Supplementary Information for 'Measuring internal forces in single-stranded DNA: Application to a DNA force clamp'

Megan C. Engel,[*,†,‡] Flavio Romano,[¶] Ard. A. Louis,[‡] and Jonathan P. K. Doye[§]

†*School of Engineering and Applied Sciences, Harvard University, 29 Oxford Street, Cambridge MA, 02138, USA*

‡*Rudolf Peierls Centre for Theoretical Physics, University of Oxford, 1 Keble Road, Oxford, OX1 3NP, UK*

¶*Dipartimento di Scienze Molecolari e Nanosistemi, Universitá Ca Foscari di Venezia, Via Torino 155, 30172, Venezia Mestre, Italy*

§*Department of Chemistry, University of Oxford, South Parks Road, Oxford, OX1 3QZ, UK*

E-mail: mcengel@seas.harvard.edu

S1 Additional Simulation Details

The 160-nucleotide poly-T strands were simulated at 21°C and [Na+]=10 M, and their end-points were held in 285 pN/nm 3D harmonic traps whose centres were 41.9 nm apart. For all simulations, the oxDNA parameters shown in Table S1 were used.

S1.1 Langevin thermostat

oxDNA provides the option of using a Langevin thermostat, which mimics coupling to a thermal bath through the addition of a frictional damping term and a random forcing term to

Table S1: oxDNA parameters used across all simulations described in this work, with the exception of those which used a Langevin thermostat (detailed in the main text), for which `thermostat = langevin`.

Parameter	Value
newtonian_steps	103
diff_coef	2.50
thermostat	john
dt	0.005
verlet_skin	0.2
use_edge	false

Newton's equations of motion.[1] The strength of the random forces is related to the magnitude of the frictional damping through the 'fluctuation-dissipation relation', which guarantees evolution in the canonical NVT ensemble.[1] Because oxDNA simulations are somewhat slower when the Langevin, rather than Andersen-like, thermostat is used, it makes only a brief appearance in the main text.

S1.2 Derivation of FENE force distribution

The expected force probability distribution for a polymer chain with only backbone interactions implemented can be derived by considering the finitely extensible nonlinear elastic (FENE) potential in oxDNA, which is given by

$$V(r) = -\frac{\epsilon}{2} \ln \left(1 - \left(\frac{r - r_0}{\Delta} \right)^2 \right), \tag{S1}$$

where $\Delta = 0.25$ (0.21 nm); $\epsilon = 2$ $k_B T^*$ with T^* a reduced temperature, 3000K; and $r_0 = 0.7564$ (0.644 nm) for oxDNA 2.0[2] in its internal unit system. This potential is shown in Figure S1. The probability of finding bases separated by r is

$$P(r) = \frac{e^{-\beta V(r)}}{\int e^{-\beta V(r)} dr} = \frac{e^{-\beta V(r)}}{Z}, \tag{S2}$$

and the free energy, A, as a function of r is

$$A(r) = -k_B T \ln(P(r)). \tag{S3}$$

The force experienced by nucleotides separated by distance r is

$$\vec{F}(r) = -\frac{\partial A}{\partial r}\hat{\mathbf{r}} = -\frac{\partial V}{\partial r}\hat{\mathbf{r}} = \frac{\frac{\epsilon}{\Delta^2}(r - r_0)}{1 - \left(\frac{r-r_0}{\Delta}\right)^2}\hat{\mathbf{r}}, \tag{S4}$$

and is directed along the backbone. If we let $\vec{F}(r) = F\hat{\mathbf{r}}$, the probability density function for observing a force F, $g(F)$, can be found by noting $g(F)\,dF = P(r(F))\,dr$, or $g(F) = P(r(F))\,\frac{dr}{dF}$, where $P(r)$ is given by equation S2 and $r(F)$ is found by inverting equation S4:

$$r(F) = -\frac{\epsilon}{2F} \pm \frac{1}{2F}\sqrt{\epsilon^2 + 4\Delta^2 F^2} + r_0. \tag{S5}$$

Only the solution featuring the negative square root is bounded. Therefore,

$$g(F) = P(r(F))\frac{dr}{dF} = \frac{1}{Z}\left[1 - \frac{1}{\Delta^2}\left(- \frac{\epsilon}{2F} - \frac{1}{2F}\sqrt{\epsilon^2 + 4\Delta^2 F^2}\right)^2\right]^{\epsilon\beta/2}\frac{dr}{dF}. \tag{S6}$$

From equation S5,

$$\frac{dr}{dF} = \frac{\epsilon}{2F^2} + \frac{2\Delta^2}{\sqrt{\epsilon^2 + 4\Delta^2 F^2}} - \frac{\sqrt{\epsilon^2 + 4\Delta^2 F^2}}{2F^2}. \tag{S7}$$

