Comparative Dynamics and Sequence Dependence of DNA and RNA Binding to Single Walled Carbon Nanotubes

Markita P. Landry,[1] Lela Vuković, [2,3] Sebastian Kruss,[1] Gili Bisker,[1] Alexandra M. Landry,[4] Shahrin Islam,[1] Rishabh Jain,[1] Klaus Schulten,[2,3] Michael S. Strano[1*]

[1] Department of Chemical Engineering, Massachusetts Institute of Technology

[2] Department of Physics, University of Illinois at Urbana-Champaign

[3] Center for the Physics of Living Cells, University of Illinois at Urbana-Champaign

[4] Department of Chemical Engineering, University of California at Berkeley

* Corresponding author: 66-570 77 Massachusetts Avenue, Cambridge, MA 02139, 617–324–4323, strano@mit.edu

Supporting Information: Supplementary Figures S1-S8, Table S1, and models S1 and S2.

Nucleotide	$\langle E_{nucl-CNT} \rangle$ (σ_E) (kcal/mol)
G (DNA)	-19.1 (1.7)
T (DNA)	-15.5 (1.6)
G (RNA)	-19.1 (1.7)
U (RNA)	-14.5 (1.5)
A (RNA)	-17.3 (2.2)
C (RNA)	-14.0 (2.8)

Table S1: Average interaction energies, E_{nucl}–SWCNT, and standard deviations, σ_E, between single nucleotides and SWCNTs. The energies were averaged over the last 10 ns of simulations for nucleotides stacked to SWCNTs.

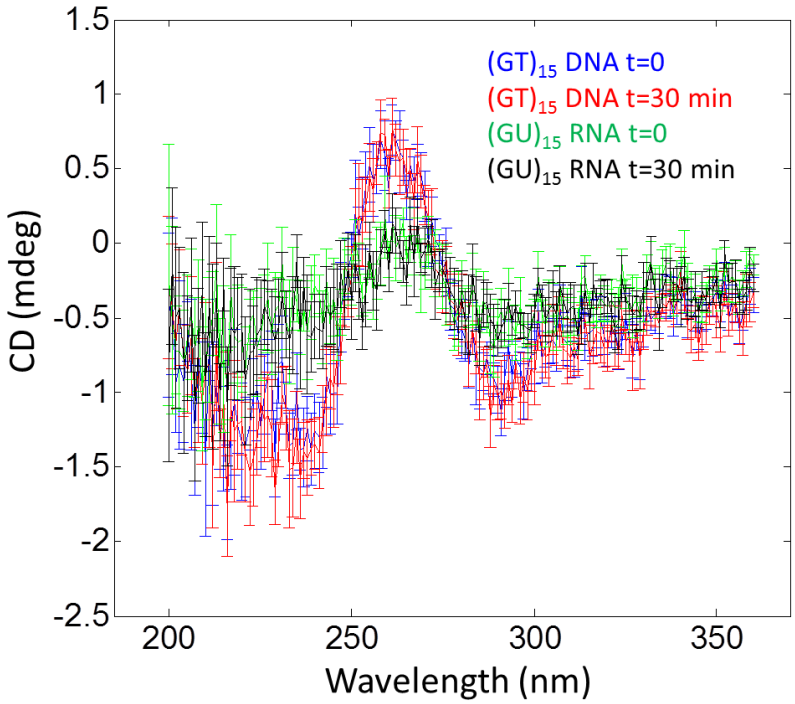

Figure S1: Circular dichroism (CD) spectra for $(GT)_{15}$ DNA- and $(GU)_{15}$RNA-SWCNT samples. Spectra were taken immediately after dilution from 100 mg/L to 2 mg/L at t = 0 and t = 30 minutes later. No difference is observed between the CD spectra of both samples within a time lapse of 30 minutes, suggesting that the DNA or RNA does not desorb from the SWCNT into solution within this time frame.

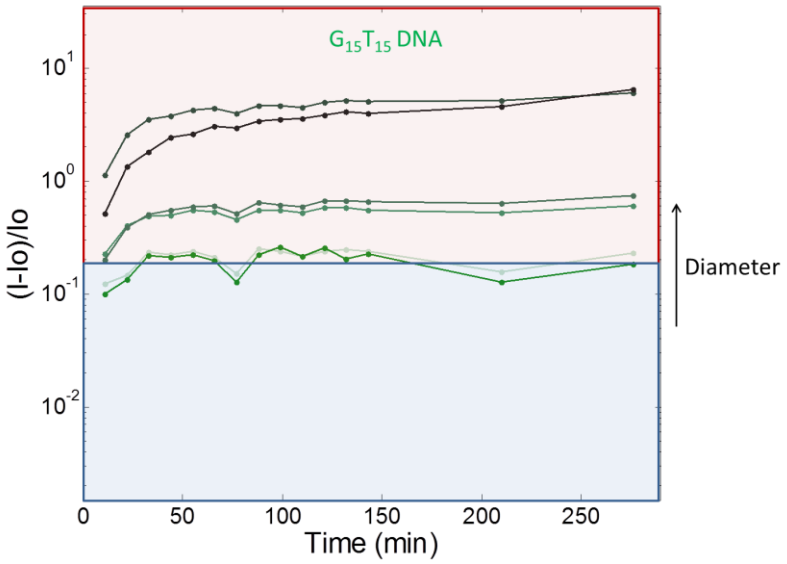

Figure S2: Fluorescence emission spectra of $G_{15}T_{15}$ DNA – SWCNT over the course of 300 minutes, 785 nm 200 mW excitation source, and 10-second acquisition time. Each peak corresponds to the emission of a particular SWCNT chirality from a multi-chirality SWCNT sample. The $G_{15}T_{15}$ DNA sequence is known to have a strong tendency to form a secondary structure. Sequences of more than 4 guanine bases can form g-quadruplex structures, therefore a $G_{15}T_{15}$ sequence is expected to vascillate between tight "DNA-like" binding to the SWCNT, and "RNA-like" binding to itself via G-quadruplex formation. As expected, the results of this equilibration experiment for $G_{15}T_{15}$ DNA-SWCNT show a much more "RNA-like" behavior of this polymer on SWCNT. This can be explained by the increased tendency of strings of guanine bases to form g-quadruplexes off the SWCNT surface.

