

SUPPLEMENTARY INFORMATION: DNA—Au (111) Interactions and Transverse Charge Transport Properties for DNA-Based Electronic Devices

Busra Demir^{1,2,3,‡}, Hashem Mohammad^{4,‡}, M. P. Anantram³ and Ersin Emre Oren^{1,2,*}

¹Department of Materials Science & Nanotechnology Engineering, TOBB University of Economics and Technology, Ankara, Turkiye

²Bionanodesign Laboratory, Department of Biomedical Engineering, TOBB University of Economics and Technology, Ankara, Turkiye

³Department of Electrical and Computer Engineering, University of Washington, 98195 Seattle, WA, USA,

⁴Department of Electrical Engineering, Kuwait University, P.O. Box 5969, Safat 13060, Kuwait.

[‡]these authors contributed equally

^{*}corresponding author

METHODS

First, we use MD simulations and clustering methods to obtain representative structures, then we use DFT calculations to generate the system Hamiltonian that contains molecular orbital energy level information and their coupling between one another. In the third step, we use the Hamiltonian with the Green's function method to calculate the transverse transmission.

Schematic representations of the modeling steps are shown in Figure S1.

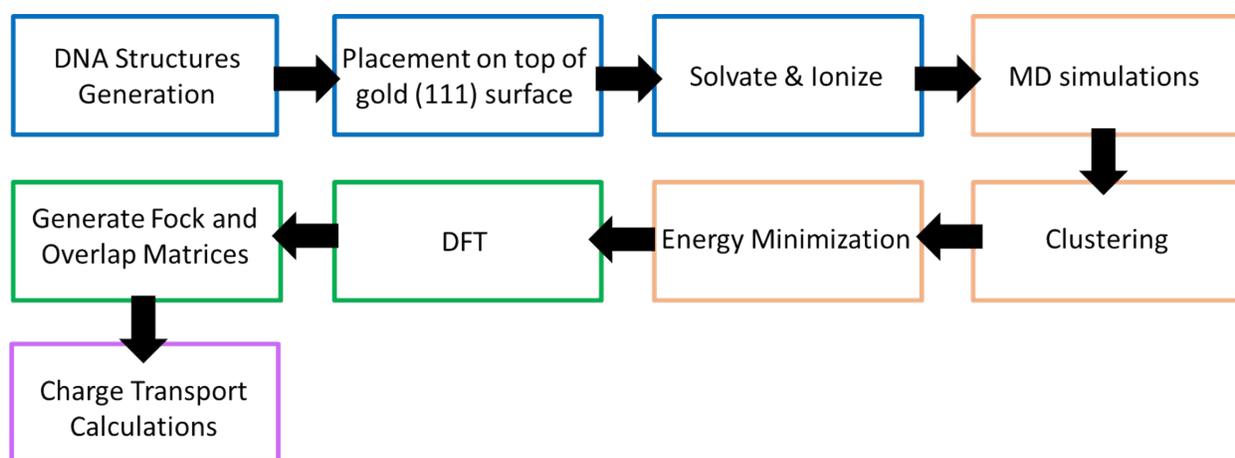


Figure S1: Modeling steps used in this study.

1. Structure preparation and MD Simulations

We used NAB Tool implemented in AMBER^{S1} and VMD's^{S2} Inorganic Builder plugin to generate the DNA structures and Au (111) surface respectively. Then, we placed the DNA structures 3.5 Å away from the Au (111) surface and we solvated the system with TIP3P^{S3} water molecules and 0.15 M KCl. CHARMM36^{S4} and INTERFACE^{S5} force fields are used for the DNA and gold substrate respectively. We used 12 Å cutoff to calculate Van der Waals potential energies and the particle-mesh Ewald (PME) method with a maximum grid spacing of 1.5 Å to compute electrostatic interactions. We use Langevin dynamics, and the simulation time-step is set to 1 fs. For all simulation steps, Au atoms were kept fixed. We first minimize the water molecules and the ions for 2000 steps at 295 K while the entire DNA molecule was kept fixed in a constant volume. After the minimization, we first let extended poly-A parts fluctuate for 2 ns while keeping the ds-DNA part of the molecule fixed. Then the entire system is equilibrated for 1 ns while only the gold substrate is fixed. Finally, the production simulations were performed for 50 ns using the Langevin piston Nose-Hoover method implemented in NAMD^{S6} to maintain atmospheric pressure. The MD simulation analysis is carried out with pytraj library, VMD plugins and tcl scripting.

