

## Revealing Thermodynamics of DNA Origami Folding via Affine Transformations

Jacob M. Majikes<sup>1</sup>, Paul N. Patrone<sup>1</sup>, Daniel Schiffels<sup>1</sup>, Michael Zwolak<sup>1</sup>, Anthony J. Kearsley<sup>1</sup>, Samuel P. Forry<sup>1</sup>, J. Alexander Liddle<sup>1,\*</sup>

<sup>1</sup> Physical Measurement Lab, National Institute of Standards and Technology, Gaithersburg, MD, 20899-6203, USA

\*J. Alexander Liddle Tel: 301-975-6050; Email: liddle@nist.gov

Present Address: J. Alexander Liddle, 100 Bureau Drive, Gaithersburg, MD 20899-6203, USA

### ABSTRACT

Structural DNA nanotechnology, as exemplified by DNA origami, has enabled the design and construction of molecularly-precise objects for a myriad of applications. However, limitations in imaging, and other characterization approaches, make a quantitative understanding of the folding process challenging. Such an understanding is necessary to determine the origins of structural defects, which constrain the practical use of these nanostructures. Here, we combine careful fluorescent reporter design with a novel affine transformation technique that, together, permit the rigorous measurement of folding thermodynamics. This method removes sources of systematic uncertainty and resolves problems with typical background-correction schemes. This in turn allows us to examine entropic corrections associated with folding and potential secondary and tertiary structure of the scaffold. Our approach also highlights the importance of heat-capacity changes during DNA melting. In addition to yielding insight into DNA origami folding, it is well-suited to probing fundamental processes in related self-assembling systems.

### INTRODUCTION

The sequence specificity inherent to DNA self-assembly provides nanofabrication options orthogonal to those of existing top-down methods, enabling the construction of multifunctional, addressable 3D nanostructures. Nanoparticles(1), quantum dots(2, 3), fluorophores(4), biomolecules(5, 6), and combinations thereof can be positioned with single-nanometer precision and accuracy(7, 8) on DNA platforms. Recent work has demonstrated production at scale(9, 10), facilitating the manufacture of drug delivery vehicles(11), calibration artefacts(12), and large area, functional surfaces(4, 13) at low cost. However, while raw material scale-up is necessary for the industrial implementation of DNA nanofabrication technology, it is not sufficient. As observed in Moore's seminal 1965 paper, the semiconductor industry has seen success in part because "*No barrier exists comparable to the thermodynamic equilibrium considerations that often limit yields in chemical reactions.*"(14) If artificial macromolecular self-assembly is to replicate the success of the semiconductor manufacturing paradigm, at a minimum we must fully understand those limiting thermodynamic equilibria and be able to predict the nature of the resulting defects and their populations. This will enable the development of design, process, and assembly strategies that ensure functionality.

Unfortunately, DNA nanostructure thermodynamics and kinetics are difficult to characterize *via* the techniques typically employed by biophysicists, as are similar complex cooperative effects relevant in biological DNA. DNA origami(15) – in which a long, single-stranded “scaffold” folds by binding many “staple” oligomers – has been the focus of much work on yield and folding. Most efforts, to date, have been theoretical,(16, 17) indirect,(18, 19) or unable to resolve individual folding events.(20–23) In addition, with few exceptions(24–26), these studies have examined structural yield, i.e., fidelity with respect to a designed shape, as opposed to functional yield, i.e., inclusion of all necessary active components. The large number of potential design configurations means that measurement techniques must be adapted to nanofabrication contexts, and they must also have sufficient throughput to adequately explore an extensive parameter space.

In this work, we investigate the effect of fold distance and persistence length on the thermodynamics of an initial folding event using high-throughput, quantitative Polymerase Chain Reaction (qPCR) equipment. These measurements are enabled by our modular fluorescent melt-curve reporter system, which minimizes extraneous signals, in combination with an affine transformation technique that significantly improves melt-curve baseline and background correction by iteratively collapsing numerous replicate measurements of a sample onto a single curve through linear parameters representing known sources of variation(27). While this study is motivated by a desire to understand the details of the origami assembly process, the methods and results are applicable to numerous types of DNA-based self-assembling systems.

Characterization of folding events in DNA origami is problematic due to the sheer number of hybridization events. One example origami structure, the Rothemund tall rectangle(15), consists of a 7.2 kilobase M13 bacteriophage viral scaffold bound together with  $\approx 200$  synthetic staples, each of which binds two or three different and disjoint sub-sequences of the scaffold(15). It must be emphasized that this results in  $\approx 600$  reversible hybridization reactions, of typically 6 to 24 bases in length, with varying degrees of independence, collectively responsible for  $\approx 200$  forced links in the topology of the scaffold. Design choices include scaffold sequence permutation, staple motifs, and scaffold routing – all of which influence the folding process. Attaining the imaging resolution required to observe individual binding events is a significant challenge and has, so far, only been performed *post mortem*, on assembled structures(25). To develop an understanding of design-assembly relationships, functional staple yields must be characterized for numerous “equivalent” structures of varying design under various processing conditions.

These design-assembly relationships for DNA origami are a function of the interplay between several cooperative energetic contributions(17, 25). The primary contributor to hybridization comes from base stacking and base pairing as described by the nearest-neighbor model(28). Additional contributions come from coaxial end-to-end base stacking with adjacent staples, and entropic effects, as sub-sequence binding events change the topology of the scaffold (16, 17, 28). As each staple binds, it reduces the entropic penalties for further folds and provides base stacking for its neighbors. It also changes the equilibrium constant for the other sub-sequences on that staple from a bimolecular to a unimolecular form. While this is not the focus of this paper, we note that the complex problem of

cooperative folding can be addressed using our programmable test system. Here we only explore zeroth order cooperative effects, i.e., changes in conformational entropy.

While the thermodynamics and kinetics of biological DNA hybridization(28–30) and the basic theory for loop entropy in short (tens-of-bases) ssDNA(28) and in long (kilobase) dsDNA(31, 32) are all established, they are not directly applicable to DNA origami. The latter begins as long ssDNA loops, likely with intrinsic secondary structure, which are folded into smaller and smaller mixed ssDNA/dsDNA loops until the entire structure comprises dsDNA helices bound together by Holiday junctions. Neither the loop entropy model for short ssDNA nor that for long dsDNA is sufficient to predict conformational entropy in these complex, dynamic systems of mixed persistence length. Additionally, the large thermal processing range for DNA nanofabrication (approximately 25 °C to 80 °C), as compared to normal physiological temperatures (36.5 °C to 37.5 °C), has an important consequence: the temperature dependence of the enthalpy and entropy, expressed as the change in heat capacity on melting,  $\Delta C_p$ , cannot be neglected. While the significance of  $\Delta C_p$  for DNA melting is established(33, 34), a general consensus on typical values and sequence dependence is lacking, and it is commonly neglected entirely in secondary structure prediction(28, 35). It is worth noting that  $\Delta C_p$  shifts the hybridization energetics in proportion to the difference between the evaluation temperature and the melt temperature and therefore becomes critical over the broad temperature ranges used in nanofabrication.

So far, we have discussed the importance of and need for a better understanding of the thermodynamics of DNA origami assembly. Below, we examine the limitations of current techniques when applied to this problem.

Experimental characterization of DNA origami systems with existing biological techniques is often as fraught as application of existing theory. For smaller ssDNA loops, it is common to use UltraViolet-Visible (UV-Vis) spectroscopy to measure melt curves or differential scanning calorimetry (DSC) to directly measure the melting energetics(36). In both cases, a significant mass of DNA is required, and both techniques measure signal from all DNA in the sample. Sufficient resolution to distinguish one approximately 80-base staple or, worse, an approximately 24-base staple sub-sequence, from the other 7.2 kilobases of DNA in the system is not yet available. For measurement of cyclization of long dsDNA, ligation reactions are often quenched with Ethylenediaminetetraacetic acid (EDTA), then gel electrophoresis is used to quantify the relative concentration of cyclized/uncyclized material for sequential snapshots of the reaction(37). In the case of origami, there is little to no gel mobility shift associated with a single staple fold of the scaffold. Intercalating dye fluorescence may be used to determine melt temperatures,(38) but there is a lack of consensus regarding baseline correction techniques(39) to which thermodynamic parameter extraction, *via* van't Hoff analysis(33, 40, 41) plotting  $\ln(K)$  vs  $1/T$ , is exceptionally sensitive. As a result, the melt temperature,  $T_m$ , is taken to be the only reliable parameter, and is often the only one reported (42). The variant of van't Hoff analysis, in which  $T_m$  is plotted against DNA concentration, common to both fluorescence and UV-Vis, is not usable for scaffold folding itself. This is because, whether a fold occurs due to the formation of secondary structure between disparate domains of the scaffold, or as the result of a second domain of an already-bound staple hybridizing, the change in topology is a unimolecular process, requiring

neither the addition (removal) of another strand to (from) the structure. Therefore, the concentration terms cancel out in the equilibrium constant.

qPCR equipment, which enables high-throughput fluorimetry, has been used with intercalating dyes to measure anneal/melt curves of whole origami(16, 17, 19), while Förster resonance energy transfer (FRET) pairs have been used to measure local melting of origami(43), inter-origami base stacking(44), and DNA tile assembly(45). However, absent a robust baseline correction method, these techniques cannot provide quantitative information. Baseline correction is difficult because fluorophores, whether FRET pairs or intercalating dyes, have a temperature-dependent fluorescence efficiency that is also sensitive to their local environment. This includes neighboring ssDNA/dsDNA transitions, as well as the DNA sequence immediately adjacent to the fluorophore.

However, if the issues associated with local environment effects can be mitigated, and the baseline correction made rigorous, the specificity and sensitivity of FRET reporters makes them attractive for measuring folding events. Here, we overcome these difficulties by implementing a modular reporter system that maintains a uniform fluorophore environment for every fold measured and introduce an analysis approach that alleviates baseline correction issues. In this analysis we make, and validate, the assumption that we can separate the effect of the variables subject to error from the intrinsic signal. In other words, the physics responsible for FRET, i.e., the relative proximity of donor to acceptor, is the same for all samples, while the measured signal depends on the sample volume, well location, detection/excitation efficiency, alignment of equipment optics per well, etc. This assumption enables us to apply an affine transformation (see Eq. 2) to the melt curve and significantly reduce subjectivity in the baseline correction process, replacing it with a more mathematically rigorous approach. With a reproducible fluorescence signal, and a robust baseline correction method in hand, we make use of the high throughput of qPCR to explore the large parameter space covering the effects of fold distance and persistence length.

## MATERIAL AND METHODS

### Test system:

Using the design capabilities of structural DNA nanotechnology, we created a minimal, single-fold test system to observe the effects of loop entropy without the occurrence of numerous simultaneous events which can significantly obfuscate analysis.

An illustration of the measured folds is shown in Fig. 1a. This system is coincidentally similar to those being used for subattomole detection of microRNA(46). Despite important distinctions between the thermodynamics of these systems, both make a strong case for the utility of simplified DNA self-assembly systems. We examined the melting of an identical sequence of DNA in a simple bimolecular no-fold system as well as with 13 different folding distances dividing the loop into figure-eights. The system which does not fold the scaffold will be referred to as the hybridization-only system. An important feature of this system is that we may vary the **effective, or apparent** persistence length,  $L_p$ (47–49). This is critical since, for the same length of biopolymer,  $L_p$  controls the effective number of freely-jointed units in the polymer chain, which in turn controls the number of possible configurational

microstates and the entropic folding penalty. The **apparent persistence length** also determines the minimal fold distance below which enthalpic bending costs become a significant factor. Since the scaffold is circular, we only probe folding distances up to half the scaffold length.

