticky ends. The stand orientation also impacts the dominant binding pathway by changing the average initial binding site $B_0$. These results are relevant for predicting hybridization efficiency of sticky ends that are rotationally restricted.

**ASSOCIATED CONTENT**

*Supporting Information*

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jpcb.1c00432.

All informal conventions used for oxDNA simulations as well as other associated input files (ZIP).

**AUTHOR INFORMATION**

Corresponding Author
Harold D. Kim — School of Physics, Georgia Institute of Technology, Atlanta, Georgia 30332, United States; Email: harold.kim@physics.gatech.edu

Authors
Nicholas M. Gravina — School of Physics, Georgia Institute of Technology, Atlanta, Georgia 30332, United States

James C. Gumbart — School of Physics, Georgia Institute of Technology, Atlanta, Georgia 30332, United States; orcid.org/0000-0002-1510-7842

Complete contact information is available at: https://pubs.acs.org/doi/10.1021/acs.jpcb.1c00432

**Notes**

The authors declare no competing financial interest.

**ACKNOWLEDGMENTS**

This work was supported by the National Science Foundation through the Division of Molecular and Cellular Biosciences (1517507) and the Physics of Living Systems Student Research Network (1806833). HDK also acknowledges support from the National Institutes of Health (R01GM112882). Computational resources were provided by the Partnership for an Advanced Computing Environment (PACE) at the Georgia Institute of Technology.

**REFERENCES**