The representative structures from the MD simulation were determined by a RMSD-based clustering algorithm within VMD software with 1.5 Å cutoff value as mentioned in the main manuscript. Representative structures were subjected to 5000 steps of energy minimization prior to DFT calculations.

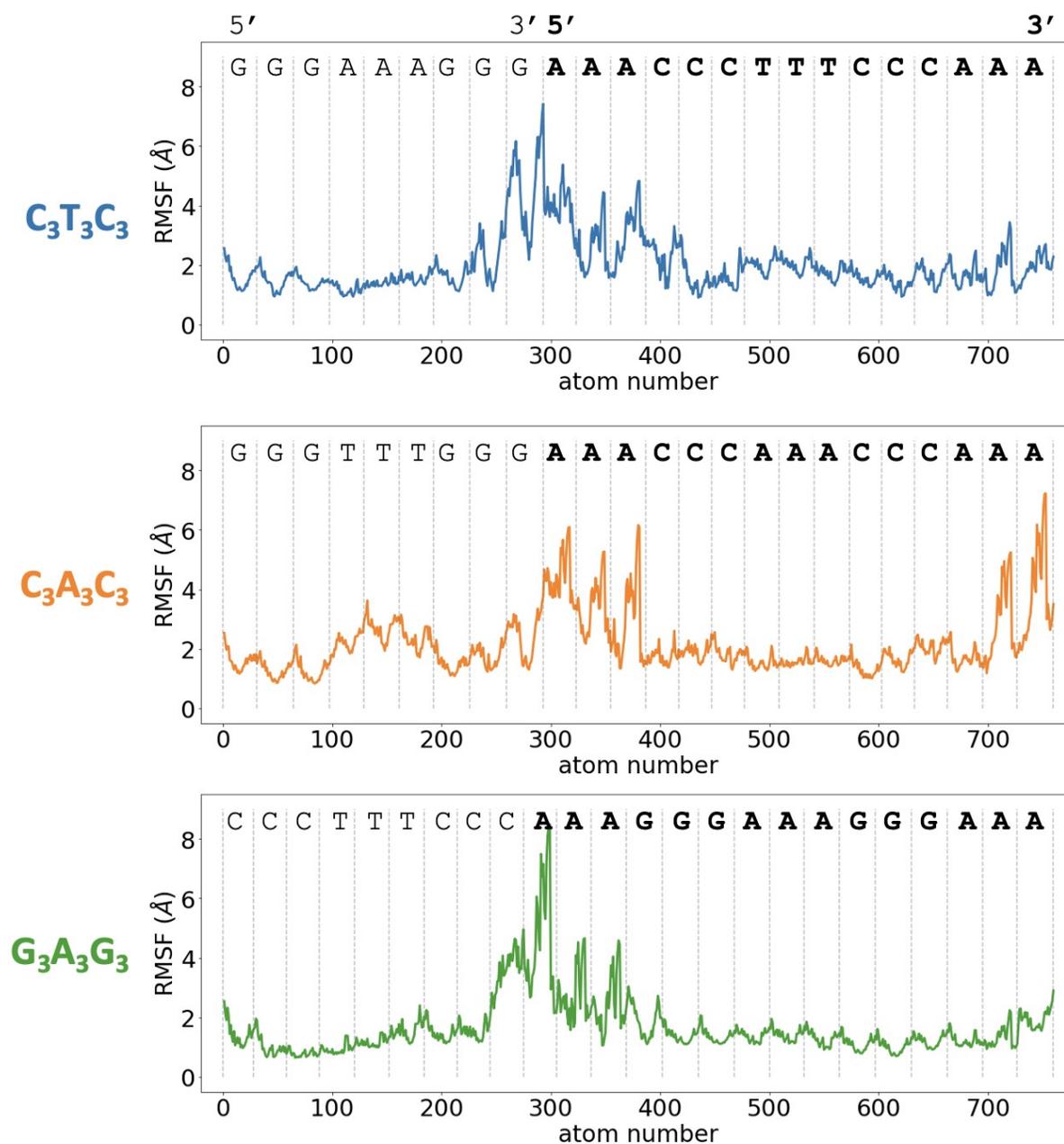


Figure S2: Plots of RMSF vs atom number for each sequence. The graph indicates the variation of RMSF for each individual atom in the molecule, denoted by its atom number. The corresponding nucleobases are indicated above each graph. The higher RMSF values indicate greater flexibility and lower RMSF values indicate greater stability.

