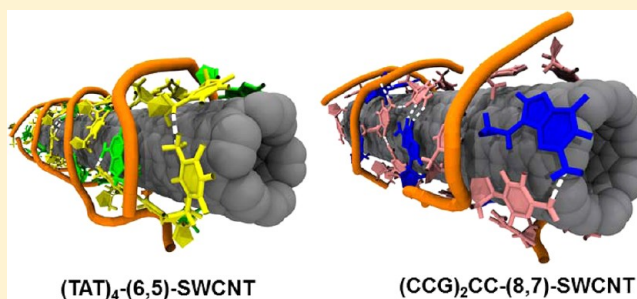


# Structural Characteristics of Oligomeric DNA Strands Adsorbed onto Single-Walled Carbon Nanotubes

Daniel Roxbury,<sup>†</sup> Anand Jagota,<sup>†,‡</sup> and Jeetain Mittal<sup>\*,†</sup><sup>†</sup>Department of Chemical Engineering, and <sup>‡</sup>Bioengineering Program, Lehigh University, Bethlehem, Pennsylvania 18015, United States

## S Supporting Information

**ABSTRACT:** The single-stranded DNA to single-walled carbon nanotube (SWCNT) hybrid continues to attract significant interest as an exemplary biological molecule–nanomaterial conjugate. In addition to their many biomedical uses, such as in vivo sensing and delivery of molecular cargo, DNA-SWCNT hybrids enable the sorting of SWCNTs according to their chirality. Current experimental methods have fallen short of identifying the actual structural ensemble of DNA adsorbed onto SWCNTs that enables and controls several of these phenomena. Molecular dynamics (MD) simulation has been a useful tool for studying the structure of these hybrid molecules. In recent studies, using replica exchange MD (REMD) simulation we have shown that novel secondary structures emerge and that these structures are DNA-sequence and SWCNT-type dependent. Here, we use REMD to investigate in detail the structural characteristics of two DNA-SWCNT recognition pairs: (TAT)<sub>4</sub>-(6,5)-SWCNT, i.e., DNA sequence TATTATTATTAT bound to the (6,5) chirality SWCNT, and (CCG)<sub>2</sub>CC-(8,7)-SWCNT as well as off-recognition pairs (TAT)<sub>4</sub>-(8,7)-SWCNT and (CCG)<sub>2</sub>CC-(6,5)-SWCNT. From a structural clustering analysis, dominant equilibrium structures are identified and show a right-handed self-stitched motif for (TAT)<sub>4</sub>-(6,5) in contrast to a left-handed  $\beta$ -barrel for (CCG)<sub>2</sub>CC-(8,7). Additionally, characteristics such as DNA end-to-end distance, solvent accessible SWCNT surface area, DNA hydrogen bonding between bases, and DNA dihedral distributions have been probed in detail as a function of the number of DNA strands adsorbed onto the nanotube. We find that the DNA structures adsorbed onto a nanotube are also stabilized by significant numbers of non-Watson–Crick hydrogen bonds (intrastrand and interstrand) in addition to  $\pi$ – $\pi$  stacking between DNA bases and nanotube surface and Watson–Crick pairs. Finally, we provide a summary of DNA structures observed for various DNA-SWCNT hybrids as a preliminary set of motifs that may be involved in the functional role of these hybrids.



## 1. INTRODUCTION

Beyond the primary sequence-dependent arrangement of biological polymers (such as carbohydrates, nucleic acids, and proteins) secondary, tertiary, and even more complex higher-order structures are known to determine their function. Secondary interactions, stemming from covalent disulfide bridges, noncovalent van der Waals, and hydrogen bonds are known to control structural rigidity in macromolecules such as ribosomes and enzymes, composed of RNA and proteins, respectively.<sup>1</sup> When these biopolymers are confined to the reduced dimensions of a nanomaterial substrate, such as graphene or carbon nanotubes, bulk structures may not be populated anymore, and instead novel ordered structures may emerge.<sup>2–8</sup>

It is known that monomer units such as nucleotides<sup>9–15</sup> and amino acids<sup>16,17</sup> can self-assemble on hydrophobic substrates to form highly ordered, crystalline monolayers. For nucleotides, the aromatic ring structure of the DNA base causes a strong energy of interaction ( $\sim 10 k_B T$ /base that varies by the base type) between a base and substrate due to overlapping  $\pi$  orbitals in aromatic graphene and carbon nanotubes.<sup>13,14,18–20</sup>

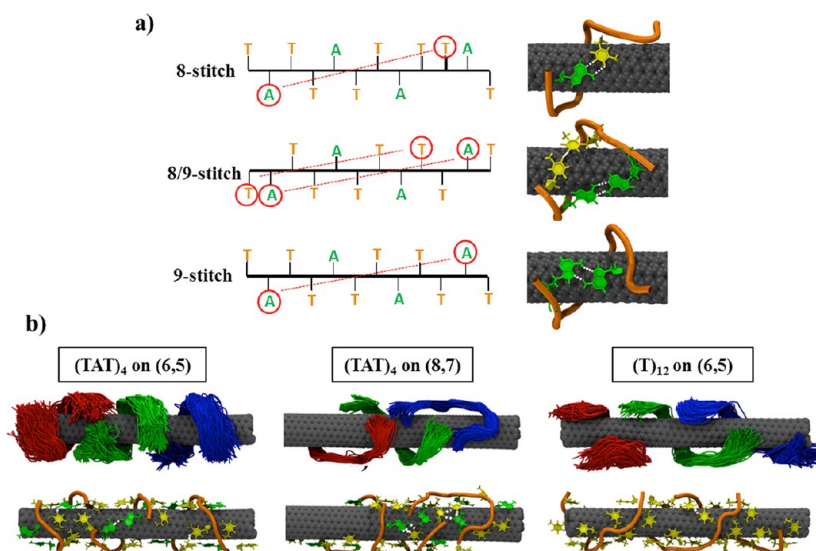
Further, lateral friction between the DNA nucleotides and carbon nanomaterials was determined to be negligible,<sup>21–25</sup> allowing adsorbed molecules to freely rearrange in order to minimize steric hindrances while maximizing the amount of energetically favorable hydrogen bonding. As such, several new structures have been observed in monolayers of adsorbed DNA nucleotides. When nucleotides are constrained by a phosphate–sugar backbone as in a DNA strand, certain monomeric arrangements are disallowed upon adsorption onto a surface.<sup>26</sup> However, molecular simulations suggest that a new class of DNA structural motifs emerges that allows for hydrogen bonding between DNA bases while the chain is confined to a 2D surface.<sup>3,4,26</sup>

Since their structural realization in 2003 by Zheng et al.,<sup>27</sup> noncovalently bound DNA-SWCNT hybrids, as well as covalently modified SWCNTs,<sup>28–31</sup> continue to elicit excitement among the scientific community. Acting as a stable

Received: September 25, 2012

Revised: November 25, 2012

Published: November 30, 2012



**Figure 1.** (a) Self-stitched equilibrium conformations for a single strand of  $(TAT)_4$  adsorbed onto a (6,5)-SWCNT. Three modes of stitching were observed, 8-stitch, 8/9-stitch, and 9-stitch, held together by at least one pair of DNA hydrogen bonds. The percentage of structures in each mode of stitching varied significantly depending on the DNA sequence. (b) Cluster and single-snapshots of multiple strand configurations of  $(TAT)_4$ -(6,5),  $(TAT)_4$ -(8,7), and  $(T)_{12}$ -(6,5)-SWCNT hybrids. The secondary structure varied greatly, depending on both DNA sequence and SWCNT type.

