SUPPLEMENTARY MATERIAL

Free Energy Landscape of Salt-Actuated Reconfigurable DNA Nanodevices

Ze Shi† and Gaurav Arya*,‡

†Department of NanoEngineering, University of California, San Diego, La Jolla, CA 92093, USA
‡Department of Mechanical Engineering and Materials Science, Duke University, Durham, NC 27708, USA

E-mail: gaurav.arya@duke.edu
Phone: 919-660-5435. Fax: 919-660-8963
Free energy of the bare hinge

The free energy landscape $G_{bh}(\theta)$ of the bare hinge was obtained via the umbrella sampling approach. This involved performing a set of “biased” MD simulations of the bare hinge (Fig. 2B), each imposing a distinct external harmonic potential $V_k = \frac{1}{2}k_{\theta}(\theta - \theta_k)^2$ of stiffness $k_{\theta}$ centered about angle $\theta_k$ to enhance sampling of hinge conformations about $\theta_k$, where $k$ is the simulation index. The distribution of hinge angles obtained from such simulations, with $\theta_k$'s spanning the open to closed hinge conformations, were analyzed using the weighted histogram analysis method (WHAM)\textsuperscript{2,3} to yield the free energy landscape. The hinge angle $\theta$ was calculated as the angle subtended by two axes fitted through the two hinge arms, and hence, $\theta$ is a function of the position coordinates of all phosphate and nucleoside groups making up the two arms. This makes the implementation of the angular restraint on the hinge quite tedious, as it involves calculating its gradient with respect to each of those coordinates. To circumvent this issue, we used biased potentials formulated in terms of the Euclidean distance $d$ between the centroids of the two arms (instead of angle), which are easier to implement, and then used the relationship between $d$ and $\theta$ to convert the resulting $G_{bh}(d)$ landscape to $G_{bh}(\theta)$ (see Fig. S3).

Free energy of the overhangs

The free energy landscape $G_{ov}(\theta)$ involves contribution from all overhangs. This term includes inter-overhang interactions both across and along the hinge arms as well as overhang interactions with the hinge arms. We made two assumptions to greatly simplify the calculation of this free energy contribution. The first is that the primary effect of the hinge arms on the overhangs is “confinement”, that is, the restriction of the configurational degrees of freedom of the overhangs through the steric bulk (excluded volume) of the arms. This assumption is reasonable because the DNA strands making up the hinge arms are all fully engaged in forming DNA helices, and thus do not offer any opportunity for base-pairing or base-stacking with the overhangs. The electrostatic interactions between the overhangs and the arms are also greatly screened out due to high salt concentrations. The remaining effect of the two arms—the steric interactions—can then be simply modeled by “repulsion” planes, representing the inner surface of the arms. The repulsion planes impose a strong and salt-independent harmonic repulsion to the phosphate and nucleoside groups of the overhangs whenever they cross the plane, effectively restricting the overhangs to within the planes. The second assumption is that the overhangs interact only with their complementary counterparts on the opposite arm, which is also reasonable considering that the separation distance between the overhangs within each arm, even the nearest neighbors, is longer than the stretched length of the overhangs. This allowed $G_{oh}$ to be further divided into independent contributions $G_{oh,i}$ from each of the $n$ pairs of overhangs attached at different locations on the arms, whereby $G_{oh}(\theta) = \sum_{i=1}^{n} G_{oh,i}(\theta)$.