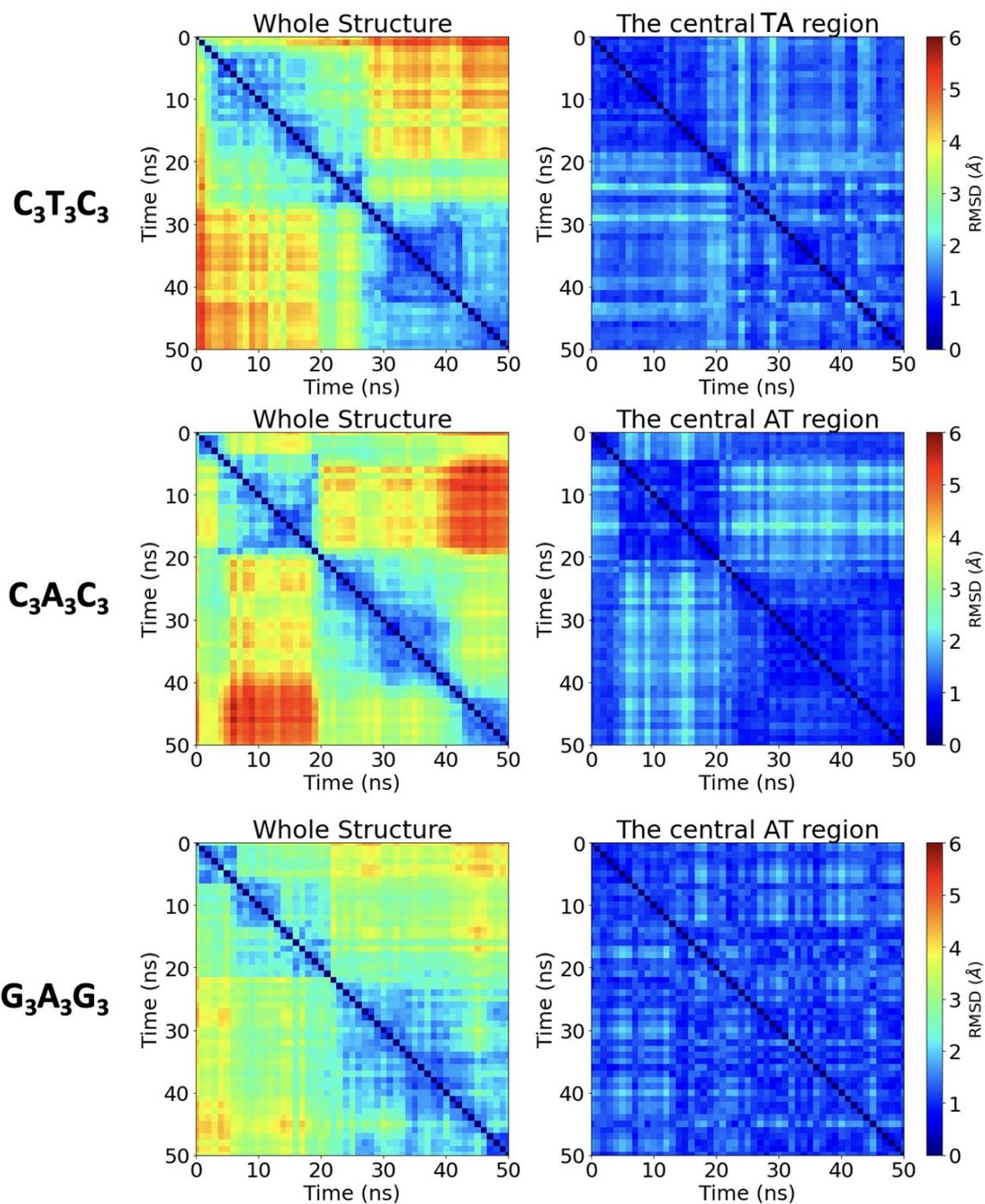


Figure S3: Pairwise RMSD analysis for the whole structure and the central triplet region. The graph indicates the variation of RMSD between every conformation saved at 1 ns time interval. Higher values of RMSD indicate a greater structural change between the two conformations.

