Characterizing the bending and flexibility induced by bulges in DNA duplexes

John S. Schreck, Thomas E. Ouldridge, Flavio Romano, Ard A. Louis, and Jonathan P. K. Doye

Citation: The Journal of Chemical Physics **142**, 165101 (2015); doi: 10.1063/1.4917199
View online: http://dx.doi.org/10.1063/1.4917199
View Table of Contents: http://scitation.aip.org/content/aip/journal/jcp/142/16?ver=pdfcov
Published by the AIP Publishing

Characterizing the bending and flexibility induced by bulges in DNA duplexes

John S. Schreck,[1,a)] Thomas E. Ouldridge,[2] Flavio Romano,[1] Ard A. Louis,[2] and Jonathan P. K. Doye[1,b)]
[1]*Physical and Theoretical Chemistry Laboratory, Department of Chemistry, University of Oxford, South Parks Road, Oxford OX1 3QZ, United Kingdom*
[2]*Rudolph Peierls Centre for Theoretical Physics, University of Oxford, 1 Keble Road, Oxford OX1 3NP, United Kingdom*

(Received 18 December 2014; accepted 27 March 2015; published online 22 April 2015)

Advances in DNA nanotechnology have stimulated the search for simple motifs that can be used to control the properties of DNA nanostructures. One such motif, which has been used extensively in structures such as polyhedral cages, two-dimensional arrays, and ribbons, is a bulged duplex, that is, two helical segments that connect at a bulge loop. We use a coarse-grained model of DNA to characterize such bulged duplexes. We find that this motif can adopt structures belonging to two main classes: one where the stacking of the helices at the center of the system is preserved, the geometry is roughly straight, and the bulge is on one side of the duplex and the other where the stacking at the center is broken, thus allowing this junction to act as a hinge and increasing flexibility. Small loops favor states where stacking at the center of the duplex is preserved, with loop bases either flipped out or incorporated into the duplex. Duplexes with longer loops show more of a tendency to unstack at the bulge and adopt an open structure. The unstacking probability, however, is highest for loops of intermediate lengths, when the rigidity of single-stranded DNA is significant and the loop resists compression. The properties of this basic structural motif clearly correlate with the structural behavior of certain nano-scale objects, where the enhanced flexibility associated with larger bulges has been used to tune the self-assembly product as well as the detailed geometry of the resulting nanostructures. We further demonstrate the role of bulges in determining the structure of a "Z-tile," a basic building block for nanostructures. © *2015 AIP Publishing LLC.* [http://dx.doi.org/10.1063/1.4917199]

I. INTRODUCTION

DNA is one of the most important molecules in biology. By virtue of the specificity of Watson-Crick base-pairing,[1] DNA is highly programmable, and beginning with the work of Nadrian Seeman in the 1980s,[2] it has been identified as a major player in nanotechnology, with many structures and devices already realized.[3,4] For example, by starting with appropriate oligonucleotides, DNA can be made to self-assemble into many structures with high yields by cooling solutions from high temperatures. The target structure is usually designed to be the global free-energy minimum by virtue of containing the largest number of base pairs. In many examples, the structures are made up of a few basic components, including double helical sections and various types of junctions where the double helices meet. Examples of these nanostructures include DNA polyhedra, such as tetrahedra,[5–7] cubes,[7,8] octahedra,[9,10] icosahedra,[11] dodecahedra,[6] buckyballs,[6] and nanoprisms.[7,12] Alternatively, the "DNA origami" or "DNA brick" techniques allow for the construction of an enormous range of shapes built from closely packed helices.[13–15] Making many of these structures is only possible because of DNA's physiochemical properties, which can be manipulated and therefore controlled. Thus, understanding the biophysical and chemical properties of DNA is of vital importance for realizing future nanodevices and nanostructures.

A common motif used in many nanostructures is a 2-way junction in which two double helices are connected by a non-complementary bulge loop of varying size.[16–21] Adjusting the bulge size in these 2-way junctions provides a way to control their flexibility. For example, a 3-arm DNA "star-tile" design,[22] illustrated in Fig. 1(a), contains a combination of 2- and 4-way structural elements linked together by double-helical sections.[23–26] Where the arms meet at the center of the tile, the inner strand (red in Fig. 1(a)) contains bulge loops in order to facilitate the bending of the arms with respect to one another. When a solution of identical star tiles is prepared and cooled, the tiles can self-assemble by linking together at the single-stranded "sticky" ends at the end of each arm. The assembly product is largely controlled by the bulge size.[12] For example, 3-arm tiles containing five nucleotides in each bulge region assemble into tetrahedra (illustrated in Fig. 1(b)), while the reduced flexibility of 3-arm tiles with three nucleotides in the bulge regions leads to the formation of dodecahedra or buckyballs, depending on the tile concentration.[6] Another example is the DNA nanoprism illustrated in Fig. 1(c), where the sizes of the bulges between the different arms can be used to control the detailed geometry of the prism, in particular the relative twist of the top and bottom faces.[12] These examples

a)Electronic mail: john.schreck@chem.ox.ac.uk
b)Electronic mail: jonathan.doye@chem.ox.ac.uk

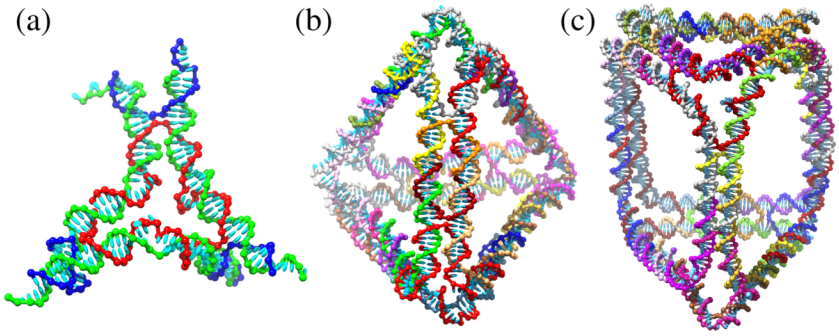

FIG. 1. Representations of nanostructures containing bulges of varying size generated with oxDNA. (a) A 3-arm star tile containing three bulge regions at the center of the tile (red strand), where each bulge contains five nucleotides. (b) A DNA tetrahedron that is formed from four of the 3-arm star tiles shown in (a). (c) A DNA nanoprism made from six 3-arm tiles, whose chirality can be controlled by varying the sizes of bulges in constituent tiles.

illustrate that small changes in sequence design can clearly have a strong influence on the final product as well as the product structure. Even though bulged duplex structures are already widely used in nanodesigns, their effects on the structure and flexibility have not been thoroughly studied and the underlying principles have not been explored. Meanwhile, strong bending of DNA duplexes in the absence of bulges continues to be an area of considerable interest.[27]

Early experimental studies showed that both DNA and RNA duplexes are subject to bending in the presence of bulges and internal loops, with bulges containing ~5–7 nucleotides inducing larger bends than smaller bulges.[16,18,28–32] Subsequent studies showed that the induced bend is not rigid and the junction can adopt a variety of conformations with similar bend angles, which signifies an increase in flexibility caused by the presence of a bulge.[33,34] Following initial computational studies of A-form RNA duplexes in two-way junctions,[35] Bailor *et al.*,[21,36–39] have recently applied three inter-helical (Euler) angles that describe the bending and twisting flexibility at the two-way junctions: one angle quantifies the bend at the bulge and the other two describe the relative twist of each of the duplex arms meeting at the bulge. Their studies of the Protein Data Base (PDB) have shown that the geometric secondary structure of 2-way junctions restricts the overall 3D orientation of helices in RNA,[21] where these constraints arise from the steric and connectivity constraints imposed by the junctions.[38,39]

The Euler angles provide a convenient and consistent way to characterize large numbers of configurations of 2-way junctions, which can be used to elucidate the level of influence the constraints imposed by the secondary structure have on the global conformation of DNA and RNA structures. In addition, several experimental groups have carried out similar analyses of two-way junctions in DNA,[29,40,41] where, for example, in recent work by Woźniak *et al.*,[42] the three Euler angles for bending and twisting have been inferred from Förster resonance energy transfer (FRET) experiments for bulged DNA duplexes with varying bulge size. However, even with these recent experimental advances, much of the literature discussing experimental and theoretical investigations of two-way junctions has to date focused primarily on RNA.

In this article, we use oxDNA, a coarse-grained model at the nucleotide level, to study in detail the structure and thermodynamics of B-DNA double helices containing a bulged loop, a motif that is henceforth referred to as a bulged duplex. We also consider tiles containing two bulges on opposite sides of a duplex ("Z-tiles") that can be used to assemble 1D structures.[43] The results of our simulations are compared with recent experiments in which the bend and twist angles of bulged duplex systems have been measured.[29,40–42] The degree of coarse-graining in oxDNA allows us to study the equilibrium ensemble of junction configurations, which involves observing substantial structural transitions at the junction repeatedly for each system. All-atom investigations of DNA and RNA systems containing bulges[44–46] have not yet been performed with this scope, due to the cost of simulating all-atom force fields.