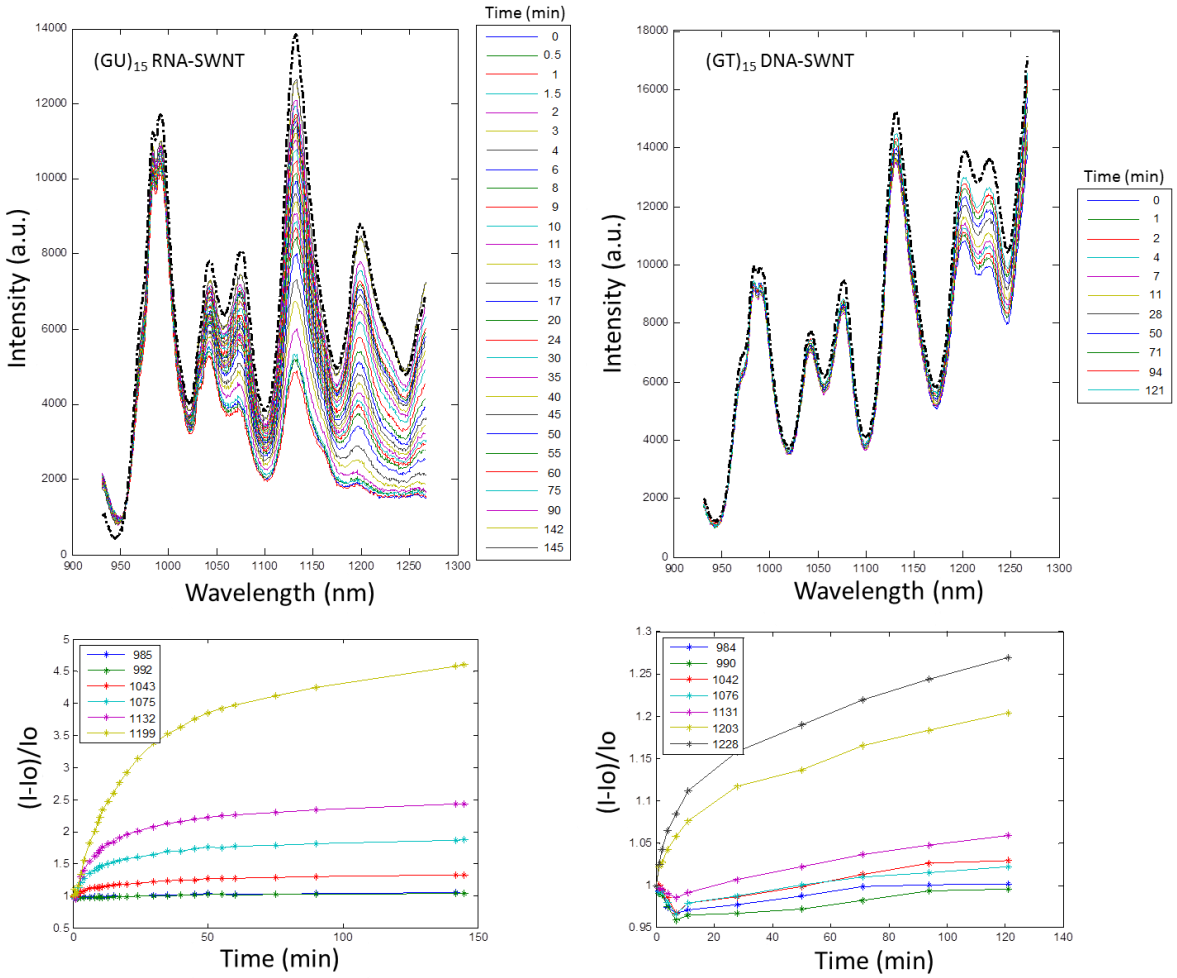

Figure S3: Fluorescence emission of $(GT)_{15}$ DNA – SWCNT and $(GU)_{15}$ RNA – SWCNT Each sample was diluted to 2 mg/L in PBS under constant laser irradiation over 145 and 121 minutes (solid lines), compared to a reference of each sample that was not laser irradiated until the final timepoint (dotted black line). Tracking the peak fluorescence over time shows that larger diameter tubes have larger fluorescence enhancement, and that this effect is independent of laser irradiation. Samples were irradiated with a 785 nm laser, where each spectrum is taken over a 20 sec exposure.