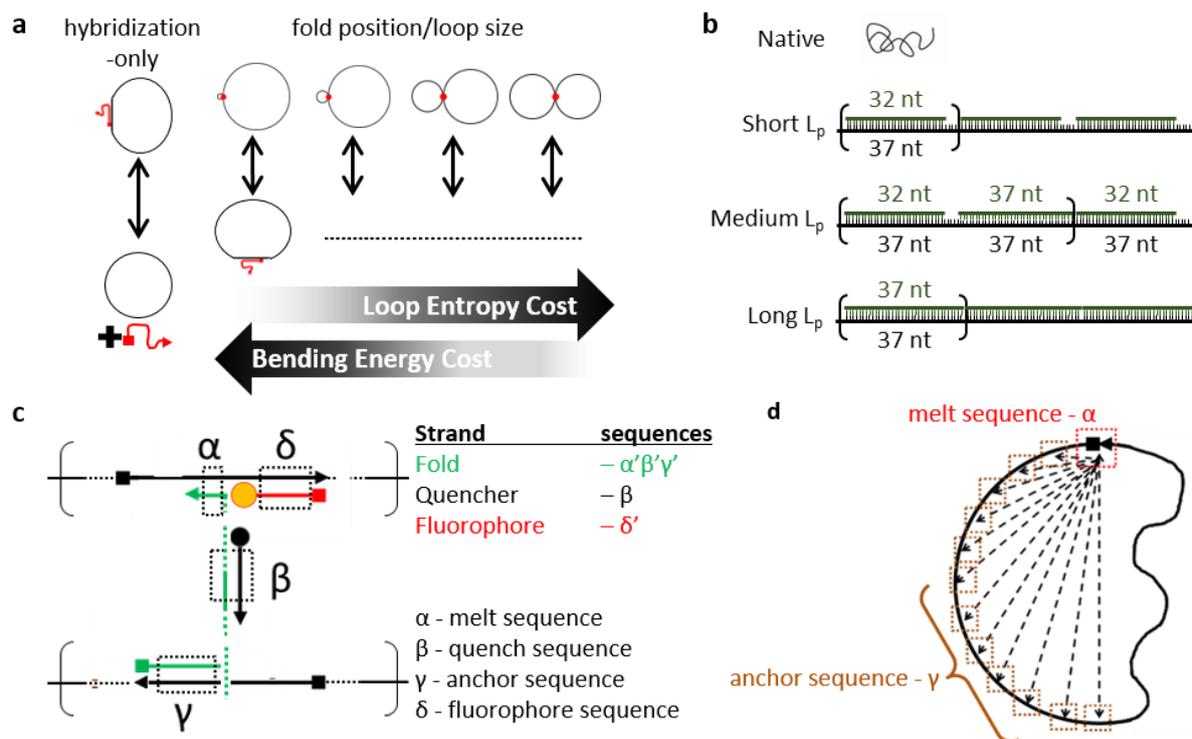


Figure 1 a) folding transition schematics for bimolecular hybridization-only (H.O.), i.e., staple binds to scaffold, and unimolecular, i.e. bound staple hybridizes fully, fold reactions. b) Persistence length strands: name, anticipated  $L_p$ , and schematic of oligomers/scaffold, c) Reporter scheme: cartoon of fold and reporter strands on the scaffold and description of sub-sequences for fold and reporter strands where X' indicates the complement to X. d) Relative positions of melt and anchor sub-sequence complements on M13 scaffold.

We adjust the **apparent** persistence length,  $L_p$ , by varying how much of the scaffold is either dsDNA and ssDNA. To systematically do so the scaffold was divided into 37 base segments; two sets of complimentary oligos for each segment were ordered, namely a 37-base full complement and a 32-base complement omitting the last 5 bases at the 3' end. The scaffold  $L_p$  was then varied by repeating motifs of these complement strands as shown in Fig. 1b. The short and long  $L_p$  systems contained only the 32 and 37-base complements respectively, while the medium system alternated between the 32 and 37-base complements across the scaffold. In addition, a native system with no complements, and a blocked system with only 32-base complements in areas of high predicted secondary structure, SI section 1, were examined. For each  $L_p$  system, a pool of complement strands was created for all segments except those complementing the anchor, melt, and fluorophore sequences for that fold. This pool was added in 5x excess of the scaffold.

While the  $L_p$  of ssDNA and regularly-nicked dsDNA(50) have been measured, determining the persistence length of these composite constructs is beyond the scope of this work, **particularly since there is some controversy regarding the persistence length of short dsDNA segments (51–54).**

However, we provide a rough estimate in SI section 3, for later use in allowing us to implement configurational entropy and bending enthalpy models.

Our reporter system must meet two criteria to effectively measure and compare fold distances. First, it must measure melting of the same sequence of DNA for each fold, to simplify comparison. Second, the local environment around the FRET pairs should be as uniform as possible across samples. To satisfy these constraints, we used the same fluorophore/quencher-labeled oligomers for all folds, both to minimize sequence-dependent changes in fluorescence as well as variability in concentration associated with measuring concentration and pipetting multiple strands.

One potential drawback of this approach is that the reporter strand does not create the Holiday junction-like structure found in typical origami. However, since the structure formed is the same between the hybridization-only control and the test samples, it has no effect on the excess energies due to changes in fold distance or persistence length.

Our reporter scheme is illustrated in Fig. 1c. The reporter comprises three strands and the scaffold, with 4 domains/sub-sequences within those strands. The melt, quench, and fluorophore sub-sequences were identical sequences for all samples, while the presence/complementary location of the anchor sub-sequence in the fold strand changed between samples to enforce fold distance. The lengths of these sub-sequences were chosen so that the melt sub-sequence  $T_m$  is more than 10 °C below that of any other sub-sequence. The right side of Fig. 1d illustrates the set of folding distances. For all 12 replicates of each folding distance at each persistence length, the sample was pipetted, mixed, annealed, and finally melted while measuring fluorescence intensity. Further details, such as the sample preparation scheme to minimize unwanted errors and fully capture pipetting variation are addressed in the methods.

### **Data analysis: baseline and background correction**

The signal in optical melt curves, whether from FRET, UV-VIS or intercalating dye measurements, does not solely measure the completion of a hybridization reaction. The source of the signal, whether absorbance or fluorescence, is temperature dependent both in the pre- and post-melt states. Additionally, various other contributors to the measured signal, such as sample/plate autofluorescence(55) and pipetting error, will lead to variability in background intensity between replicates of the measured curves. Inherently, the methods used to remove both temperature dependent baseline and background signal will change the data in the melt transition, making rigorous correction critical.

Baseline correction of optical melt curves typically consist of a two-window function fit across user-selected windows for the high and low temperature shoulders of the sigmoid(40, 56). These fits are then extrapolated across the transition of interest, and the measured signal is corrected via Eq. 1, where  $h(T)$ ,  $l(T)$ ,  $m(T)$ , and  $\sigma(T)$  are the high-temperature plateau fit, low-temperature plateau fit, measured signal, and intrinsic signal, respectively.

$$\sigma(T) = \frac{m(T)-l(T)}{h(T)-l(T)} \quad [1]$$

The appeal of the standard two-window fit is its simplicity and nominal ability to address both temperature dependence and background signal. However, this approach is only valid if the fit function is truly representative, if the fit window does not include any points from the transition, and if the sources of measured signal variation only shift or magnify the melt curve rather than distorting it.

While these assumptions may be true for UV-Vis measurements, they are problematic for fluorescence measurements. Although examples of non-linear baseline functions have been used(39), it is not clear that any of these functions, linear or otherwise, accurately represent the complex mix of factors contributing to the baseline fluorescence. Additionally, the final analysis result can be highly dependent on user-selected fitting windows, especially if there is noise in the measurement. Given that there are not clear objective figures of merit for this window selection, the subjective nature of the two-window fitting approach creates significant uncertainty.

To address these shortcomings, we apply an approach based on affine transformations. Critically, we do not assume the functional form of the baseline or background. Rather, we assume that the physical processes contributing to the baseline are the same for every replicate of a sample, but we allow the scale of the intrinsic signal to be a variable, as a result of, e.g., pipetting error. In other words, we assume that the measured signal is related to the intrinsic signal through linear combinations of a finite number of physical processes, whose relative contributions are determined as part of the analysis. With a sufficient number of sample replicates, the figure of merit for optimization in the analysis is how well all replicate melt curves collapse onto a single intrinsic melt curve via affine transformation, as shown in Fig. 2(27).

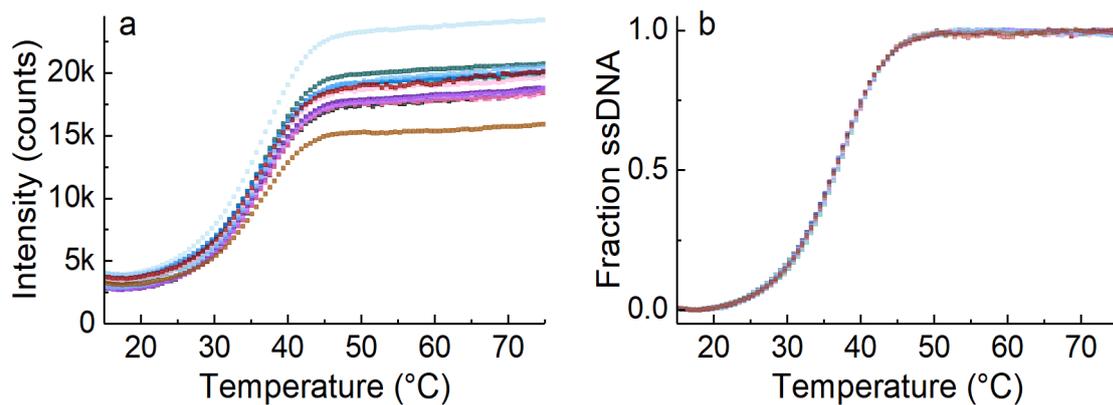


Figure 2. Hybridization-only (H.O.) control for the short  $L_p$  scaffold before a), and after b), affine transformation. The empty plate background,  $B(T)$ , and fluorescence temperature dependence baseline,  $R(T)$ , are subtracted from a) to b), but are separately measured in control experiments.

Unprocessed melt curves, acquired as described in the methods, are shown in Figure 2a. The samples share the same characteristics, i.e. melt temperature and functional form, arising from the intrinsic signal. Affine transformations were applied to collapse these curves and reveal the intrinsic signal in Figure 2b.

A detailed description of this process is included in SI. Section 2. In brief, we first use a set of empty wells to determine the plate autofluorescence and optics alignment-related background,  $B(T)$ . Next, we use a set of wells containing all strands, bar the folding strand, to determine the fluorescence-temperature-dependent baseline,  $R(T)$ . The functional forms of the baseline and background determined from the control experiments are then used in combination with the measured melt curves,  $m_i(T)$ , to extract the intrinsic melt signal,  $\sigma(T)$ . This is done by finding  $a_i$ ,  $b_i$ , and  $c_i$  affine parameters for each replicate,  $i$ , such that the measured replicate melt curves,  $m_i(T)$ , collapse to the

single intrinsic melt signal,  $\sigma(T)$  [Eq. 2]. In all such cases, the relative error on the collapse is a useful metric for the uncertainty of the intrinsic melt signal.

As the transformations are applied directly to experimental data we never need to, nor do we, specify functional forms for the baseline or background. This frees us from making poor modeling choices about the actual behavior of  $\sigma$ ,  $R$ , and  $B$ , aside from needing to satisfy an equation having the general form of Eq. 2. *For this analysis, an additional optimization constraint is imposed, specifically requiring a monotonic slope at the high temperature plateau.* While there are enough data points to represent the entire melt curve, there were insufficient low-temperature points on all curves to impose an additional low-temperature monotonicity constraint.

$$m_i(T) = a_i\sigma(T)R(T) + b_iB(T) + c_i \quad [2]$$

For the analysis presented here, we assume that FRET quenching reduces the fluorescence intensity of the folded state to a sufficiently low level that its temperature dependence is negligible. This may not be true for every system and buffer condition.

One subjective choice in the affine transformation is application of a monotonicity constraint, and the window over which it is applied. The data presented here, Fig 7 in particular, was analyzed with only a high temperature monotonicity constraint. There are insufficient data points on the low temperature plateau of the control and long fold samples to apply a low temperature monotonicity constraint without some sensitivity to window choice. This is illustrated in SI section 12.

Eq. 2 may be easily modified for other systems. For example, for intercalating dyes the free-dye fluorescence temperature dependence cannot be neglected. Additionally, terms for free-dye fluorescence would be supplemented with terms for free-dye fluorescence in the presence of ssDNA. These possibilities are discussed in our best-practices paper for this analysis(27).

One limitation of this method is that the curve-collapse optimization is most effective when only one physical process dominates the measured signal.

As a counter-example, if variations in pipetted volume contribute a similar change in signal as the background, the optimization can become inefficient or even infeasible. Additionally, while this approach removes much subjectivity from the analysis, the choice of initial equation, e.g. Eq. 1, and the linearity constraints at high and low temperatures are still choices for the experimentalist.