Table S1: Cluster sizes and their percentages

	CTC		CAC		GAG	
	# of structures	percentage	# of structures	percentage	# of structures	percentage
C1	22592	45.183%	15033	30.065%	23244	46.487%
C2	13179	26.357%	13281	26.561%	8478	16.956%
C3	8613	17.226%	7706	15.412%	5256	10.512%
C4	1552	3.104%	4442	8.884%	4737	9.474%
C5	1086	2.172%	4182	8.364%	2769	5.538%
C6	1034	2.068%	2577	5.154%	2070	4.140%
C7	501	1.002%	905	1.810%	1253	2.506%
C8	393	0.786%	801	1.602%	707	1.414%
C9	374	0.748%	390	0.780%	585	1.170%
C10	234	0.468%	99	0.198%	243	0.486%
Unclustered	443	0.886%	585	1.170%	659	1.318%

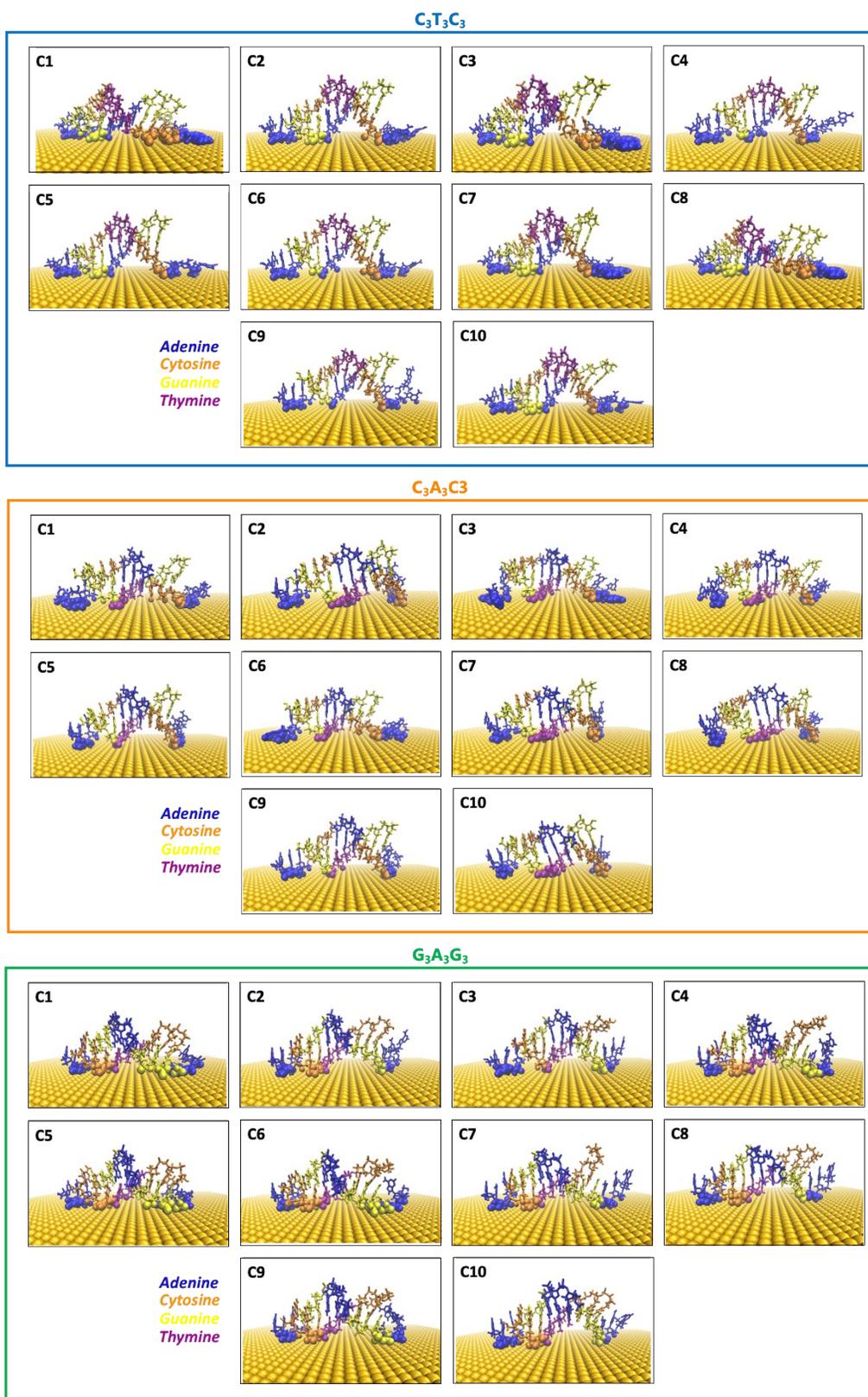


Figure S4: Representative structures for all sequences, surface interacting atoms showed as a ball shape. Blue represents Adenines, orange represents cytosine, yellow represents Guanines and pink represents Thymines.