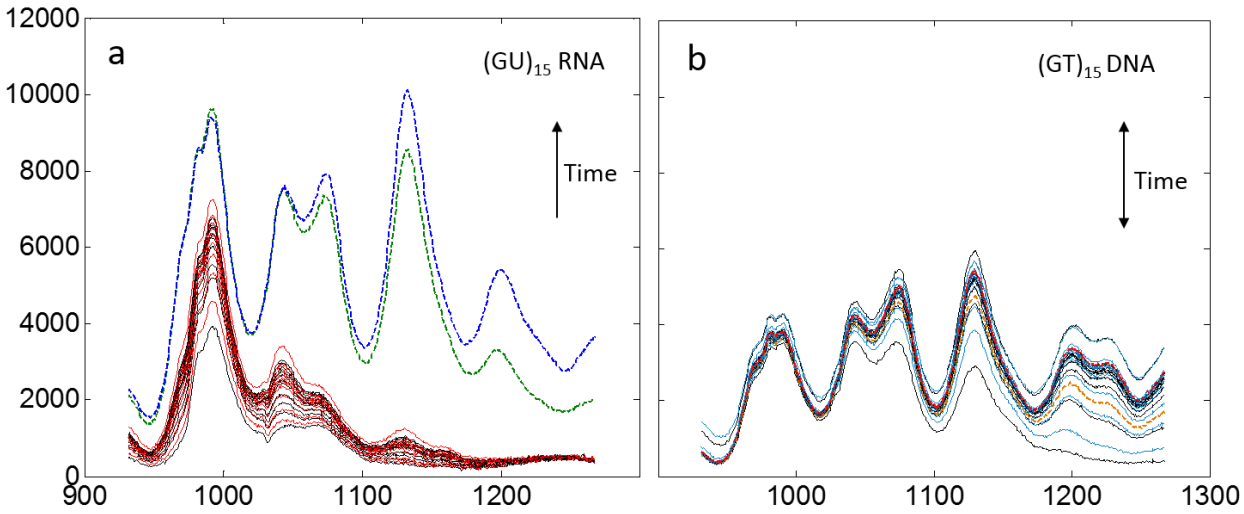

Figure S4: 2-day fluorescence modulation of $(GT)_{15}$ DNA – SWCNT and $(GU)_{15}$ RNA – SWCNT. The fluorescence emission of a 5 mg/L SWCNT of (a) $(GU)_{15}$ RNA – SWCNT was monitored at 5- minute intervals (alternating red and black traces), at 24 hours (dashed green trace), and at 48 hours (dashed blue trace), and a (b) $(GT)_{15}$ DNA – SWCNT sample was monitored at 5 minute intervals (alternating blue and black traces), at 24 hours (dashed red trace), and at 48 hours (dashed orange trace). The fluorescence of the $(GU)_{15}$ RNA – SWCNT sample shows a monotonic increase as a function of time, whereas the $(GT)_{15}$ DNA – SWCNT sample fluctuates in its fluorescence modulation over the same time course.

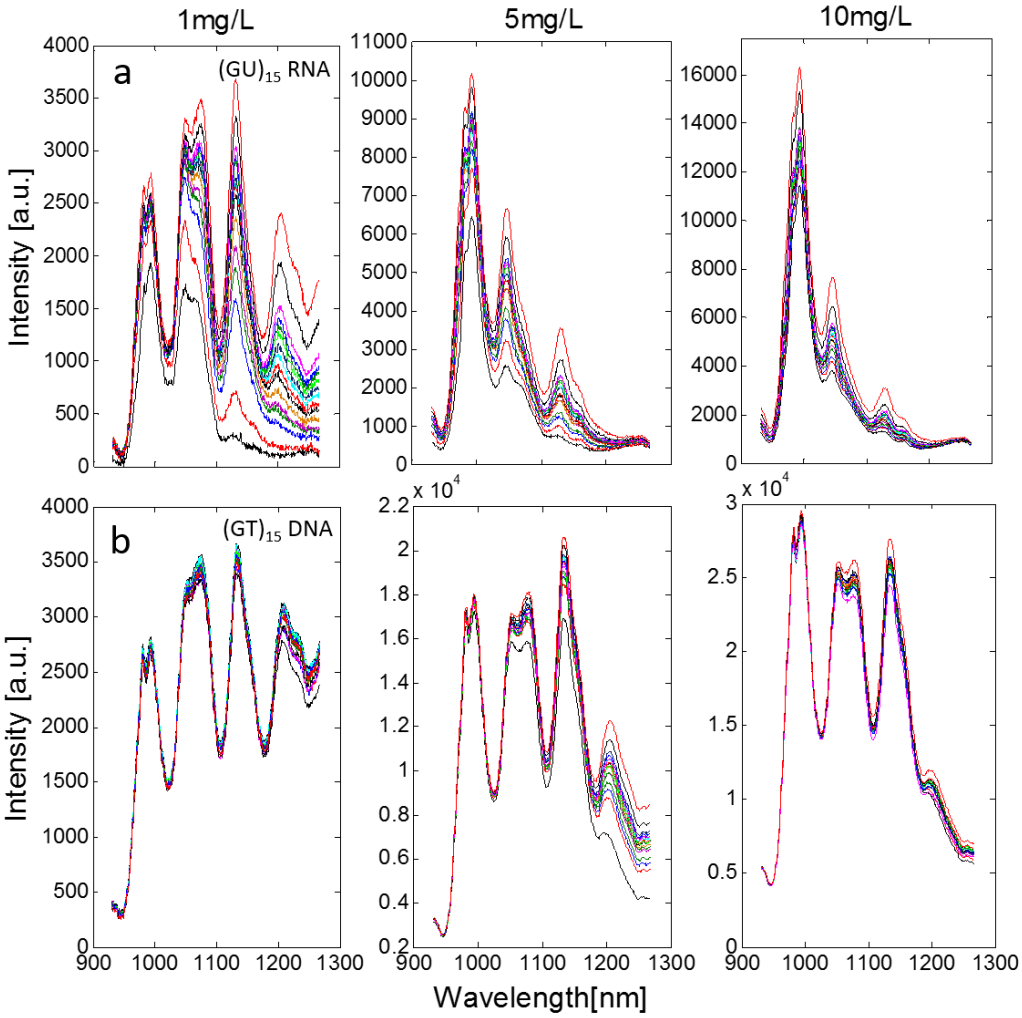

Figure S5: Time-dependent fluorescence modulation of $(GU)_{15}$ RNA – SWCNT and $(GT)_{15}$ DNA – SWCNT is also concentration dependent. The fluorescence intensity of $(GU)_{15}$ RNA – SWCNT and $(GT)_{15}$ DNA – SWCNT solutions were monitored in 11-minute intervals for the course of 143 minutes, 210 minutes, and again at 276 minutes for 1mg/L SWCNT, 5 mg/L SWCNT, and 10 mg/L SWCNT, respectively. Results show that the fluorescence of (a) $(GU)_{15}$ RNA –SWCNT solutions are more susceptible to modulation at lower SWCNT concentrations than higher SWCNT concentrations over the course of the 4.6 hours. (b) A $(GT)_{15}$ DNA – SWCNT sample does not show a noticeable susceptibility to SWCNT concentration over the same time-course.