*Ultimately, implementation of affine transformations in analysis of optical melt curves allows us to simultaneously perform baseline and background corrections with reduced subjective, model-choice errors.*

### **Data analysis: thermodynamics extraction**

Calculation of  $\Delta H$  and  $\Delta S$  is performed *via* van't Hoff analysis; in which the equilibrium constant,  $K$ , is calculated and fitted to Eq. 3 as a function of temperature.

$$\ln(K) = \frac{-\Delta H}{RT} + \frac{\Delta S}{R} \quad [3]$$

Equilibrium constants are calculated using Eq. 4 & 5 for the hybridization-only control (bimolecular) and folded samples (unimolecular) reactions, respectively. Where [ssDNA] and [dsDNA] refers to the concentration of the scaffold in the melted, or dehybridized, and the annealed, or hybridized, state respectively, while [excess] refers to the relative excess of the fold strand to the scaffold. Each of

these is treated as a fraction multiplied by the total scaffold concentration, which in Eq. 5 cancels out, as one would expect a unimolecular reaction to be concentration independent. As previously described, this is because the reporter signals when the melt sub-sequence of the fold strand hybridizes, Fig. 1c, at which point the more stable anchor sub-sequence of that strand is already bound to the scaffold. As such, the folding (unfolding) of the scaffold neither consumes (generates) a free staple strand, and is effectively unimolecular in nature. The exception to this is the hybridization-only control, which does not fold the scaffold and is bimolecular

$$K = \frac{[ssDNA][foldStrand]}{[dsDNA]} = \frac{[ssDNA][excess-dsDNA]}{[dsDNA]} = \frac{[ssDNA][excess-1+ssDNA]}{[1-ssDNA]} \quad [4]$$

$$K = \frac{[ssDNA]}{[dsDNA]} = \frac{[ssDNA]}{[1-ssDNA]} \quad [5]$$

Estimating  $\Delta S$  requires extrapolation to  $1/T=0$ , which is to say, infinite temperature. Small uncertainties in the extrapolation of  $\ln(K)$  as a function of  $1/T$ , give rise to large uncertainties in the value of the intercept at  $1/T = 0$ , i.e.,  $\Delta S$ . While a linear relationship between  $\ln(K)$  and  $1/T$  is often assumed in the van't Hoff analysis, the change in heat capacity,  $\Delta C_p$ , results in a nonlinearity. The relationship between  $\Delta C_p$ ,  $\Delta H(T)$  and  $\Delta S(T)$  is described in Eq. 6.

$$\Delta C_p = \Delta H/\Delta T = T \cdot \Delta S/\Delta T \quad [6]$$

$$\Delta T = T_{ref} - T_m \quad [7]$$

$T_{ref}$  an arbitrarily chosen temperature at which  $\Delta H$  and  $\Delta S$  are evaluated. It is also worth noting that fitting  $\Delta C_p$  from curvature in van't Hoff analysis of optical melt curves is not always successful(35). Limited data sampling, in units of  $1/T$ , makes the typical fitting of the curvature in that space highly dependent on the quality of the data. For a thorough discussion of this, we direct the reader to the relevant section in the review by Mikulecky and Feig(33).

### Loop entropy and worm-like chain bending model

The change in entropy resulting from converting a linear polymer to a loop can be modelled using the Jacobson-Stockmayer entropy extrapolation(28). This may be readily rearranged, SI section 3, to give an estimate of the loop-to-figure-eight transition entropy shown in Eq. 8. Where, in number of bases of ssDNA,  $N_0$  is the length of the original loop, and  $N_1$  and  $N_2$  are the lengths of the two-resulting figure-eight portions illustrated in Fig. 1.  $\Delta S_{ref}$  is the entropy lost in forming an ssDNA loop comprising  $C$  nucleotides. This value was experimentally determined for small hairpins, and is extrapolated to larger loops. In our predictions we use the values for a 30 nucleotide hairpin(28). The relationship between  $\Delta S$  and the j-factor terminology used for long dsDNA cyclization is shown in Eq. 9.

It is worth noting that the units for  $C$  are effectively "nucleotides of ssDNA." The value of  $C$ , together with the prefactor, empirically represent the effective number of bases per segment length. Applying Eq. 8 to DNA nanofabrication systems requires a unit conversion from bases of the scaffold systems

shown in Fig. 1 to bases of ssDNA. The tentative conversion factors we used are discussed in SI section 3.

$$\Delta S_{fold} = \Delta S_{ref} + 2.44R \ln \left( \frac{N_1 N_2}{30 N_0 C} \right) \quad [8]$$

$$\Delta S_{fold} = -R \ln(j) \quad [9]$$

Implementation of the worm-like chain (WLC) model for the energy of bending,  $E_b$ , per unit length of the arc length of the polymer is given in Eq. 10, where  $R$  is the radius of curvature of the bend(57). As such the  $\Delta H$  associated with bending is given by Eq. 11, where  $L_0$  is the contour length of the DNA.

$$E_b = \frac{L_p k_b T}{2R^2} \quad [10]$$

$$\Delta H_{bend} = \frac{L_p k_b T}{2R^2} * L_0 = \frac{L_p k_b T}{2 \left( \frac{L_0}{2\pi} \right)^2} * L_0 \quad [11]$$

As noted above, our mixed dsDNA/ssDNA system is complex, and a detailed microscopic treatment, beyond the scope of this paper, would be needed to accurately describe its mechanical properties. We therefore emphasize that Eqs. 8-11 serve only as a rough guide to the relevant energetics.

### FRET Reporter and folding strands

As noted above, it is vital to maintain a uniform environment for the fluorophores for all measured systems, including DNA sequence. We achieve this by implementing the fluorophore-quencher system illustrated in Fig. 1c, in which the sequences adjacent to the fluorophore (JOE), and the quencher (Iowa BlackF), do not change. The folding distance (Fig. 1a) is controlled by a 42-base anchor sequence that is adjusted to bind at different locations around the loop of the scaffold strand. To minimize non-fold-related fluorescence background, we control the stoichiometry of our system (Table 1) to ensure that, at low temperature, all fluorophores are bound to a scaffold resulting in a consistent local environment.

To explore the effect of zeroth-order topology changes on melting behavior, we use 13 different folding distances, controlled by the binding location of the anchor sequence (Fig. 1d). The fluorophore, quencher, and anchor sub-sequences all have a predicted  $T_m$  of  $\approx 70$  °C under our buffer conditions as evaluated by typical nearest neighbor models(30, 58). These temperatures are well above the predicted 35 °C  $T_m$  of the melt sub-sequence, ensuring that the fluorophore is unquenched only when the melt sub-sequence dehybridizes.

Buffer conditions were chosen to match standard DNA nanofabrication protocols: 50 mMol/L Cacodylate buffer, 12.5 mMol/L Mg Acetate, 1 mMol/L EDTA, pH 7.2. Cacodylate was chosen as it buffers pH well across a wide temperature range. DNA sequences are given in section 11 of the SI.

### Persistence Length

The systems with different apparent persistence lengths illustrated in Fig. 1b are the native scaffold, short (repeating 32 base dsDNA interspersed by 5 bases ssDNA or free scaffold), medium (repeating 32 bases dsDNA, 5 bases ssDNA, 37 bases dsDNA, followed by 32 bases dsDNA), and long (repeating 37 bases dsDNA separated by nicks) This varies  $L_p$  between that of ssDNA ( $L_p \approx 1$  nm to 3

nm(59)) and regularly nicked dsDNA ( $L_p \approx 14$  nm to 43 nm depending on counterion charge and concentration(50)).

The lowest persistence length, ssDNA native M13 scaffold, will in practice contain a great deal of secondary structure. We identified some scaffold locations which could form > 10-base secondary structure across long scaffold distances (SI. section 1), with  $T_m$  on the same order as that of the melt sub-sequence. As such, we included an additional scaffold system to investigate the effect of this secondary structure. In this system, we introduce a number of “blocking” strands, designed to hybridize with and out compete this long-distance secondary structure.

We denote the variants of our scaffold, in order of persistence length, as “native”, “blocked”, short, medium, and long. A description of the blocked system and potential secondary structure is given in SI section1, while the sequences for strands creating different apparent persistence lengths,  $L_p$ , are given in SI section 14.

## Melt experiments

Our experimental matrix consists of 13 fold distances across scaffolds with 5 different persistence lengths. In addition, it is desirable to replicate each configuration a sufficient number of times to assess the variation due to pipetting while also identifying and removing spurious data sets such as those compromised by bubbles in the wells. We therefore choose to perform 12 replicates of each experiment to ensure high-quality data.

The melt protocol consisted of an initial denaturing step (80 °C for 1 min) an annealing sequence (75 °C to 15 °C, 0.61 °C steps, 3.5 min hold) and finally melting (15 °C to 75 °C, 0.61 °C steps, 3.5 min hold) in which the fluorescence was measured. The raw fluorescence intensity in the green channel, rather than the calculated multicomponent intensity, was used, as uncertainties in the instrument’s multicomponent intensity calculation were unknown.

	<i>Absolute Concentration</i>	<i>Relative excess (to scaffold)</i>
<i>Scaffold</i>	45 nMol/L	--
<i>Fold Strand</i>	225 nMol/L	5x
<i><math>L_p</math> Strands</i>	225 nMol/L	5x
<i>Quencher</i>	450 nMol/L	10x
<i>Fluorophore</i>	30 nMol/L	0.67x

Table 1. Strand concentration, absolute and relative to scaffold.

Samples were distributed over two plates for each persistence length set. These plates consisted of 12 replicates of each fold distance, 12 replicates of a fluorescence baseline control containing no fold strand, and 24 empty background wells.

To ensure a consistent sampling of pipetting error simultaneously with manageable sample preparation, 12 replicates of all buffer and strands except the fold strand were independently prepared. These master replicates were used as a base stock for the 12 replicates of each fold distance by multichannel pipettor through the two plates for each scaffold persistence length. As such,

pipetting variability within any fold distance is uncorrelated, but pipetting variation is correlated between the N<sup>th</sup> replicate of all folding distances.

The qPCR equipment had a certified temperature precision of 0.31 °C, and an optical shot noise of approximately 50 to 100 intensity counts with experiments measuring an intensity change of  $\approx 0$  to  $\approx 20,000$  intensity counts.

### **Differential Scanning Calorimetry**

DSC scans were performed on equipment with a 300  $\mu\text{L}$  active volume, baseline repeatability of 0.028  $\mu\text{Watts}$ , and short-term noise of 0.015  $\mu\text{Watts}$ , courtesy of the University of Maryland/NIST Institute for Bioscience & Biotechnology Research (IBBR).

While the buffer conditions were identical to the melt experiments performed in qPCR equipment, the strand concentrations and excesses were modified to obtain sufficient signal. Specifically, strand concentrations were much higher and were kept at a 1:1 ratio to avoid measuring the formation of secondary structure in excess strands.

To identify the annealing peaks, scans were performed on subsets of the reporter with increasing numbers of strands. Initial scans were run with equimolar melt and truncated scaffold strands at 90  $\mu\text{Mol/L}$ . This sample was removed from the DSC, measured, and equimolar fluorophore strand was added for a final concentration of  $\approx 75 \mu\text{Mol/L}$ . The sample removal was repeated, and the quencher strand was added for a final concentration of  $\approx 62 \mu\text{Mol/L}$ . Reuse of the same sample allowed confirmation of individual peak identities.

This final sample containing equimolar concentrations of all strands was then pipetted to a 96 well plate and run in the qPCR for comparison. Unfortunately, running a full 10  $\mu\text{L}$ , the minimal sample volume, at the same concentration as the DSC sample would create signal well outside the dynamic range of the detector in the qPCR unit, which would result in significant measurement error. As dilution could significantly change the melt temperature, we chose to take 0.76  $\mu\text{L}$  of DSC sample and pipetted it into 9.25  $\mu\text{L}$  of mineral oil (a mixture of alkanes and cycloalkanes) in the PCR plate. As the mineral oil is not miscible with the water, and floats to form a capping layer, this allowed us to measure the DSC sample without losing water to evaporation. All scans were performed from 5 °C to 100 °C at a rate of 1 °C/min.

## **RESULTS**

### DSC confirmation of fluorescent reporter:

We use DSC as an orthogonal measurement technique to confirm that the fluorescence reporter system is measuring the anticipated reaction. We do this using a reduced system that allows direct comparison between DSC and fluorescence measurements.

Given that the scaffold is approximately seven hundred times more massive than the 10-base melt sub-sequence, a DSC sample with sufficient melt sub-sequences to provide measurable signal would strain solubility, in addition to being exorbitant in cost. It would also be difficult to analyze due to the additional noise associated with 7.2 kilobases of DNA. DSC was therefore run against a synthetic model system identical in sequence to the reporter, but with a truncated scaffold 100 bases in length. Additionally, the fold strand contained no anchor sub-sequence. All sequences and fluorophores were otherwise identical. In other words, the model system differs from the bimolecular hybridization-only system by omission of 7.15 kilobases of scaffold. Details of the DSC scans are addressed in the methods.