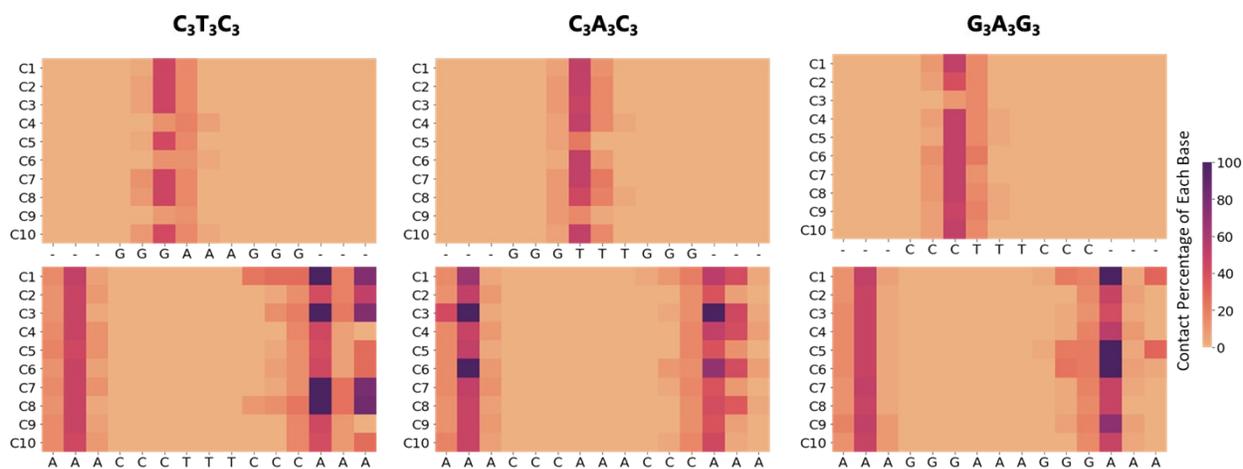


Figure S5: Heatmap of the percentage of gold contacting atoms for each representative conformation in CTC, CAC, and GAG cases. The colour gradient represents the percentage of contact atoms in each DNA nucleobase in separate strands, with purple indicating 100% and orange indicating 0% percentage. Each row corresponds to the representative conformation selected from MD simulations, and each column represents residues.

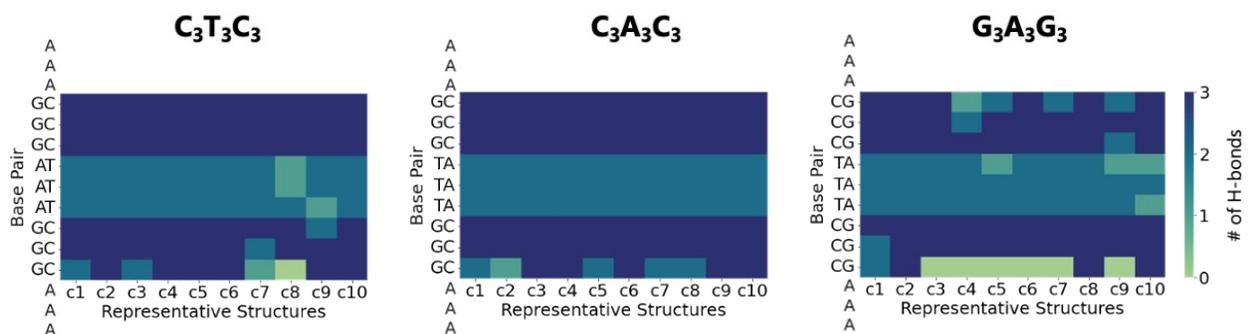


Figure S6: Number of hydrogen bonds for the representative structures of each sequence which corresponds to center of each cluster as described in the main text.