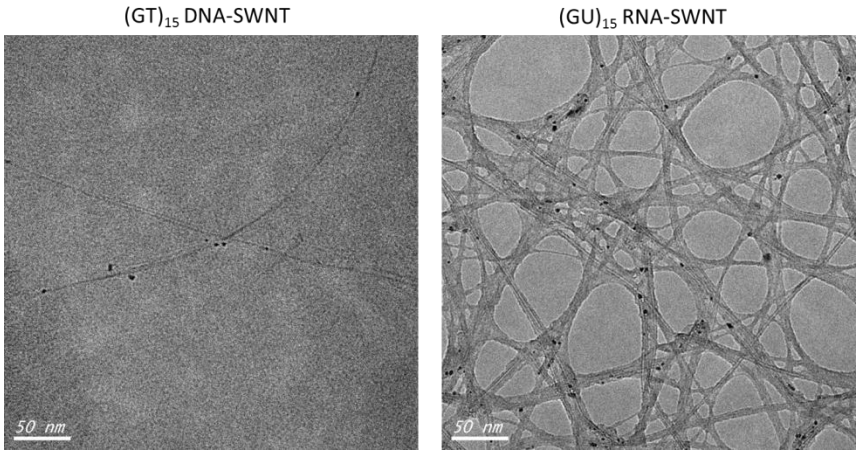

Figure S6: Representative transmission electron microscopy (TEM) of $(GT)_{15}$ DNA-SWCNT and $(GU)_{15}$ RNA-SWCNT. We observe primarily singly-dispersed $(GT)_{15}$ DNA-SWCNT and primarily parallel bundles of $(GU)_{15}$ RNA-SWCNT after depositing and drying equal concentrations of each sample. TEM images were acquired by a JEOL JEM 2010 high-resolution TEM with a LaB6 electron gun operated between 80 – 200 kV. For our TEM imaging experiments, 1 µl of 10 mg/L SWCNT solutions were dropped onto HC200-Cu Holey Carbon Film (Electron Microscopy Sciences), and dried in air for a few minutes prior to imaging.

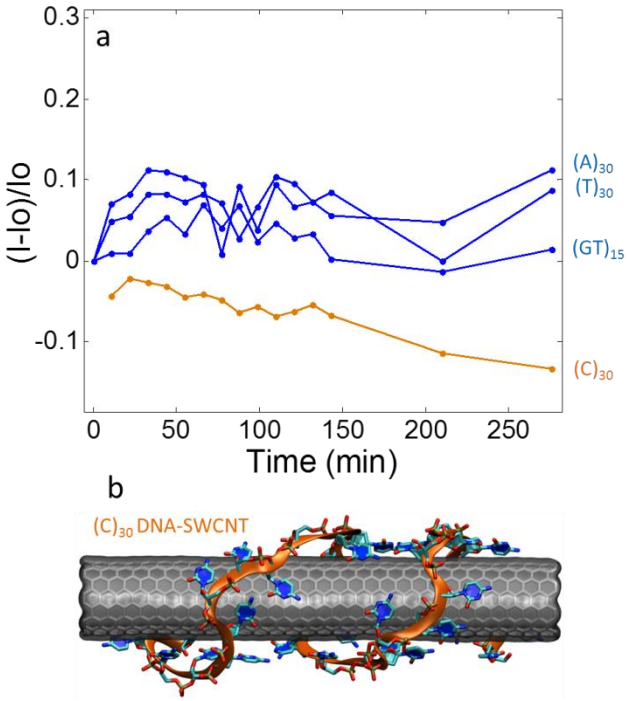

Figure S7: Time-dependent fluorescence for stable C_{30} DNA sequence on SWCNT (a) Emission peak for (9,4) chirality SWCNT, normalized to the initial sample intensity for $(C)_{30}$ DNA -SWCNT (orange) as compared to other DNA-SWCNT complexes (blue). (b) The $(C)_{30}$ DNA strand helically wraps SWCNT in MD simulations. A snapshot of $(C)_{30}$ DNA and SWCNT at the end of 250 ns MD simulations. SWCNT is shown in charcoal, nucleic acid strand is shown in licorice representations, and the phosphate backbone is highlighted in orange or gold; water and ions are not shown for clarity. Nucleic acid atoms are shown in teal (C), red (O), tan (P), and blue (N) colors.