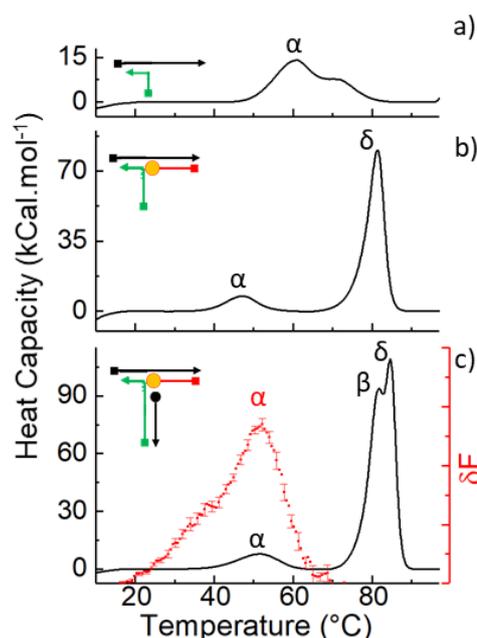


Figure 3. DSC scans of reporter analog: a) only fold and scaffold strands. b) fold, scaffold, and fluorophore strands. c) all reporter strands. Red curve, derivative of fluorescence for the same sample. Peaks and sub-sequences labeled to match

	<b>Peak</b>		
	$\alpha$	$\beta$	$\delta$
<b>Nearest-Neighbor Model</b>	54.9 °C	79.8 °C	80.9 °C
<b>DSC</b>	51.0 ± 2.3 °C	81.2 ± 0.5 °C	84.7 ± 0.3 °C
<b>Fluorescence (maximum of <math>\delta F/\delta T</math>)</b>	54.4 ± 0.25 °C	N/A	N/A

Table 2.  $T_m$  for reporter from Nearest-Neighbor, DSC, and Fluorescence. Uncertainties are one standard deviation derived from three repeat runs on each sample.

The results of DSC scans are shown in Fig. 3. From a) through c), the sample was removed from the DSC and an additional strand was added, as indicated by the inset schematics in Fig. 3. As the strands were added sequentially, we can readily identify the peaks associated with the dissociation of the sub-sequences illustrated in Fig. 1c:  $\alpha$  (melt),  $\beta$  (fluorophore),  $\gamma$  (quencher). Table 2 shows the calculated  $T_m$  for each sub-sequence. As strands are added the shift in the  $\alpha$  peak and loss of the secondary shoulder apparent in Fig. 3a versus Figs. 3b and 3c suggests that there is some unintended interaction between the nominally non-binding regions of the fold and scaffold strands, enhancing the stability of the intended fold-scaffold duplex. The shifts in both  $\alpha$  and  $\delta$  peaks between Figs. 3b and 3c are likely due to minor base stacking between the melt and fluorophore strands.

The  $T_m$  shown in Table 2 display a reasonable match with nearest-neighbour model predictions. The progression of, and the peaks in Figs. 3a-c, compared to the fluorescence measurements shown in Fig.

3c (red), confirm that the FRET reporter is probing the intended melt sequence illustrated in Fig. 1c. It is reasonable to infer that the fluorescence melt curves measured here are coming from the melt sub-sequence,  $\alpha$ , rather than from some spurious side reaction.

$\Delta H$  extracted from the DSC data shown in Fig. 3 is addressed in SI section 4. These values are not discussed here due to the significant uncertainty introduced by the ambiguity of the DSC baseline correction arising from the close spacing of the peaks.

### **Melt experiments**

The melt curves shown in Fig. 4a, demonstrate the low noise resulting from affine transformation across 12 replicates. By design, the melt sub-sequence responsible for these melt curves does not change. As such, all differences between these samples should result from the shift between unimolecular and bimolecular reaction types, entropic folding penalties, and potential enthalpic helix-bending penalties, where the latter two result from, e.g., different conformational microstates.

The trends shown in Fig. 4a, for the short persistence length, match with thermodynamic intuition. The hybridization-only sample is less stable than the others as it involves a bimolecular reaction, at low concentration, rather than a unimolecular reaction. For the folded samples, increasing fold distances shift the melt curve to lower temperatures indicating an increasing entropy change on melting, as

expected. The error bars shown in Fig. 4, and all other figures presented, represent one standard deviation above and below the mean of the replicates.

Figure 4b shows van't Hoff plots for the folded samples. The hybridization-only sample is not shown here as its equilibrium constant is calculated for a bimolecular reaction, and is therefore outside the plotted range. The van't Hoff plots exhibit a slight curvature associated with  $\Delta C_p$ .

The  $T_m$  from these curves is reported on a semi- $\log_2$  plot in Fig. 5. While changes in  $T_m$  are specific to this system, they effectively illustrate the energetics associated with these changes in fold distance and  $L_p$ .

It is apparent from Fig. 5 that the melt temperature of the native scaffold systems generally occurs at a higher temperature than the other  $L_p$  systems. This is consistent with the varied and extensive secondary structure expected in the native scaffold. Due to this local secondary structure, in the form of hairpins, the native scaffold may have much smaller effective looping lengths than the other  $L_p$  systems, resulting in a lower entropy penalty on folding.

To examine whether the controlling secondary structure comprises many relatively short dsDNA regions or a small number of long-distance folds, an additional "blocked" scaffold system was tested. In this, regions of potential long-distance intra-scaffold folds were blocked with ssDNA oligos, SI section 1. This system follows the same trend as the native scaffold, suggesting that the long-distance secondary structure is not responsible for the increased  $T_m$ . However, the blocked scaffold has a higher  $T_m$  for the

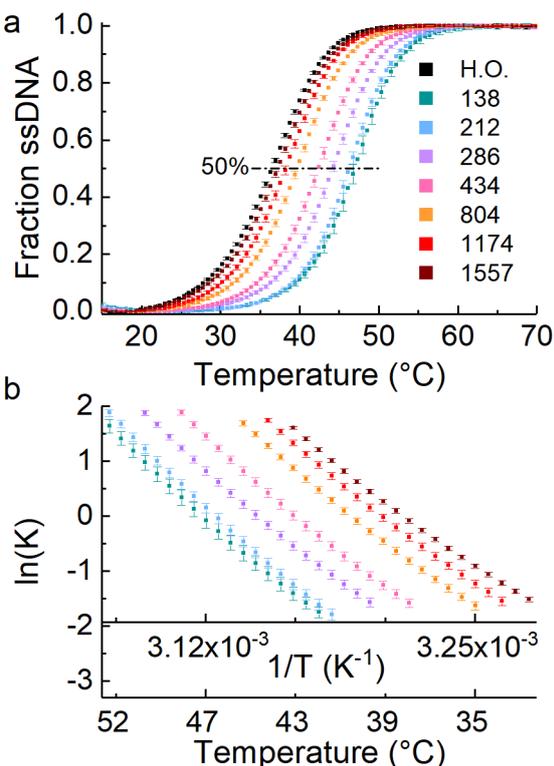


Figure 4. Averaged melt data and van't Hoff plots from the short  $L_p$  scaffold. H.O. indicates the hybridization-only control. The numbers in the legend refer to the number of bases between the melt and anchor subsequences for that curve. Error bars are one standard deviation above and below the mean of the replicate values.

large folds, suggesting that some competition with native secondary structure has been removed, as intended, by the introduction of the blocking staples.

Excepting the shortest fold for the most rigid scaffold, the remaining samples follow two clear trends. Increasing loop distance depresses the  $T_m$ , as does increasing persistence length, primarily at short fold distances. The 138-base fold for the fully dsDNA scaffold has only 4 jointed sections of dsDNA interspersed with nicks. As such, it may be that the low number of sections allows for some unusually stable conformation(60) and a  $T_m$  higher than expected, a feature also seen for the short and long persistence lengths.

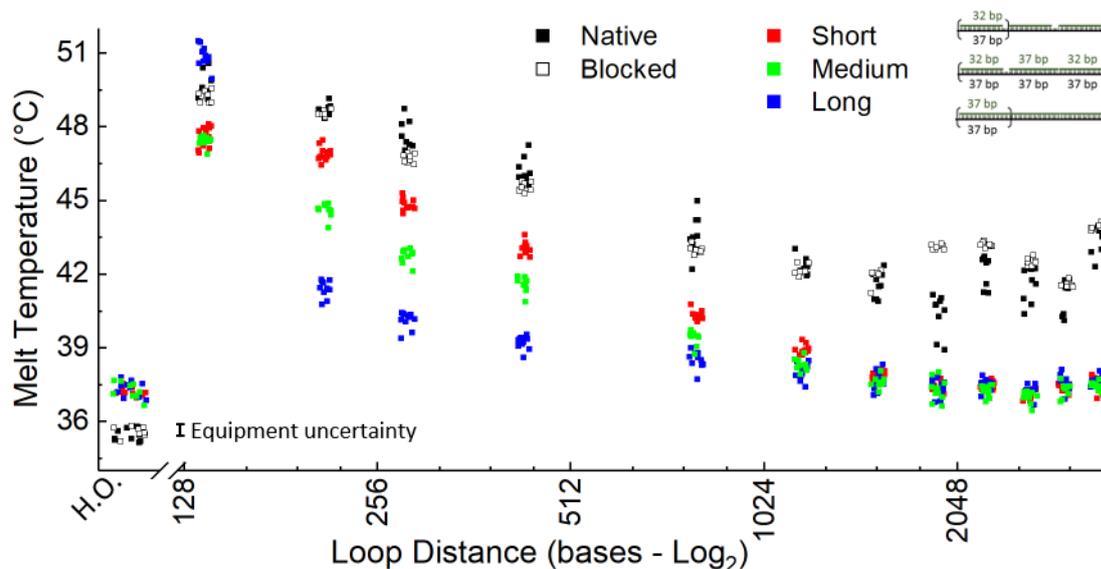


Figure 5.  $T_m$  as a function of the distance along the scaffold between the fold and melt subsequences ( $\text{Log}_2$ ) and system  $L_p$ . H.O. indicates the hybridization-only control. The single error bar represents the 0.3 °C temperature uncertainty as supplied by the manufacturer. The scatter in loop distance was added to the plot to allow the individual  $T_m$  values to be visualized.

As it relates to DNA origami, the melt curves in Fig. 4, and the melt temperatures in Fig. 5, **tell an interesting story which corroborates those developed in previous work(16–18, 21)**. The relatively low melt temperature for the hybridization-only control, combined with the temperature increase for unimolecular reactions suggests that full origami folding follows a three-step process, although confirming this process will require additional testing in systems with multiple folds. First, the long staple sub-sequences ( $\geq 16$  bases) would bind through bimolecular reactions. Second, the free ends of these partially-bound staples would hybridize in a unimolecular reaction. Third, these folds should dramatically increase the rate of subsequent unimolecular folding events by effectively decreasing the remaining fold distances. We note that, as can be seen in Fig. 5, the increase in melt temperature due to a short versus a long fold distance can be  $> 10$  °C. As this difference is due to an increase in  $\Delta S$ , the change

in  $T_m$  will depend on the magnitude of  $\Delta H$  and/or  $\Delta S$  for that reaction, and will be more pronounced for shorter sub-sequences.

### van't Hoff analysis & Heat Capacity Corrections

Fig. 6 illustrates the van't Hoff analysis of optical melt curves, and the effect of neglecting  $\Delta C_p$  terms. It is worth reiterating that extraction of  $\Delta C_p$ , and  $\Delta H$  and/or  $\Delta S$  temperature dependence are equivalent actions as per Eq. 6.

The curvature in the van't Hoff plot can be subtle, making extraction of reliable values for  $\Delta C_p$  problematic. This is exemplified by earlier work (61) which indicated that it was not possible to extract  $\Delta C_p$  from optical melt curves. Our ability to do so in this case can be attributed both to our affine transformations as well as to improvements in equipment and experimental protocols since the 1990s. While there is a debate in the literature regarding whether it is acceptable to neglect  $\Delta C_p$  in thermodynamic parameter prediction/extraction(33, 45), we consider these results to be evidence of the non-negligible effect of  $\Delta C_p$  in DNA melting.

Additional considerations regarding fitting choices, such as fixing individual thermodynamic parameters during fitting, or the role of pipetting errors on van't Hoff analysis uncertainties, are discussed in SI sections 5-9. Entropy-enthalpy compensation is addressed in SI section 10, where we give evidence that the apparent compensation in this system is likely due to neglect of  $\Delta C_p$ (62).

The results of van't Hoff analysis for the short  $L_p$  system are shown in Fig. 7. Here, we use a three-parameter fit, and  $\Delta H$  and  $\Delta S$  are reported referenced to 44 °C, as this is approximately the mid-point of the  $T_m$  range. A discussion of this choice of convention is in SI section 13. A benefit of this choice is that it better represents the uncertainty in the system. The error in the fit is amplified as the reference temperature diverges from the range over which the data is fitted.