While keeping track of force components F_x, F_y, F_z (and thus force magnitude, $|F|$) is straightforward in oxDNA, the vector direction $\hat{\mathbf{r}}$ is not as easily accessible, and indeed will be different for all nucleotides along the chain. Therefore, for simplicity, instead of working with $g(F)$ (where F is defined by $\vec{F}(r) = F\hat{\mathbf{r}}$), we derive $P(F_z)$, the probability distribution for a single component of the force. To do so, we note that for a random variable A defined as the product of two other random variables, $A = BC$, the probability density distributions

of each of the three variables are related thus:[3]

$$P_A(a) = \int_{-\infty}^{\infty} P_B(b) P_C(a/b) \frac{1}{|b|}\, db. \tag{S8}$$

Since $F_z = \vec{F} \cdot \hat{\mathbf{z}} = F\cos\theta$, we can use equation S8 with $A = F_z$, $B = F$, and $C = \cos\theta$ to obtain

$$P_{F_z}(f_z) = \int_{-\infty}^{\infty} g_F(f') P_{cos\theta}(f_z/f') \frac{1}{|f'|}\, df' \tag{S9}$$

We assume for simplicity that the instantaneous force vector is equally likely to point in all directions; namely, we assume $P(\phi, \theta) = \frac{1}{4\pi}$. In this isotropic case, $P(\theta) = \frac{1}{2}\sin\theta$, and the probability density distribution for $\cos\theta$ is given by

$$P_{cos\theta}(\cos\theta) = P(\theta)\left|\frac{d\theta}{d(\cos\theta)}\right| = \frac{1}{2}. \tag{S10}$$

Now, we adjust the integration limits of Eq. S9 to reflect the fact that $F_z \leq F$ by definition:

$$P(F_z) = \int_{-\infty}^{-F_z} \frac{g_F(f')}{2|f'|}\, df' + \int_{F_z}^{\infty} \frac{g_F(f')}{2|f'|}\, df' \tag{S11}$$

Eq. S11 is plotted in black in Fig. 3 (c) in the main text. The agreement with simulation data is excellent where stacking interactions are suppressed, as expected.

S1.3 Stacking potential

To discover why the stacking interaction broadens the force distributions in Fig. 3 (c) in the main text, we can make a simple argument based on the form of the stacking potential in oxDNA.

For simplicity, we consider only the radial part of the stacking potential:

$$V(r) = \epsilon_s\Big(1 - e^{-a(r-r_{0s})}\Big)^2 - \epsilon_s\Big(1 - e^{-a(r_c-r_{0s})}\Big)^2, \tag{S12}$$

where $\epsilon_s = 1.3523 + 2.6717 k_B T$, $r_c = 0.9$, $r_{0s} = 0.4$, and $a = 6$ for oxDNA 2.0[2] in its internal unit system.