dispersant, single-stranded DNA has effectively been used noncovalently to accomplish solution-based tasks such as sorting by diameter,<sup>32</sup> length,<sup>33</sup> and species<sup>34,35</sup> and for the aligned placement of SWCNTs onto surfaces.<sup>36,37</sup> Such hybrids have also been used in the creation of nanostructured constructs.<sup>38,39</sup> Due to their intrinsic electronic properties, DNA-SWCNTs have been used as field effect transistors (FET) for detection of various analytes in solution and in gaseous phase.<sup>40–43</sup> The highly sensitive band gap photoluminescence in DNA-SWCNT hybrids, strongly dependent on their local environment,<sup>44</sup> has been utilized for molecular detection.<sup>45–48</sup>

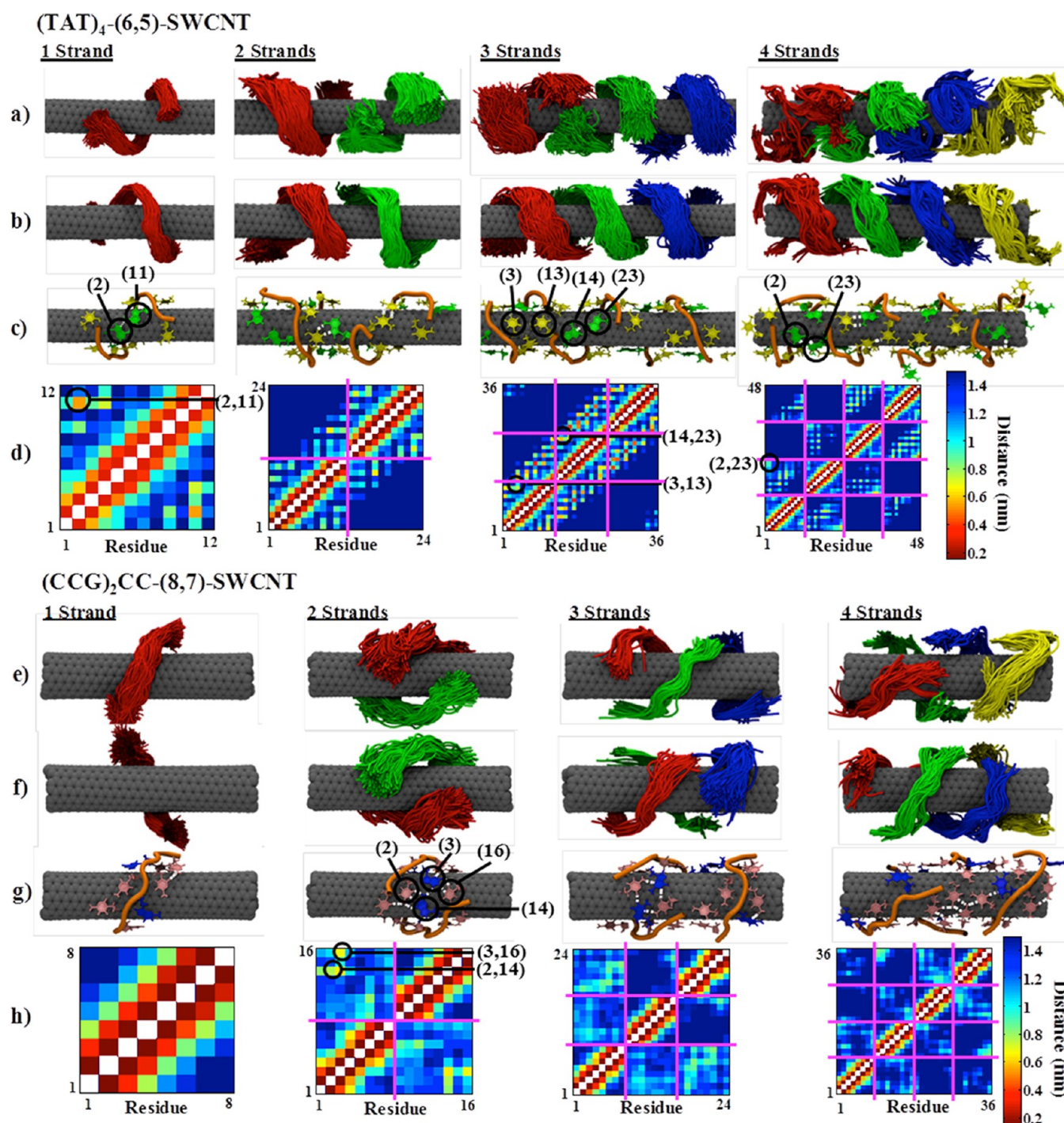
It has been determined experimentally that certain short strands of DNA have the ability to recognize and select for specific species (chiralities) of single-walled carbon nanotubes (SWCNTs).<sup>5</sup> These DNA “recognition sequences” are highly sequence-specific oligomers, approximately 8–20 nucleotides in length, which have been used to sort chirality-diverse mixtures of SWCNTs.<sup>2,5</sup> The recognition ability of a DNA strand for a particular target SWCNT directly correlates with its experimentally measured binding affinity.<sup>49</sup> As a characteristic example, the DNA sequence  $(TAT)_4$ , i.e., 5'-TATTATTAT-TAT-3', binds to its target (6,5)-SWCNT an order of magnitude stronger than the closely related sequences  $(TAT)_4T$  and  $(TAT)_3TA$ . This difference due to the addition or subtraction of a single residue is too large to be explained by nonspecific length effects.

To probe the presumed differences in molecular structure arising from small variations in DNA sequence, molecular dynamics (MD) simulations have been used recently. Johnson et al. showed the emergence of a helically wrapped motif for long strands of single-stranded DNA on a single SWCNT.<sup>50</sup> Karachevtsev et al. further examined these interactions by simulating 25-mer homo-oligonucleotides placed near a (16,0)-SWCNT, obtaining time scales of adsorption, to rank the relative binding free energies.<sup>51</sup> With short MD simulation, double-stranded DNA appears to retain its native structure in the proximity of an SWCNT.<sup>52</sup> The application of advanced sampling techniques in MD simulations of DNA-SWCNT hybrid systems produced results that were significantly better at

representing the equilibrium structural ensemble. One such technique, known as replica exchange molecular dynamics (REMD),<sup>53,54</sup> uses higher temperature simulations (in addition to the temperature of interest), where the system can easily surmount energy barriers separating the various conformational states, to achieve robust configurational sampling. The REMD technique was first applied in the context of the DNA-SWCNT hybrid by Martin et al. where free energy landscapes were found to be complex with multiple minima separated by high barriers.<sup>55</sup> Johnson et al. then applied this MD technique to the DNA sequence  $(GT)_7$  simulated for 100 ns on an (11,0)-SWCNT. They showed that single DNA strands form a broad spectrum of structures (with similar free energies) while adsorbed to the surface of a SWCNT.<sup>56</sup> More recently, Xiao et al. have investigated equilibrium structures of dinucleotides on various SWCNT species through REMD simulation.<sup>57</sup> They reported variations in structure dependent upon the dinucleotide composition as well as diameter and chirality of the SWCNT. To investigate the sequence-dependent nature of DNA-SWCNT hybrids, we have previously simulated, using 200 ns-long REMD simulations, the (6,5)-SWCNT and single strands of its DNA recognition sequence,  $(TAT)_4$ , as well as closely related sequences  $T_{12}$ ,  $A_{12}$ , and  $T_4A_4T_4$ .<sup>3</sup> We found that when the length of a single strand is reduced such that the DNA's contour length (outstretched length) is on the same order as that of the SWCNT circumference, many interesting structural properties of DNA emerge in simulation. We showed that a right-handed helical structure is the dominant conformation in equilibrium for these strands and further that it is stabilized by hydrogen bonded “self-stitches” between distant DNA bases, i.e., a 12-mer DNA strand wraps around the (6,5)-SWCNT and hydrogen bonds to itself (Figure 1a). We also found that a motif in which consecutive DNA bases occupy alternate sides of the backbone is preferred. Finally, we showed that variations in structure were dependent upon DNA sequence, a finding not previously established.