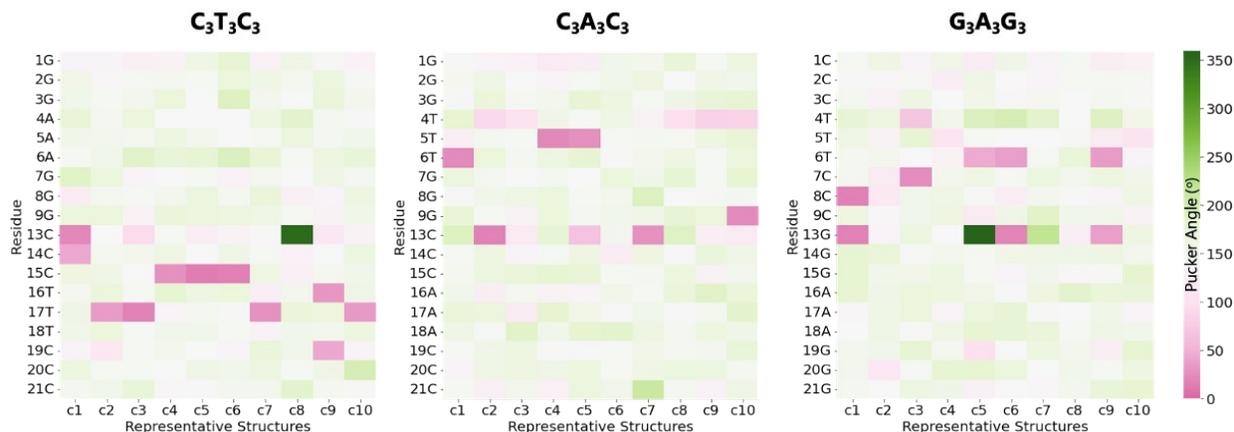


Figure S7: Pucker angle analysis for all residues in each cluster for every sequence. The graph indicates the variation of sugar pucker angle. Ideally, the values between 0 and 36° correspond to A-form DNA and 144 and 180° correspond to B-form DNA.

2. DFT Calculations

We removed the water molecules, ions, and the gold substrate from the system to enable the convergence of DFT calculations, which were carried out using Gaussian 16^{S7} with B3LYP/6-31G(d,p) basis set. We used the polarizable continuum model (PCM) with a dielectric constant of 78.36 to represent the water solvent effect. The total charge of the system was set to the number of phosphate groups in the DNA molecule, which was -22 (the terminal bases do not have any phosphate groups). After achieving convergence, the Fock and Overlap matrices from the DFT calculations were used for the next step.

3. Charge Transport Calculations

The quantum transport calculations were carried out using the Green's function method with decoherence probes added to model phase decoherence^{S8}. The probe current at each energy were kept zero meaning that the electron does not gain or lose energy while traversing the system. As mentioned in the main manuscript, we assume the contact locations to be at the contacting atoms which are 5 Å away from the surface and the central triplet's backbone atoms. The contact self-energy to model both Au (111) substrate and top contact was set to 600 meV to represent strong coupling. The decoherence scattering rate was set to 100 meV to mimic large decoherence, with an energy decaying factor of 50 meV, limiting the decoherence effect on the onsite potentials (or energy levels) of the nucleotides.

After obtaining the Fock, H_0 , and overlap matrices, S_0 from DFT calculations, we used the Löwdin transformation to convert H_0 into a Hamiltonian, $H_{orthogonal}$, in an orthogonal basis set via the following equation:

$$H_{orthogonal} = S_0^{-\frac{1}{2}} H_0 S_0^{-\frac{1}{2}} \quad (1)$$

Here, the diagonal elements of $H_{orthogonal}$ represent the energy levels at each atomic orbital, and the off-diagonal elements correspond to the coupling between the different atomic orbitals. Then, we partitioned the $H_{orthogonal}$ into nucleotides (except the contact locations) and diagonalized the $H_{orthogonal}$ using the following transformation.

$$H = U^\dagger H_{orthogonal} U \quad (2)$$

Here, each block of H is a nucleotide and the diagonal blocks of H are now diagonal matrices. The diagonal elements of the diagonal blocks represent the eigenvalues of the corresponding nucleotide. The off-diagonal blocks of H represent the hopping parameters between the molecular orbitals of the equivalent nucleotides.