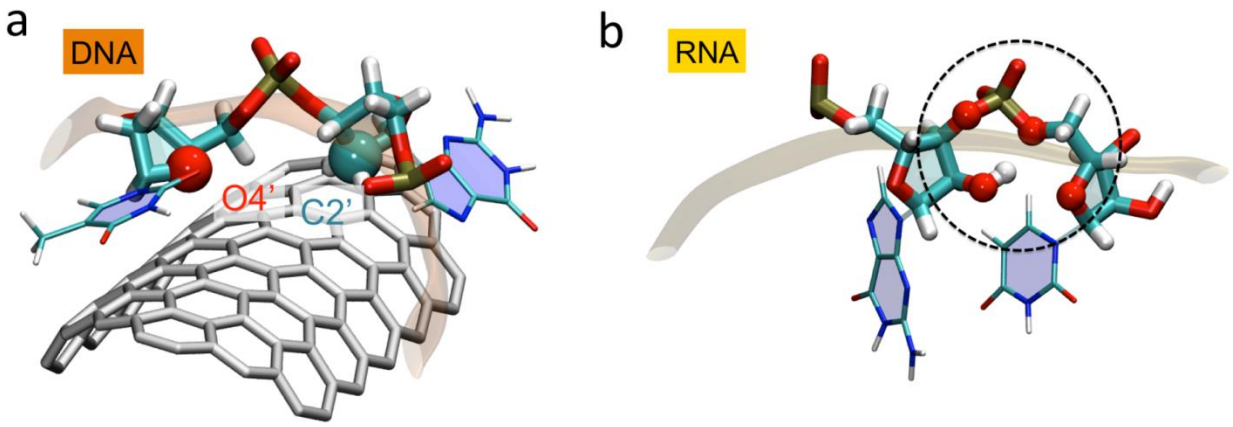

Figure S8. MD Simulations of sugar ring puckering for polynucleotide bases on SWCNT (a) A representative simulation snapshot of DNA backbone and sugar rings when DNA bases are stacking to the SWCNT. Either O4' or C2' group of the DNA nucleotide can point towards the SWCNT surface. (b) Simulation snapshot of two RNA bases stacking on each other in solution. 2' OH group of one base can form hydrogen bond to several nearby oxygen atoms of the neighboring nucleotide and further stabilize the stacked conformation. In (a-b), SWCNT atoms are shown in light grey, and nucleic acid atoms are shown in teal (C), red (O), tan (P), blue (N), and white (H) colors. Transparent orange and yellow lines show DNA and RNA backbones, respectively, in new cartoon representation.

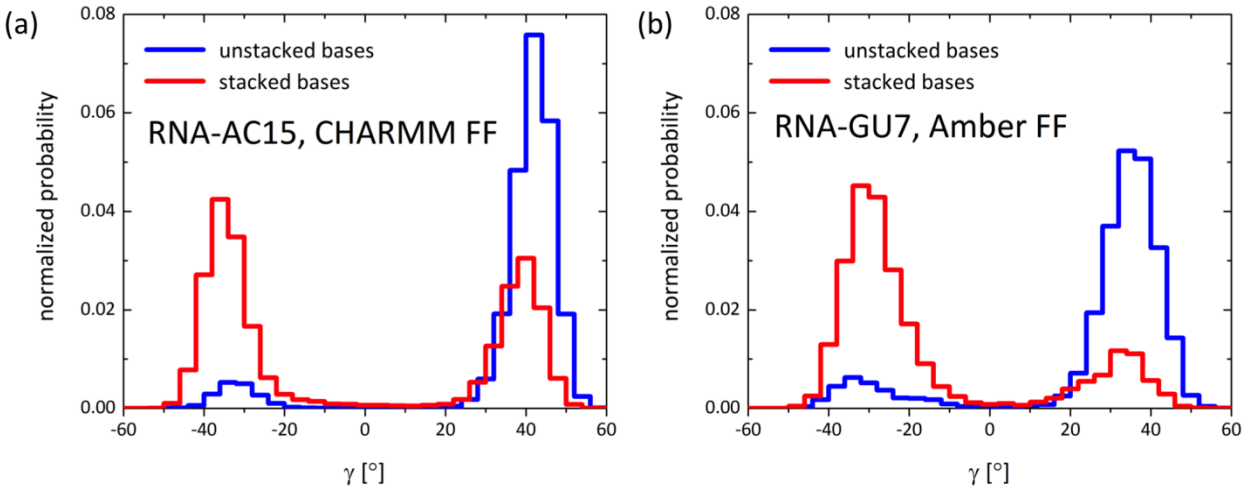

Figure S9: Sugar conformations in nucleic acid strands interacting with SWCNT (a) $(AC)_{15}$ RNA-SWCNT described with the CHARMM force field; (b) $(GU)_7$ RNA-SWCNT described with the Amber force field. Histogram distributions of ° (C4'-C3'-C2'-C1') are shown separately for bases that are stacked on SWCNT and bases that are not stacked on the SWCNT. The histograms distributions are obtained from complete trajectories of $(AC)_{15}$ RNA-SWCNT and $(GU)_7$ RNA-SWCNT systems. This suggests that this phenomenon is primarily due to the rigidity of the RNA backbone in lieu of the nucleotide bases.