After identifying single strand equilibrium structures for a variety of DNA-SWCNT hybrids, we then extended our study to probe the interactions between multiple strands (three) of





**Figure 2.** Clustering of simulated structures based on DNA backbone atom positions. (a,e) Snapshots of largest clusters shown with overlaid backbones and (b,f) rotated 180° about SWCNT axis. (c,g) Nucleic representations shown for each configuration extracted from largest clusters. (d,h) Mean smallest distance maps for DNA residues in the largest cluster for each configuration. Strands have been separated by the magenta horizontal and vertical lines. Residue distances with respect to themselves are marked in white. Additionally, certain residues have been marked to show the emergence of secondary structure in a given configuration. Note that the scale varies as the number of residues is different for every configuration.

DNA adsorbed onto a single SWCNT. By using nearly sufficient DNA to cover the SWCNT surface, this system more closely mimics experimentally synthesized hybrids.<sup>4</sup> The main conclusions from this study were that the adsorbed DNA structure and interstrand interactions are again highly dependent upon DNA sequence as well as on SWCNT chirality. Three strands of (TAT)<sub>4</sub> on a (6,5)-SWCNT gave rise to a right-handed helical structure in which consecutive DNA strand ends

conjoined to effectively mimic one long strand. In contrast, three strands of T<sub>12</sub> on the same (6,5)-SWCNT formed a left-handed helical structure arranged in an overlapping fashion reminiscent of previously proposed  $\beta$ -barrel structures.<sup>26</sup> Furthermore, the ordered helical structures of (TAT)<sub>4</sub> disappeared completely when simulated on the significantly larger (8,7)-SWCNT (Figure 1b).

In this paper, we report in greater detail on the intrinsic properties associated with various DNA-SWCNT hybrid combinations. We wish to answer the two main questions: (1a) are DNA-SWCNT properties dependent upon surface crowding? (1b) What is the effect of added DNA strands on a given SWCNT? (2) Additionally, are these attributes also dependent upon DNA sequence? To answer the first question, we show how the progressive addition of DNA strands in a simulation on a single SWCNT affects many of the properties intrinsic to the hybrid molecule. These include trends in DNA backbone structure, hydrogen bonding, DNA backbone dihedral angles, and solvent accessible SWCNT surface area. To answer the second question, we have examined in detail two different DNA-SWCNT recognition pairs,<sup>5</sup> (TAT)<sub>4</sub>-(6,5)-SWCNT and (CCG)<sub>2</sub>CC-(8,7)-SWCNT, and their off-recognition pairs, (TAT)<sub>4</sub>-(8,7)-SWCNT and (CCG)<sub>2</sub>CC-(6,5)-SWCNT. The DNA sequence and SWCNT chiralities were chosen from known recognition pairs at the two ends of the range of diameters studied experimentally<sup>5,49</sup> and theoretically.<sup>3,4</sup>

## 2. METHODOLOGIES

All of the simulation methodologies and parameters are nearly identical to those we have described previously.<sup>3,4</sup> Here, we provide the essentials; the readers can refer to these previous papers for more detail. Configurations composed of one, two, three, or four strands of (TAT)<sub>4</sub> or (CCG)<sub>2</sub>CC are placed in desorbed states around (6,5) or (8,7)-SWCNTs, respectively (see the Supporting Information, section S1). The DNA sequences and SWCNTs are chosen for being known recognition pairs.<sup>5</sup> The SWCNTs used are of length 79.7 and 55.3 Å, with diameters of 7.46 and 10.18 Å for (6,5) and (8,7)-SWCNT chiralities, respectively. The nanotubes are frozen throughout the duration of the simulations and extended to the edges of the simulation box effectively to mimic an infinitely long cylinder under periodic boundary conditions. All structures are visualized in VMD.<sup>58</sup>

The GROMACS 4.5.3 simulation package<sup>59–61</sup> is used in conjunction with the CHARMM27 force field<sup>62,63</sup> for REMD simulations. DNA-SWCNTs hybrids are solvated in 79.7 × 34.6 × 34.6 Å and 55.3 × 40.0 × 40.0 Å water-boxes for (6,5) and (8,7)-SWCNT configurations, respectively. These contained approximately 2500 TIP3P model<sup>64</sup> water molecules with the appropriate number of sodium counterions to balance the negative phosphate charges. The total system size is ~10 000 atoms. Periodic boundary conditions are applied in all directions with long-range electrostatics calculated using the particle mesh Ewald method (PME).<sup>65</sup> Forty replicas of each configuration are created for REMD NVT simulation, having temperatures ranging from 296 to 587 K. Replica temperatures are chosen such that exchange acceptance ratios remain around 20% with an exchange time of 2 ps. The time step of the simulation is 2 fs. With the exception of the single strand cases, each simulation is run for a minimum of 400 ns, for a total aggregate time of 40 × 400 = 16 μs for each simulation. The trajectories are saved at every 10 ps, yielding a total of 40 000 snapshots for production analysis. The single strand configurations are run for 200 ns, needing less time to equilibrate since interstrand interactions are nonexistent. See the Supporting Information, section S2, for more information on equilibration.

For the two, three, and four strand configurations, data analysis is performed on the last 300 ns of available data using

the 300 K replica trajectory. The first 100 ns of trajectory are discarded as an initial equilibration period. For the single strand configurations, data analysis is performed on the last 150 ns of the trajectory. In clustering analysis,<sup>66</sup> trajectories are subjected to a user-defined root-mean squared deviation (RMSD) cutoff value, applied to the DNA backbone atoms, while removing rigid body rotations and translations. Snapshots of the clusters, created using VMD,<sup>58</sup> contain the overlaid DNA backbone atoms. The mean smallest distances between consecutively numbered DNA residues, e.g., 1–36 for three 12-mer strands of (TAT)<sub>4</sub>, are calculated and plotted for each configuration cluster in the form of a two-dimensional contact map.

The number of hydrogen bonds in a given snapshot of a trajectory are calculated using a cutoff angle of 30° (angle made between acceptor atom and donor atom pair) and a cutoff distance of 0.35 nm (distance between acceptor atom and hydrogen of the donor pair). Hydrogen bonding fractions are then reported as an average number over the equilibrium ensemble.

For dihedral angle analysis, a potential of mean torque (PMT) plot is constructed using simulated equilibrium probabilities of the five major backbone dihedral angles in DNA. For comparison, backbone dihedral probabilities are extracted from simulations of single and double-stranded DNA, ssDNA and dsDNA, respectively. For dsDNA, a standard MD simulation of 50 ns is performed on a pre-equilibrated double-helix of the sequence CGCGTTAAGCGC and its complement. In the case of ssDNA, 50 ns of REMD is performed in order for the DNA configurations to be sampled effectively without being trapped in any one metastable state.