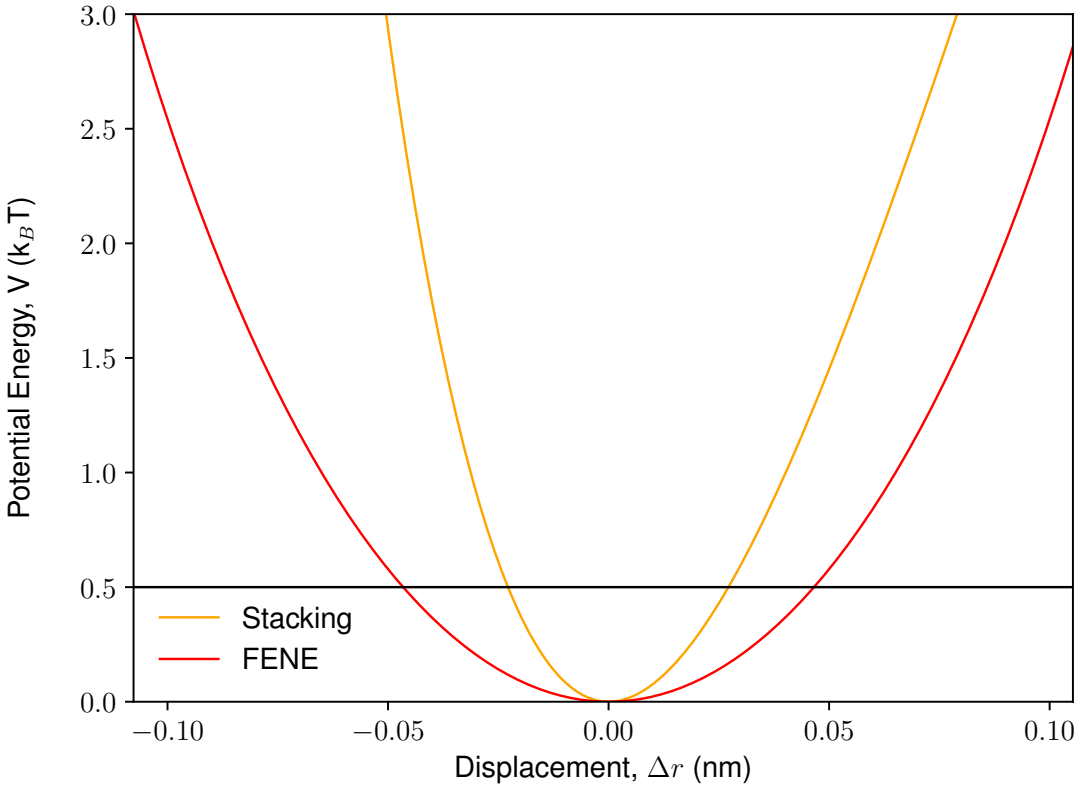

Figure S1: oxDNA2.0 radial FENE (red) and stacking (orange) potentials, with the minima of the potentials shifted to the origin. The stacking potential is clearly much steeper than the FENE at the thermal energy (black line), giving rise to the large force distribution width for the stacking curves in Fig. 3(c) of the main text.

Figure S1 shows this potential alongside the radial part of the FENE potential, with the minima of the potentials shifted to the origin. The black line indicates the thermal energy that the system is expected to possess in equilibrium according to equipartition, i.e. $k_B T/2$, at T=21 °C. We can evaluate the force, $F(\Delta r) = -\frac{dV}{dr}$, at the intersection of each potential with the thermal energy, which occurs at radial displacements from equilibrium of $\Delta r = 0.13$ nm and $\Delta r = 0.07$ nm for the FENE and stacking potentials, respectively. $F_{FENE}^{thermal} \sim 89$ pN and $F_{stack}^{thermal} \sim 193$ pN for these displacements, which, after taking into account that only one component of the force is plotted in Fig. 3(c) of the main text, match

the observed distribution widths well. We can thus confidently attribute the form of the observed force distributions to the oxDNA potentials, and the substantial force distribution width associated with stacking interactions to the steepness – and thus, large instantaneous forces – of the stacking potential.

S1.4 Calculation of effective trap stiffness

Consider two nucleotides separated along the z axis – the long axis of the force clamp – by d. We assume each nucleotide moves in an isotropic 3D harmonic potential for simplicity, such that the Boltzmann probability of observing position r_1 for nucleotide 1, positioned at the origin, is

$$\begin{aligned} P(\mathbf{r_1}) = P(x_1, y_1, z_1) &= \frac{e^{-\beta k(x_1^2+y_1^2+z_1^2)/2}}{\int_{-\infty}^{\infty} e^{-\beta k(x_1^2+y_1^2+z_1^2)/2} dx_1\, dy_1\, dz_1} \\ &= \left(\frac{\beta k}{2\pi}\right)^{3/2} e^{-\beta k(x_1^2+y_1^2+z_1^2)/2} \end{aligned} \tag{S13}$$

and for nucleotide 2 is

$$P(\mathbf{r_2}) = P(x_2, y_2, z_2) = \left(\frac{\beta k}{2\pi}\right)^{3/2} e^{-\beta k(x_2^2+y_2^2+(z_2-d)^2)/2}, \tag{S14}$$

where β is the inverse thermal energy and k is the stiffness of each trap. The distance between the two nucleotides is $r = \sqrt{(x_2-x_1)^2 + (y_2-y_1)^2 + (z_2-z_1)^2}$, so the mean r and r^2 values are given by (where $dV_i = dx_i dy_i dz_i$):

$$\langle r \rangle = \int_{-\infty}^{\infty} \sqrt{(x_2-x_1)^2 + (y_2-y_1)^2 + (z_2-z_1)^2} P(x_1, y_1, z_1) P(x_2, y_2, z_2) dV_1 dV_2 \tag{S15}$$

$$\langle r^2 \rangle = \int_{-\infty}^{\infty} \big((x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2\big) P(x_1, y_1, z_1) P(x_2, y_2, z_2) dV_1 dV_2 \qquad \text{(S16)}$$

where the joint probability distribution $P(x_1, y_1, z_1, x_2, y_2, z_2)$ is simply the product $P(x_1, y_1, z_1)P(x_2, y_2, z_2)$ since the variables are independent. Equations S15 and S16 give the mean and variance ($\langle r^2 \rangle - \langle r \rangle^2$) of the $P(r)$ distribution, and can be numerically integrated with different values of k until the observed parameters are obtained. Doing so yields an effective trap stiffness of $k = 16.32\,\text{pN/nm}$ and separation of $d = 42.76\,\text{nm}$. These values are only slightly different from the results obtained by neglecting position fluctuations in x and y altogether, assuming $r = \Delta z = z_2 - z_1$, and taking the variance of $P(r)$ to be the sum $\sigma_r^2 = \sigma_{z_1}^2 + \sigma_{z_2}^2 = \frac{1}{\beta k} + \frac{1}{\beta k}$, with the individual variances $\sigma_{z_i}^2$ equal and determined using the equipartition theorem. Under these assumptions, $k = 16.23\,\text{pN/nm}$ and $d = 42.77\,\text{nm}$. For origami geometries with smaller $z : x$ ratios, however, this more accurate method for calculating effective stiffness may prove necessary.