The oxDNA model has been highly successful at reproducing structural, mechanical, and thermodynamic properties for single- and double-stranded DNA.[47] Moreover, applications to study the fundamental biophysics of DNA, including the kinetics of hybridization,[48,49] toehold-mediated strand displacement,[50] the response to mechanical stress such as the over-stretching transition of dsDNA under tension,[51] the formation of cruciform structures under negative twist,[52] and the role of topology in the formation of kissing hairpin complexes,[53] have confirmed the robustness of the model. Furthermore, the model has proved useful in providing physical insight into the action of DNA nanodevices, such as nanotweezers[54] and walkers,[55,56] and is starting to be applied to characterize large DNA nanostructures.[57] oxDNA captures the relative flexibility of single strands, which can adopt a variety of helical and non-helical structures due to a strand's ability to stack and unstack,[47,58] and the comparatively stiff duplexes. Because oxDNA simultaneously captures these thermodynamic and geometric effects for both single and double strands, it is well-suited to studying how the interplay of such fundamental factors shapes the overall behavior of bulges.

II. MODEL AND METHODS

A. oxDNA model

In oxDNA, a single strand of DNA is modeled as a chain of rigid nucleotides, where each nucleotide contains one interaction site for the backbone and two more interaction sites associated with the stacking, coaxial stacking, and cross-stacking interactions. The interactions in the model are illustrated in Fig. 2. Base-pairing interactions contribute to the overall potential energy only when bases obey Watson-Crick specificity, e.g., A–T or G–C pairs. In addition to base pairing and stacking, there are interactions associated with backbone connectivity and excluded volume. A detailed description of

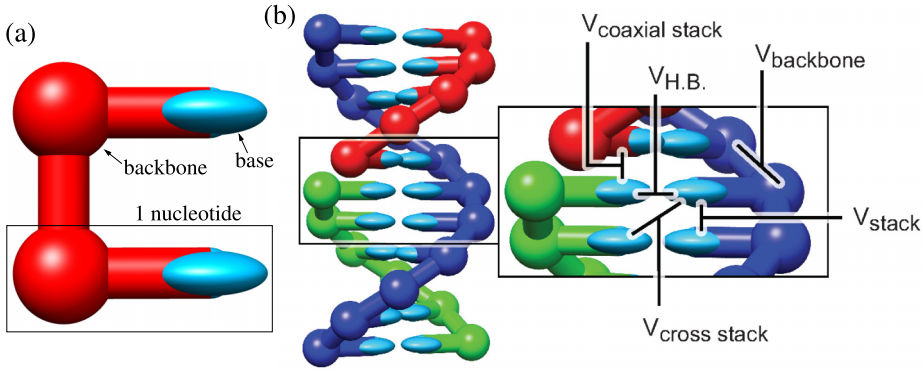

FIG. 2. Simplified representation of (a) the rigid nucleotides that are the basic unit in oxDNA and (b) an 11 base-pair double helix that illustrates the various interactions in the oxDNA model. Reproduced with permission from Doye *et al.*, Phys. Chem. Chem. Phys. **15**, 20395 (2013). Copyright 2013 by the PCCP Owner Societies.

each interaction can be found in Ref. 47. Simulation codes for oxDNA are publicly available from the oxDNA website.[59]

The model does have some simplifications that are important for the current study. First, the model was fit to experiments using a salt concentration of $[Na^+] = 0.5M$, where the electrostatic interactions are strongly screened. Even though some studies have discussed the effect of salt concentration on the overall flexibility of bulged duplexes, the concentration $[Na^+] = 0.5M$ is in the high-salt regime that is relevant to most DNA nanotechnology experiments, and which is our primary interest. Second, the oxDNA double-helix is symmetrical with major and minor grooves being of the same size. While this may be important for certain motifs, we do not expect this approximation to change the dominant physics underlying bulge behavior or the generic trends as a function of bulge size. Third, non-canonical base pairing interactions (non-A–T/G–C interactions between the Watson-Crick edges of bases), and other interactions involving edges such as sugar edge bonds and Hoogsteen pairing, are neglected in the oxDNA potential.[60] These interactions likely affect the flexibility of RNA bulged systems[46] and may affect DNA bulged systems as well. Additionally, we choose to use the "average-base" parametrization of oxDNA, in which the hydrogen bonding associated with base pairing and the stacking interactions has the same strength independent of the bases involved, rather than its sequence-dependent parametrization.[58] This parametrization is advantageous for studying general properties of DNA unmodulated by sequence-dependent effects.

B. DNA bulged duplex systems studied

In our simulations, we first consider a system where two DNA strands associate to form a bulged duplex,

- 5′ – **CTA GCC TTGC** $(T)_M$ *GGAT GCT ACC* – 3′,
- 5′ – *GGT AGC ATCC* **GCAA GGC TAG** – 3′,

where each duplex arm flanking either side of the bulge contains 10 hybridized base pairs (bold regions are complementary to each other, as are italic regions), and the bulge region contains M consecutive thymine (T) bases. This structure is very similar to the bulged duplexes found in the star tiles discussed above,[6] which contain 10 and 11 nucleotides in the arms, respectively. To investigate the effects of the size of the bulge loop on the structural properties, we will consider M ranging from zero to fifteen in our simulations. Additionally, the simulation temperature was set at 23 °C, which is near the temperature (25 °C) where the DNA nanoprism was found at high yield in experiments.[12]

In addition, we also consider Z-tiles,[43] which are duplexes containing two bulge loops that can self-assemble into 1D structures. The following strand can be used to form a Z-tile:

- 5′ – **CTAACC**ACTGG*TGTCCGGACA***GGTTAG** CCAGT – 3′,

where the bold regions are complementary, as are the underlined regions. The central italic region is a palindrome. Two of these strands can hybridize through the bold and italic regions, where the final structure contains one (italic) duplex section, two (bold) duplex sections, and 4 single-stranded (underlined) sections (two bulges and two sticky ends). Note that the italic duplex section is about one turn in length. The sticky ends are specifically designed to be complementary to the nucleotides in the bulged loop. Under certain conditions, the sticky ends may bind with loops from nearby tiles and can form a T-junction, that is, a region in which 3 helical arms meet at a bulge that resembles a T shape, which facilitates the self-assembly of the Z tiles into 1D nanostructures. The Z-tile is further discussed in Sec. III B. The simulation temperature for Z-tiles was set at 22 °C, the temperature at which the assembled tiles produced 1D nanostructures in high yield.[43]

C. Simulation details

To calculate free energies of structures and investigate flexibility, bulged duplexes for several bulge sizes M are simulated by employing virtual-move Monte Carlo (VMMC), as introduced by Whitelam and co-workers.[61] We use the variant introduced in the Appendix of Ref. 61. VMMC is a cluster-move algorithm that efficiently samples from the canonical ensemble for systems of strongly interacting particles, which we found particularly useful for DNA.[57]

The relative free-energy of the bulged duplexes was sampled as a function of two order parameters: (1) the total number of base pairs in the system and (2) the end-to-end distance, R_{ee}, defined as the distance between the center of mass of the bases at the 5′ and 3′ ends of the strand not containing the bulge. Two bases are considered to be paired when the hydrogen bonding interaction between them is 0.093 times its well-depth (0.596 kcal mol^{-1} at 23 °C). This cutoff was chosen because the distribution of base-pairing energies is strongly bimodal, with interactions being typically very negative or zero. We use a finite value (rather than zero) as a criterion to

avoid counting fleeting and extremely weak interactions that could obscure real effects. Small variations in this criterion do not significantly affect the results. OxDNA predicts that the base pairs at the ends of double helices may temporarily break, a process termed fraying, even for systems well below their melting temperatures. Since this phenomenon may occur at either terminal end of a bulged duplex, or in the middle near the bulge loop, the effects that fraying may have on bending can be monitored by following changes in the number of base pairs. The end-to-end distance, R_{ee}, is useful for monitoring how much the system bends, with strong bending occurring when R_{ee} is small compared to the length of a relaxed duplex.

D. Stacking and bulge classes

In our simulations, the bulged duplex can adopt a variety of conformations for different bulge sizes, some of which are illustrated in Fig. 3, where the stacking at the center of the duplex system may or may not be interrupted. For example, in Fig. 3(a), the stacking interaction between the two bases that flank either side of the bulge, which are not neighbors along the sequence, is intact. The stacking interaction between the two bases directly opposite to the bulge, which are neighbors, is also intact and the system is mainly straight. In the model, the stacking interaction between the bases flanking the bulge is referred to as a coaxial stack to distinguish it from the stacking interaction between two neighboring bases, such as the stack opposite to the bulge. Fig. 3(b) illustrates a configuration where a 1-base bulge has become inserted into the helix and stacks with its neighboring bases that are adjacent to the bulge. The stack opposite the bulge is intact and the system is mainly straight. In Fig. 3(c), a configuration is shown where the coaxial stack is broken and a base from the loop is inserted into the helix while the rest of the bulge bases are outside of the helix. The stack opposite from the bulge is intact, and only marginal bending of the system away from the bulge is observed. In Fig. 3(d), both the coaxial stack and the stack opposite to the bulge are broken and the system is bent away from the bulge. Other illustrated states of the system include a configuration in which fraying disrupts the base pairs flanking the bulge as well as interrupting stacking along the helix (Fig. 3(e)) and a configuration which bends towards the bulge (Fig. 3(f)). Similar observations are made for the Z-tile at each of the two bulged regions.