The transverse transmission along the molecule was then calculated using Green's function method. The retarded Green's function (G^r) was found by solving the following equation:

$$[E - (H + \Sigma_L + \Sigma_R + \Sigma_B)]G^r = I \quad (3)$$

Where E is the energy level, and H is the Hamiltonian defined in Eq. 2. $\Sigma_{L(R)}$ is the left (right) contact self-energy, which represents the coupling strength of the DNA to the bottom (up) contacts by which charge enters and leaves the DNA. The self-energy of the decoherence probe is defined as Σ_B , which also represents the coupling strength between the DNA and the decoherence probes.

The self-energy of the contacts is defined as $\Sigma_{L(R)} = -\frac{i}{2}\Gamma_{L(R)}$, where i is the imaginary unit. The decoherence probe self-energy is defined as $\Sigma_B(E) = -\frac{\Gamma_k(E)}{2}$, where k represents the k^{th} energy level, and Γ_k represents the coupling strength between the probe and the energy level k , which is taken as an energy-dependent parameter as follows:

$$\Gamma_k(E) = \Gamma_B \times \exp\left[-\frac{|E - \epsilon_k|}{\lambda}\right] \quad (4)$$

where Γ_B determines the value of the decoherence strength, and λ is a decay parameter that determines how quickly the decoherence decays away from an energy level.

The decoherence probes were attached to each nucleotide (backbone + base) excluding the contact atoms (top and bottom), where the total number of decoherence probes is 22 (24 nucleotide – 2 contact groups) in the low-bias region, the current at the i^{th} probe calculated with:

$$I_i = \frac{2e}{h} \sum_{j=1}^N T_{ij}(\mu_i - \mu_j), \quad i = 1, 2, 3, \dots, N \quad (5)$$

where $T_{ij} = \Gamma_i G^r \Gamma_j G^a$ is the transmission probability between the i^{th} and j^{th} probes, and $T_{ij} = G^a = (G^r)^\dagger$ is the advanced Green's function. The net current at each decoherence probe

should be zero, this yields N_b independent equations from which the following relation can be derived,

$$\mu_i - \mu_L = \left(\sum_{j=1}^{N_b} W_{ij}^{-1} T_{jR} \right) (\mu_R - \mu_L), \quad i = 1, 2, 3, \dots, N_b \quad (6)$$

Here, W_{ij}^{-1} is the inverse of $W_{ij} = (1 - R_{ii})\delta_{ij} - T_{ij}(1 - \delta_{ij})$, where R_{ii} is the reflection probability at

probe i , and is given by $R_{ii} = 1 - \sum_{i \neq j} T_{ij}$. The currents at the top I_L and bottom I_R contacts are not zero because they are governed by the conservation of electron number, $I_L + I_R = 0$. This yields the equation for the current at the left contact as

$$I_L = \frac{2e}{h} T_{eff} (\mu_L - \mu_R) \quad (7)$$

Comparing equations 5 to 7 yields an effective transmission term:

$$T_{eff} = T_{LR} + \sum_{i=1}^{N_b} \sum_{j=1}^{N_b} T_{Li} W_{ij}^{-1} T_{jR} \quad (8)$$

In Eq 8, T_{LR} is the coherent transmission from the top electrode to the bottom electrode. The second term is the decoherence contribution to the transmission via decoherence probes. From Eq 7, the zero bias conductance can be approximated as $G = G_0 T_{eff}$, where the quantum of

conductance G_0 , can be calculated as $G_0 = \frac{2e^2}{h} \approx 7.75 \times 10^{-5} \Omega^{-1}$.