S2 Additional Results

As explained in the main text, it is important to ensure that our simulations are sampling all relevant secondary structures. To gauge whether we had sampled secondary structures sufficiently, we split our simulation output for each system in half and computed histograms of which bases formed hydrogen bonds throughout the simulation, shown in Figures S3-S5 for all three designed forces. Only IsoII results at the highest salt concentration simulated, 5 M [Na^+], are pictured, as secondary structures become less favourable at lower ionic strengths due to electrostatic repulsion and secondary structure sampling is therefore most difficult at high salt. Though well over 10 000 possible pairs of bases exist for these strands, the figures show only those hydrogen bonds which were actually observed and assigns a numeric ID to each unique bond. The amount of overlap between blue and red histograms is taken

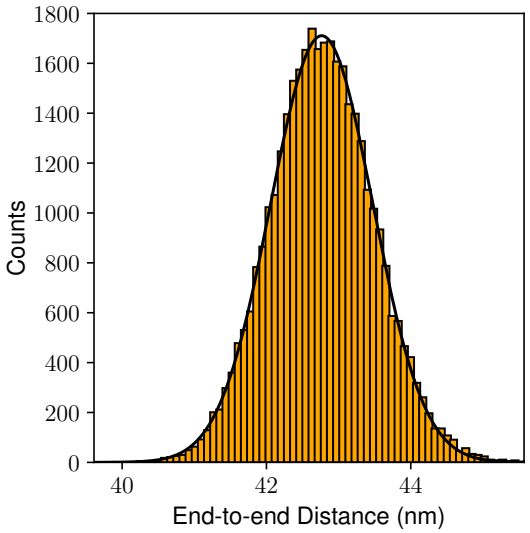

Figure S2: End-to-end distribution for scaffold nucleotides on either edge of the central gap in the structure of Nickels et al.[4]. The width of this distribution can be used to determine an effective stiffness for the traps that are used to represent the constraints that the origami places on the tension-bearing single strand, as detailed in the main text.

as a proxy of how well secondary structures were sampled. For the 4.0 pN design, the main patterns of secondary structures visited are the same for both histograms, and we conclude that $\sim 1 \times 10^9$ MD steps ($\times$ 24 replicas) is an appropriate simulation length for this system. We expect equilibration over secondary structures to be easier at higher internal forces, since force helps to destabilize secondary structures. Indeed, the histogram overlap is less for the larger 2.5 pN and 1.2 pN designs, and more different pairs of bases are seen to interact via hydrogen bonding as the designed force decreases (there are also combinatorially more possibilities for the longer strands). The results reported for these low-force systems should be treated correspondingly more cautiously. While the accuracy of the results can be improved with longer simulations, we observed the same qualitative trends for all three force designs, and thus have confidence in the results even for the 2.5 pN and 1.2 pN systems.

The complete results for all systems simulated can be found in Table S2. Figure S6 contains the forces measured at various interfaces along the strand containing the Holliday junction at different salt concentrations, for all three force designs, and with and without