The free energy landscape $G_{oh,i}(\theta)$ for each overhang pair was also obtained via umbrella sampling. The bare hinge was substituted by two repulsion planes to mimic the steric constraint imposed by the hinge arms (Fig. 2C). Without loss of generality, one of the repulsion planes was held fixed along the $x$-$z$ axes, while the other plane was allowed to rotate freely about the hinge vertex represented by the $y$-axis. The phosphate group of the
overhang covalently linked to the stationary hinge arm was held fixed at position \((x, y, z) = (l, 0, 0)\), where \(l\) denotes the distance of the overhang attachment point from the vertex. The phosphate group of the overhang linked to the rotating hinge arm was subjected to two external potentials: a stiff spring that confined its motion along the arc \((l \cos \theta, l \sin \theta, 0)\), where \(\theta\) is the hinge angle, and an external harmonic potential \(V_k = \frac{1}{2}k_\theta(\theta - \theta_k)^2\) similar to that used in bare-hinge simulations, which biased the angle between the overhangs towards \(\theta_k\). Note that the repulsion plane rotates in sync with this phosphate group, so, by construct, the two repulsion planes also subtend the angle \(\theta\). Unlike the bare-hinge simulations, the hinge angle \(\theta\) here is only a function of the position of this rotatable phosphate group, and hence, the bias potential \(V_k\) can be directly implemented in terms of \(\theta\). This system however presents a new challenge, namely, the large free energy differences between the various hybridization states of the overhangs, that is, with varying number of base pairs \(\xi\) formed between the two strands, which causes severe sampling issues. Hence, we used the more efficient VMMC simulations for sampling the overhang conformations, with a bias in the transition probability between hybridization states to further enhance their sampling. A modified WHAM was used to analyze the distributions in hinge angles and hybridization states obtained from simulations performed using different values of \(\theta_k\) to obtain the 2D free energy landscape \(G_{\text{oh},i}(\theta, \xi)\) for each overhang pair. Summing up free energies from the different hybridization states, via \(\exp \left[ -G_{\text{oh},i}(\theta)/k_B T \right] = \sum_\xi \exp \left[ -G_{\text{oh},i}(\theta, \xi)/k_B T \right]\), then yields the free energy landscape we seek.

**Simulation details**

The MD simulations of the bare hinge were performed using the oxDNA package at a temperature of 300 K and a monovalent salt concentration \([\text{Na}^+]\) of 500 mM. The temperature was held constant via an Anderson-like thermostat. As in previous studies,\(^1\) we used the average-base parametrization option in oxDNA and a simulation time step of 15.15 fs. The umbrella simulation “windows” were spaced at equal intervals of \(\sim 1^\circ\); each window was simulated for 20 million time steps, and the stiffness of the harmonic restraint implemented between the centers of masses of the hinge arms in each window was set to 0.16 kcal/mol/Å\(^2\). Due to the coarse-grained nature of the oxDNA model,\(^1\) each time step represents roughly 5 ps of real time, and therefore each simulation effectively sampled \(\sim 100\) µs of real time. Given that this time scale is even longer than the characteristic rotational relaxation time of the free hinge arms (Fig. S4), we expect each window to be well sampled.

The VMMC simulations of the overhang pairs were also performed using the oxDNA package at 300 K and at three different monovalent salt concentrations: 0.2 M, 0.4 M, and 1.2 M. The umbrella-sampling windows here were also spaced at 1° intervals and each window was simulated for 1 billion VMMC steps. This sampling time was more than adequate to generate smooth angle distributions about the angular restraint that exhibited strong overlap between adjacent windows (see Figs. S5 and S6). As explained earlier, the free energy of the overhang strands was sampled as a function of the total number of base pairs formed between the two strands and the end-to-end distance between them defined as the separation distance between the center of mass of the 5’ and 3’ terminal nucleotides (those attached to the repulsion planes) of the two strands. Two bases were considered to be paired when the hydrogen-bonding interaction between them exceeded 0.596 kcal/mol,
the characteristic thermal energy at 300 K. The steric confinement of the overhangs by the repulsion planes, the attachment of the terminal nucleotides to the repulsion plane, and the umbrella potential on the angle between the repulsion planes were all implemented as customized harmonic restraints in oxDNA with stiffnesses of 160 kcal/mol/Å², 0.8 kcal/mol/Å², and 6 kcal/mol/rad², respectively.