The trends reported in Fig. 7 all agree with thermodynamic intuition. For  $\Delta H(T)$ , above some critical maximum curvature one would expect to pay an enthalpic penalty for bending the helix. One would expect this to bending term to be large initially, and to diminish quickly with increasing loop size. For  $\Delta S(T)$ , one would expect a long-distance fold to reduce the number of configurational microstates more than a short fold. Overall, we would not anticipate  $\Delta C_p$  to change with topology to a measurable extent,

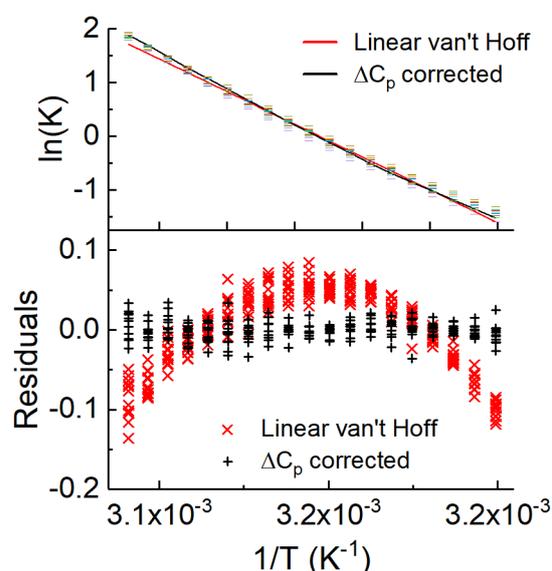


Figure 6. van't Hoff analysis and residuals for the short  $L_p$ , 286 base fold. The top fold data points are individual measurements; the uncertainty is captured in the spread of these points.

and this is what we observe, within error, and the values are consistent with the wide range of per-base  $\Delta C_p$  values found in the literature(33).

The composite construct is much stiffer than ssDNA, and calculating its behaviour is non-trivial. This affects the estimation of both  $\Delta H(T)$  and  $\Delta S(T)$ .

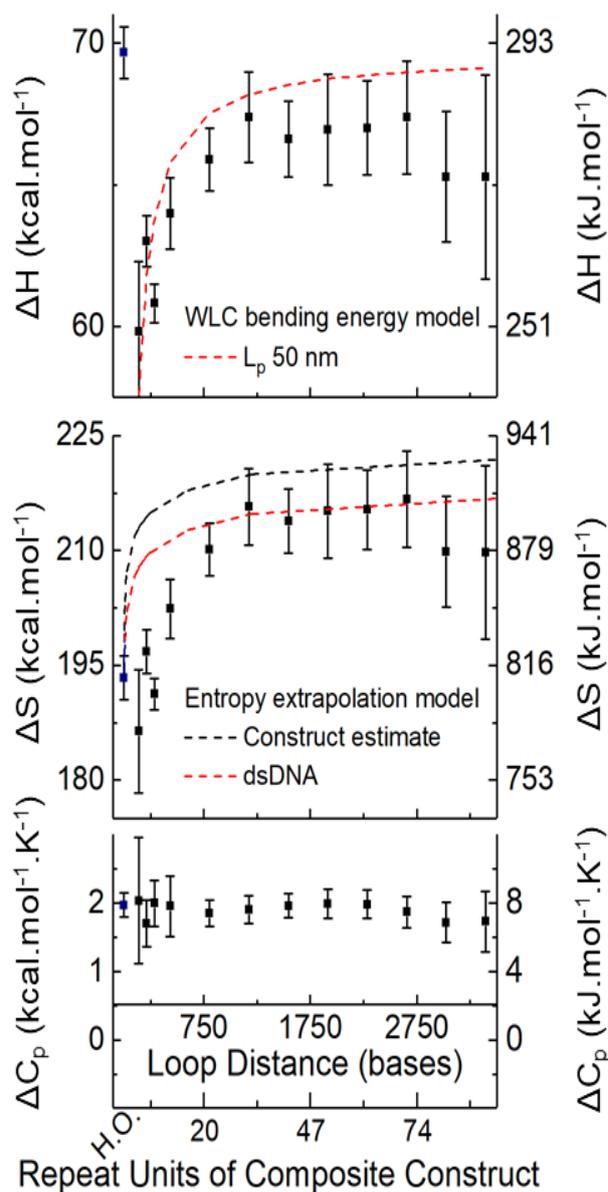


Figure 7. Fitted  $\Delta H^{44^\circ\text{C}}$ ,  $\Delta S^{44^\circ\text{C}}$  and  $\Delta C_p$  for the short  $L_p$  scaffold. Dotted lines indicate the minimal literature models for the energetic effects of folding a circular worm-like-chain. Error bars are one standard deviation above and below the means of the replicate values. H.O. indicates the hybridization-only control, and is plotted at a loop distance of 0. The x axes (Loop Distance/Repeat Units of Composite Construct) are the same for all three plots.

as an indication of the energetic effects that should be borne in mind when considering the observed trends. Further details of the approximations are discussed in section 3 of the SI.

However, even with these limitations, our extracted  $\Delta H(T)$  and  $\Delta S(T)$  show similar trends compared to our estimates. It is thus reasonable to interpret the dip in  $\Delta H(T)$  relative to the hybridization-only control

For excess  $\Delta H(T)$ , the composite will have a higher enthalpy of bending at a given radius of curvature than ssDNA. Unfortunately, a rigorous model for this composite system is significantly beyond the scope of this work, particularly for folds with few composite repeat units, where the ssDNA and dsDNA segments would have to be treated independently.

In order to estimate the excess  $\Delta H(T)$  we used the bending energy from the Worm-Like-Chain (WLC) model and the  $L_p$  of dsDNA, 50 nm(57). This choice requires a caveat for two reasons. First, the short fold distances, where excess  $\Delta H(T)$  is highest, comprise only a small number of composite repeat units, limiting the applicability of the WLC model. Second, the  $L_p$  of these composite constructs at lengths long enough for this continuum model to apply is unknown.

To estimate  $\Delta S(T)$ , we use the Jacobson-Stockmayer extrapolation for ssDNA, but we need to account for the fact that the same length of ssDNA, as compared to our composite, will contain many more Kuhn lengths and thus be able to sample many more configurational states. We use the Kuhn lengths of ssDNA and dsDNA to derive an appropriate conversion factor, SI section 3, assuming that the number of repeat units is sufficiently large to behave statistically.

The dotted lines in Fig. 7 are derived using these approximations, and therefore serve only

at short folds as being consistent with a bending energy penalty. Similarly, it is reasonable to assign the increase in  $\Delta S(T)$  as a function of fold distance to configurational entropy costs.

We note here that, while the melt curves for medium and long  $L_p$  systems, shown in SI section 8, appear by eye to be of equally high quality as that of the short  $L_p$  scaffold, their slightly higher uncertainty propagates aggressively through the van't Hoff analysis. This, combined with the relatively minimal difference in anticipated  $L_p$ , prevent us from being able to evaluate trends between the various  $L_p$  systems in the same way we could for melt temperatures (Fig. 5).

The complexity of predicting both  $\Delta H(T)$  and  $\Delta S(T)$  as a function of topology change in a partially folded DNA origami will likely require sufficiently nuanced modelling to capture not only changes in folding, but also changes in  $L_p$  associated with single staple-subsequence binding events. The uncertainties in  $\Delta H(T)$ ,  $\Delta S(T)$ , and  $\Delta C_p$  are correlated, and their variations as a function of  $L_p$  and fold distance are small compared to the hybridization-only control. Higher resolution data is needed to detect these subtle effects, but would require equipment with a temperature control precision better than 0.3 °C, and with a data sampling density greater than one measurement per 0.6 °C.

The improvements provided by the affine transformation and the high number of replicates is dramatic, allowing us to fit  $\Delta C_p$  then correct for it, a task often not possible with optical data(61). Even though the uncertainties in  $\Delta S(T)$  are inherently greater than those in  $\Delta G(T)$  (62), we are able to measure a < 10 % change in  $\Delta S(T)$ .

## DISCUSSION

Structural DNA nanotechnology's recent success can be attributed to the unusual ease with which even novices can design, fabricate, and image new nanostructures. However, imaging resolution limits the field to purely qualitative measurements of yield. Bridging the gap to implementation demands quantitative measures of the yield for all necessary, active components. If a plasmonic structure fails to contain functional strands addressing metallic nanoparticles to their correct locations, or a drug delivery system is missing a targeting moiety, it will be of little solace that the structure is otherwise of the correct shape. It is critical to understand the folding process sufficiently to predict such functional yields. An essential first step in predicting functional yield is to understand the driving thermodynamics.

We have performed an initial exploration of conformational entropy and bending enthalpy in a model DNA origami fold system. Our results indicate that, while existing theoretical models for loop energy, when modified for the composite construct stiffness, show a similar trend and magnitude to the observed folds, much more sophisticated modeling is needed to predict multi-step folding in full origami systems. However, our results suggest that folding proceeds along the following lines. Among the sub-sequences on a staple, the longest binds first in a bimolecular reaction, consistent with the nucleation seed subsequence concept(63). After this first event, the subsequent subsequences bind unimolecularly, sequentially, as controlled by loop entropy (Fig. 4), strongly favoring the short folds. These short folds further reduce the loop entropy for nearby staples and the system assembles progressively. As loop entropy and other cooperative energetics are sequence-agnostic, and capable of overwhelming the sequence-dependent energetics, this would explain why

qualitative experimental work has found origami folding to proceed like a traditional nucleation and growth system with predetermined nucleation sites(19, 21, 22).

We obtained these results by applying our newly-developed tools, which allow high-throughput, quantitative thermodynamic measurement of the smallest unit process of DNA origami folding. These tools enable exploration of the extensive parameter space associated with origami design and the many coupled reversible reactions which define the folding process for each possible design. The approach presented here may also be applied to the measurement of relatively small energy changes arising in biological DNA systems. While this experiment, run on basic qPCR equipment, did not resolve the even smaller differences in  $\Delta H$ ,  $\Delta S$ , and  $\Delta C_p$  associated with small changes in persistence length, these changes were clearly visible in melt temperature trends. Further reductions in the uncertainties of the raw data should enable the extraction of these small thermodynamic differences. In the future, we will conduct a sensitivity analysis to determine the quality and number of replicates required to extract entropic and enthalpic penalties with sufficient precision to simultaneously measure and distinguish persistence length effects and loop distance effects.

Our ability to make these measurements relies on careful design of the fluorescence reporter, sample stoichiometry, and the application of affine transformations. The latter, by removing systematic errors associated with typical fluorescence baseline corrections, enables the use of replicate data sets to significantly reduce measurement uncertainty. We are thus able to measure changes in  $\Delta S(T)$  well below 10 % of the total value, all with cost-effective qPCR equipment and reagents.

*Importantly, the potential to extract high-quality thermodynamic data that we demonstrate here is a positive indication that, by using the entire dataset obtained in standard fluorescence DNA melting experiments useful additional information can be extracted with minimal extra effort.*

By improving *operando* characterization of origami folding, we hope to complement existing approaches to measurement of functional yield in order to develop design heuristics for DNA nanotechnology as a whole. Finally, we note that the type of variable persistence-length system we introduce here could be valuable as a model to explore the control of stiffness and topology in polymer physics. Supplementary Data are available at NAR online.

## ACKNOWLEDGEMENT

Special thanks to the University of Maryland/NIST Institute for Bioscience & Biotechnology Research (IBBR) for access to their DSC equipment.

## REFERENCES

1. Liu,Q., Song,C., Wang,Z.G., Li,N. and Ding,B. (2014) Precise organization of metal nanoparticles on DNA origami template. *Methods*, **67**, 205–214.
2. Bui,H., Onodera,C., Kidwell,C., Tan,Y., Graugnard,E., Kuang,W., Lee,J., Knowlton,W.B., Yurke,B. and Hughes,W.L. (2010) Programmable periodicity of quantum dot arrays with DNA origami nanotubes. *Nano Lett.*, **10**, 3367–3372.