Preliminary simulations indicated the presence of four main bulge conformations. The classifications are as follows:

A: The bases in the loop disrupt the duplex as little as possible and are flipped out, which results in almost no bending.
B: The bases from the bulge loop are inserted into helix while maintaining stacking opposite the bulge, resulting in some degree of static bending away from the bulge.
C: A combination of **A** and **B** in which some bases are inserted into the helix, while others are flipped out.
D: The stacks opposite and across from the bulge are broken, resulting in increased flexibility and a large static bend.

In order to apply these classifications, we focus on the stacking interactions between the pairs of bases near the bulge at the center of the system as well as base pairs that are adjacent to the bulge, as illustrated in Fig. 4. To determine whether a base stacks or coaxially stacks with another base, a lower bound for the stacking interaction between two bases is defined to be the same as the hydrogen bonding cutoff value of 0.596 kcal mol^{-1} at 23 °C. The same reasoning behind the base pair energy cutoff applies to the cutoff for the stacking interactions.

In Fig. 4(a), the state of the coaxial stacking interaction between bases across the bulge is denoted by the symbol i_N, while the stacking states of the two bases on either side of the bulge are denoted by i_{N-1} and i_{N+1}, respectively. Likewise, the state of the stack opposite the bulge is denoted by j_N, while the state of the stacks between neighboring bases on the same strand on either sides of the bulge is denoted by j_{N-1} and j_{N+1}, respectively. Finally, the state of the stacks between the bases with squares and the bases with triangles on either sides of the bulge is denoted by k_1 and k_{M+1}, respectively, and is referred to as "stem-loop" stacks, while the stacks between the bases in the loop are denoted by $k_2, k_3, \ldots, k_M$. If the magnitude of the stacking interaction is greater than or equal to the lower bound, $i_m = 1$, otherwise the stack is taken to be broken and $i_m = 0$, and similarly for j and k.

If the stack opposite the bulge, j_N, is present, we take the two duplex arms to be "stacked" and group these configurations

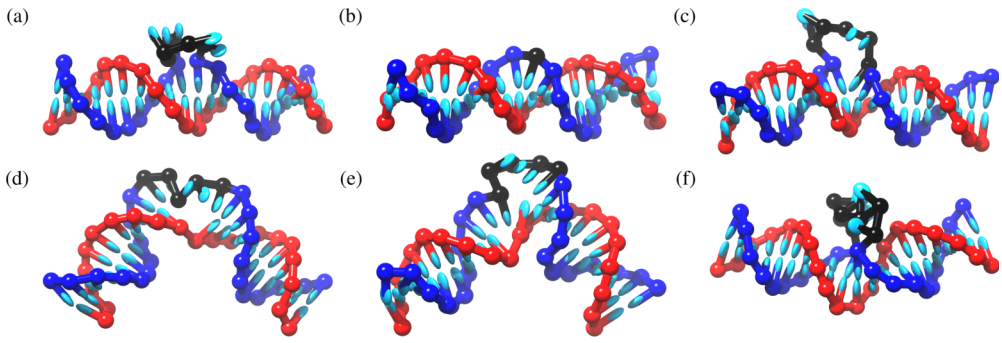

FIG. 3. Configurations of bulged duplex: (a) a configuration classified as **A** in which bases stack across and opposite to the bulge ($M = 5$); (b) a configuration classified as **B** where a bulge loop of size $M = 1$ is inserted into the duplex; (c) a configuration ($M = 5$) classified as **C** where the stacking across the bulge is broken, but not opposite to the bulge, and one base from the loop is inserted inside the duplex; (d) a configuration ($M = 5$) classified as **D**, where the stacking across and opposite to the bulge is broken. Both stem-loop stacks are intact; (e) a configuration ($M = 5$) that has its helix interrupted due to the fraying of a base pair near the bulge; and (f) a configuration ($M = 5$) that is bent into the bulge. For clarity, all backbone elements in the bulge loop are colored black.

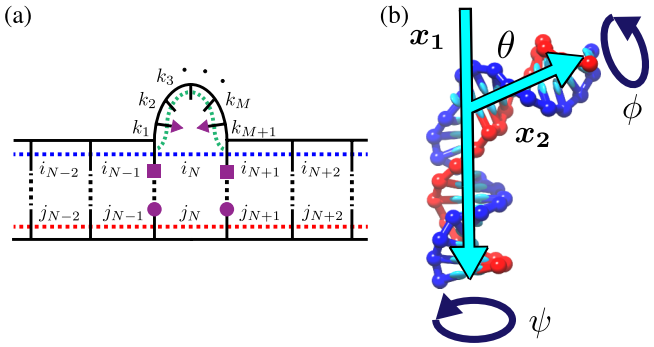

FIG. 4. (a) Illustration of the interactions near the bulge. Short, solid black line segments represent bases, dashed blue lines represent stacking interactions between neighboring bases on the strand containing the bulge (top strand, bases next to the bulge are indicated by squares) and are labeled using the symbol i, while dashed red lines represent stacking interactions between neighboring bases on the opposite strand (bottom strand, bases opposite the bulge are indicated by circles) and are labeled with the symbol j. Dashed green lines that are indicated by the symbol k represent the stacks that involve bases from the bulge loop. Dashed black lines represent H-bond interactions between bases. (b) The bend angle θ and twist angles ϕ and ψ are illustrated along with vectors $\boldsymbol{x_1}$ and $\boldsymbol{x_2}$ that are used to define the angles.

into a stacked set. Configurations where j_N is disrupted are taken to be "unstacked" and are grouped into an unstacked set. According to our classification scheme, a stacked configuration may fall into either **A**, **B**, or **C**, while an unstacked configuration is classified as **D**. The stacked configurations can be subdivided by taking into account the status of the stem-loop stacks that flank both sides of the bulge (k_1 and k_{M+1}) and the stacks between bases in the loop ($k_2, \ldots, k_M$). A stacked configuration falls into **A** if both stacks at the stem-loop interfaces are broken (i.e., k_1 and k_{M+1} are broken); **B** if one or fewer stacking interactions are broken in the strand containing the bulge, counting all stacks from one stem-loop stack to the other, i.e., $k_1 \ldots k_{M+1}$; and **C** if greater than one stack in the loop is broken but stacking is intact at one or both of the stem-loop stacks (more than one of $k_2, \ldots, k_M$ broken and at least one of k_1 or k_{M+1} intact). We note that configurations that have frayed base pairs at the junction can be difficult to classify using our scheme. In the supplementary material, we discuss several extensions to the classification scheme to properly deal with frayed configurations.[62]

E. Bend and twist angles of bulged duplexes

We define the bend and twist angles using a scheme similar to that used in Ref. 21. To measure the bend angle, θ, for a given configuration, we place unit vectors labeled $\boldsymbol{x_1}$ and $\boldsymbol{x_2}$ along the helical axes of each of the duplex arms flanking the bulge. These vectors are illustrated by blue arrows in Fig. 4(a) and in more detail in supplementary material Fig. S1(a).[62] The duplex arms have the freedom to twist about the vectors $\boldsymbol{x_1}$ and $\boldsymbol{x_2}$ as characterized by the angles ϕ and ψ, respectively, as illustrated in Fig. S1(a).[62] The vectors $\boldsymbol{x_1}$ and $\boldsymbol{x_2}$ are defined by finding the longest stretch of base pairs in each arm and then drawing a line from the center-of-mass of the base pair at one end of the duplex arm to the center-of-mass of the base pair at the opposite end of the same duplex arm. For normalized $\boldsymbol{x_1}$ and $\boldsymbol{x_2}$, the bend angle θ is then calculated using

$$\boldsymbol{x_1} \cdot \boldsymbol{x_2} = -\cos(\theta), \tag{1}$$

a definition that follows the convention for the bend angle in the recent literature.[21] In the supplementary material,[62] we develop a simple convention to determine whether the system is bent away from ($0° < \theta \leq 180°$) or bent into the bulge ($-180° \leq \theta < 0°$).

Finally, we explicitly define the duplex twisting angles ϕ and ψ. Each angle can be calculated by first computing a vector that points from the base flanking the bulge to its complementary partner directly across from the bulge. The two vectors are referred to as $\boldsymbol{d_1}$ and $\boldsymbol{d_2}$ and point from square to circle in Fig. 4(a) and are illustrated in Fig. S1(a).[62] The twist angles ϕ and ψ can be calculated using

$$\boldsymbol{d_1} \cdot \boldsymbol{z} = \cos(\phi), \tag{2}$$

$$\boldsymbol{d_2} \cdot \boldsymbol{z} = \cos(\psi), \tag{3}$$

respectively, where $\boldsymbol{z} = \boldsymbol{x_1} \times \boldsymbol{x_2}$ is a vector normal to the plane of the bulged duplexes. Similar to θ, we develop a convention for determining when the angles ϕ and ψ take on the values between $0° < \phi \leq 180°$ ($0° < \psi \leq 180°$) and $180° < \phi \leq 360°$ ($180° < \psi \leq 360°$). This is discussed in section S1 in the supplementary material.[62] The relative twist between the duplex arms flanking the bulge is taken as $\phi - \psi$. For reference, in a relaxed duplex as represented by oxDNA, $\phi - \psi \approx 32°$, i.e., the twist per base pair rises in a duplex.

