Biophysical Journal, Volume *110*

Supplemental Information

Graphene Symmetry Amplified by Designed Peptide Self-Assembly

Gina-Mirela Mustata, Yong Ho Kim, Jian Zhang, William F. DeGrado, Gevorg Grigoryan, and Meni Wanunu

Supplementary Materials and Methods

Assembly modelling. The assembly optimization framework was implemented using the Molecular Software Library(1) in conjunction with the EEF1.1 implicit-solvent force-field(2, 3). Solvation parameters in the force-field are derived from experimental measurements for model compounds(2), with electrostatic interactions adjusted based on potentials of mean force from explicit-solvent calculations(3, 4). Here and in all implicit-water calculations, graphene carbon atoms were modelled as electrostatically neutral and with well depth and van der Waals radius parameters borrowed from the aromatic atom carbon type "CA" in CHARMM parameter set 22. In stage-one calculations, graphene-bound poses were explored by sampling backbone φ and ψ angles, side-chain χ angles, elevation, and orientation relative to graphene. Given that the modelled sequences were periodic, we assumed the backbone geometry to be periodic as well—i.e., with repeating φ and ψ angles (except termini, see below). Thus, one φ-angle value and one ψ-angle value defined the entire backbone. To account for amino-acid backbone conformational preferences, φ/ψ-angle combinations were extracted from the Protein Coil Library(5) for each of the two amino acids comprising the modelled sequence. The two resulting φ/ψ-pair sets where then combined (with the larger randomly subsampled to match the size of the smaller) and the corresponding Ramachandran distribution was optimally covered with tiles of 5° x 5°. We found that 655 tiles were sufficient to cover 95% of φ/ψ observations for Lys/Phe and 579 tiles covered the same fraction for Lys/Val, and we used the centres of these tiles to restrict backbone conformations in sampling peptide-graphene poses.

For a given backbone conformation, peptide orientation relative to graphene was defined in terms of two angles: α controlled rotation of the peptide around its own "helical" axis (repeating φ/ψ angles always define a helix), and β set the angle between this axis and graphene's surface. For every backbone conformation considered, we varied α between 0 and 360° in 90 increments and β between -30° and 30° in 15 increments, reasoning that sharp angles of attack (i.e., ones where a peptide terminus is aimed at graphene and peptide side-chains have little opportunity for graphene contacts) would not result in productive assemblies. Thus, over 880,000 φ/ψ/α/β combinations were considered for $(KF)_4$ and over 780,000 for $(KV)_4$. For each of these combinations, elevation with respect to graphene was optimized. To this end, we performed a series focusing iterations over the initial range of elevations between 0 and 10 Å. At each iteration, the current range of elevations was sampled in 20 linear increments, with side-chains repacked at each increment in the EEF1.1 force-field using the energy-based rotamer library by Subramaniam and Senes(6). The range for the subsequent iteration was reduced by a factor of three (i.e., the focusing scale factor was 3) and centered around the increment with the best post-repacking EEF1.1 energy in the current iteration. Four focusing iterations were performed, such that final resolving power of the search was ~0.1 Å.

Graphene-bound peptide poses from stage one, in the order of ascending energy, were considered in stage-two calculations, where P1 and P2 lattice parameters were optimized (Figure 1). The P1 lattice was defined in terms of its two 2D translational symmetry vectors. So as to avoid optimization in the full four-dimensional parameter space, we first found the single best vector of translational symmetry, followed by a search for the optimal second symmetry vector in the context of the first. Specifically, given a graphene-bound peptide pose from stage one, we first sought the energetically optimal placement of a second peptide image, defining the principal symmetry axis. This search was performed with 2D focusing iterations to find the optimal displacements in the XY plane, Δx_1 and Δy_1. The initial range for the two parameters was set to