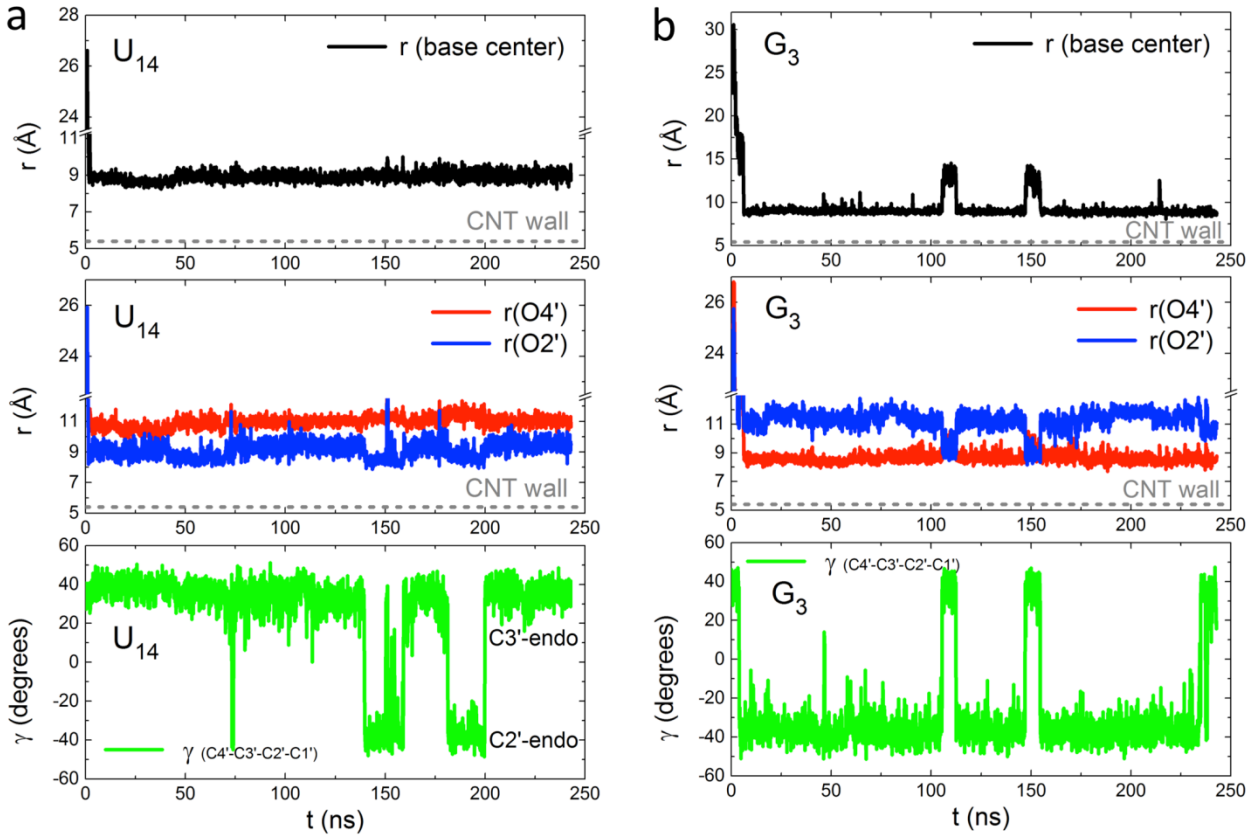

Figure S10. Conformational changes of sugar rings of $(GU)_{15}$ RNA upon nucleotide stacking on SWCNT. Time dependence of radial position of the base with respect to the SWCNT axis, time dependence of radial positions of oxygen atoms O4' and O2' of the nucleotide sugar ring, and time dependence of the dihedral angle γ for U14 (a) and G3 (b). The plots show representative examples of two possible binding modes of RNA nucleotide sugar rings and SWCNT. Radial position of a base was determined by evaluating the distance of base center of mass (COM of N1, N3, and C5 atom of U base, or N3, C6, and C8 atoms of G base) from the SWCNT axis. Nucleotides G3 and U14 were chosen because they remained stacked to SWCNT most of the time in MD simulations. O2' atom of U14 sugar ring stays close to the SWCNT surface, whereas O4' atom of G3 sugar ring stays close to the SWCNT surface.

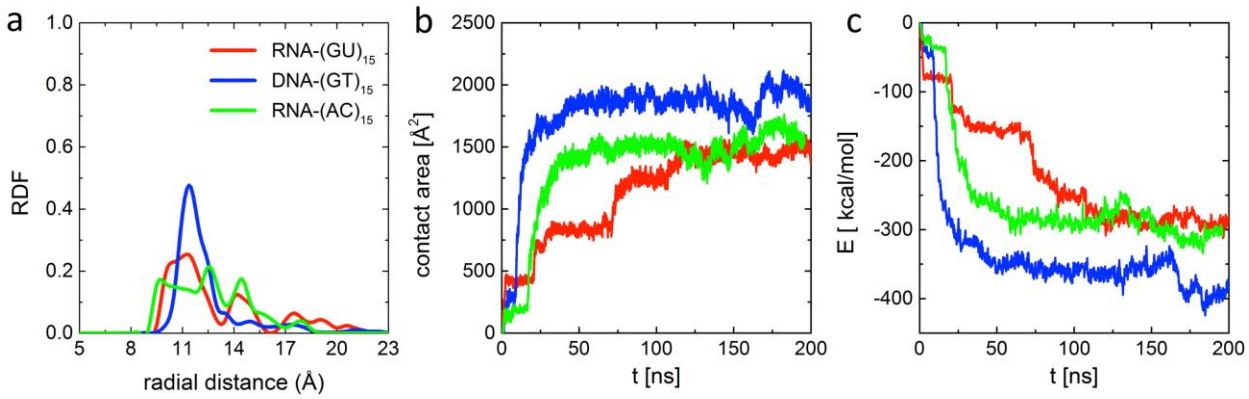

Figure S11. DNA and RNA binding to SWCNTs in a duplicate MD simulation. (a) Radial distribution of phosphate groups (P-atoms) of nucleic acid strands at distance r from the central axis of SWCNT, evaluated for the last 50 ns of simulations. While most of DNA remains close to the SWCNT surface (r < 14 Å), a significant fraction of RNA extends further away from the SWCNT surface (up to ~ 23 Å). (b) Contact areas between SWCNT and nucleic acid strands, evaluated every 20 ps. (c) Interaction energies between SWCNT and nucleic acid strands evaluated every 200 ps. The plots shown in (a-c) are representative of the subset of possible nucleic acid configurations, captured in performed MD simulations. At simulation ends, contact areas and interaction energies of the second independent simulation exhibit the same trend as in the first MD simulation (Fig. 5), with DNA strand having the largest contact area and the most favorable interaction energy with the SWCNT.