3. Ko, S.H., Gallatin, G.M. and Liddle, J.A. (2012) Nanomanufacturing with DNA origami: Factors affecting the kinetics and yield of quantum dot binding. *Adv. Funct. Mater.*, **22**, 1015–1023.
4. Kuzyk, A., Jungmann, R., Acuna, G.P. and Liu, N. (2018) DNA Origami Route for Nanophotonics. 10.1021/acsp Photonics.7b01580.
5. Linko, V., Nummelin, S., Aarnos, L., Tapio, K., Toppari, J. and Kostianen, M. (2016) DNA-Based Enzyme Reactors and Systems. *Nanomaterials*, **6**, 139.
6. Aslan, H., Krissanaprasit, A., Besenbacher, F., Gothelf, K. V. and Dong, M. (2016) Protein patterning by a DNA origami framework. *Nanoscale*, **289**, 1760–1763.
7. Schnitzbauer, J., Strauss, M.T., Schlichthaerle, T., Schueder, F. and Jungmann, R. (2017) Super-resolution microscopy with DNA-PAINT. *Nat. Protoc.*, **12**, 1198–1228.
8. Gopinath, A., Miyazono, E., Faraon, A. and Rothmund, P.W.K. (2016) Engineering and mapping nanocavity emission via precision placement of DNA origami. *Nature*, **535**, 401–405.
9. Kick, B., Praetorius, F., Dietz, H. and Weuster-Botz, D. (2015) Efficient Production of Single-Stranded Phage DNA as Scaffolds for DNA Origami. *Nano Lett.*, **15**, 4672–4676.
10. Halley, P.D., Patton, R.A., Chowdhury, A., Byrd, J.C. and Castro, C.E. (2019) Low-cost, simple, and scalable self-assembly of DNA origami nanostructures. *Nano Res.*, 10.1007/s12274-019-2384-x.
11. Zahid, M., Kim, B., Hussain, R., Amin, R. and Park, S.H. (2013) DNA nanotechnology: a future perspective. *Nanoscale Res. Lett.*, **8**, 119.
12. Steinhauer, C., Jungmann, R., Sobey, T.L., Simmel, F.C. and Tinnefeld, P. (2009) DNA origami as a nanoscopic ruler for superresolution microscopy. *Angew. Chemie - Int. Ed.*, **48**, 8870–8873.
13. Daems, D., Pfeifer, W., Rutten, I., Sacca, B., Spasic, D. and Lammertyn, J. (2018) Three-Dimensional DNA Origami as Programmable Anchoring Points for Bioreceptors in Fiber Optic Surface Plasmon Resonance Biosensing. 10.1021/acscami.8b04757.
14. Moore, G.M. (1965) Cramming more components onto integrated circuits. **38**, 114.
15. Rothmund, P.W.K. (2006) Folding DNA to create nanoscale shapes and patterns. *Nature*, **440**, 297–302.
16. Dannenberg, F., Dunn, K.E., Bath, J., Turberfield, A.J., Ouldrige, T.E., Kwiatkowska, M., Turberfield, A.J. and Ouldrige, T.E. (2015) Modelling DNA origami self-assembly at the domain level. *J. Chem. Phys.*, **143**, 1–19.
17. Arbona, J.-M., Aimé, J.-P. and Elezgaray, J. (2013) Cooperativity in the annealing of DNA origamis. *J. Chem. Phys.*, **138**, 015105.
18. Dunn, K.E., Dannenberg, F., Ouldrige, T.E., Kwiatkowska, M., Turberfield, A.J. and Bath, J. (2015) Guiding the folding pathway of DNA origami. *Nature*, 10.1038/nature14860.
19. Majikes, J.M., Nash, J.A. and LaBean, T.H. (2016) Competitive annealing of multiple DNA origami: formation of chimeric origami. *New J. Phys.*, **18**, 115001.
20. Song, J., Arbona, J.-M., Zhang, Z., Liu, L., Xie, E., Elezgaray, J., Aime, J.-P., Gothelf, K.V., Besenbacher, F. and Dong, M. (2012) Direct visualization of transient thermal response of a DNA origami. *J. Am. Chem. Soc.*, **134**, 9844–7.
21. Lee Tin Wah, J., David, C., Rudiuk, S., Baigl, D., Estevez-Torres, A.A., Wah, J.L.T., David, C., Rudiuk, S., Baigl, D. and Estevez-Torres, A.A. (2016) Observing and Controlling the Folding

- Pathway of DNA Origami at the Nanoscale. *ACS Nano*, **10**, 1978–1987.
22. Majikes, J.M., Nash, J.A. and LaBean, T.H. (2017) Search for effective chemical quenching to arrest molecular assembly and directly monitor DNA nanostructure formation. *Nanoscale*, **9**, 1637–1644.
  23. Kosinski, R., Mukhortava, A., Pfeifer, W., Candelli, A., Rauch, P. and Saccà, B. (2019) Sites of high local frustration in DNA origami. *Nat. Commun.*, **10**, 1–12.
  24. Wagenbauer, K.F., Wachauf, C.H. and Dietz, H. (2014) Quantifying quality in DNA self-assembly. *Nat. Commun.*, **5**, 3691.
  25. Strauss, M.T., Schueder, F., Haas, D., Nickels, P.C. and Jungmann, R. (2018) Quantifying absolute addressability in DNA origami with molecular resolution. *Nat. Commun.*, **9**, 1–7.
  26. Schneider, F., Möritz, N. and Dietz, H. (2019) The sequence of events during folding of a DNA origami. *Sci. Adv.*, **5**.
  27. Patrone, P., Kearsley, A., Majikes, J. and Liddle, J.A. (Under Review) Analysis and Uncertainty Quantification of DNA Fluorescence Melt Data: Applications of Affine Transformations. *Anal. Biochem.*
  28. SantaLucia, J. and Hicks, D. (2004) The thermodynamics of DNA structural motifs. *Annu. Rev. Biophys. Biomol. Struct.*, **33**, 415–440.
  29. SantaLucia, J. (1998) A unified view of polymer, dumbbell, and oligonucleotide DNA nearest-neighbor thermodynamics. *Proc. Natl. Acad. Sci. U. S. A.*, **95**, 1460–5.
  30. Owczarzy, R., Tataurov, A. V., Wu, Y., Manthey, J.A., McQuisten, K.A., Almabrazi, H.G., Pedersen, K.F., Lin, Y., Garretson, J., McEntaggart, N.O., *et al.* (2008) IDT SciTools: a suite for analysis and design of nucleic acid oligomers. *Nucleic Acids Res.*, **36**.
  31. Allemand, J.F., Cocco, S., Douarache, N. and Lia, G. (2006) Loops in DNA: An overview of experimental and theoretical approaches. *Eur. Phys. J. E*, **19**, 293–302.
  32. Rippe, K., von Hippel, P.H. and Langowski, J. (1995) Action at a distance: DNA-looping and initiation of transcription. *Trends Biochem. Sci.*, **20**, 500–506.
  33. Mikulecky, P.J. and Feig, A.L. (2006) Heat Capacity Changes Associated with Nucleic Acid Folding. *Biopolymers*, **81**, 38–58.
  34. Rouzina, I. and Bloomfield, V.A. (1999) Heat Capacity Effects on the Melting of DNA. 1. General Aspects. **77**, 3242–3251.
  35. Breslauer, K.J., Frank, R., Blöcker, H. and Marky, L. a (1986) Predicting DNA duplex stability from the base sequence. *Proc. Natl. Acad. Sci. U. S. A.*, **83**, 3746–50.
  36. Rozners, E., Pilch, D.S. and Egli, M. (2015) Calorimetry of Nucleic Acids. 10.1002/0471142700.nc0704s63.
  37. Du, Q., Smith, C., Shiffeldrim, N., Vologodskaya, M. and Vologodskii, A. (2005) Cyclization of short DNA fragments and bending fluctuations of the double helix. *Proc. Natl. Acad. Sci.*, **102**, 5397–5402.
  38. Sobczak, J.-P.J., Martin, T.G., Gerling, T. and Dietz, H. (2012) Rapid folding of DNA into nanoscale shapes at constant temperature. *Science*, **338**, 1458–61.
  39. Sanford, L.N., Kent, J.O. and Wittwer, C.T. (2013) Quantum Method for Fluorescence Background

- Removal in DNA Melting Analysis. *Anal. Chem.*, 10.1021/ac4024928.
40. Marky, L.A. and Breslauer, K.J. (1987) Calculating Thermodynamic Data for Transitions of any Molecularity from Equilibrium Melting Curves. **26**, 1601–1620.
  41. Bloomfield, V.A., Crothers, D.M. and Tinoco, I. (2000) *Nucleic Acids: Structures, Properties, and Functions* University Science Books, Sausalito, CA 94965.
  42. You, Y., Tataurov, A. V and Owczarzy, R. (2011) Measuring Thermodynamic Details of DNA Hybridization Using Fluorescence. *Biopolymers*, **95**, 472–486.
  43. Wei, X., Nangreave, J., Jiang, S., Yan, H. and Liu, Y. (2013) Mapping the thermal behavior of DNA origami nanostructures. *J. Am. Chem. Soc.*, **135**, 6165–6176.
  44. Fern, J., Lu, J. and Schulman, R. (2016) The Energy Landscape for the Self-Assembly of a Two-Dimensional DNA Origami Complex. *ACS Nano*, 10.1021/acsnano.5b05309.
  45. Nanostructures, D.N.A., Saccà, B., Meyer, R., Feldkamp, U., Schroeder, H. and Niemeyer, C.M. (2008) High-Throughput , Real-Time Monitoring of the Self-Assembly of DNA Nanostructures by FRET Spectroscopy \*\*. 10.1002/anie.200704836.
  46. Chandrasekaran, A.R., Maclsaac, M., Dey, P., Levchenko, O., Zhou, L., Andres, M., Dey, B.K. and Halvorsen, K. (2019) Cellular microRNA detection with miRacles: microRNA- activated conditional looping of engineered switches. *Sci. Adv.*, **5**, eaau9443.
  47. Rivetti, C., Walker, C. and Bustamante, C. (1998) Polymer chain statistics and conformational analysis of DNA molecules with bends or sections of different flexibility. *J. Mol. Biol.*, **280**, 41–59.
  48. Du, Q., Vologodskaya, M., Kuhn, H., Frank-Kamenetskii, M. and Vologodskii, A. (2005) Gapped DNA and cyclization of short DNA fragments. *Biophys. J.*, **88**, 4137–4145.
  49. Roll, C., Ketterlé, C., Faibis, V., Fazakerley, G.V. and Boulard, Y. (1998) Conformations of Nicked and Gapped DNA Structures by NMR and Molecular Dynamic Simulations in Water. *Biochemistry*, **37**, 4059–4070.
  50. Furrer, P., Bednar, J., Stasiak, A.Z., Katritch, V., Michoud, D., Stasiak, A. and Dubochet, J. (1997) Opposite Effect of Counterions on the Persistence Length of Nicked and Non-nicked DNA.
  51. Weber, G., Essex, J.W. and Neylon, C. (2009) Probing the microscopic flexibility of DNA from melting temperatures. *Nat. Phys.*, **5**, 769–773.
  52. Vafabakhsh, R. and Ha, T. (2012) Extreme Bendability of DNA Less than 100 Base Pairs Long Revealed by Single-Molecule Cyclization. *Science (80-. )*, **1097**, 1–6.
  53. Le, T.T. and Kim, H.D. (2014) Probing the elastic limit of DNA bending. *Nucleic Acids Res.*, **42**, 10786–10794.
  54. Vologodskii, A. and Frank-kamenetskii, M.D. (2013) Strong bending of the DNA double helix. **41**, 6785–6792.
  55. Reiter, M. and Pfaffl, M.W. (2008) Effects of plate position, plate type and sealing systems on real-time pcr results. *Biotechnol. Biotechnol. Equip.*, **22**, 824–828.
  56. Wittwer, C.T., Reed, G.H., Gundry, C.N., Vandersteen, J.G. and Pryor, R.J. (2003) High-Resolution Genotyping by Amplicon Melting Analysis Using LCGreen. **860**, 853–860.
  57. Measurements, F. (1999) Estimating the Persistence Length of a Worm-Like Chain Molecule from Force-Extension Measurements. **76**, 409–413.

58. Owczarzy,R. (2008) IDT Biophysics: DNA Thermodynamics Predictor.
59. Pedersen,R., Marchi,A.N., Majikes,J.M., Nash,J.A., Estrich,N.A., Courson,D.S., Hall,C.K., Craig,S.L. and LaBean,T.H. (2015) Properties of DNA. In *Handbook of Nanomaterials properties*.pp. 1125–1157.
60. Cloutier,T.E. and Widom,J. (2005) DNA twisting flexibility and the formation of sharply looped protein-DNA complexes. *Proc. Natl. Acad. Sci.*, **102**, 3645–3650.
61. Chaires,J.B. (1997) Possible origin of differences between van't Hoff and calorimetric enthalpy estimates. *Biophys. Chem.*, **64**, 15–23.
62. Sharp,K. (2001) Entropy--enthalpy compensation: Fact or artifact? *Protein Sci.*, **10**, 661–667.
63. Ke,Y., Bellot,G., Voigt,N. V., Fradkov,E. and Shih,W.M. (2012) Two design strategies for enhancement of multilayer–DNA-origami folding: underwinding for specific intercalator rescue and staple-break positioning. *Chem. Sci.*, **3**, 2587.

## TABLE AND FIGURES LEGENDS

*Table 1. Melt experiment concentrations*

*Table 2. DSC, nearest-neighbour, and fluorescence melt temperatures*

Figure 1. a) folding transition schematics for bimolecular hybridization-only (H.O.), i.e., staple binds to scaffold, and unimolecular, i.e. bound staple hybridizes fully, fold reactions. b) Persistence length strands: name, anticipated  $L_p$ , and schematic of oligomers/scaffold, c) Reporter scheme: cartoon of fold and reporter strands on the scaffold and description of sub-sequences for fold and reporter strands where  $X'$  indicates the complement to  $X$ . d) Relative positions of melt and anchor sub-sequence complements on M13 scaffold.