[0; 40] Å and [-20; 20] Å, respectively, and four focusing iterations were performed, with a focusing factor of 3. The first iteration partitioned the initial range for each parameter into 20 increments (for a total of 400 sampled Δx_1,Δy_1 combinations), while subsequent three iterations used 10 increments (100 sampled combinations). At each of the sampled points, side-chains were repacked in the presence of graphene via the same procedure as above, except that structural symmetry was imposed (i.e., equivalent sites of the symmetry-related peptides were restricted to have the same rotamer). Once the optimal displacement was found, five peptide images were generated on graphene along the corresponding symmetry axis, and an identical focusing procedure was used to find a second pair optimal displacements (and thus second symmetry axis), Δx_2 and Δy_2, in the context of these images. Finally, the two best symmetry axes found were used as input into a continuous minimization procedure via the Simplex algorithm of Nelder and Mead, seeking to minimize the total energy of a 5x5 peptide assembly fragment. The optimal assembly geometry resultant from this minimization was taken as the final best assembly geometry for the input graphene-bound pose, and the per-chain energy of the 5x5 assembly was used to find the best assembly.

In addition to the two translational symmetry vectors, the P2 lattice also had a center of rotational symmetry relating the two peptide images in the asymmetric (Figure 1). Thus, to avoid direct optimization in the full six-dimensional parameter space, the optimal rotational symmetry point (and thus asymmetric unit) was found first, after which the same procedure as above was used to find two translational symmetry vectors given the asymmetric unit. The rotational symmetry point was represented as a displacement in the XY-plane, Δx_r and Δy_r, relative to the center of mass of the first peptide image. Thus, optimization of this rotational center proceeded similarly as for symmetry vectors above, except that instead of a translation at each sampled point a rotation was applied about the corresponding point.

In all calculations, peptide N-termini were acetylated and C-termini amidated. Due to this loss of strict backbone periodicity, and because termini would be expected to have additional flexibility, during repacking terminal patches were treated as pseudo-sidechains with 12 rotamers corresponding to the terminal φ and ψ angles sampled in increments of 30°.

All implicit-solvent MD simulations were run using CHARMM 38b2(7), at 298.15 K, with the EEF1.1 force-field. Graphene atoms were fixed for the sake of efficiency. VALOCIDY calculations were performed as described earlier(8). Integration was carried out in the bond-angle-torsion coordinate system associated with the simulated peptide, while graphene atoms were fixed. Initial configurations for side-on and end-on peptide poses were taken from the modelling procedure above as the optimal single-peptide bound conformation and the conformation in the optimal assembly, respectively. For the purposes of free-energy calculations, thermodynamic states were defined around these two poses as ensembles of conformations with φ/ψ backbone angles within 30° of their starting values and distances between peptide Cα atoms and the graphene plane within 2 Å of their starting values. During MD, these states were sampled by restraining dihedral angles and Cα-to-graphene distances with flat-bottom potentials using the MMFP module in CHARMM(7). The free energy of each state was computed by averaging estimates from 100 independent simulations with 1-ns of sampling time after 100 ps of equilibration. The side-on state was found to be preferred by 15.4 kcal/mol, while the standard deviations of the 100 estimates were 1.6 kcal/mol and 1.7 kcal/mol for the side-on and end-on poses, respectively, demonstrating good convergence.

Multi-canonical MC for modeling of DNA orientational preferences. Multi-canonical Monte Carlo (MCMC) was implemented in an external module that communicated with CHARMM to