## 3. RESULTS

**3.1. Structure-Based Clustering Reveals Dominant Equilibrium Structures.** The simulated equilibrium trajectories are clustered based upon the backbone atom positions of all available DNA residues, allowing one to extract the dominant structure for each stable configuration. From user-defined RMSD cutoff values that depend on the total number of DNA residues (see the Supporting Information, section S3), Figure 2 illustrates the largest clusters for all of the DNA-SWCNT recognition pair combinations that represent ~40% of each respective trajectory. It is clear that variations in structure arise at the single-strand level. The one strand (TAT)<sub>4</sub>-(6,5) configuration formed the previously reported right-handed helical self-stitched structure,<sup>3</sup> stabilized by the pair of hydrogen bonds between distant DNA bases on a strand that has fully wrapped the SWCNT. This is confirmed by the close contacts between distant residues in its mean smallest distance map (e.g., coordinates (2,11), corresponding to a separation of 9 residues shows much lower distance than actual separation distance along the DNA contour length). In contrast, a single strand of (CCG)<sub>2</sub>CC has a contour length much shorter than the circumference of the (8,7)-SWCNT. Thus, complete wrapping with hydrogen-bonded self-stitching is absent in the one strand (CCG)<sub>2</sub>CC-(8,7) configuration and there is a slight left-handed pitch of the (CCG)<sub>2</sub>CC DNA on the (8,7)-SWCNT. In addition, there is evidence that bases occupy alternate sides of the DNA backbone, shown by a checkerboard pattern in the distance maps.

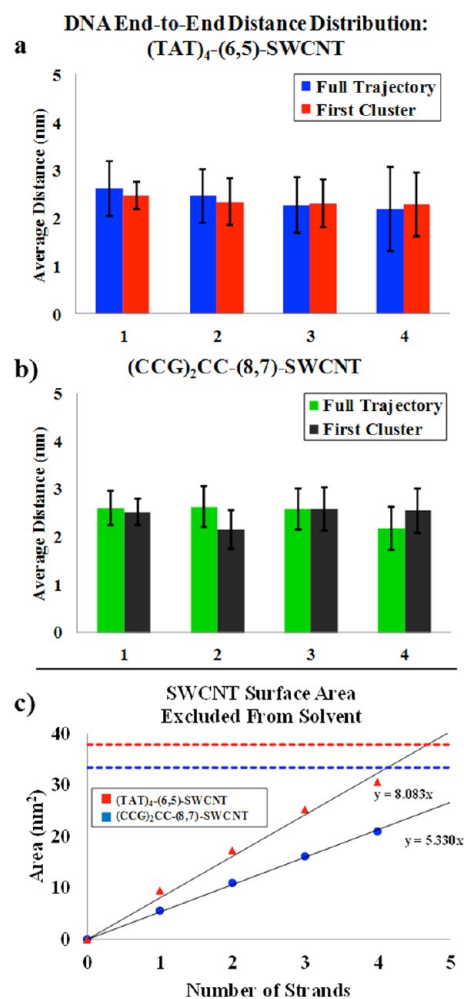
As DNA strands are added to the surface of the SWCNT, two very different structural motifs are observed. The (TAT)<sub>4</sub>-(6,5) configurations of two, three, and four strands appear to retain the individual self-stitched structure, at the same time



arranging such that additional strands dock near existing strand ends. The mean smallest distance maps show a secondary band which is 8–9 bases away from the 45° line, for example, the (3,13) and (14,23) pairs. Particularly in the three strand configuration, the structure effectively appears to form one long 36 residue-long, right-handed single-stranded helix, as if the strands had ligated. In this case, intrastrand as well as interstrand interactions, shown as close contacts between representative residues (14,23) and (3,13), respectively, become rather significant. By visual observation, the four strand (TAT)<sub>4</sub>–(6,5) configuration appears to be crowded. For example, when displayed in its nucleic representation, we find multiple DNA bases are desorbed from the SWCNT surface, indicating the crowded nature of the configuration. The resultant linear charge density of 4.5–6.0 e<sup>−</sup>/nm correlates well with the experimentally determined value of 6.0 e<sup>−</sup>/nm (albeit for a different DNA sequence: poly(GT)).<sup>2</sup> In addition, somewhat random arrangement found in the mean smallest distant map for the four strand cluster is due to the fact that consecutive strands have changed directionality. For instance, a 5′ → 3′ (residues 1–12) adjoined to a 3′ → 5′ strand (residues 13–24) would show close contacts between 12 and 24, rather than 12 and 13, as seen in the two and three strand cases. See the Supporting Information, section S4, for more detailed analysis on DNA directionality in the configurations.

In the (CCG)<sub>2</sub>CC-(8,7) configurations, it appears that an ordered unit is composed of two strands. Essentially, the second strand hydrogen bond stitches to the first strand at both ends to form a doubly stitched clasp. Moreover, the hydrogen bonded DNA strand ends are generally forming guanine–cytosine quartets.<sup>9</sup> This structure is evident in the two strand configuration cluster as well as the mean smallest distance map, where there is indication of residues from different strands coming into close contact, (2,14) and (3, 16). Two strands appear to make a self-contained structure so that when a third strand is placed on the SWCNT, it adopts a somewhat outstretched conformation similar to that of a single strand. This is again seen in the mean smallest distance map; notice the regions of low contact in the upper right quadrant. The four strand configuration possessed certain aspects of the clasp motif (i.e., two consecutive clasps); it can equivalently also be thought of as a double-stranded left-handed helix. Essentially, this is a less ordered form of the previously proposed  $\beta$ -barrel structure for DNA adsorbed to SWCNTs.<sup>26</sup> In this case, any two individual strands did not complete one wrap of the SWCNT but rather cooperatively formed a more elaborate double-stranded structure. The mean smallest distance map for the four strand configuration again shows the prevalence of distant residues in close contact in the formation of this double-stranded structure.

**3.2. Geometrical Properties of DNA-SWCNT Hybrids: DNA End-to-End Distances and Solvent Accessible SWCNT Surface Area.** The three-dimensional DNA end-to-end distances, i.e., 3′ terminal hydrogen to 5′ terminal hydrogen, are compared among many configurations. As shown in Figure 3a,b, the average end-to-end distances do not vary substantially when the number of DNA strands is increased. Deviations, calculated from the probability distribution of the distances, show that all of the configurations of a given DNA-SWCNT pair are within one standard deviation. These plots are shown in the Supporting Information, section S5. In general, the probability distributions of the trajectory and first cluster vary in the degree of overlap, suggesting that the