III. RESULTS

A. Bulged duplex systems

1. Stacking classes

We first discuss how the balance between structural classes **A**, **B**, **C**, and **D** changes as a function of the length of the bulge, M, in oxDNA. Each bulged duplex system with a given bulge length was simulated at least 10 times and configurations were collected until the standard error of the mean for the points computed in the average bend angle versus bulge size was less than 1% of the computed mean value. The results are shown in Fig. 5.

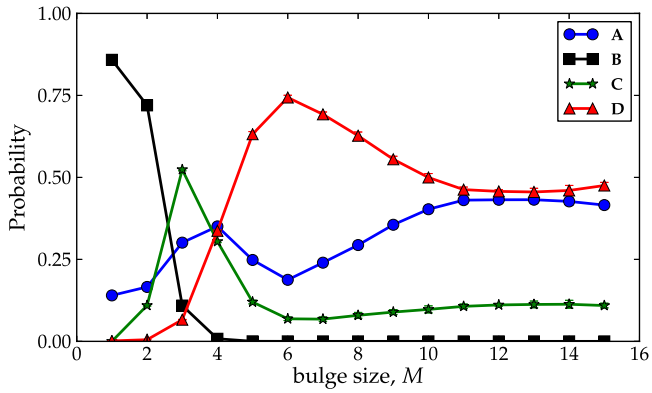

FIG. 5. The fraction of configurations found in structural classes **A**, **B**, **C**, and **D** are plotted as a function of the bulge size M. In the plot the error bar on each point are smaller than the symbol size.

Systems with small bulge loops ($M = 1$ and $M = 2$) are dominated by class **B**. In this scenario, the bases in the bulge are inserted into the helix, as was illustrated in Fig. 3(b). Insertion is favored over flipping the bases outside of the helix because more stacking interactions are preserved. For $M = 3$, it is no longer feasible to insert 3 bases into the duplex and maintain stacking opposite the bulge. Therefore, the probability of **B** decreases dramatically, and instead configurations where one or two of the bases in the bulge are inserted into the duplex, while the others predominately flip out, become prevalent. Additionally, $M = 3$ is the first case where unstacked states make a significant contribution. For $M > 4$, class **D** (e.g., unstacked configurations) becomes the most probable class and of the stacked states, **A** becomes increasingly dominant over class **C** as M increases. Also very apparent is that the probability of stacked states goes through a maximum at $M = 6$ before plateauing off at $M \gtrsim 11$ where cases **A** and **D** are roughly equally probable with some contributions still arising from case **C**.

The most obvious trend is that although unstacked configurations are rare for small M, they become common at larger M. In stacked configurations, the endpoints of the bulge loop are strongly constrained; unstacked configurations offer greater freedom. Whether or not unstacking occurs depends on the relative benefit of this freedom compared to the cost of disrupting junction stacking.

Loops with small M do not benefit as much from the additional freedom of the unstacked ensemble for two reasons. First, they are too short for all relative orientations of the duplex arms to be explored. As the loop gets longer, the duplex arms have more freedom and the unstacked state becomes more favorable. Second, steric penalties associated with the stacked configuration are smaller for small M.

The bulged duplex systems that we study are somewhat similar to a system with one or two single-stranded dangling ends at a junction in a duplex, except that in our cases, the strands are connected and form a bulge loop. Duplexes with such dangling ends can occur during toehold-mediated strand displacement when an invader strand is displacing an incumbent strand during branch migration, as described in Ref. 50. In these systems, once the invading strand is bound to the toehold, there is a free-energy penalty to initiating displacement, even though the number of base pairs is unchanged, because it is unfavorable to have two single-stranded overhangs at the junction. The free-energy penalty arises from the overcrowding of nucleotides at the junction and saturates once both overhangs have at least 3 or 4 nucleotides, because further bases are sufficiently far enough away from the junction that their contribution to overcrowding is minimal.

Similar to the displacement system, for the bulged duplex to maintain a stacked state (i.e., one of cases **A**, **B**, and **C**) the bulge loop must arrange itself to minimize steric clashing amongst the bulge bases and also minimize duplex disruption caused by inserted bulge bases which can compete with the coaxial stack. Alternatively, a bulged duplex may unstack and bend away from the bulge gap, freeing the bulge bases to spread out into space and decrease overcrowding. As with displacement intermediates, the benefit of spreading out is greater when there are more nucleotides at the junction, providing the second cause for the increase in unstacking with increasing M.

Several aspects of the data, however, are not explained by this analysis. First, why does case **C** become less favorable compared to case **A** as loop length increases? Second, why does case **D** increase with respect to case **A** and then subsequently decrease (which will subsequently give a non-monotonic variation of the bend angle as shown in Fig. 6(b))? Both of these questions can be understood in terms of how changes in configuration at the junction are related to the typical physical length of the bulge. In order for the duplex arms to be stacked (in class **A**, **B**, or **C**), any bases not incorporated into the helix must be compressed so that the single-stranded loop region adopts a conformation with a short end-to-end distance. When the bulge loops are short ($M \leq 4$), inserting one or two bases into the duplex helps to reduce the compression substantially: a 2-base loop is less constrained than a 3-base loop because it needs to be compressed less. As the loop gets larger, however, this difference becomes less substantial and so it is not as advantageous for a stacked duplex to incorporate bases. Thus, case **C** is favorable for short loops, but less so for longer loops.

Next, we address the competition between cases **A** and **D**. Medium-sized loops (i.e., $5 \leq M \leq 10$) benefit more from unstacking than longer loops ($M \geq 11$) which resist compression less. This is because medium sized loops are only slightly longer than the persistence length of ssDNA (~1–2 nm inferred in several studies[63–65]), and unstacking of duplexes at the junction allows them to stretch out more easily. By contrast, for longer loops, the relative cost of being bent is lower, and so, they benefit a lot less from being unstacked at the junction. Further, for moderately sized loops, the increase in end-to-end distance upon unstacking is a much larger fraction of the loop contour length than for the larger loops, again implying a greater benefit to unstacking for moderately sized loops. With this in mind, it is perhaps unsurprising that the plateau seen in Fig. 5 is reached when the bulge size is a few times the persistence length of ssDNA and the typical increase in end-to-end extension of the loop upon unstacking.

This typical behavior of polymers can be seen, for example, in the end-to-end probability distribution of a worm-like chain.[66] The analytical formula for the probability density of the end-to-end distance in Ref. 66 clearly shows that a shorter polymer benefits a lot more from having its end-to-end extension increased by a fixed absolute distance when compared with a longer polymer, as illustrated in Fig. S2,[62] where chains with contour lengths that are two and four times the persistence length are compared. Similar results can also be obtained for a freely jointed chain.

We checked whether the non-monotonic behavior was peculiar to our model of ssDNA, or a generic polymer effect, by switching off nearest-neighbor stacking interactions between loop nucleotides (while maintaining them in the stem and at the interface of stem and loop). The results, shown in Fig. 6(a), which plots the free-energy difference between the combined classifications **A**, **B**, **C** and classification **D**, show that stacking between bases in the loop does not significantly change our results and therefore that our results are a robust consequence of generic polymer behavior.

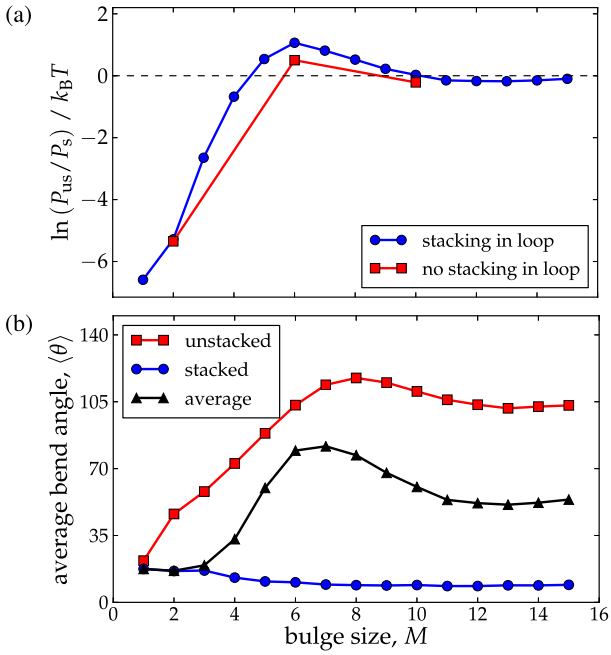

FIG. 6. (a) The relative difference in free-energy between unstacked and stacked configurations. P_{us} and P_s are the probabilities that a bulged duplex is found in the unstacked and stacked states, respectively. (b) The average bend angle for stacked (red circles) and unstacked (blue squares) populations is plotted against bulge size M. The average bend angle for the complete set of configurations is also plotted (black triangles). In both figures, the error bar on each data point is smaller than the symbol size.

2. Effects of stacking on bending angle θ

Fig. 6(b) shows the average bend angle $\langle\theta\rangle$ as a function of bulge size, as well as the bend angle for stacked (classes **A**, **B**, and **C**) and unstacked states (class **D**). The average bend angle for a bulged duplex for small bulge sizes is small when most configurations are stacked, but it quickly grows with M as unstacked configurations become more favorable. The average bend angle peaks at 80° at $M = 7$ before decreasing again and finally plateauing at about 55° for $M \geq 11$, this behavior mainly reflecting the variation of the probability of being unstacked (Fig. 6(a)).