evaluate energies of different poses. The aim of the setup phase of the simulation was to derive a biasing function that achieved close-to-uniform sampling of energies within a pre-specified range. The range was chosen manually, based on the results of initial unbiased simulations, to span 70 kcal/mol from the lowest-encountered energy minus an offset of ~10 kcal/mol. The initially guessed weight function was updated in three iterations, each constituting 100 independent MCMC trajectories sampling 50,000 conformations (starting from random poses of DNA above the assembly). The first 10,000 points from each trajectory were discarded as equilibration, and the remaining $4 \cdot 10^7$ points were used to build the observed energy histogram. Given this histogram, the biasing function from the previous iteration was updated to produce a more uniform distribution in the next iteration. In all cases, three iterations were sufficient to reach a point where the entire 70-kcal/mol range of energies was well sampled (except below the lowest-found energy). Upon completion of the setup phase, one final iteration was performed to collect data. To build the energy distributions as a function of angle, shown in Figure S7, all collected points were broken into 5° bins by angle of DNA rotation, and the average energy was computed in each bin (with most containing hundreds of thousands of points) by removing the contribution of the biasing function via standard inverse reweighting(9).

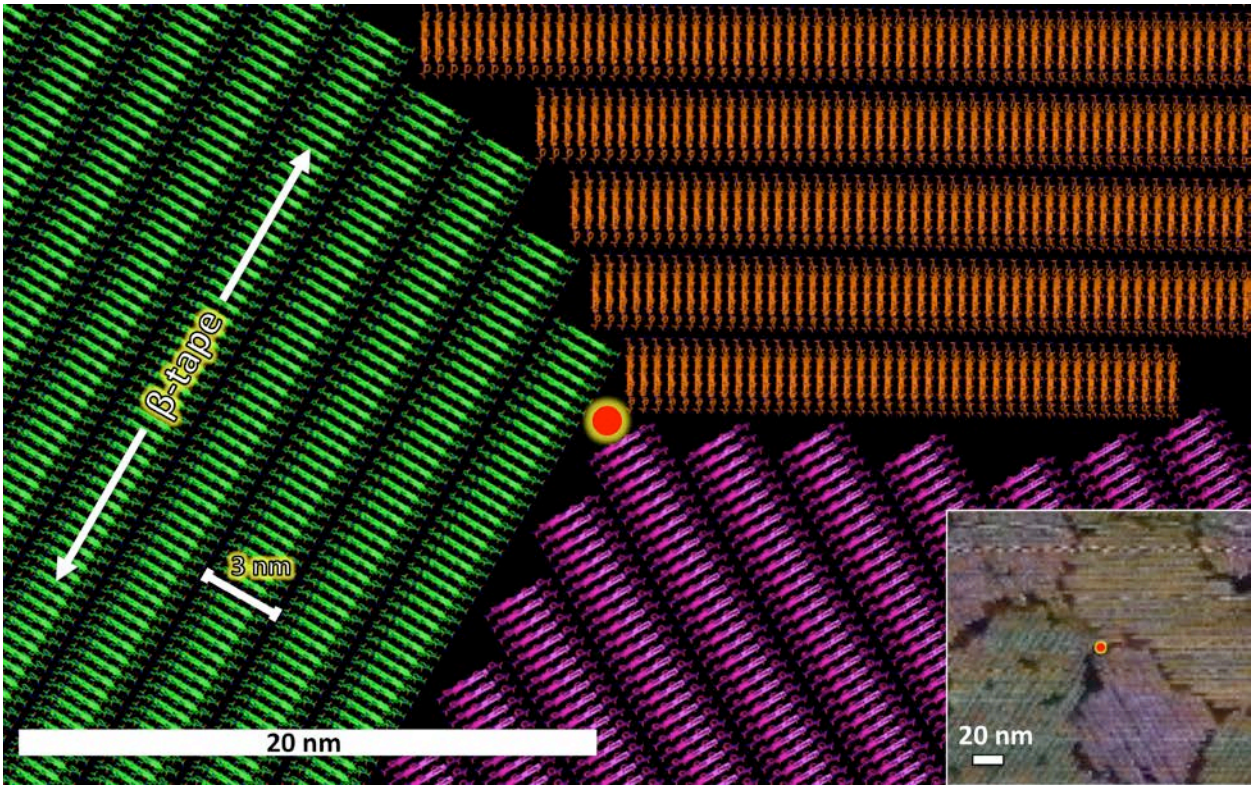

Figure S1: Predicted optimal assembly geometry for $(KF)_4$. Shown is a hypothetical example where three assembly domains of optimal geometry (shown in green, orange, and magenta), each aligned along one of six preferred directions, converge at a point (red dot). The inset is a representative AFM image, showing a similar convergence of three domains, colored equivalently for clarity. A β-tape and its dimension corresponding to the grooves in the AFM image are shown.