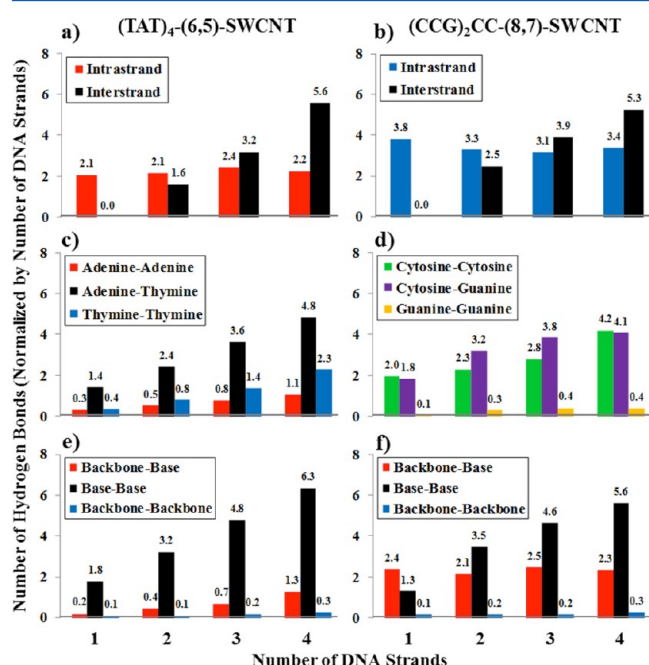
**Figure 3.** Average DNA end-to-end distances for (a) (TAT)<sub>4</sub>-(6,5) and (b) (CCG)<sub>2</sub>CC-(8,7) configurations obtained from the full trajectories as well as the first clusters. Error bars shown are based on standard deviation in the probability distributions. (c) Calculations on the amount of SWCNT surface excluded from contact with the solvent due to the presence of adsorbed DNA. Linear trends suggest that the entire amount of DNA remains adsorbed to the SWCNT. The red and blue dotted horizontal lines are the solvent accessible surface areas for the (6,5) and (8,7)-SWCNTs, respectively. Note that if 5 strands of (TAT)<sub>4</sub> were simulated on the (6,5)-SWCNT, there would not be enough surface area to allow all of the DNA to adsorb.

dominant cluster is not always representative of the entire trajectory. Note that upon the addition of the fourth strand to the (TAT)<sub>4</sub>-(6,5) configuration, broadening of the distribution is seen (larger standard deviation), representing partially desorbed DNA strands.

Since SWCNT optical properties are greatly affected by their local dielectric environment, knowledge of the solvent accessible SWCNT surface area is important for DNA-SWCNT hybrid design considerations for use as analyte sensing devices. For all configurations, the amount of SWCNT surface area excluded from water by the DNA is calculated (by placing a sphere the size of a water molecule near the surface of the SWCNT) as a function of the number of adsorbed DNA strands, Figure 3c. This was found to be a linearly increasing function for both (TAT)<sub>4</sub>-(6,5) and (CCG)<sub>2</sub>CC-(8,7) DNA-SWCNT pairs. Evidence of deviation from the linear fit in the (TAT)<sub>4</sub>-(6,5) at the four strand data point suggests that a

crowded state is reached. Additionally, the trend line from the other set of configurations suggests that  $(CCG)_2CC$  has not reached a crowded state on the (8,7)-SWCNT.

**3.3. Hydrogen Bonding Analysis.** It is known that hydrogen bonding plays a significant role in the development of secondary structures in biological molecules, e.g., the canonical double-stranded DNA is partially stabilized by Watson–Crick hydrogen bonding of bases. Additionally, when DNA is constrained to the surface of a SWCNT, we have shown that hydrogen bonding is crucial for determining which structures are stabilized.<sup>3,4</sup> Figure 4 presents analysis of several types of



**Figure 4.** Hydrogen bonding trends that are observed in the simulated configurations. Reported numbers are normalized by the number of DNA strands in the simulation.

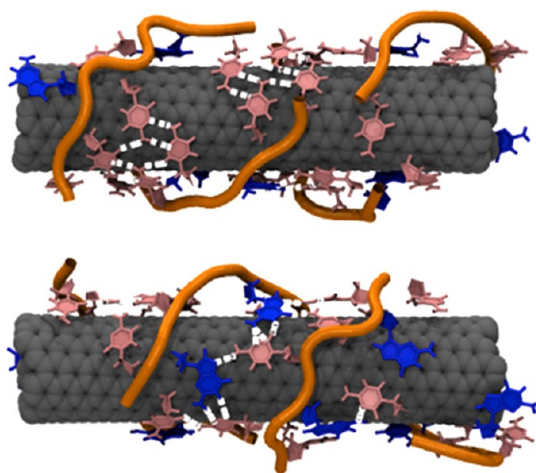
hydrogen bonding schemes. Figure 4a,b show the difference between intrastrand (DNA hydrogen bonds within the same strand encompassing base–base, backbone–base, and backbone–backbone motifs) and interstrand hydrogen bonding. For the two DNA sequences used in this paper, it appears that the number of intrastrand hydrogen bonds is a sequence-dependent quantity that changes little when additional strands are added. For example,  $(CCG)_2CC$  has some 3–4 hydrogen bonds per strand, whereas  $(TAT)_4$  has only about 2, despite the fact that the former has only 8 residues compared to the 12 of the latter. By contrast, the number of interstrand hydrogen bonds increases approximately linearly with additional strands. For both DNA sequences, the normalized number of interstrand hydrogen bonds exceeds the number of intrastrand hydrogen bonds on the addition of a third DNA strand to the simulation.

Figures 4c and d show how hydrogen bonding varies by base type. Since either DNA sequence contains only two types of bases, there are three possible combinations available (adenine–adenine, adenine–thymine, and thymine–thymine for  $(TAT)_4$  and guanine–guanine, guanine–cytosine, and cytosine–cytosine for  $(CCG)_2CC$ .) Note that there is a built-in bias in both sets of simulations because the frequency of base types is not uniform. For example, because there are two “Ts”

for each “A” in the  $(TAT)_4$  sequence, if bases were picked independently (ignoring constraints imposed by the backbone connectivity), the expected probability of AA, “A bonding with A”, would be  $(1/3)^2 = 1/9$ , TT would be 4/9, and AT would be 4/9. Figure 4a shows clear deviation from these values, for example, AT hydrogen bonds are more than twice as likely as TT, indicating the emergence of structures stabilized by these bonds. In general, as the number of DNA strands increases, so does the normalized number of hydrogen bonds in each configuration. In the  $(TAT)_4$ -(6,5) configurations, the Watson–Crick base pair, adenine–thymine, has the largest normalized number throughout all of the configurations of differing numbers of DNA strands. Moreover, non-Watson–Crick base pairings, adenine–adenine and thymine–thymine, are still prevalent in all of the configurations. In the case of  $(CCG)_2CC$ -(8,7), the cytosine–cytosine and cytosine–guanine base pairs have significantly higher fractions than those of guanine–guanine. If bases were picked independently, combinations would appear with probability 1/16, 9/16, 6/16 for GG, CC, and CG, respectively. Indeed, on the addition of a fourth strand to the configuration, cytosine–cytosine and the canonical cytosine–guanine are about equally likely hydrogen bonding partners. Note that for surface adsorbed DNA structures, non-Watson–Crick base pairing (i.e., CC and GG in this instance) is as important as the canonical Watson–Crick pairing.