Model S1. Model for time-dependent polymer rearrangement on SWCNT.

The change in SWCNT fluorescence intensity correlates with the stacking of individual nucleotide bases along the surface of the SWCNT, as suggested in **Figure 6**. To model this system, we consider each SWCNT divided into θ segments, each representing a segment of the SWCNT that increases in luminescence with ordered DNA or RNA polymer wrapping, as shown previously [48].

The change in fluorescence increases with each base N that stacks onto the SWCNT surface.

$$\frac{(I-Io)}{Io} \propto [N\Theta] \quad (1)$$

Here, the total number of available SWCNT stacking sites can be determined by a simple sum of sites upon which nucleotides have stacked and free sites:

$$[\Theta]_{tot} = [N\Theta] + [\Theta] \quad (2)$$

The equilibrium constant K_A for the base stacking interaction reaction stated in Eqn. 1 can be represented by:

$$K_A = \frac{[N\Theta]}{[N][\Theta]} \quad (3)$$

At any given timepoint, the number of sites is a sum of available sites and sites occupied by nucleotide stacking as determined by the stacking rate:

$$[\Theta]_{tot} = [N\Theta] + \frac{[N\Theta]}{[N]\,K_A} \quad (4)$$

In turn, the stacking rate is determined by the accessibility of individual bases when they are in polymer form. We propose, based on our experimental results and our molecular dynamics

simulations that the following factors contribute to the propensity of base-stacking to SWCNT. First, the persistence length of the polymer will contribute to the ease of nucleotide stacking. Previous studies show that the ratio of single-stranded DNA/RNA persistence lengths at our experimental conditions of 50 mM NaCl is approximately 0.75 [58]. Therefore, we expect that DNA nucleotides will stack onto the SWCNT more readily than the RNA nucleotides constrained into a stiffer polymer. Second, the formation of off-SWCNT secondary structures can inhibit base stacking. Our molecular dynamics simulations show a higher propensity of off-SWCNT secondary structure formation for RNA over DNA; a smaller fraction of RNA polymer is found to stack to SWCNT, as shown in **Figure 5c-d,** where ssRNA forms larger folded regions in the solution than DNA, as highlighted in **Figure 5a**. Third, the radius of the SWCNT affects the ease of polymer stacking. As the persistence length of both DNA and RNA is longer than the diameter of all assayed SWCNT by at least two-fold [33,58], DNA and RNA polymers are expected to wrap larger SWCNT more easily than smaller SWCNT diameters, as we show experimentally in **Figure 1c**.

Therefore, the SWCNT diameter, d, and the polymer persistence length, Lp, are determinants in the ability of nucleotides to stack on the SWCNT. The SWCNT diameter and polymer persistence length are inversely proportional to the SWCNT intensity response, and both contribute a resistance to nucleotide stacking. We can define the turn-on response of the SWCNT due to nucleotide stacking as the cooperative binding of nucleotides to a multi-site SWCNT via a variant of the hill equation:

$$\frac{(I-Io)}{Io} = \frac{[\Theta]^n}{(K_A)+[N]^n} \quad (5)$$

$$\frac{(I-Io)}{Io} = \alpha \frac{[N\Theta]}{[\Theta_{tot}]} + \gamma = \alpha \frac{([N]K_A)^n}{([N]K_A)^n+1} + \gamma \quad (6)$$

Where (I-Io)/Io is the normalized change in SWCNT fluorescence, α is a factor that combines our three effects representing nucleotide resistance to stacking on the SWCNT, and ɣ is a background constant.

We model the base-stacking kinetics with an undefined reaction order, since the stacking of an individual nucleotide is affected by its neighboring bases, or by secondary structure formation in the case of bases beyond the polymer's persistence length. As our experimental results show no net desorption of whole polymers from the SWCNT, we estimate that polynucleotides rarely desorb from the SWCNT once adsorbed:

$$N + \Theta \rightarrow N\Theta \quad (7)$$

$$rate = \frac{d[N\Theta]}{dt} = \alpha[N\Theta]^n + \gamma \quad (8)$$

It is assumed that the SWCNT fluorescence increase is instantaneous upon base stacking.

$$\frac{(I-Io)}{Io} = \alpha[N\Theta]^n + \gamma \quad (9)$$

We then use this model to fit our data for the 7,5 chirality SWCNT for both $(GT)_{15}$ DNA-SWCNT and $(GU)_{15}$ RNA-SWCNT (**Figure 1c, dotted lines**). As expected, for $(GT)_{15}$ DNA-SWCNT, the contribution of α, the resistance to nucleotide stacking, is negligible with $\alpha = 1.6 \times 10^{-6}$. The fit

yields a cooperativity of n = 2.2, and a background constant of ɣ = 0.03. The positive cooperativity is reflective of the short DNA persistence length that enables adjacent nucleotides to wrap the SWCNT circumference without requiring neighboring bases to adopt energetically unfavorable conformations to stack on the SWCNT surface. In contrast, for $(GU)_{15}$ RNA-SWCNT, the contribution of the resistance to nucleotide stacking is influential to the SWCNT fluorescence response with α = 2.4. The fit yields a negative cooperativity of n = 0.3, and a background constant of ɣ = -1.6, suggesting that the longer RNA persistence length induces negative cooperativity amongst RNA nucleotides. Our model combines our experimental and theoretical results to help elucidate the time-dependent fluorescence modulation we observe amongst polynucleotide-SWCNT samples.