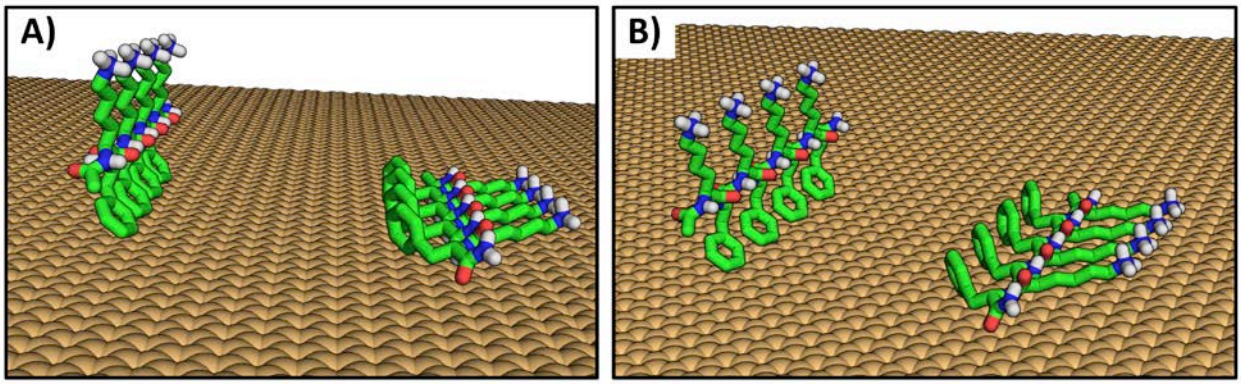

Figure S2: Shown in **A)** and **B)** are two different views of the end-on (left in both panels) and the side-on (right in both panels) peptide binding poses.

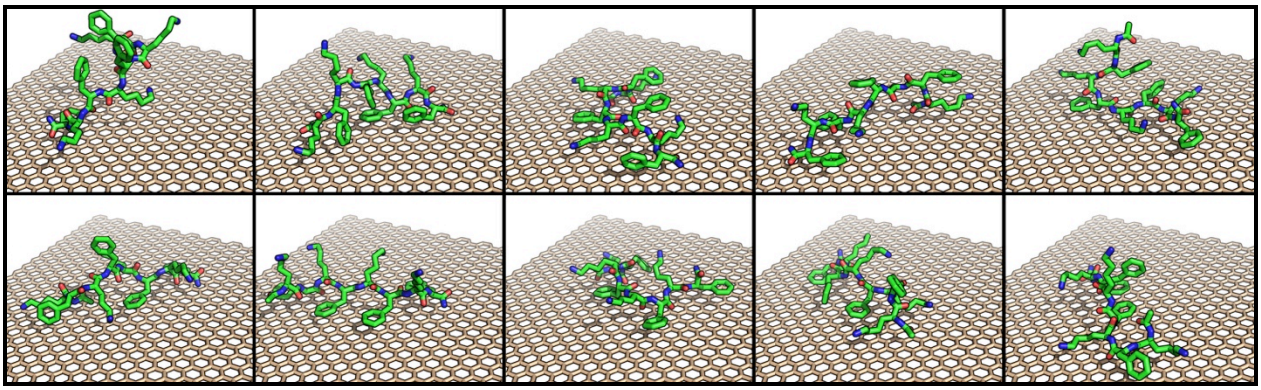

Figure S3: End states of ten independent 10-ns explicit-water MD simulations of $(KF)_4$ over graphene (waters hidden for clarity; see Materials and Methods), showing that the side-on conformation, with side-chains of both Lys and Phe making substantial hydrophobic contacts with graphene, is preferentially sampled.