As a final measure to probe the nature of hydrogen bonding in DNA-SWCNT hybrid structures, the physical location of the hydrogen bond, as it relates to the DNA molecule, is explored. An oligomer DNA strand can be separated into two basic components, bases and a backbone. The backbone consists of phosphate groups and ribose sugars. Furthermore, there are a few oxygen atoms located on the backbone that are able to participate in hydrogen bonding. The normalized fraction of hydrogen bonding is classified as backbone–base, base–base, or backbone–backbone. Base–base hydrogen bonding is dominant in the  $(TAT)_4$ -(6,5) hybrid (Figure 4e). By contrast, the  $(CCG)_2CC$ -(8,7) configurations possess a significantly higher fraction of backbone–base hydrogen bonding (Figure 4f). This appears to be due to guanine and cytosine’s third hydrogen bond donor atoms which enable them to bond to backbone phosphate oxygen atoms with much greater ease than adenine or thymine, see the Supporting Information, section S6.<sup>67</sup>

**3.4. Observed Structures of Interest.** While our main focus has been on the more prevalent structures, it is also useful to note interesting structures that may be forming in low percentages. The main reason for this is that Tu et al. have reported that DNA recognition sequences sometimes select their target SWCNT with less than 1% efficacy.<sup>5</sup> Therefore, the actual structure that has allowed for their chirality-based separation motif may well exist in very low percentages in the equilibrium ensemble. In the case of the  $(TAT)_4$ -(6,5) pair, the dominance of the right-handed single-stranded structure strongly suggests that it plays a pivotal role in allowing for chirality separation. Essentially all  $(TAT)_4$  DNA strands form this ligated helical structure. However, things are quite different in the  $(CCG)_2CC$ -(8,7) hybrid, as the addition of DNA strands to the SWCNT was shown to alter the overall DNA structure. Figure 5 displays a snapshot of the trajectory for the four strand  $(CCG)_2CC$ -(8,7) configuration. In this particular snapshot, a double-stranded  $\beta$ -barrel structure is seen with two distinct grooves. There exists a cytosine–cytosine groove and a



**Figure 5.** (top) Snapshot of the four strand  $(\text{CCG})_2\text{CC}-(8,7)$  configuration and (bottom) rotated by  $180^\circ$  along the SWCNT axis. In this particular snapshot two distinct grooves became apparent: a cytosine–guanine quartet groove and a cytosine–guanine network groove.

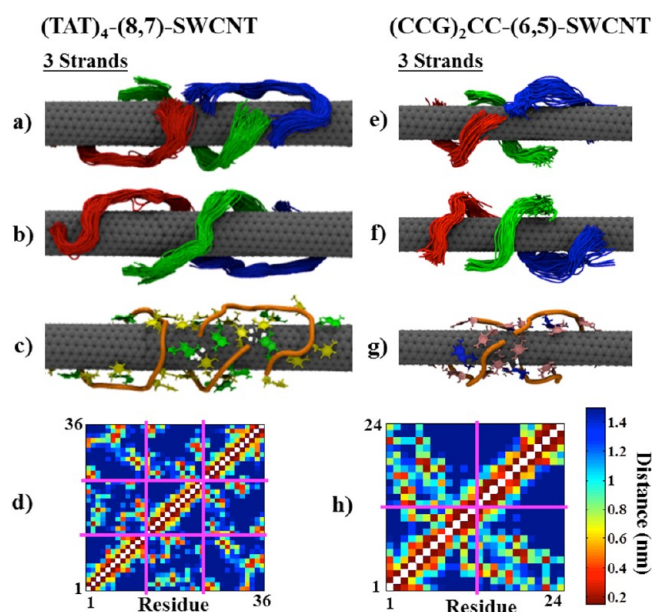
cytosine–guanine groove. In the former of the two grooves, cytosines arrange in quartet formations, greatly increasing the total number of hydrogen bonds. On the opposite side of the SWCNT, the cytosine–guanine groove is additionally rich in hydrogen bonds but forms a network of bases rather than individual quartets, similar to a previous study where we found the emergence of guanine–guanine and thymine–thymine grooves in configurations of  $(\text{GT})_{30}-(6,5)\text{-SWCNT}$ .<sup>26</sup>

**3.5. Off-Recognition DNA-SWCNT Hybrids.** Previous sections all involved a SWCNT and its DNA recognition sequence.<sup>5</sup> Here, we investigate the effect of switching the DNA strands such that two off-recognition DNA-SWCNT pairs are created. Three strand simulations of  $(\text{TAT})_4$  and  $(\text{CCG})_2\text{CC}$  on  $(8,7)$  and  $(6,5)\text{-SWCNTs}$ , respectively, were run and the largest clusters extracted based upon DNA backbone atom positions. The largest clusters are presented in Figure 6. We find that the DNA sequence  $(\text{TAT})_4$  loses the ability to fully wrap the much larger  $(8,7)\text{-SWCNT}$  and form hydrogen bonded self-stitches. Therefore, the idealized right-handed helical symmetry, found previously in Figure 2a, vanishes completely. Instead, the DNA strand forms a clasp on the SWCNT surface (red and green strands in Figure 6a). The remaining blue strand does not participate in the clasp formation and instead forms a random loop. This is further confirmed by a generally disordered pattern in the mean smallest distance map for the cluster as compared to the ordered secondary band seen in Figure 2d.

In addition,  $(\text{CCG})_2\text{CC}$  forms a similar, yet more out-stretched, docked clasp structure on the  $(6,5)\text{-SWCNT}$  when compared to the  $(8,7)$ . Two of the DNA strands adjoin at their ends, whereas the third completed SWCNT coverage on the opposite side. An ordered structure is observed in the mean smallest distance map for this particular cluster. In general, it can be concluded that DNA strands form significantly different structures dependent upon the chirality of SWCNT that they are adsorbed onto.

#### 4. DISCUSSION AND CONCLUSIONS

In this equilibrium simulation study, we have investigated several of the properties intrinsic to the DNA-SWCNT hybrid



**Figure 6.** Simulated structures for off-recognition DNA-SWCNT hybrid molecules. (a,b,e,f) Largest cluster based upon DNA backbone atom position. (c,g) Nucleic representation illustrating bases for one typical structure in the cluster. (d,h) Mean smallest distance maps for the clusters based upon consecutively numbered residues. Strands have been separated by the magenta horizontal and vertical lines.

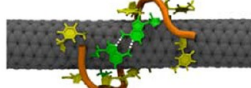
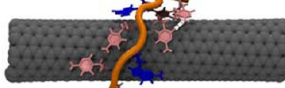
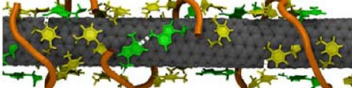
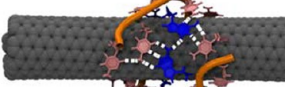
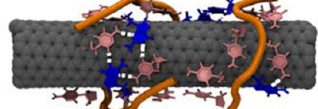
molecule and have correlated these to a systematic increase in the number of adsorbed DNA strands. Significant variation is observed in the equilibrium structures when comparing different combinations of DNA and SWCNT species. Many newly discovered DNA motifs have been observed in the simulations. It was found that DNA behaves significantly differently when confined to the surface of a SWCNT than in free solution (a DNA backbone dihedral analysis is presented in the Supporting Information, section S7). A summary of the 5 observed novel motifs is provided in Table 1. Note the hierarchy embedded in the structures. For example, the “self-stitched” motif is the basic unit of the “docked self-stitch”. Likewise, this applies to the “clasp” and the “docked clasp” motifs. This nomenclature may be useful for future studies in the design of DNA-SWCNT hybrids.

In addition to motifs, we have shown that certain aspects, such as intrastrand hydrogen bonding are determined by the DNA sequence rather than the amount of DNA strands in the simulation. Moreover, we show that non-Watson–Crick base pairing plays a significant role in establishing secondary structure in SWCNT-adsorbed DNA.