Model S2. Model of concentration-dependent relationship between the concentration, the phase of the DNA or RNA on the SWCNT corona, and the resulting intensity of the sample. We model the total intensity of our polynucleotide-SWCNT sample as a sum of the intensities provided by the stacking of DNA or RNA bases on the SWCNT surface. These entities are represented as species A or B, where A is a nucleotide tightly bound to the SWCNT surface, and B is a nucleotide that is loosely bound to the SWCNT surface, in equilibrium with water. Each polynucleotide species, A or B can be associated to a SWCNT fluorescent site θ on the SWCNT surface, where A produces a strong fluorescent response, and B produces a weak fluorescent response due to time sharing its position on θ with water. When water occupies the SWCNT, it quenches that fluorescent site into $[\theta]_u$.

$$I_{tot} = [A\theta] + (B + [\theta]_u) + [\mathrm{B}\theta] \tag{10}$$

The data show three clear regimes; therefore we adapted our model to reflect two reversible equilibria between species at the interface of three intensity regimes (**Figure 3a, green, white, purple**). The association, and therefore the normalized total SWCNT fluorescence intensity $\frac{I_{tot}}{\mathrm{I_o}}$, is strong between A and θ below 10^{-1} mg/L SWCNT, is low between B and θ between 10^{-1} mg/L and 10^{0} mg/L SWCNT, and recovers for the complexation of B and θ above 10^{0} mg/L SWCNT.

$$\frac{I_{tot}}{\mathrm{I_o}} = [A\theta] \rightleftharpoons B + [\theta]_u \qquad \text{and} \qquad \frac{I_{tot}}{\mathrm{I_o}} = B + [\theta]_u \rightleftharpoons [B\theta] \tag{11}$$

$$K_1 = \frac{B+[\theta]_u}{[A\theta]} \tag{12}$$

$$K_2 = \frac{[B\theta]}{[B]+[\theta]_u} \tag{13}$$

For the high SWCNT-concentration regime, species [Bθ] increases in number and produces detectable fluorescence above 10^{0} mg/L SWCNT. Here, the total number of available SWCNT surface sites can either be fluorescent and occupied by a stacked nucleotide, [Bθ], or non-fluorescent and occupied by water via an un-stacked nucleotide $\mathrm{B} + [\theta]_\mathrm{u}$. The equilibrium constant for this regime can be represented as:

$$[\theta]_{tot} = [B\theta] + B + [\theta]_u = [B\theta]\frac{[B\theta]}{[B][K_2]} + [\theta]_u = [B\theta]\left(1 + \frac{1}{[B][K_2]}\right) + \mathrm{B} + [\theta]_u \tag{14}$$

Once again, $[\theta]_u$ does not contribute to the fluorescence of the SWCNT, therefore, the normalized intensity is proportional to the ratio of properly stacked bases to unstacked bases $I_B \propto [B\theta]/B$, which is directly proportional to the concentration of SWCNT.

$$\frac{I_B}{I_o} = \alpha \frac{[B\theta]}{[B]} = \alpha \frac{([B\theta]K_2)^n}{([B]K_2)^n+1} + c \quad (15)$$

For DNA, fitting this model to our data (R^2 = 0.991) yields no cooperativity between DNA nucleotides with n = -0.006. For RNA, fitting this model to our data (R^2 = 0.997) yields a proportionality factor of $\alpha = -1.68e^4$, with a negative cooperativity between RNA nucleotides of n = -0.38. The higher cooperativity between RNA bases is expected due to the increased rigidity of the RNA backbone, and its interference with forming a properly stacked structure on the SWCNT surface, whereas the relationship for DNA as a function of SWCNT concentration is linear in this regime.

Likewise, for the low SWCNT-concentration regime, $A[\theta] \rightleftharpoons B + [\theta]_u$, species A[θ] predominates, where the total number of available SWCNT surface sites can either be highly fluorescent and occupied by a stacked nucleotide, A[θ], or lowly-fluorescent and occupied interchangeably by a loosely-stacked nucleotide or water, $B + [\theta]_u$. The equilibrium constant for this regime can be represented as:

$$[\theta]_{tot} = [A\theta] + B + [\theta]_u = \frac{[A\theta]K_1}{[A]} + B + [\theta]_u = [A\theta]\left(1 + \frac{K_1}{A}\right) + B + [\theta]_u \quad (16)$$

$[\theta]_u$ does not contribute to the normalized fluorescence of the SWCNT, therefore, the normalized intensity is proportional to the ratio of properly stacked bases to unstacked bases $\frac{I_A}{I_o} \propto B/[A\theta]$, which is inversely proportional to concentration of SWCNT.

$$I_A = \alpha \frac{[B]}{[A\theta]} = \alpha \frac{([B]K_1)^n+1}{([A\theta]K_1)^n} + c \quad (17)$$

For DNA, fitting this model to our data ($R^2 = 0.683$) yields negative cooperativity between DNA nucleotides of n = 0.758. For RNA, fitting this model to our data ($R^2 = 0.705$) shows negative cooperativity between RNA nucleotides with n = 0.784. The higher cooperativity between RNA bases is expected due to the increased rigidity of the RNA backbone, and its interference with forming a properly stacked structure on the SWCNT surface.

The convolution of normalized intensity functions I_A and I_B provides the description of the SWCNT fluorescence for the entire range of SWCNT concentration domain, τ:

$$(I_A * I_B)(c) = \int_{10^{-3}}^{10^1} I_A(\tau)\, I_B(c-\tau) d\tau = \alpha \left\{ \frac{([B]K_1)^n+1}{([A\theta]K_1)^n} + \frac{([B\theta]K_2)^n}{([B]K_2)^n+1} \right\} + c \quad (18)$$

These fits for DNA and RNA are plotted in **Figure 3a** of the main manuscript as grey dotted lines.