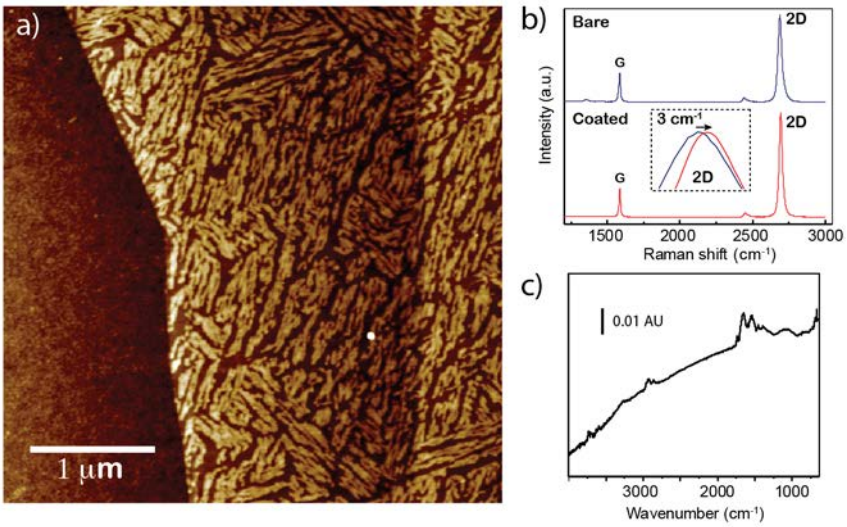

Figure S4: a) AFM image in air of peptide **2** adsorbed onto a graphene flake deposited on Si/SiO_2 using the Scotch-tape method. The peptide layer is clearly seen on the graphene portion of the image, whereas the SiO_2 substrate is free of adsorbed peptide. **b)** Effect of peptide incubation on the Raman spectra of the graphene sheets. As expected, bare monolayers of graphene exhibit strong Raman peaks at approximately 1580 cm^{-1} (due to its G band) and at 2680 cm^{-1} (2D-band). Other Raman modes are at approximately 1350 cm^{-1} (the D-mode), 1620 cm^{-1} (D'- mode) and 2947 cm^{-1} (D+G-mode). The intensity represented is the ratio I_D/I_G, proportional to the defect density. While the G band remains unchanged by the peptide incubation, the 2D band is slightly shifted, indicating that the peptide attaching to the graphene has a slight doping effect. **c)** FTIR absorption spectrum for the deposited layer, displaying absorption band peaks near 1655 and 1550–1540 cm^{-1} for the adsorbed peptide layer. These bands at 1655 and 1540 cm^{-1} are consistent with the amide I and amide II bands of the layer, respectively, which confirms the presence of peptides on the graphene.