The average bend angle for unstacked states increases monotonically from $M = 1$ and reaches a maximum at $M = 8$ where $\langle\theta\rangle = 117°$, but then starts to decrease as M grows, eventually leveling off at $\langle\theta\rangle \approx 105°$ for $M \geq 11$. The drop in the average bend angle before leveling off is also a signal of the compression effect. The configurations belonging to the stacked population start off slightly bent at 20° due to the high probability that the coaxial stack is broken because a few bases have inserted into the helix, but the bending angle tends towards 10° for longer loops because there is less incentive for the bases to insert into the helix.

Another noteworthy feature is that the bend angle in stacked and unstacked junction configurations is essentially independent of the nearest-neighbor stacking between the bases in the loop (see Fig. S3).[62] The interactions at the stem-loop interface, however, are important. These stem-loop stacks influence the bending angle θ when configurations are classified as **D** (configurations in which duplex arms are unstacked at the junction). We demonstrated this by switching off the nearest-neighbor stacking interactions between bulge loop bases, the stem-loop stacks, the coaxial stack, and the stack opposite to the bulge. We compared this system to systems in which the stem-loop interactions were not removed. The results, presented in Fig. S3, show that for unstacked configurations the stem-loop interactions cause the system to bend significantly more than if the stacks were not present.[62] This is because such stem-loop interactions direct the loop bases to carry on in the direction of the stems, thus requiring a larger bend angle to avoid steric repulsion, as is clear from the configuration in Fig. 3(d).

For a more detailed look at how changes in bulge size influence the flexibility of bulged duplexes, in Fig. 7, we plot the bend angle and end-to-end distance distributions for various M. For stacked configurations, there is a clear asymmetry between bending into and away from the bulge loop, with configurations that bend away from the bulge more favored because these configurations reduce steric clashing between the bulge loop and the duplex. This asymmetry is most pronounced for small M where base insertion is prevalent.

The separation of the stacked and unstacked states is clear from the scatter plots in Fig. 7. As unstacked configurations have a slightly increased effective contour length relative to their stacked counterparts, the two distributions overlap more when projected onto the end-to-end distance, R_{ee}, than θ. The range of bend angles that are available to unstacked states becomes wider as the bulge size increases and underlies the increase of $\langle\theta\rangle$ with bulge size for $M = 1 - 8$. Comparing the bend angle distributions for $M = 8$ with $M = 10$ for the unstacked states, it is clear that the distribution for $M = 10$ is less sharply peaked than the distribution for $M = 8$ and underlies the slight decrease in the average bend angle $\langle\theta\rangle$ for the unstacked states when increasing the bulge size from $M = 8$ to $M = 10$.

3. Effects of stacking on duplex twist angle ϕ – ψ

To quantify the relative twisting of the duplex arms, $\phi - \psi$ is calculated for the same set of configurations. In Fig. 8, the probability that a configuration occupies a state with a given value of $(\theta, \phi - \psi)$ is plotted as a 2D histogram. As with the analysis of the bend angle, θ, we split the configurations into stacked and unstacked sets. A very pronounced (and expected) feature common in all bulged duplexes we considered is a stacked duplex population with a relative twist that is found to lie in a very narrow range, centered at approximately 32°, and is due to configurations in class **A**.

As illustrated in Fig. 8, once a bulge is introduced into the duplex, the relative twist at the bulge maintains a signal at 32° for all values of M studied; however, a second signal at approximately $\phi - \psi \approx 70°$ appears for stacked configurations that are mainly bent away from the bulge. These configurations mostly correspond to the scenario where some number of the bases from the loop are inserted into the duplex at the bulge, as illustrated in Figs. 3(b) and 3(e). As we had shown earlier for all bulged duplexes, one, and sometimes two bases can be inserted into the duplex. Clearly, extra bases at the center of the duplex cause the bulged duplex to twist more at the bulge than stacked configurations belonging to class **A**. The spread for the relative

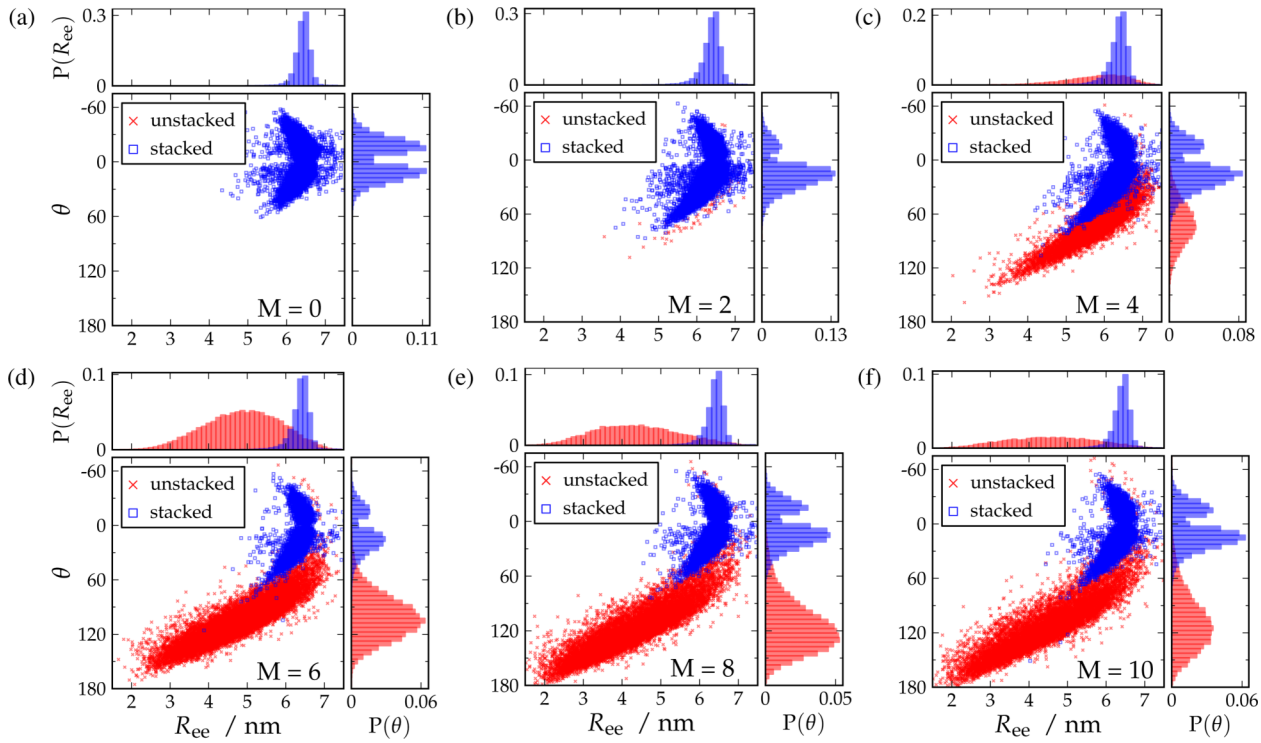

FIG. 7. In each figure, there are three plots that illustrate the relationship between bend angle θ and end-to-end distance R_{ee}: a scatter plot in (R_{ee}, θ) for sets of configurations generated in our simulations, the probability of a state occupying a particular value of R_{ee} (top panel), and the probability of a state occupying some value of the bend angle, θ (right panel). Data are plotted for a selection of bulge sizes in the range $M = 0$–$M = 10$. In the scatter plots, red crosses indicate unstacked states and blue squares represent stacked states. The probability plots retain the same color scheme as the scatter plots. Equivalent plots for other bulge sizes can be found in the supplementary material.[62]

twist angles in these states is also quite narrow and the effects due to base insertion are greatest for $M = 1$ loops when the base in the bulge is much more likely to be found inserted rather than flipped outside of the duplex. Base insertion significantly persists up to $M = 4$, then the effect decreases for $M > 4$ and plateaus. These plots nicely illustrate the fact that there remains a finite probability at large M that a base from a longer loop can still become inserted into a duplex, as can be seen in Fig. 3(c).