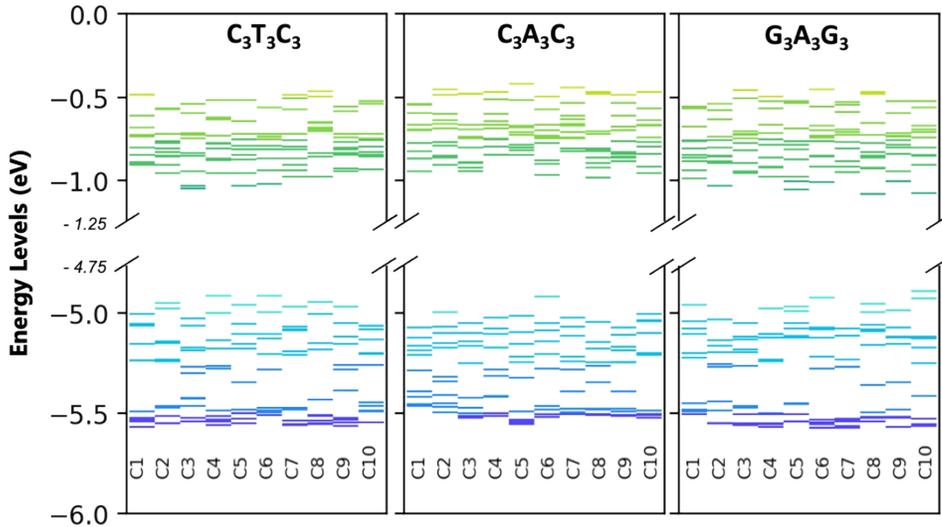


Figure S8: Energy band diagram showing only the first 10 orbitals from occupied and unoccupied states for representative structures of $C_3T_3C_3$, $C_3A_3C_3$, and $G_3A_3G_3$ cases.

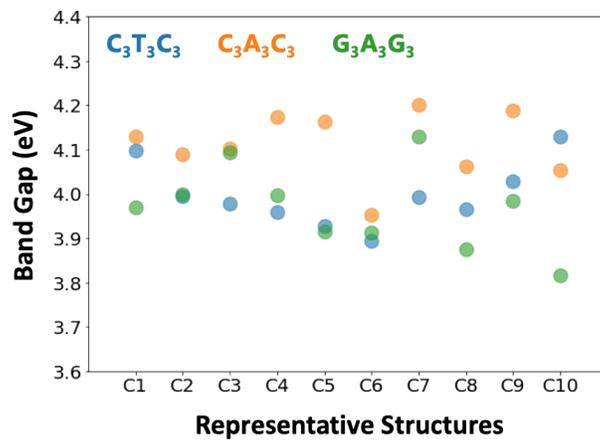


Figure S9: Band gap distribution for representative conformations of $C_3T_3C_3$, $C_3A_3C_3$, and $G_3A_3G_3$ cases.



Figure S10: Molecular orbitals (for iso value 0.02) for each conformation. Red & Green: HOMO; Blue & Orange: HOMO-1; Yellow & Navy: HOMO-2; Pink & Purple: HOMO-3.

Supplementary References

- S1 D. A. Case, R. M. Betz, D. S. Cerutti, C. I. T.E., T. A. Darden, R. E. Duke, T. J. Giese, H. Gohlke, A. W. Goetz, N. Homeyer, S. Izadi, P. Janowski, J. Kaus, A. Kovalenko, T. S. Lee, S. LeGrand, P. Li, C. Lin, T. Luchko, R. Luo, B. Madej, D. Mermelstein, K. M. Merz, G. Monard, H. Nguyen, H. T. Nguyen, I. Omelyan, A. Onufriev, D. R. Roe, A. Roitberg, C. Sagui, C. L. Simmerling, W. M. Botello-Smith, J. Swails, R. C. Walker, J. Wang, R. M. Wolf, X. Wu, L. Xiao and P. A. Kollman, *AMBER 2016, University of California, San Francisco*.
- S2 W. Humphrey, A. Dalke and K. Schulten, *J. Mol. Graph.*, 1996, **14**, 33–38.
- S3 W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey and M. L. Klein, *J. Chem. Phys.*, 1983, **79**, 926–935.
- S4 J. Huang and A. D. Mackerell, *J. Comput. Chem.*, 2013, **34**, 2135–2145.
- S5 H. Heinz, T. J. Lin, R. K. Mishra and F. S. Emami, *Langmuir*, 2013, **29**, 1754–1765.
- S6 J. C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R. D. Skeel, L. Kalé and K. Schulten, *J. Comput. Chem.*, 2005, **26**, 1781–1802.
- S7 D. J. Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Petersson, G. A.; Nakatsuji, H.; Li, X.; Caricato, M.; Marenich, A. V.; Bloino, J.; Janesko, B. G.; Gomperts, R.; Mennucci, B.; Hratch, 2016, *Gaussian, Inc., Wallingford CT*, 16.
- S8 J. Qi, N. Edirisinghe, M. G. Rabbani and M. P. Anantram, *Phys. Rev. B.*, 2013, **87**, 085404.