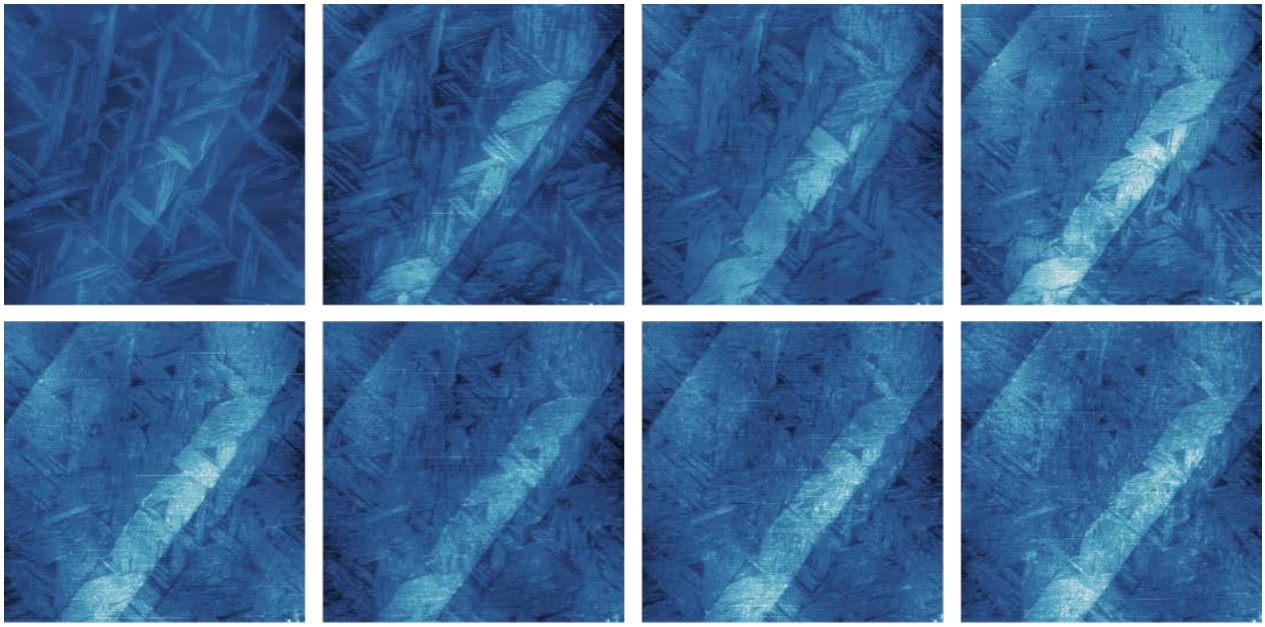

Figure S5: First three minutes of assembly of peptide **7** on HOPG captured using AFM. This peptide layer did not bind DNA, in contrast to peptides **2** and **7**.