Fig. 8 also shows some interesting features of the unstacked bent states. Initially, when the unstacked population first appears with a significant probability compared to the stacked state ($M = 4$), the relative twist angles roughly fall in a circular distribution in the $\theta - (\phi - \psi)$-plane, centered

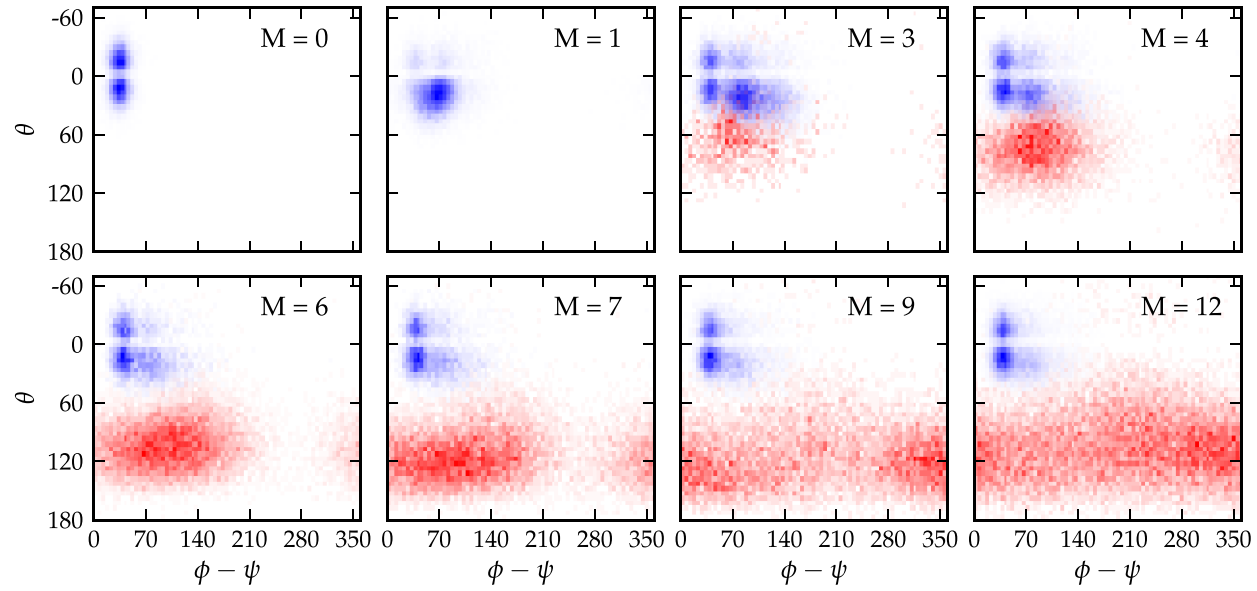

FIG. 8. The probability that a configuration assumes some value of (θ, ϕ-ψ) is plotted for various bulge sizes M. Configurations were separated based on the state of the stacking between the double stranded regions, with the stacked set in blue and the unstacked set in red. The probability distributions for the stacked and unstacked sets were normalized to one separately to enhance clarity. Equivalent plots for other bulge sizes can be found in the supplementary material.[62]

around $\theta \approx 70°$ and $\phi - \psi \approx 100°$. However, deviations from the center are modest and are increasingly less probable due to the constraint of the short loop. There is also some overlap between stacked and unstacked populations for $M = 4$, which corresponds to those configurations where the coaxial stack is broken, but not the stack opposite the bulge, and those configurations where both stacks are broken, but the bending angle is still constrained by the geometry of the loop, respectively. Upon increasing the bulge size further from $M = 6$, the circular distribution elongates along the $\phi - \psi$ axis, because the longer loop allows the bulged duplex arms considerably more freedom to twist relative to each other. Continuing this trend, once the bulge size increases up to $M \geq 11$, the duplex arms can be oriented at almost any relative twist angle with relative ease.

B. Z-tile structure

As mentioned in Sec. II, the Z-tile, illustrated in Fig. 9(a), is a symmetric DNA building block containing two bulge regions, three short duplex regions, and two sticky regions at the terminal ends of the tile, which are complementary to the bulges. The bulge size can be exploited to control the angle in between the duplex regions. Under certain conditions, many identical copies of the Z-tile may link together where the sticky regions of one tile hybridize with the loops from another tile to form T-junctions.[43] An example of a 1D ribbon built using oxDNA and containing 20 Z-tiles is illustrated in Fig. 9(b). Assembled T-junctions can be seen in the structure in Fig. 9(c).

The bending and twisting angles for each bulged region in the Z-tile, as well as end-to-end distance for each duplex region flanking a bulged region, can be defined similarly to the same quantities for the bulged duplex system. A junction is defined to be unstacked if the stack opposite the bulge is broken and stacked otherwise. We use the stacked/unstacked convention to determine if the Z-tile has 0, 1, or 2 bends. The angles and end-to-end distances are defined and illustrated in Fig. S1(b) in the supplementary material.[62] We show the results for the angles by computing the average bend angle at the bulge in each of the two bulge regions, $\frac{1}{2}(\theta_1 + \theta_2)$, and the total average twist measured at the bulges as $\frac{1}{2}(\phi_1 - \psi_1 + \phi_2 - \psi_2)$.

The results for the bend and twist angles for the experimental Z-tile that we consider here,[43] which contains 5 bases in each bulge loop, are plotted in Fig. 10. The plots illustrate that the bending and twisting features of the Z-tiles are similar to the same features found in bulged duplexes, namely, a Z-tile can be approximately straight (indicated by "I" in the figure) with a frequency of 16%, one arm can be bent, while the other arm is approximately straight and vice versa ("L") with

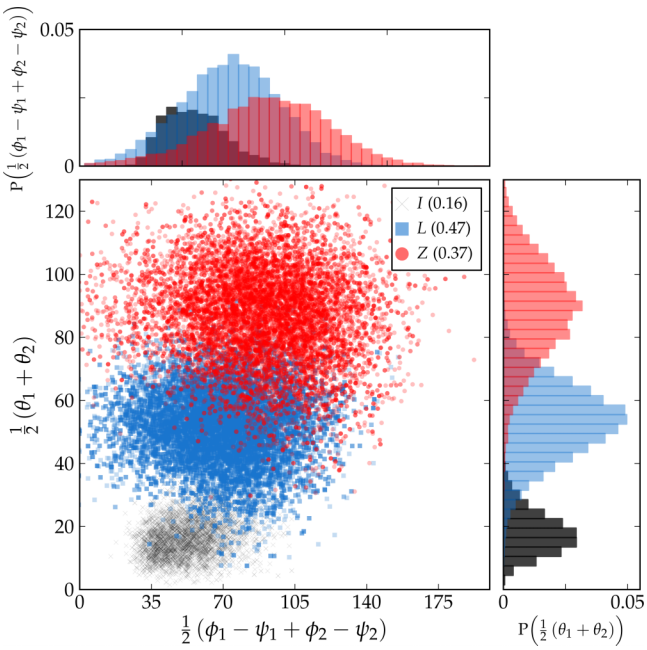

FIG. 10. A scatter plot of the average bend angle versus the average twist at the bulge for different classes of the Z-tile configurations. "I" refers to Z-tile configurations in which both bulged regions are negligibly bent, "L" refers to those configurations in which one region is bent, while the other region is straight, and "Z" refers to those configurations in which both regions are bent at the bulge. Top panel: distributions of the average twist angles for I, L, and Z configurations. Right panel: distributions of the average bend angle of the Z-tile for I, L, and Z configurations.

a frequency of 47%, and finally, both arms can be bent ("Z") with a frequency 37%. These values are consistent with two independent bulges each with bending probability ~62%, as observed in Fig. 5. Predictably, the "Z" configurations also display the most amount of flexibility as is evidenced by the widths of the distributions for the bend and twist angles, while the bend and twisting angles for the "I" configurations cluster more tightly around an average value when compared with the "Z" configurations. Thus, when both arms are straight, the structure is quite stiff, but it can gain a significant amount of flexibility when the system breaks the stacks opposite to both of the bulge loop regions. The configurations where both arms are bent have a similar geometry to Z-tiles that have self-assembled into larger structures.

C. Comparison with experiments and all-atom simulations

FRET experiments carried out by Woźniak *et al.*,[42] yielded bend angles of 32°, 56°, and 73°, for $M = 1, 3,$ and 5,

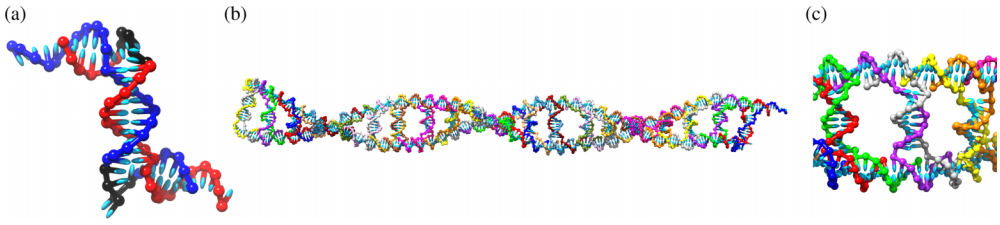

FIG. 9. (a) Illustration of a Z-tile where the structure is bent at both bulges. For clarity, the two bulge regions are colored black. (b) An assembly of twenty Z-tiles that are linked together at T-junctions. (c) Assembled structure of T-junctions.

respectively, which are broadly consistent with previous experimental results.[29,40,41] We see the same trend in the average bend angle predicted by oxDNA, namely, 18°, 20°, and 60° for the same bulge sizes, but the angles are noticeably smaller. Interestingly, the experimental values are actually closer to the average oxDNA bend angle for unstacked states, namely, 21°, 58°, and 88°, for $M = 1, 3,$ and 5, respectively. This perhaps suggests that oxDNA somewhat overestimates the probability of being in the unstacked state and could be due to overestimating the coaxial stacking strength or a neglect of local geometric factors that hinder stacking across the bulge, although we note that there are also significant uncertainties in the experimental measurements. Another possibility is that we underestimate the bend caused by base insertion, or the number of bases that can be inserted. A recent study of bulged duplexes in DNA observed an increase in flexibility of the motifs (a wider range of bend angles was observed) upon an increase in bulge loop size, findings which are broadly consistent with the predictions of oxDNA for the average bend angle distributions.[67] Interestingly, Bailor and co-workers[21,39] have reported results for RNA where the bend angle in bulged RNA duplexes monotonically increased up to about $M = 7$ and then plateaued. This is again similar to the behavior we see for the oxDNA bend angles for the unstacked configurations, which rises and then plateaus at $M = 8$.