Figure 2. Hybridization-only (H.O.) control for the short  $L_p$  scaffold before a), and after b), affine transformation. The empty plate background,  $B(T)$ , and fluorescence temperature dependence baseline,  $R(T)$ , are subtracted from a to b, but are separately measured in control experiments.

Figure 3. DSC scans of reporter analog: a) only fold and scaffold strands. b) fold, scaffold, and fluorophore strands. c) all reporter strands. Red curve, derivative of fluorescence for the same sample. Peaks and sub-sequences labeled to match

Figure 4. Averaged melt data and van't Hoff plots from the short  $L_p$  scaffold. H.O. indicates the hybridization-only control. The numbers in the legend refer to the number of bases between the melt and anchor subsequences for that curve. Error bars are one standard deviation above and below the mean of the replicate values.

Figure 5.  $T_m$  as a function of fold distance ( $\log_2$ ) and system  $L_p$ . H.O. indicates the hybridization-only control. The single error bar represents the 0.3 °C temperature uncertainty as supplied by the manufacturer. The scatter in loop distance was added to the plot to allow the individual  $T_m$  values to be visualized.

*Figure 6. van't Hoff analysis and residuals for the short Lp, 286 base fold. The top fold data points are individual measurements; the uncertainty is captured in the spread of these points.*

*Figure 7. Fitted  $\Delta H^{44}$  °C ,  $\Delta S^{44}$  °C and  $\Delta C_p$  for the short Lp scaffold. Dotted lines indicate the minimal literature models for the energetic effects of folding a circular worm-like-chain. Error bars are one standard deviation above and below the mean of the replicate values. H.O. indicates the hybridization-only control, and is plotted at a loop distance of 0.*

Supplementary Information:

## Contents

Section 1: Anticipated Secondary Structure .....	27
Section 2: Affine Transformation Algorithm:.....	28
Section 3: Entropy model for folding hybrid systems, estimation of persistence length, & L <sub>p</sub> Strands	30
Section 4: DSC data & uncertainty .....	31
Section 5: Hybridization-Only system comparison .....	35
Section 6: Pipetting variation & Hybridization-Only systems .....	35
Section 7: $\Delta H^0$ vs. $\Delta H^0(T)$ .....	36
Section 8: Data from All L <sub>p</sub> Systems .....	37
Section 9: Entropy Enthalpy Compensation .....	40
Section 10: uncertainty in van't Hoff analysis of optical melt curves.....	41
Section 11: fold strand sequences .....	41
Section 12: Extracted thermodynamics with and without low temperature monotonicity constraint	42
Section 13: Discussion of heat capacity, enthalpy and entropy conventions .....	43
Section 14: L <sub>p</sub> strand sequences .....	44

## Section 1: Anticipated Secondary Structure

Given the unexpected pattern of melt temperatures for the native scaffold, a preliminary examination of the M13MP18 scaffold secondary structure was performed using ViennaRNA<sup>1</sup> using parameters derived by Turner and Mathews<sup>2</sup>. This identified a countable number of regions where the scaffold could form >9 base secondary structure.

Figure S1 shows a map of this secondary structure, where the black circle represents the M13, blue arcs indicate a subset of the engineered fold distances, while yellow, orange and red arcs indicate 9, 11, and 12 base regions of possible long-distance secondary structure.

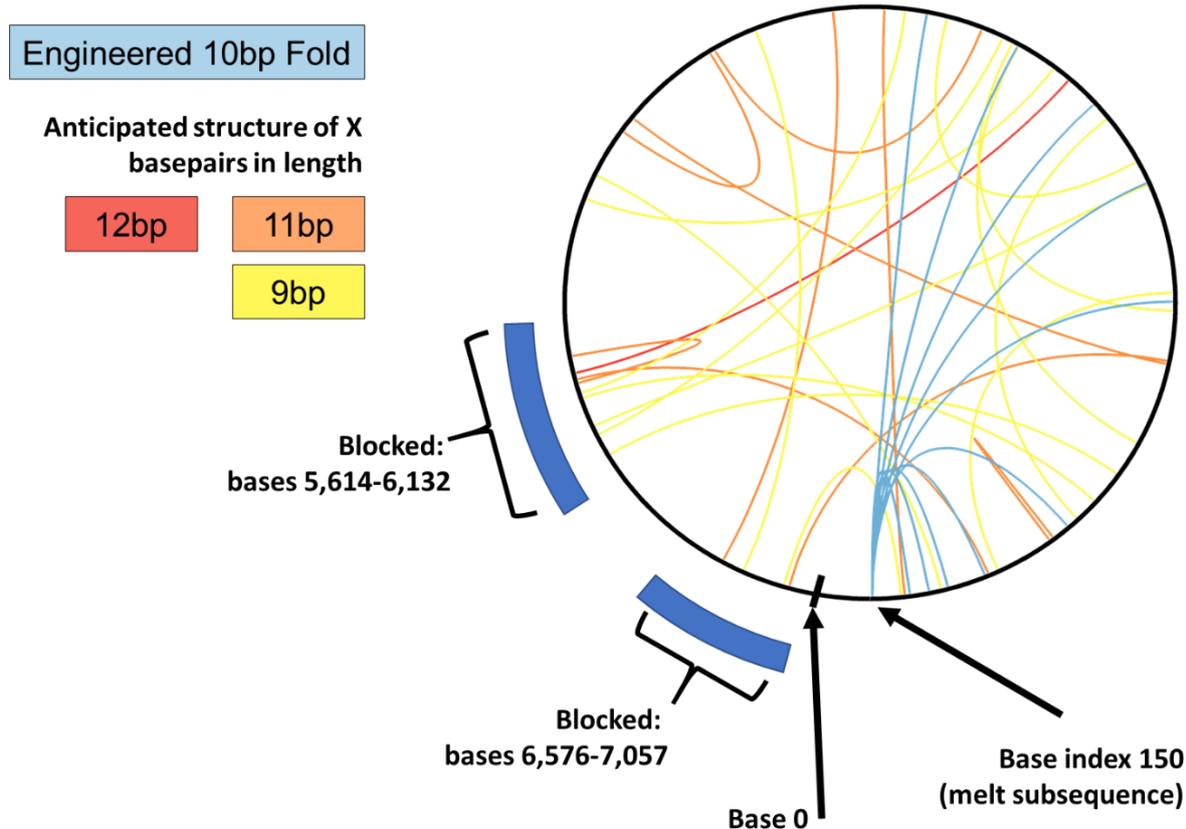
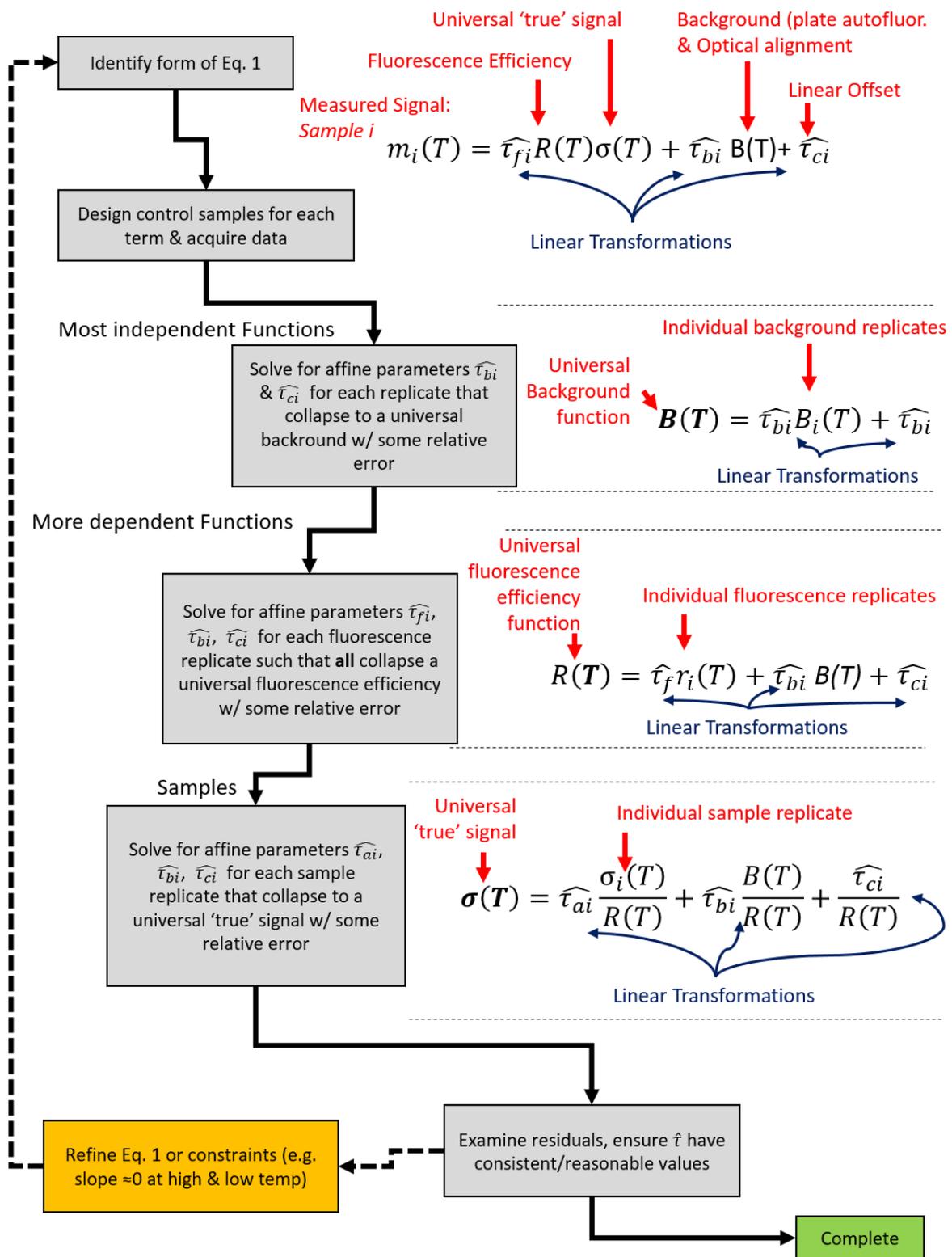


Figure S8. Anticipated secondary structure and blocking strands

We identified regions on the 'left' side of the M13 (bases 3,500 to 7200) with the 11 or 12 base possible matches. We then identified a subset of the repeating 37 base oligos to block these regions from forming secondary structure.

As such, the blocking oligomers consisted of the 37 base oligomers from the A-side, which refers to the left side of the scaffold, bases 3,407 to 7,240. See section 15 for more details. The blocked positions shown above in Fig. S1 corresponds to positions E5-F5, G7-H7 in the A-side plate.

## Section 2: Affine Transformation Algorithm:



An alternative to equation 1, for systems with significantly different fluorescence efficiency temperature dependence in the quenched and unquenched state might be:

Universal 'true' signal (ssDNA content)

Background (plate autofluor. & optics alignment)

Linear Offset

Measured Signal:  
Sample  $i$

$$m_i(T) = \widehat{\tau}_{fi} R(T) \sigma(T) + \widehat{\tau}_{qi} Q(T) [1 - \sigma(T)] + \widehat{\tau}_{bi} B(T) + \widehat{\tau}_{ci}$$

Quenched Fluorescence Efficiency

Universal 'true' signal (dsDNA content)

In this case, another subjective user decision would be whether the linear transformations for quenched and unquenched efficiency are identical or not. Given that both are linked to pipetting variation on fluorophore content, this might be a reasonable choice.