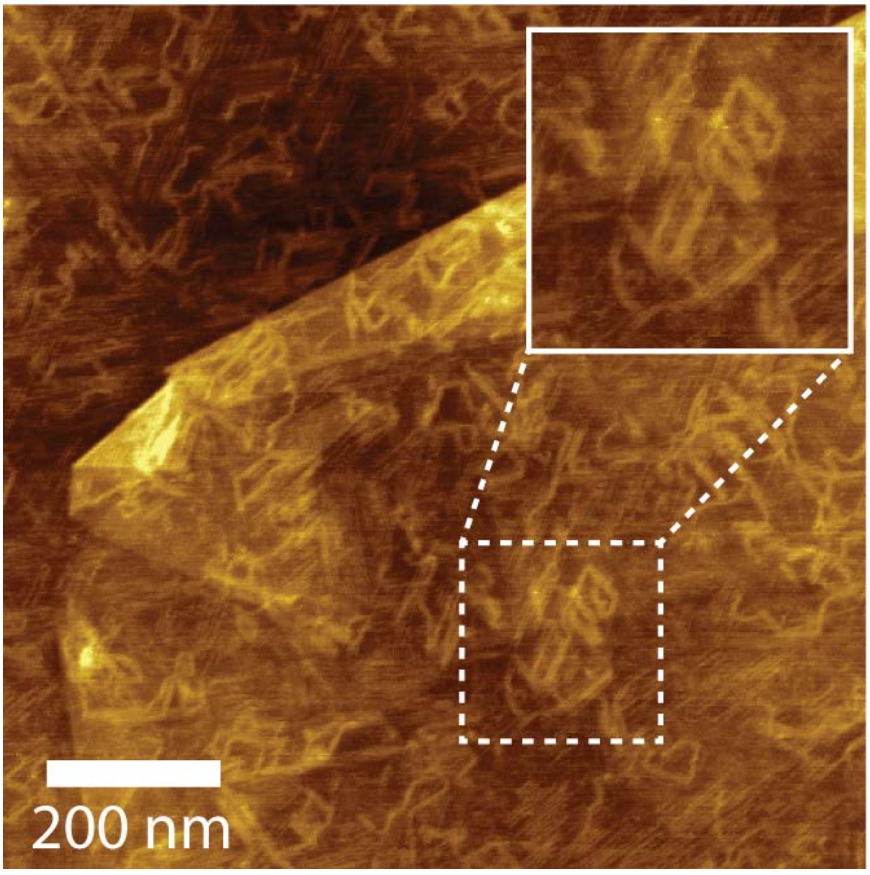

Figure S6: DNA (2,000 bp) strongly aligns with peptide **2** domains on graphene flakes "Scotch-taped" onto an SiO_2 substrate. Area shown in the image is a large multilayered graphene flake with folded flake regions that compact DNA structure).