The generic states that we have obtained for bulged duplexes of varying size are physically plausible and are also broadly consistent with structures obtained from several all-atom MD studies of bulged systems in RNA and DNA. In Ref. 44, simulations of a well-known DNA octamer containing a one-base bulge found that the bulge base remained either stacked with adjacent bases in the helix, or it was flipped out, though no complete transitions between the stacked and unstacked structures were observed in any single simulation. These conformations are analogous to cases **A** and **B**, and case **C** is essentially a combination of A and B expected to arise in longer loops. Several subpopulations were also reported in the simulations, but with such a short simulation timescale, it is difficult to determine if conformations belonging to the reported subpopulations are thermodynamically representative. Similar short-comings apply to an all-atom study of the flanking bases in RNA kissing-complexes, which noted some transitions from configurations where the bulge bases are inserted into the helix to configurations where the bases are flipped out.[45]

A noteworthy all-atom study of RNA molecules containing varying bulge sizes observed the presence of a wobble G–U base pair flanking the bulge that increased flexibility of the structure when compared with a Watson-Crick G–C base pair in the same location.[46] As noted earlier, the oxDNA potential cannot capture the influence of non-Watson-Crick interactions on flexibility in bulged systems.

Our results are also very relevant to the self-assembly of multi-arm star tiles into polyhedra that was mentioned in Sec. I, and examples of oxDNA representations of these structures are shown in Fig. 1. In particular, the assembly products can depend sensitively on the bulge size with larger bulges favoring structures with larger bend angles at the vertices,[6] and the bulge size can also be used to influence the angles at the vertices of the final polyhedra when not fully constrained by the topology of the polyhedron, as is the case for the prism in Fig. 1(c). First, the observation of well-formed polyhedra in experiment, which would require the breaking of stacking at the bulge, is evidence for the existence of bulge states corresponding to our case **D**. Second, our results show that the free-energy cost of breaking stacking is relatively small for $M \geq 3$ (i.e., $< 2k_BT$), allowing the inter-tile assembly to take place with reasonable ease. Finally, as the bulge loop gets longer, the range of bend angles available clearly increases until a point is reached where the bulge is sufficiently big that all bend angles are feasible. Simply put, larger bulges give rise to greater flexibility once the stacking at the junction is broken. For example, in the unstacked state, the free-energy cost of having a bond angle of 120° (as is required for the triangles in the tetrahedron) is significantly smaller for the $M = 5$ than the $M = 3$ bulge (by roughly 4 k_BT), and so helps to explain why tetrahedra are not formed from 3-arm tiles with bulges of size $M = 3$ but only for $M = 5$.

Even though the bulged duplex systems studied here are simple in comparison to star-tile assemblies, preliminary investigations into large structures such as Mao's tetrahedra and dodecahedra[6] show that the four structure classifications observed in a bulged duplex system with a given M are also observed in these much larger structures that contain bulges. The prevalence of each type within a nanostructure will of course be influenced by geometric and steric constraints arising from the overall polyhedral structure. Additionally, in future work, we intend to study how bulge size explicitly controls the rate of closure of linear trimers to forming triangles. Similarly, in the nanoprism, the lower free-energy cost of bending for larger bulges can explain how the bulge size can be used to control the angles at the vertices and hence the relative twist of the two triangular faces. A more extreme example is that 3-, 4-, 5-, and 6-arm motifs have been found to fully dimerize to form nanotubes with well-defined diameters and lengths when the bulge size is large ($M = 9$), and when the tile contains hairpins in the tile arms, which further increases flexibility.[68] In such structures, the bend angle is likely to be close to 180°.

IV. CONCLUDING REMARKS

We have studied the structural properties of bulged duplexes and Z-tiles using the coarse-grained oxDNA model, with the aim of understanding and characterizing the flexibility of these basic motifs that are widely used in DNA nanotechnology experiments as a means to tune the self-assembly and the equilibrium structure of the final product. We find that bulged duplexes typically adopt one of the four configurations, three of which involve the duplex arms stacking at the junction and one of which does not. When loop sizes are small, the bulged duplex systems mainly prefer to be in a stacked state with some or all of the bases in the loop inserted into the helix. Base insertion has the effect of increasing the amount of twist in the duplex because of the added stacking interaction sites at the center of the system. The resultant average bending angle for small loops is modest and does not deviate too far away from

the average bend angle of duplexes in the absence of bulges, which is about 10°–20°.

Once the bulge loops are about four or more nucleotides in length, configurations are increasingly found to be unstacked. In both stacked and unstacked configurations, the loop tends to be found on one side of the bulged duplex and with either the coaxial stack alone broken or both stacks at the center of the system broken. The latter case allows the system to access greater bending angles. The model also predicts that medium sized bulges that are on the order of the persistence length of ssDNA will resist compression more strongly than longer bulges. This effect causes the system to significantly favor the unstacked configurations over the stacked configurations, with the unstacked configurations exhibiting a large static bend that is partially driven by stacking of the bulge loop with the stems. However, the range of the twisting between the duplex arms is somewhat restricted. Systems with large bulge sizes are less affected by a constrained loop and were found to have the greatest flexibility in which the bending angle θ can assume values over a wide range, and also the relative twist between two duplex arms is free to take on nearly any value from 0° to 360°. We also studied the Z-tile and found similar bending and twisting features that were seen in the duplex system.

oxDNA is a coarse-grained model that was derived to represent generic properties of DNA in a computationally efficient way. As such, it would not be expected to quantitatively reproduce all data from experiments on bulges, to which it was not explicitly parameterized. For example, the simplicity of the excluded volume interaction between nucleotides and the absence of non-Watson-Crick base pairing limit accuracy at the finest levels of detail. It does, however, serve to highlight the underlying physics that is of relevance to real systems. For example, we have identified generic factors that drive changes in occupancy and properties of the four identified configuration types with bulge length. These factors are related to basic polymer properties and geometrical/steric constraints and therefore are also expected to be relevant in experiment. However, given the simplified representation of the local DNA structure in our model, more experiments and all-atom simulations would be desirable to gain a fuller picture of structures containing bulges and to directly test some of our predictions. Specifically, as computational resources improve, all-atom approaches could be used to explore the coexistence of qualitatively distinct states as identified here and perhaps explore their relative prevalence as a function of loop length. These techniques would potentially provide a more quantitatively accurate picture, as well as allowing for neglected features such as non-Watson-Crick base pairing. Experimentally, single-molecule FRET experiments of sufficiently high time resolution could potentially resolve the coexistence of stacked and unstacked configurations. Simpler bulk FRET experiments could be used to observe the plateau in bend angle as a function of bulge length and perhaps also detect the non-monotonicity predicted here.

The properties that we have observed for bulged duplexes also provide insights into the use of this motif in DNA nanotechnology. In particular, the relatively small free-energy cost for unstacking for $M \geq 3$ and the greater flexibility in the unstacked state allowed by increasing bulge size can help to rationalize further the design rules for controlling the self-assembly product and final structure of polyhedra assembled from multi-arm star tiles. As is illustrated in Fig. 1, oxDNA is efficient enough to allow us to study these nanostructures, and understanding their self-assembly and structure will be the subjects of future work.

ACKNOWLEDGMENTS

The authors are grateful to the Engineering and Physical Sciences Research Council for financial support and to Oxford's Advanced Research Computing and E-infrastructure South for computing support. T.E.O. acknowledges funding from University College, Oxford. J.S.S. thanks Dr. Majid Mosayebi and Dr. Petr Šulc for helpful discussions.