Although we have attempted to examine comprehensively some aspects of DNA-sequence/SWCNT pairs, our study is far from exhaustive. Out of the vast sequence space, we have chosen two specific DNA strands and two species of SWCNTs to explore in detail. A coarse-grained DNA model, accounting for complexities such as non-Watson–Crick base–base hydrogen bonding, would be crucial to investigate this vast two-dimensional sequence space of equilibrium DNA-SWCNT hybrid structures in a high-throughput manner. Work on this has already commenced.<sup>68</sup> Additionally, questions remain as to how the DNA strand length will affect its SWCNT-adsorbed structure. End effects and localized defects are also important and must be considered as these are inherent to prepared SWCNTs.



Table 1. Novel Motifs Assumed by DNA when Adsorbed to SWCNTs

Type	Properties	DNA-SWCNT Combination	Snapshots
Self-Stitched <sup>3</sup>	Single strand forms complete wrap around the SWCNT and hydrogen bond stitches to itself	(TAT) <sub>4</sub> -(6,5) 1 Strand	
Incomplete Wrap	Single strand with a contour length shorter than the circumference of the SWCNT	(CCG) <sub>2</sub> CC-(8,7) 1 Strand	
Docked Self-Stitched <sup>4</sup>	Additional strands placed adjacent will adjoin to ends of a self-stitched strand, mimicking one long strand	(TAT) <sub>4</sub> -(6,5) 2,3,4 Strands	
Clasp	For sequences requiring at least two strands to make a complete wrap around the SWCNT. Stitching occurs at two separate locations on opposite sides of the SWCNT	(CCG) <sub>2</sub> CC-(8,7) 2 Strands	
Docked Clasp	Similar to previous docked example, additional strands adjoin to first clasp to form an extended structure. This is similar in appearance to double-stranded β-barrels formed by longer strands of DNA <sup>26</sup>	(CCG) <sub>2</sub> CC-(8,7) 3,4 Strands	

## ■ ASSOCIATED CONTENT

### ■ Supporting Information

Many additional modes of structural analysis of DNA-SWCNT simulations, including DNA directionality, DNA end-to-end distances, and DNA backbone dihedrals. This material is available free of charge via the Internet at <http://pubs.acs.org>.

## ■ AUTHOR INFORMATION

### Corresponding Author

\*E-mail: [jeetain@lehigh.edu](mailto:jeetain@lehigh.edu).

### Notes

The authors declare no competing financial interest.

## ■ ACKNOWLEDGMENTS

This work has been supported by the National Science Foundation through grant CMMI-1014960 and by a Faculty Innovation Grant (FIG) to A.J. from Lehigh University. This research was also supported in part by the National Science Foundation through XSEDE resources provided by the Texas Advanced Computing Center (TACC) under Grant No. TG-MCB100049. The authors have also benefitted from discussions with Drs. Ming Zheng, Xiaomin Tu, and Constantine Khripin at the National Institute for Standards and Technologies (NIST) in Gaithersburg, MD.

## ■ REFERENCES

- (1) Alberts, B.; Johnson, A.; Lewis, J.; Raff, M.; Roberts, K.; Walter, P. *Molecular Biology of the Cell*, Fourth Edition; Garland Science: New York, 2002.
- (2) Khripin, C. Y.; Manohar, S.; Zheng, M.; Jagota, A. *J. Phys. Chem. C* **2009**, *113* (31), 13616–13621.
- (3) Roxbury, D.; Jagota, A.; Mittal, J. *J. Am. Chem. Soc.* **2011**, *133* (34), 13545–13550.
- (4) Roxbury, D.; Mittal, J.; Jagota, A. *Nano Lett.* **2012**, *12* (3), 1464–1469.
- (5) Tu, X. M.; Manohar, S.; Jagota, A.; Zheng, M. *Nature* **2009**, *460* (7252), 250–253.
- (6) Yang, Z.; Wang, Z.; Tian, X.; Xiu, P.; Zhou, R. *J. Chem. Phys.* **2012**, *136* (2), 025103–10.
- (7) Zuo, G.; Huang, Q.; Wei, G.; Zhou, R.; Fang, H. *ACS Nano* **2010**, *4* (12), 7508–7514.
- (8) Grigoryan, G.; Kim, Y. H.; Acharya, R.; Axelrod, K.; Jain, R. M.; Willis, L.; Drndic, M.; Kikkawa, J. M.; DeGrado, W. F. *Science* **2011**, *332* (6033), 1071–1076.
- (9) Xu, S. L.; Dong, M. D.; Rauls, E.; Otero, R.; Linderroth, T. R.; Besenbacher, F. *Nano Lett.* **2006**, *6* (7), 1434–1438.
- (10) Edelwirth, M.; Freund, J.; Sowerby, S. J.; Heckl, W. M. *Surf. Sci.* **1998**, *417* (2–3), 201–209.
- (11) Mamdouh, W.; Dong, M. D.; Xu, S. L.; Rauls, E.; Besenbacher, F. *J. Am. Chem. Soc.* **2006**, *128* (40), 13305–13311.
- (12) Mamdouh, W.; Kelly, R. E. A.; Dong, M. D.; Kantorovich, L. N.; Besenbacher, F. *J. Am. Chem. Soc.* **2008**, *130* (2), 695–702.
- (13) Tao, N. J.; Shi, Z. *J. Phys. Chem.* **1994**, *98* (5), 1464–1471.
- (14) Sowerby, S. J.; Cohn, C. A.; Heckl, W. M.; Holm, N. G. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98* (3), 820–822.