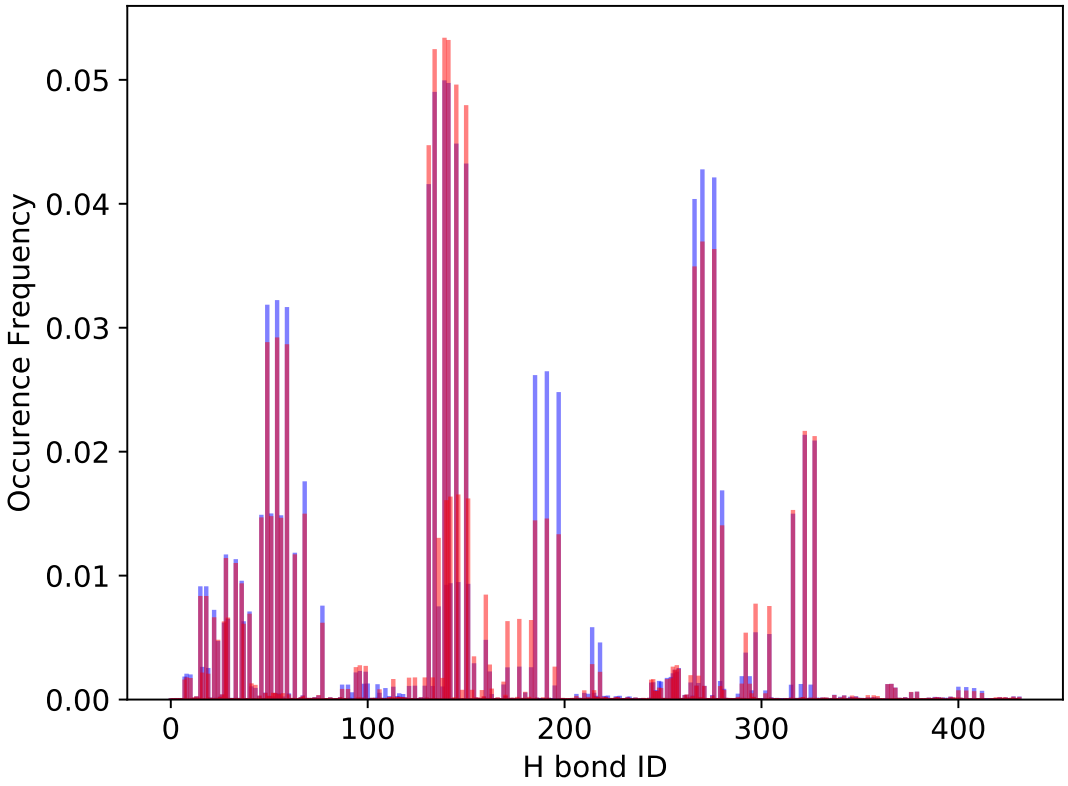

Figure S3: Map of all unique base pairs observed in oxDNA simulations of the 4.0 pN IsoII structure at $[Na^+]$=5 M. The data were split in half (12 identical replicas of 1.3×10^9 MD steps each), indicated by red and blue histograms. The amount of overlap gives a rough indication of whether the important secondary structures are being sampled sufficiently.

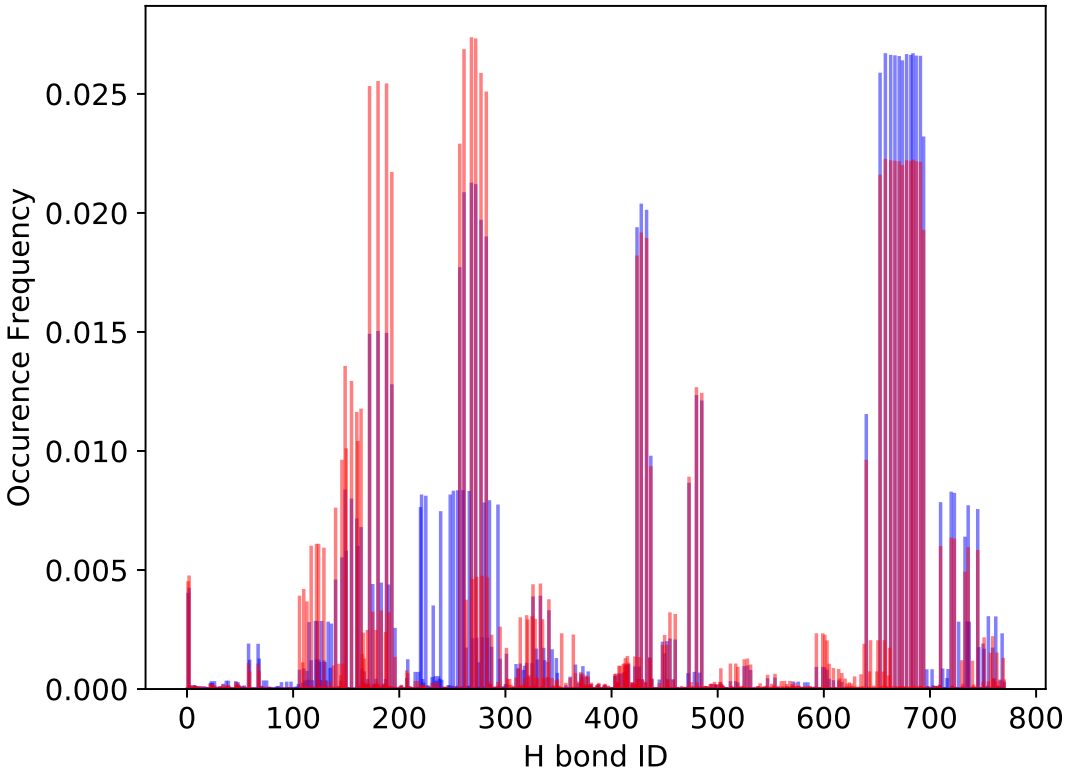

Figure S4: Map of all unique base pairs observed in oxDNA simulations of the 2.5 pN IsoII structure at $[Na^+]$=5 M. The data were split in half (12 identical replicas of 1.4×10^9 MD steps each), indicated by red and blue histograms. The amount of overlap gives a rough indication of whether the important secondary structures are being sampled sufficiently.