[1]J. Watson and F. H. Crick, Nature **171**, 737 (1953).
[2]N. C. Seeman, J. Theor. Biol. **99**, 237 (1982).
[3]J. Bath and A. J. Turberfield, Nat. Nanotechnol. **2**, 275 (2007).
[4]V. Linko and H. Dietz, Curr. Opin. Biotechnol. **24**, 555 (2013).
[5]R. P. Goodman, I. A. Schaap, C. F. Tardin, C. M. Erben, R. M. Berry, C. F. Schmidt, and A. J. Turberfield, Science **310**, 1661 (2005).
[6]Y. He, T. Ye, M. Su, C. Zhang, A. E. Ribbe, W. Jiang, and C. Mao, Nature **452**, 198 (2008).
[7]R. Iinuma, Y. Ke, R. Jungmann, T. Schlichthaerle, J. B. Woehrstein, and P. Yin, Science **344**, 65 (2014).
[8]C. Zhang, S. H. Ko, M. Su, Y. Leng, A. E. Ribbe, W. Jiang, and C. Mao, J. Am. Chem. Soc. **131**, 1413 (2009).
[9]W. M. Shih, J. D. Quispe, and G. F. Joyce, Nature **427**, 618 (2004).
[10]Y. He, M. Su, P. Fang, C. Zhang, A. E. Ribbe, W. Jiang, and C. Mao, Angew. Chem., Int. Ed. **122**, 760 (2010).
[11]C. Zhang, M. Su, Y. He, X. Zhao, P. Fang, A. E. Ribbe, W. Jiang, and C. Mao, Proc. Natl. Acad. Sci. U. S. A. **105**, 10665 (2008).
[12]C. Zhang, W. Wu, X. Li, C. Tian, H. Qian, G. Wang, W. Jiang, and C. Mao, Angew. Chem., Int. Ed. **51**, 7999 (2012).
[13]P. W. Rothemund, Nature **440**, 297 (2006).
[14]S. M. Douglas, H. Dietz, T. Liedl, B. Högberg, F. Graf, and W. M. Shih, Nature **459**, 414 (2009).
[15]Y. Ke, L. L. Ong, W. M. Shih, and P. Yin, Science **338**, 1177 (2012).
[16]F. A. Riordan, A. Bhattacharyya, S. McAteer, and D. M. Lilley, J. Mol. Biol. **226**, 305 (1992).
[17]D. Lilley, Proc. Natl. Acad. Sci. U. S. A. **92**, 7140 (1995).
[18]D. M. Lilley, Q. Rev. Biophys. **33**, 109 (2000).
[19]E. Bindewald, R. Hayes, Y. G. Yingling, W. Kasprzak, and B. A. Shapiro, Nucleic Acids Res. **36**, D392 (2008).
[20]J. A. Cruz and E. Westhof, Cell **136**, 604 (2009).
[21]M. H. Bailor, X. Sun, and H. M. Al-Hashimi, Science **327**, 202 (2010).
[22]C. Zhang, Y. He, M. Su, S. H. Ko, T. Ye, Y. Leng, X. Sun, A. E. Ribbe, W. Jiang, and C. Mao, Faraday Discuss. **143**, 221 (2009).
[23]M. Ortiz-Lombardía, A. González, R. Eritja, J. Aymamí, F. Azorín, and M. Coll, Nat. Struct. Mol. Biol. **6**, 913 (1999).
[24]D. M. Lilley, in *Mathematics of DNA Structure, Function and Interactions* (Springer, 2009), pp. 213–224.
[25]C. Laing and T. Schlick, J. Mol. Biol. **390**, 547 (2009).
[26]D. Han, S. Pal, Y. Yang, S. Jiang, J. Nangreave, Y. Liu, and H. Yan, Science **339**, 1412 (2013).
[27]A. Vologodskii and M. D. Frank-Kamenetskii, Nucleic Acids Res. **41**, 6785 (2013).
[28]C.-H. Hsieh and J. D. Griffith, Proc. Natl. Acad. Sci. U. S. A. **86**, 4833 (1989).
[29]C. Gohlke, A. Murchie, D. Lilley, and R. M. Clegg, Proc. Natl. Acad. Sci. U. S. A. **91**, 11660 (1994).
[30]D. M. Lilley, R. M. Clegg, S. Diekmann, N. C. Seeman, E. Von Kitzing, and P. J. Hagerman, Nucleic Acids Res. **23**, 3363 (1995).
[31]M. Zacharias and P. J. Hagerman, J. Mol. Biol. **247**, 486 (1995).
[32]M. Zacharias and P. J. Hagerman, J. Mol. Biol. **257**, 276 (1996).
[33]H. M. Al-Hashimi, Y. Gosser, A. Gorin, W. Hu, A. Majumdar, and D. J. Patel, J. Mol. Biol. **315**, 95 (2002).
[34]H. D. Kim, G. U. Nienhaus, T. Ha, J. W. Orr, J. R. Williamson, and S. Chu, Proc. Natl. Acad. Sci. U. S. A. **99**, 4284 (2002).
[35]V. B. Chu, J. Lipfert, Y. Bai, V. S. Pande, S. Doniach, and D. Herschlag, RNA **15**, 2195 (2009).

[36]M. H. Bailor, A. M. Mustoe, C. L. Brooks III, and H. M. Al-Hashimi, Curr. Opin. Struct. Biol. **21**, 296 (2011).
[37]M. H. Bailor, A. M. Mustoe, C. L. Brooks III, and H. M. Al-Hashimi, Nat. Protoc. **6**, 1536 (2011).
[38]A. M. Mustoe, M. H. Bailor, R. M. Teixeira, C. L. Brooks III, and H. M. Al-Hashimi, Nucleic Acids Res. **40**, 892 (2012).
[39]A. M. Mustoe, H. M. Al-Hashimi, and C. L. Brooks III, J. Phys. Chem. B **118**, 2615 (2014).
[40]U. Dornberger, A. Hillisch, F. A. Gollmick, H. Fritzsche, and S. Diekmann, Biochemistry **38**, 12860 (1999).
[41]F. Stühmeier, A. Hillisch, R. M. Clegg, and S. Diekmann, J. Mol. Biol. **302**, 1081 (2000).
[42]A. K. Woźniak, G. F. Schröder, H. Grubmüller, C. A. Seidel, and F. Oesterhelt, Proc. Natl. Acad. Sci. U. S. A. **105**, 18337 (2008).
[43]C. Tian, C. Zhang, X. Li, C. Hao, S. Ye, and C. Mao, Langmuir **30**, 5859 (2013).
[44]M. Feig, M. Zacharias, and B. M. Pettitt, Biophys. J. **81**, 352 (2001).
[45]K. Réblová, E. Fadrná, J. Sarzynska, T. Kulinski, P. Kulhánek, E. Ennifar, J. Koča, and J. Šponer, Biophys. J. **93**, 3932 (2007).
[46]B. N. Macchion, R. Strömberg, and L. Nilsson, J. Biomol. Struct. Dyn. **26**, 163 (2008).
[47]T. E. Ouldridge, A. A. Louis, and J. P. K. Doye, J. Chem. Phys. **134**, 085101 (2011).
[48]T. E. Ouldridge, P. Šulc, F. Romano, J. P. K. Doye, and A. A. Louis, Nucleic Acids Res. **41**, 8886 (2013).
[49]J. S. Schreck, T. E. Ouldridge, F. Romano, P. Sulc, L. Shaw, A. A. Louis, and J. P. K. Doye, "DNA hairpins destabilize duplexes primarily by promoting melting rather than by inhibiting hybridization," e-print arXiv:1408.4401 (2014).
[50]N. Srinivas, T. E. Ouldridge, P. Šulc, J. M. Schaeffer, B. Yurke, A. A. Louis, J. P. K. Doye, and E. Winfree, Nucleic Acids Res. **41**, 10641 (2013).
[51]F. Romano, D. Chakraborty, J. P. K. Doye, T. E. Ouldridge, and A. A. Louis, J. Chem. Phys. **138**, 085101 (2013).
[52]C. Matek, T. E. Ouldridge, A. Levy, J. P. K. Doye, and A. A. Louis, J. Phys. Chem. B **116**, 11616 (2012).
[53]F. Romano, A. Hudson, J. P. K. Doye, T. E. Ouldridge, and A. A. Louis, J. Chem. Phys. **136**, 215102 (2012).
[54]T. E. Ouldridge, A. A. Louis, and J. P. K. Doye, Phys. Rev. Lett. **104**, 178101 (2010).
[55]T. E. Ouldridge, R. L. Hoare, A. A. Louis, J. P. K. Doye, J. Bath, and A. J. Turberfield, ACS Nano **7**, 2479 (2013).
[56]P. Šulc, T. E. Ouldridge, F. Romano, J. P. K. Doye, and A. A. Louis, Nat. Comput. **13**, 535 (2012).
[57]J. P. K. Doye, T. E. Ouldridge, A. A. Louis, F. Romano, P. Šulc, C. Matek, B. E. K. Snodin, L. Rovigatti, J. S. Schreck, R. M. Harrison *et al.*, Phys. Chem. Chem. Phys. **15**, 20395 (2013).
[58]P. Šulc, F. Romano, T. E. Ouldridge, L. Rovigatti, J. P. K. Doye, and A. A. Louis, J. Chem. Phys. **137**, 135101 (2012).
[59]See https://dna.physics.ox.ac.uk for more details about the oxDNA model, including instructions on how to download and use the simulation code.
[60]N. B. Leontis, J. Stombaugh, and E. Westhof, Nucleic Acids Res. **30**, 3497 (2002).
[61]S. Whitelam, E. H. Feng, M. F. Hagan, and P. L. Geissler, Soft Matter **5**, 1251 (2009).
[62]See supplementary material at http://dx.doi.org/10.1063/1.4917199 for further details of methods and results.
[63]M. Murphy, I. Rasnik, W. Cheng, T. M. Lohman, and T. Ha, Biophys. J. **86**, 2530 (2004).
[64]A. Y. Sim, J. Lipfert, D. Herschlag, and S. Doniach, Phys. Rev. E **86**, 021901 (2012).
[65]H. Chen, S. P. Meisburger, S. A. Pabit, J. L. Sutton, W. W. Webb, and L. Pollack, Proc. Natl. Acad. Sci. U. S. A. **109**, 799 (2012).
[66]N. Becker, A. Rosa, and R. Everaers, Eur. Phys. J. E **32**, 53 (2010).
[67]X. Shi, K. A. Beauchamp, P. B. Harbury, and D. Herschlag, Proc. Natl. Acad. Sci. U. S. A. **111**, E1473 (2014).
[68]H. Qian, C. Tian, J. Yu, F. Guo, M.-S. Zheng, W. Jiang, Q.-F. Dong, and C. Mao, Small **10**, 855 (2014).