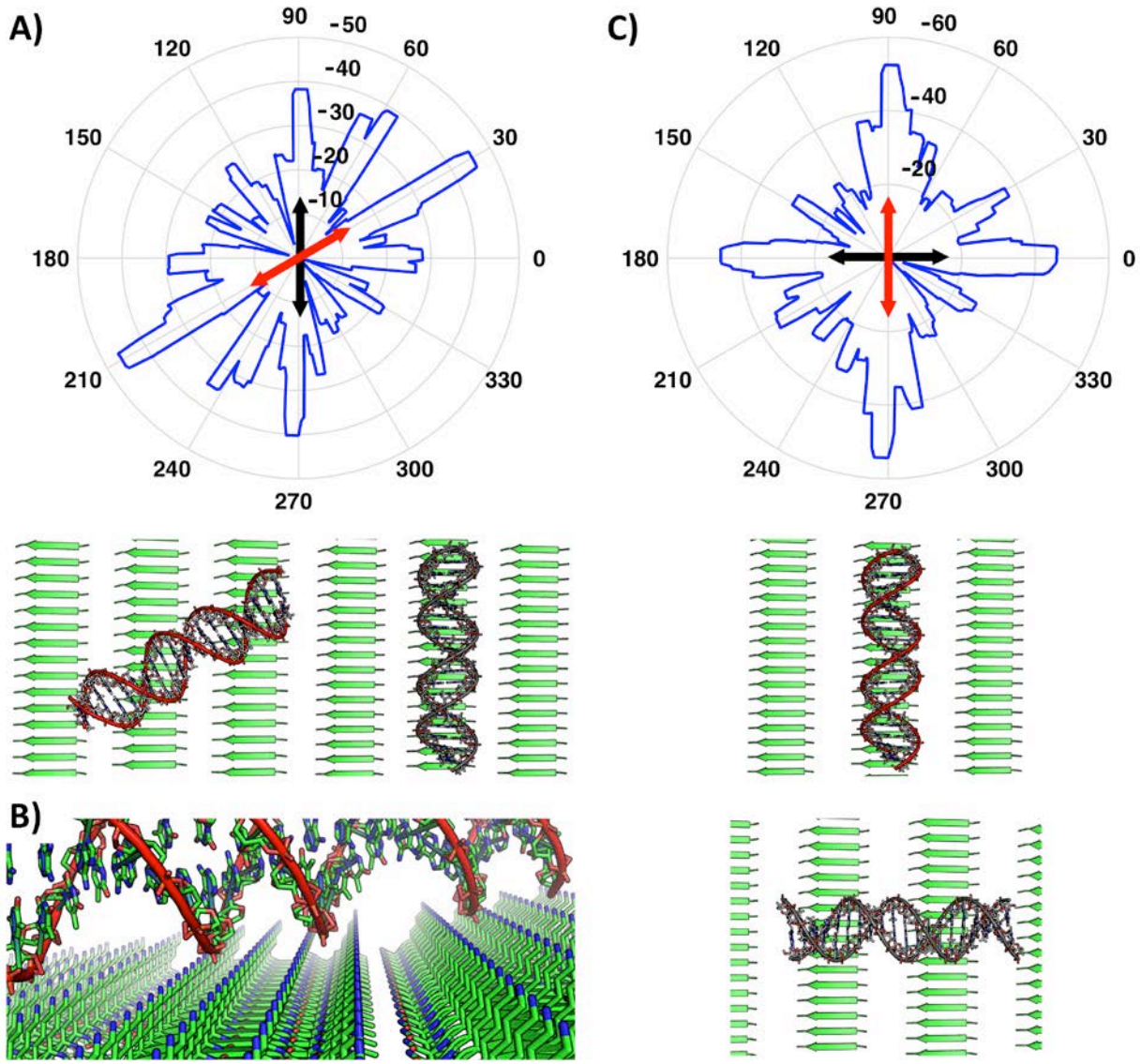

Figure S7: Computed DNA-assembly alignment preferences. Polar plots in **A)** and **C)** show the average energy (on the radial axis in units of kcal/mol) as a function of DNA rotation in the plane of the assembly (the angular coordinate; the 90°/270° direction corresponds to the β-tape axis). Whereas **A)** represents the prediction for the ground state structure, **B)** corresponds to the perturbed model, where inter-β-tape spacing is increased by 0.5 nm. The most and the second most preferred alignment directions are designated with red and black arrows, respectively, in both cases. Shown under each polar plot are the most energetically favorable poses encountered among those aligned in either the optimal or second optimal directions (DNA coloring corresponds to the axes above). **B)** presents a detailed view of the optimal conformation from **A)**. The pose scores well as it allows for very intimate contacts between the phosphate backbone and amine groups of lysines. In fact, the angle of the alignment is such that at the point of contact with the surface, the phosphate backbones intercalate between adjacent rows of Lys resides.

References

1. Kulp, D. W., S. Subramaniam, J. E. Donald, B. T. Hannigan, B. K. Mueller, G. Grigoryan, and A. Senes. 2012. Structural informatics, modeling, and design with an open-source Molecular Software Library (MSL). J Comput Chem 33:1645-1661.
2. Lazaridis, T., and M. Karplus. 1999. Effective energy function for proteins in solution. Proteins 35:133-152.
3. Lazaridis, T. 2003. Effective energy function for proteins in lipid membranes. Proteins 52:176-192.
4. Masunov, A., and T. Lazaridis. 2003. Potentials of mean force between ionizable amino acid side chains in water. J Am Chem Soc 125:1722-1730.
5. Fitzkee, N. C., P. J. Fleming, and G. D. Rose. 2005. The Protein Coil Library: a structural database of nonhelix, nonstrand fragments derived from the PDB. Proteins 58:852-854.
6. Subramaniam, S., and A. Senes. 2012. An energy-based conformer library for side chain optimization: improved prediction and adjustable sampling. Proteins 80:2218-2234.
7. Brooks, B. R., C. L. Brooks, 3rd, A. D. Mackerell, Jr., L. Nilsson, R. J. Petrella, B. Roux, Y. Won, G. Archontis, C. Bartels, S. Boresch, A. Caflisch, L. Caves, Q. Cui, A. R. Dinner, M. Feig, S. Fischer, J. Gao, M. Hodoscek, W. Im, K. Kuczera, T. Lazaridis, J. Ma, V. Ovchinnikov, E. Paci, R. W. Pastor, C. B. Post, J. Z. Pu, M. Schaefer, B. Tidor, R. M. Venable, H. L. Woodcock, X. Wu, W. Yang, D. M. York, and M. Karplus. 2009. CHARMM: the biomolecular simulation program. J Comput Chem 30:1545-1614.
8. Grigoryan, G. 2013. Absolute free energies of biomolecules from unperturbed ensembles. J Comput Chem 34:2726-2741.
9. Janke, W. 1998. Multicanonical Monte Carlo simulations. Physica A: Statistical Mechanics and its Applications 254:164-178.