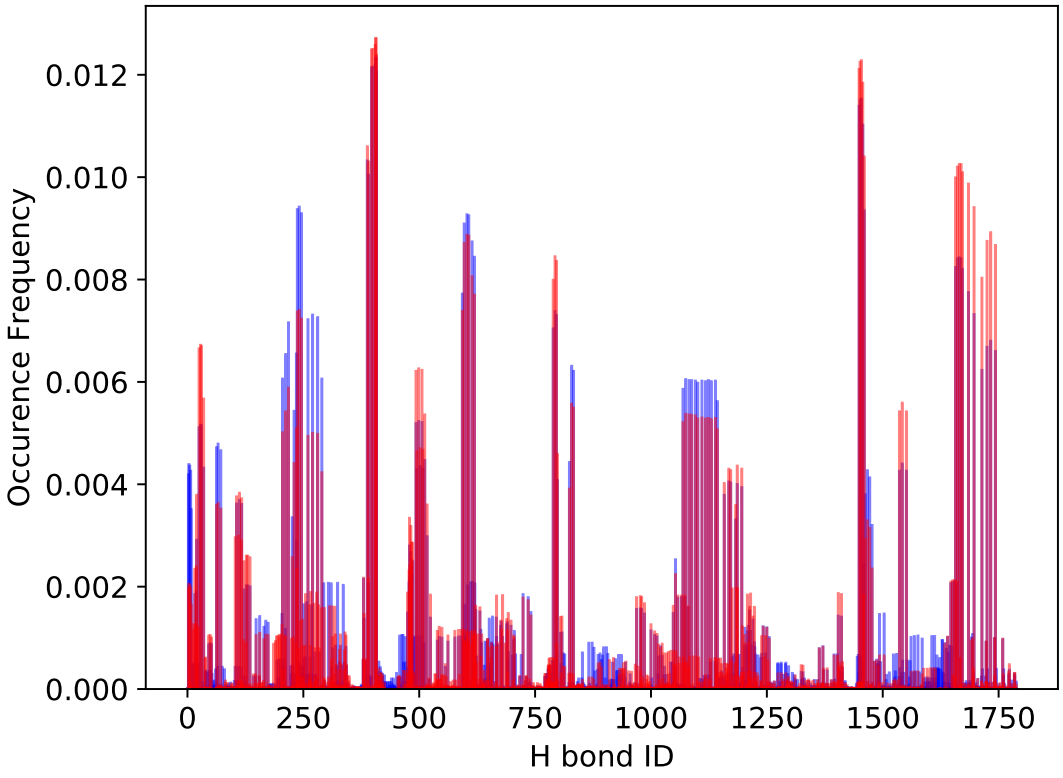

Figure S5: Map of all unique base pairs observed in oxDNA simulations of the 1.2 pN IsoII structure at $[Na^+]$=5 M. The data were split in half (12 identical replicas of 8×10^8 MD steps each), indicated by red and blue histograms. The amount of overlap gives a rough indication of whether the important secondary structures are being sampled sufficiently.

secondary structures permitted. The results of simulations of the 4.0pN design with different harmonic trap stiffnesses are summarised in Figure S7; above $\sim$ 20pN nm, the stiffness doesn't appreciably affect the force, so the assumption of fixed endpoints is reasonable. However, for other origami structures, the compliance may be relevant.

References

(1) Allen, M.; Tildesley, D. *Computer Simulation of Liquids*; OUP Oxford, 2017.

(2) Snodin, B. E. K.; Randisi, F.; Mosayebi, M.; Šulc, P.; Schreck, J. S.; Romano, F.; Ouldridge, T. E.; Tsukanov, R.; Nir, E.; Louis, A. A.; Doye, J. P. K. Introducing Improved Structural Properties and Salt Dependence into a Coarse-Grained Model of DNA. *J. Chem. Phys.* **2015**, *142*, 234901.

(3) Rohatgi, V. *An Introduction to Probability Theory and Mathematical Statistics*; Wiley Series in Probability and Mathematical Statistics; Wiley, 1976.

(4) Nickels, P. C.; Wunsch, B.; Holzmeister, P.; Bae, W.; Kneer, L. M.; Grohmann, D.; Tinnefeld, P.; Liedl, T. Molecular Force Spectroscopy with a DNA Origami–Based Nanoscopic Force Clamp. *Science* **2016**, *354*, 305–307.

Table S2: Summary of the simulations performed on the system of Nickels et al. [4], including the average ssDNA internal force in each case. Errors in force are the standard errors of a weighted mean across a number of different nucleotides at which forces were measured, as shown in Fig. S6. For almost all sets of parameters, 24 identical replica simulations were run; where fewer replicas are reported below, some runs were discarded due to the (rare) occurrence of a transition between isomer states.