**Modified WHAM**

Each umbrella-sampling VMMC simulation \( k \) of an overhang pair yielded a “biased” 2D distribution \( \rho_k^{(b)}(\theta, \xi) \) in hinge angle \( \theta \) (angle between repulsion planes) and hybridization state \( \xi \) (number of base pairs formed between the overhangs). To recover the overhang free energy landscape from a collection of such biased 2D distributions, we used a modified version of the 2D weighted histogram analysis method. While 2D WHAM traditionally involves harmonic biases in *both* order parameters, our VMMC simulations implemented a harmonic bias in only order parameter, \( \xi \), the potential \( V_k(\theta) = \frac{1}{2}(\theta - \theta_0)^2 \) in hinge angle \( \theta \). The bias in the other order parameter, \( \xi \), was given by a conditional function \( B_k(\xi) \) specifying the bias factor for each hybridization state:

\[
B_k(\xi) = \begin{cases} 
    b_k, & \text{if } \xi = \xi_j, j = 0, 1, \ldots, 6 \\
    0, & \text{otherwise.} 
\end{cases} 
\]

(1)

For a given simulation \( k \), the biased probability distribution \( \rho_k^{(b)}(\theta, \xi) \) in hinge angle and hybridization state should be proportional to \( B_k(\xi)e^{-[U_0(\theta, \xi)+V_k(\theta)]/k_BT} \), where \( U_0(\theta, \xi) \) is the intrinsic (unbiased) free energy landscape. Hence, the absolute value of the probability distribution can be written as

\[
\rho_k^{(b)}(\theta, \xi) = \frac{B_k(\xi)e^{-[U_0(\theta, \xi)+V_k(\theta)]/k_BT}}{\int B_k(\xi)e^{-[U_0(\theta, \xi)+V_k(\theta)]/k_BT}d\theta d\xi} \times \frac{\int e^{-U_0(\theta, \xi)/k_BT}B_k(\xi)e^{-V_k(\theta)/k_BT}d\theta d\xi}{\int B_k(\xi)e^{-V_k(\theta)/k_BT}d\theta d\xi} 
\]

(2)

By defining \( \langle B_k(\xi)e^{-V_k(\theta)/k_BT}\rangle_{U_0} \) as \( e^{-F_k/k_BT} \), the true (unbiased) probability distribution \( \rho_k(\theta, \xi) \) is then given by

\[
\rho_k(\theta, \xi) = \rho_k^{(b)}(\theta, \xi) \left[ \frac{e^{(V_k(\theta)-F_k)/k_BT}}{B_k(\xi)} \right]. 
\]

(3)

The \( e^{(V_k(\theta)-F_k)/k_BT}/B_k(\xi) \) term represents the bias correction factor. Note that from the definition of \( B_k(\xi) \) (Eq. 2), a simulation run that goes through all the hybridization states \( (\xi) \) is considered as one biased simulation, therefore the bias correction factor should be the same for hinge angle distributions with the same bias potential \( V_k(\theta) \) regardless of the hybridization status.
Having established this new form of the bias correction factor, the remaining procedure for determining the free energy profile is similar to canonical WHAM.\textsuperscript{2,3} This involves solving for the true probability distribution in hinge angle and hybridization state via an iterative procedure involving the following two calculations:

$$\rho_{\theta_j,\xi} = \frac{\sum_k M_k \rho_{k,\theta_j,\xi}^{(b)}}{\sum_k M_k b_{k,\xi} e^{-(V_{k}(\theta_j) - F_k)/k_B T}},$$  \hspace{1cm} (4)

$$e^{-F_k/k_B T} = \sum_j \sum_{\xi} \rho_{\theta_j,\xi} b_{k,\xi} e^{-V_k(\theta_j)/k_B T},$$  \hspace{1cm} (5)

where \(\rho_{\theta_j,\xi}\) and \(\rho_{k,\theta_j,\xi}^{(b)}\) are the true and biased probability distributions discretized into histogram bins \(j\) centered about angles \(\theta_j\) and \(M_k\) is the total number of angles and hybridization states used for building histograms in each simulation \(k\). To obtain the true probability distribution, bias correction factor is updated with the estimated probability distribution. Thus obtained true probability distribution then yields the free energy landscape via \(G_{oh,i}(\theta_j,\xi) = -k_B T \ln \rho_{\theta_j,\xi}\).