- (15) Sowerby, S. J.; Holm, N. G.; Petersen, G. B. *Biosystems* **2001**, *61* (1), 69–78.
- (16) Mou, J.; Yang, W. S. *Ultramicroscopy* **1992**, *42–44*, Part 2 (0), 1025–1030.
- (17) Roman, T.; Diño, W. A.; Nakanishi, H.; Kasai, H. *Eur. Phys. J. D* **2006**, *38* (1), 117–120.
- (18) Manohar, S.; Tang, T.; Jagota, A. J. *Phys. Chem. C* **2007**, *111* (48), 17835–17845.
- (19) McCarthy, B.; Coleman, J. N.; Czerw, R.; Dalton, A. B.; Panhuis, M. I. H.; Maiti, A.; Drury, A.; Bernier, P.; Nagy, J. B.; Lahr, B.; Byrne, H. J.; Carroll, D. L.; Blau, W. J. *J. Phys. Chem. B* **2002**, *106* (9), 2210–2216.
- (20) Sowerby, S. J.; Edelwirth, M.; Heckl, W. M. *J. Phys. Chem. B* **1998**, *102* (30), 5914–5922.
- (21) Coleman, J. N.; Ferreira, M. S. *Appl. Phys. Lett.* **2004**, *84* (5), 798–800.
- (22) Wall, A.; Ferreira, M. S. *Phys. Rev. B* **2006**, *74* (23), 233401.
- (23) Shi, X.; Kong, Y.; Zhao, Y.; Gao, H. *Acta Mech. Sin.* **2005**, *21* (3), 249–256.
- (24) Manohar, S.; Mantz, A. R.; Bancroft, K. E.; Hui, C. Y.; Jagota, A.; Vezenov, D. V. *Nano Lett.* **2008**, *8* (12), 4365–4372.
- (25) Iliafar, S.; Wagner, K.; Manohar, S.; Jagota, A.; Vezenov, D. J. *Phys. Chem. C* **2012**, *116* (26), 13896–13903.
- (26) Roxbury, D.; Manohar, S.; Jagota, A. J. *Phys. Chem. C* **2010**, *114* (31), 13267–13276.
- (27) Zheng, M.; Jagota, A.; Semke, E. D.; Diner, B. A.; Mclean, S. R. L.; Richardson, R. E.; Tassi, N. G. *Nat. Mater.* **2003**, *2*, 338–343.
- (28) Karousis, N. *Chem. Rev.* **2010**, *110* (9), 5366.
- (29) Banerjee, S.; Hemraj-Benny, T.; Wong, S. S. *Adv. Mater.* **2005**, *17* (1), 17–29.
- (30) Villa, C. H.; Dao, T.; Ahearn, I.; Fehrenbacher, N.; Casey, E.; Rey, D. A.; Korontsvit, T.; Zakhaleva, V.; Batt, C. A.; Philips, M. R.; Scheinberg, D. A. *ACS Nano* **2011**, *5* (7), 5300–5311.
- (31) Villa, C. H.; McDevitt, M. R.; Escorcia, F. E.; Rey, D. A.; Bergkvist, M.; Batt, C. A.; Scheinberg, D. A. *Nano Lett.* **2008**, *8* (12), 4221–4228.
- (32) Zheng, M.; Jagota, A.; Strano, M. S.; Santos, A. P.; Barone, P.; Chou, S. G.; Diner, B. A.; Dresselhaus, M. S.; Mclean, S. R. L.; Onoa, G. B.; Samsonidze, G. G.; Semke, E. D.; Usrey, M.; Walls, D. J. *Science* **2003**, *302* (5650), 1545–1548.
- (33) Huang, X.; Mclean, S. R. L.; Zheng, M. *Anal. Chem.* **2005**, *77* (19), 6225–6228.
- (34) Kim, S. N.; Kuang, Z.; Grote, J. G.; Farmer, B. L.; Naik, R. R. *Nano Lett.* **2008**, *8* (12), 4415–4420.
- (35) Tu, X.; Hight Walker, A. R.; Khripin, C. Y.; Zheng, M. *J. Am. Chem. Soc.* **2011**, *133* (33), 12998–13001.
- (36) McLean, R. S.; Huang, X.; Khripin, C.; Jagota, A.; Zheng, M. *Nano Lett.* **2005**, *6* (1), 55–60.
- (37) Khripin, C. Y.; Zheng, M.; Jagota, A. *J. Colloid Interface Sci.* **2009**, *330* (2), 255–265.
- (38) Müller, K.; Malik, S.; Richert, C. *ACS Nano* **2010**, *4* (2), 649–656.
- (39) Han, S.-p.; Maune, H. T.; Barish, R. D.; Bockrath, M.; Goddard, W. A. *Nano Lett.* **2012**, *12* (3), 1129–1135.
- (40) Keren, K.; Berman, R. S.; Buchstab, E.; Sivan, U.; Braun, E. *Science* **2003**, *302* (5649), 1380–1382.
- (41) Staii, C.; Johnson, A. T.; Chen, M.; Gelperin, A. *Nano Lett.* **2005**, *5* (9), 1774–1778.
- (42) Gelperin, A.; Johnson, A. T. C. *J. Breath Res.* **2008**, *2* (3), 037015.
- (43) Khamis, S. M.; Jones, R. A.; Johnson, A. T. C.; Preti, G.; Kwak, J.; Gelperin, A. *AIP Adv.* **2012**, *2* (022110), 1–11.
- (44) Choi, J. H.; Strano, M. S. *Appl. Phys. Lett.* **2007**, *90*, 223114.
- (45) Barone, P. W.; Baik, S.; Heller, D. A.; Strano, M. S. *Nat. Mater.* **2005**, *4*, 86–92.
- (46) Heller, D. A.; Jeng, E. S.; Yeung, T. K.; Martinez, B. M.; Moll, A. E.; Gastala, J. B.; Strano, M. S. *Science* **2006**, *311*, 508–511.
- (47) Heller, D. A.; Jin, H.; Martinez, B. M.; Patel, D.; Miller, B. M.; Yeung, T. K.; Jena, P. V.; Hobartner, C.; Ha, T.; Silverman, S. K.; Strano, M. S. *Nat. Nanotechnol.* **2008**, *4*, 114–120.
- (48) Jin, H.; Heller, D. A.; Kalbacova, M.; Kim, J.-H.; Zhang, J.; Boghossian, A. A.; Maheshri, N.; Strano, M. S. *Nat. Nanotechnol.* **2010**, *5* (4), 302–309.
- (49) Roxbury, D.; Tu, X.; Zheng, M.; Jagota, A. *Langmuir* **2011**, *27* (13), 8282–8293.
- (50) Johnson, R. R.; Johnson, A. T. C.; Klein, M. L. *Nano Lett.* **2008**, *8* (1), 69–75.
- (51) Karachevtsev, M. V.; Karachevtsev, V. A. *J. Phys. Chem. B* **2011**, *115* (29), 9271–9279.
- (52) Zhao; Johnson, J. K. *J. Am. Chem. Soc.* **2007**, *129* (34), 10438–10445.
- (53) Sugita, Y.; Kitao, A.; Okamoto, Y. *J. Chem. Phys.* **2000**, *113*, 6042–6051.
- (54) Sugita, Y.; Okamoto, Y. *Chem. Phys. Lett.* **1999**, *314* (1–2), 141–151.
- (55) Martin, W.; Zhu, W.; Krilov, G. *J. Phys. Chem. B* **2008**, *112* (50), 16076–16089.
- (56) Johnson, R. R.; Kohlmeyer, A.; Johnson, A. T. C.; Klein, M. L. *Nano Lett.* **2009**, *9* (2), 537–541.
- (57) Xiao, Z.; Wang, X.; Xu, X.; Zhang, H.; Li, Y.; Wang, Y. *J. Phys. Chem. C* **2011**, *115* (44), 21546–21558.
- (58) Humphrey, W.; Dalke, A.; Schulten, K. *J. Mol. Graphics* **1996**, *14* (1), 33–38.
- (59) Berendsen, H. J. C.; van der Spoel, D.; van Drunen, R. *Comput. Phys. Commun.* **1995**, *91* (1–3), 43–56.
- (60) Lindahl, E.; Hess, B.; van der Spoel, D. *J. Mol. Model* **2001**, *7* (8), 306–317.
- (61) van der Spoel, D.; Lindahl, E.; Hess, B.; Groenhof, G.; Mark, A. E.; Berendsen, H. J. C. *J. Comput. Chem.* **2005**, *26* (16), 1701–1718.
- (62) Foloppe, N.; MacKerell, A. D. *J. Comput. Chem.* **2000**, *21* (2), 86–104.
- (63) MacKerell, A. D., Jr.; Banavali, N. K. *J. Comput. Chem.* **2000**, *21* (2), 105–120.
- (64) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926.
- (65) York, D. M.; Darden, T. A.; Pedersen, L. G. *J. Chem. Phys.* **1993**, *99*, 10.
- (66) Fred, A. H.; Christine, P.; Xavier, D.; Walter, T.; Wilfred, F. V. G., A Strategy for Analysis of (molecular) Equilibrium Simulations: Configuration Space Density Estimation, Clustering and Visualization.
- (67) Saenger, W. *Principles of Nucleic Acid Structure*; Springer-Verlag: New York, 1984.
- (68) Shankar, A.; Jagota, A.; Mittal, J. *J. Phys. Chem. B* **2012**, *116* (40), 12088–12094.