Design	**Isomer**	**Secondary structures**	**k (pN/nm)**	**Salt (M)**	**# MD steps (x24)**	**approx. CPU days**	**Force (pN)**
4.0 pN	IsoI	On	16.32	5	1×10^9 ($\times$23)	85	6.54±0.05
				1	9×10^8	113	6.46±0.06
				0.5	5.8×10^8	96	5.71±0.07
				0.15	4.8×10^8 ($\times$19)	102	3.6±0.1
		Off	16.32	5	1.3×10^9	104	2.91±0.03
				0.15	5×10^8 ($\times$19)	103	2.33±0.09
	IsoII	On	16.32	5	1.3×10^9	113	6.35±0.04
				1	9×10^8	113	6.52±0.05
				0.5	5.8×10^8	96	5.43±0.06
				0.15	4.8×10^8	130	3.4±0.1
		Off	1.632	5	7.8×10^8	64	2.23±0.05
			16.32	5	7.8×10^8	64	2.61±0.05
				0.15	5.7×10^8	131	1.88±0.07
			163.2	5	8.8×10^8	64	2.61±0.05
		Off, no stacking	16.32	5	7.8×10^8	59	3.26±0.04
2.5 pN	IsoI	On	16.32	5	1.2×10^9	148	5.07±0.05
		Off	16.32	5	1.4×10^9 ($\times$23)	131	1.38±0.03
	IsoII	On	16.32	5	1.4×10^9	169	5.01±0.05
		Off	16.32	5	1.4×10^9 ($\times$23)	131	1.19±0.03
1.2 pN	IsoI	On	16.32	5	8×10^8	180	3.83±0.08
		Off	16.32	5	1×10^9 ($\times$23)	155	0.43±0.03
	IsoII	On	16.32	5	8×10^8	180	3.71±0.07
		Off	16.32	5	7.5×10^8	40	0.36±0.03