References

(1) Shi, Z.; Castro, C. E.; Arya, G. Conformational dynamics of mechanically compliant DNA nanostructures from coarse-grained molecular dynamics simulations. \textit{ACS Nano} \textbf{2017}, \textit{11}, 4617–4630

(2) Souaille, M.; Roux, B. Extension to the weighted histogram analysis method: combining umbrella sampling with free energy calculations. \textit{Comput. Phys. Commun.} \textbf{2001}, \textit{135}, 40–57

(3) Kumar, S; Rosenberg, J. M.; Bouzida, D. and Swendsen, R. H.; Kollman, P. A. The weighted histogram analysis method for free-energy calculations on biomolecules. I. The method. \textit{J. Comput. Chem.}, \textbf{1992}, \textit{13}, 1011–1021
Figure S1: CaDNAno design of the actuatable hinge with the single-stranded scaffold (that also connects the two hinge arms) and staple strands shown in blue and red, respectively.
Figure S2: Zoomed image of the hinge joint in oxDNA representation showing the 2-base and 16-base single-stranded DNA connections (colored blue and red colors) between the hinge arms. Also shown is a schematic of the hinge joint at roughly similar orientation.
Figure S3: Autocorrelation $\langle (\theta(t) - \theta_0)(\theta(0) - \theta_0) \rangle$ in hinge angle fluctuations computed from MD simulation of a bare hinge (solid black line), where $t$ is time and $\theta_0$ is the average measured angle. Also shown is an exponential fit to the decaying portion of the curve (dashed red lines). The characteristic decay time constant obtained from the fit is specified in each figure.
Figure S4: Frequency plot showing the hinge angle against the corresponding distance between the centers of mass (COMs) of the two hinge arms, as obtained from multiple umbrella-sampling MD simulations of the bare hinge restrained at different COM distances via harmonic potentials. Solid line indicates good linear fit between hinge angle and distance.
Figure S5: Histograms of hinge angles sampled in each umbrella sampling window spaced at 1° intervals for the C1 overhang pair at Na$^+$ = 1.2 M. The histograms are cumulated across the five possible hybridization states: (a) $\xi = 0$, (b) $\xi = 1$, (c) $\xi = 2$, (d) $\xi = 3$, and (e) $\xi = 4$ base pairs.
Figure S6: Histograms of hinge angles sampled in each umbrella sampling simulation spaced at 1° intervals for the C10 overhang pair at \( \text{Na}^+ = 1.2 \) M. The histograms are cumulated across the five possible hybridization states: (a) \( \xi = 0 \), (b) \( \xi = 1 \), (c) \( \xi = 2 \), (d) \( \xi = 3 \), and (e) \( \xi = 4 \) base pairs.
Figure S7: 2D free energy landscapes $G_{\text{oh,i}}(\theta, \xi)$ for overhangs attached at ten different sites (C1–C10) at three different Na$^+$ concentration (0.2 M, 0.4 M, and 1.2 M). All results are shown for overhangs with 4 complementary bases and 3-base linkers at 300 K. Empty regions in the plots indicate states with extremely low probabilities of occurrence.
Figure S8: Free energy landscape of the bare hinge obtained from experiments (filled circles) along with a harmonic fit to the data (solid line). The landscape was obtained via $G_{bh}(\theta) = -k_B T \ln \rho(\theta)$, where $\rho(\theta)$ is the distribution of hinge angles obtained from analyzing hundreds of TEM images of the hinges.