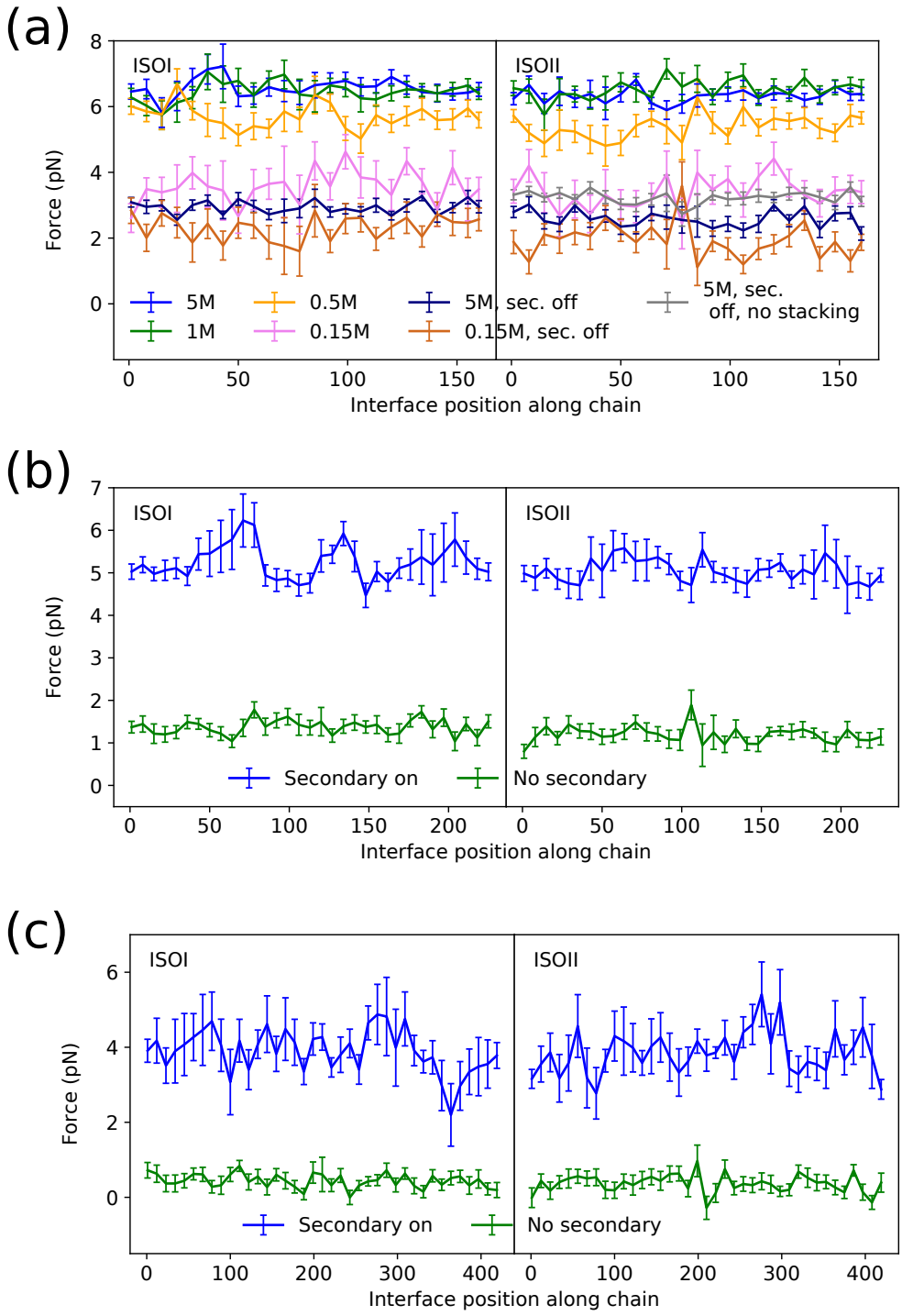

Figure S6: Net forces crossing interfaces along the ssDNA strand for Nickels et al.[4]'s (a) 4.0 pN; (b) 2.5 pN; and (c) 1.2 pN designs. Errors are standard deviations over ~ 24 identical replica simulations of $\sim 10^9$ MD steps. Results for the 4.0 pN system are shown under various salt conditions, and additionally for simulations in which both secondary structure and stacking interactions are suppressed (grey curve); the forces in this case are higher than when stacking interactions are permitted (indigo curve) because a stacked strand has a longer Kuhn length than an unstacked one, and thus has fewer conformations available to it, lowering the internal force.

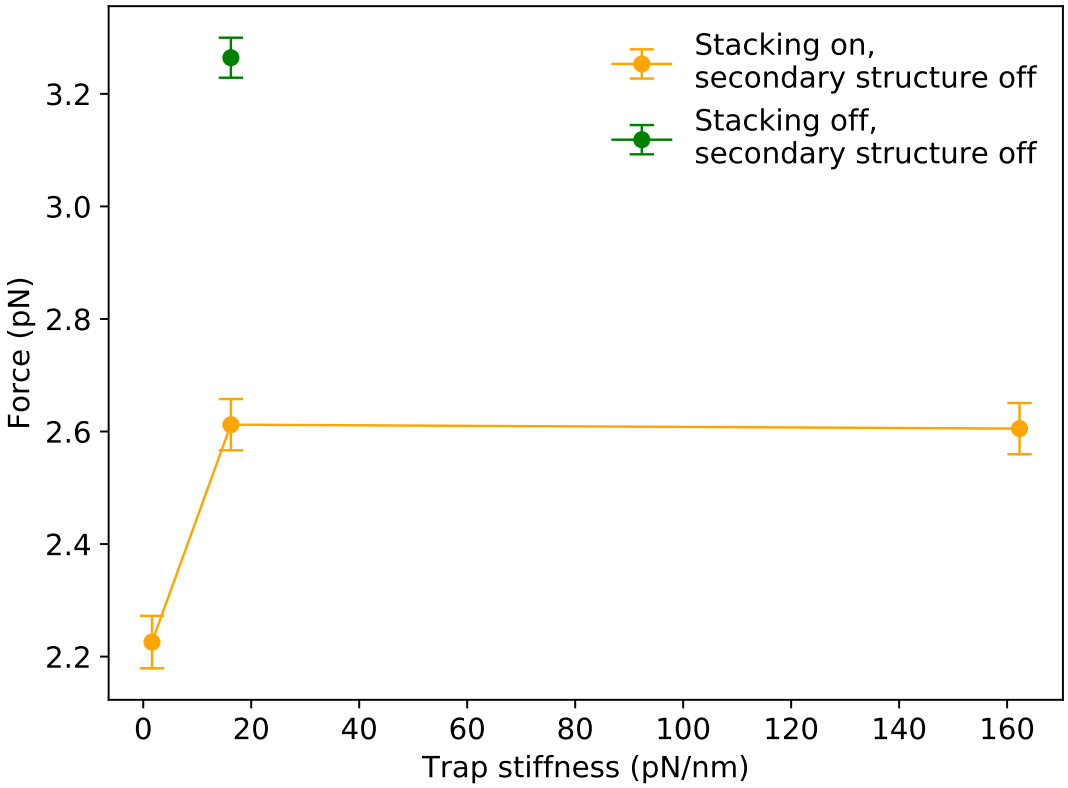

Figure S7: Effect of trap stiffness on internal force in ssDNA, for the 4.0 pN design of Nickels et al.[4]. Our simulations are performed in the regime where the force has plateaued, and origami stiffness is therefore not expected to be relevant.