Another alternative to equation 1, for systems with intercalating dyes might be:

Intercalated Fluorescence Efficiency (dsDNA)

Universal 'true' signal (ssDNA content)

Background (plate autofluor. & optics alignment)

Linear Offset

Measured Signal:  
Sample  $i$

$$m_i(T) = \widehat{\tau}_{fi} R_{ds}(T) \sigma(T) + \widehat{\tau}_{qi} R_{ss}(T) [1 - \sigma(T)] + \widehat{\tau}_{fi} R_{free}(T) + \widehat{\tau}_{bi} B(T) + \widehat{\tau}_{ci}$$

Intercalated Fluorescence Efficiency (ssDNA)

Universal 'true' signal (ssDNA content)

Free Dye Fluorescence Efficiency

## Section 3: Entropy model for folding hybrid systems, estimation of persistence length, & $L_p$ Strands

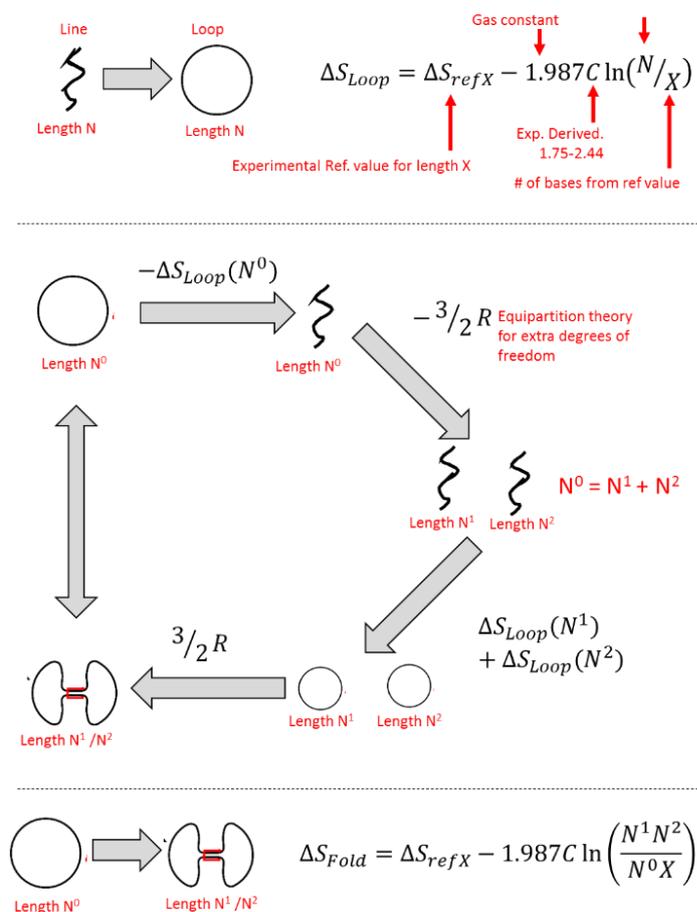


Figure S9: Entropy model for folding

It is worth emphasizing that the apparent  $L_p$  is only a simple placeholder for the exact behavior of such a complex system.

While the  $L_p$  systems afforded us both freedom from secondary structure as well as control over the polymer flexibility, it significantly complicated application of any model to interpret the trends in excess entropy and enthalpy.

The Jacobson-Stockmayer entropy extrapolation, described in the top of Figure S2, models the entropy change from a line/loop. This is used to model a loop to fold transition as described in the middle/bottom of Figure S2.

It is critical to note that the units for  $N$  are in bases of ssDNA. However, our hybrid systems have **both** ssDNA and dsDNA.

To address this, we needed to create an estimated conversion factor between the units 'bases ssDNA' and 'bases of hybrid system.'

To resolve this difference, we made the aggressive assumption that the alternating ssDNA and dsDNA subsections were so small (37 bases) compared to the scaffold that they behaved homogeneously.

In order to remove the potential effects of secondary structure, as well as to determine if differences in scaffold persistence length,  $L_p$ , resulted in measurable differences in melt temperature,  $T_m$ , the persistence length of the scaffold was varied.

The scaffold, excepting the melt, anchor, and quencher subsequences, was approximately 7,151 bases in length. This in turn was divided into  $\approx 190$  sections, each 37 bases in length. For each section, a 37 base oligo complement and a 32 base oligo complement were ordered, SI Section 15. These oligos were selectively added to the samples to force the scaffold to have repeating lengths of ssDNA/dsDNA, the content of which would change the *apparent*  $L_p$  of the polymer as a whole. It is

If this was true, the number of Kuhn lengths in a single repeat of the hybrid would equal the number of Kuhn lengths of ssDNA in a single repeat plus the number of Kuhn lengths of dsDNA in that repeat. As such, within the repeat unit of the composite construct we could take the number of bases of the dsDNA segment and multiply it by the ratio of the Kuhn lengths, or persistence lengths, of the two polymers to determine the number of equivalent ssDNA bases that segment of dsDNA contributed. This is then added to the number of bases of ssDNA in the repeat unit to find the equivalent number of ssDNA bases in each repeat. This is then divided by the total number of bases in the composite construct to make a conversion factor between the units “bases ssDNA” and “bases composite construct” as in Eq. S1.

$$\frac{\text{bases of ssDNA}}{\text{base of hybrid repeat}} = \frac{\text{bases}_{\text{ssDNA-perRepeat}} + \text{bases}_{\text{dsDNA-Repeat}} \frac{L_p\text{-ssDNA}}{L_p\text{-dsDNA}}}{\#\text{bases per repeat}} \quad [\text{S1}].$$

For this, the persistence length of dsDNA in Mg buffer was taken to be 50 nm, and the persistence length of ssDNA was taken to be 3 nm.<sup>3</sup>

It is worth explicitly stating that this is obviously a suboptimal modeling approach. However, as there is no simple and tested model for enthalpies of bending and configurational entropies of systems with repeating changes in persistence length, and as the purpose of this study is measurement of small changes in energy due to topological changes, a precise and thorough theoretical framework is well beyond the scope of this work.

System	Bases ssDNA per Hybrid repeat
'Short'	0.168
'Medium'	0.159
'Long'	0.0448

Table S3. Conversion factors for entropy model

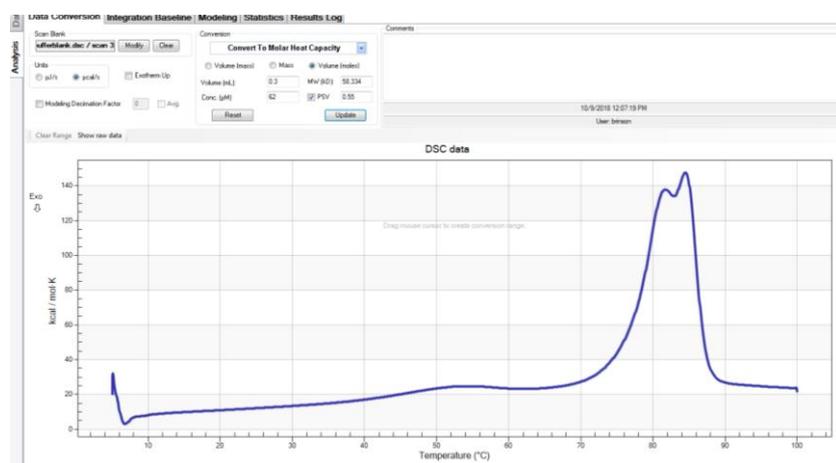
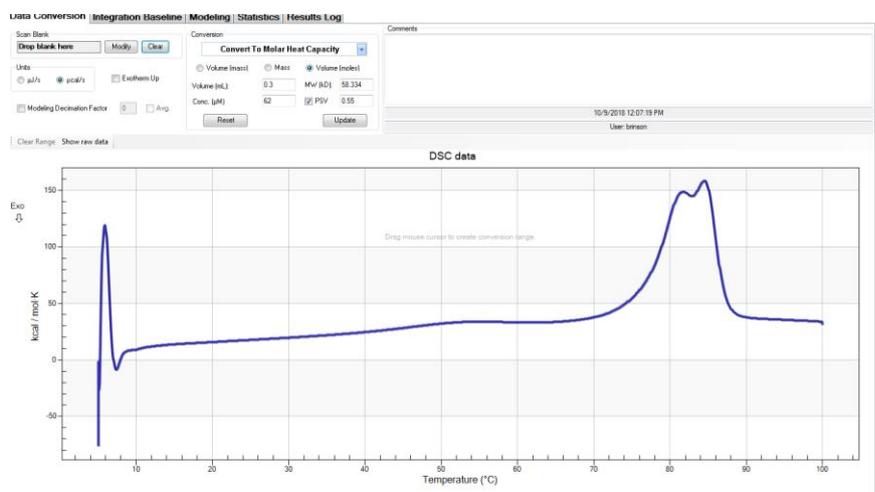
Enthalpy was modeled by the worm-like-chain model, assuming the energy of bending a rigid segment is entirely enthalpic in nature<sup>4</sup>.  $L_p$  is the persistence length,  $L_0$  is the contour length for loop being measured ( $N^0$ ),  $R$  is the radius of curvature, and  $T$  is the temperature.

$$E_b = \frac{L_p K_b T}{2R^2} \Rightarrow \Delta H_{\text{bend}} = \frac{L_p K_b T}{2R^2} * L_0 = \frac{L_p K_b T}{2\left(\frac{L_0}{2\pi}\right)^2} * L_0 \quad [\text{S2}]$$

## Section 4: DSC data & uncertainty

While DSC is a highly useful tool for comparing  $\Delta H$  for a reaction between samples, there were several reasons the data reported in the main text did not include analysis of DSC data.

The error associated with baseline correction in attempting to extract  $\Delta H$  for comparison to van't Hoff analysis is significant. Below are the raw, and buffer scan corrected data for the full DSC model system.



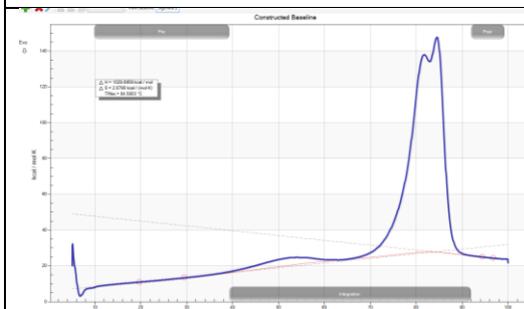
While the software can analyze the data without baseline correction, this involves 5 fit parameters per peak (offset, slope,  $\Delta H$ ,  $\Delta C_p$ ,  $T_m$ ) with two of the parameters. Due to the overlap in first two terms, *attempting to fit using these general models was infeasible for our reporter system.*

Less general 2-state models can be fit to the data, however these require *neglect of  $\Delta C_p$  and baseline correction.* While the former is unfortunate, the latter is prohibitive, as we will show below.

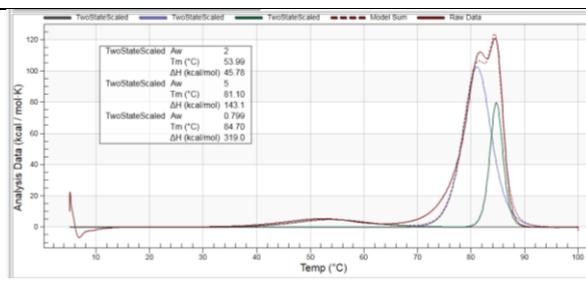
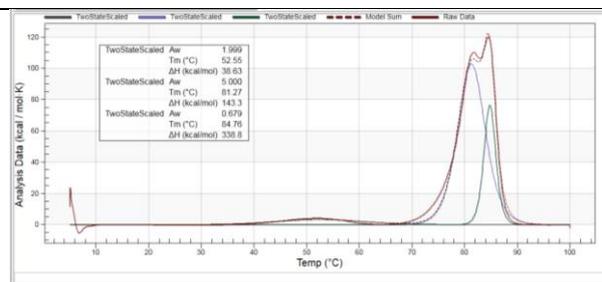
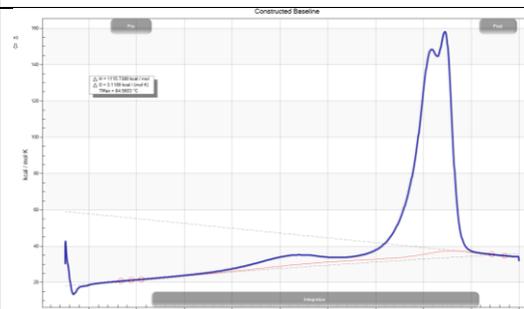
Numerous potential functions for baseline correction are available, for this illustration it was sufficient to use the most linear baselines and a 2<sup>nd</sup> order/linear baseline pair. Table S2 shows several different baseline type and window choices and their result on the calculated values.

The variation between  $\Delta H$  values, combined with the inability to fit a general function (and thus obtain  $\Delta C_p$ ) is problematic. A reporter with more distinct melt peaks would likely have been much more tractable to analysis.

1<sup>st</sup> order/1<sup>st</sup> order



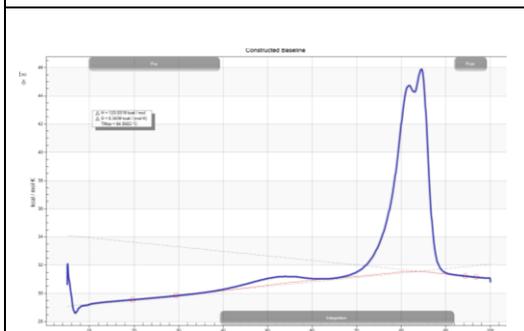
2<sup>nd</sup> order/1<sup>st</sup> order



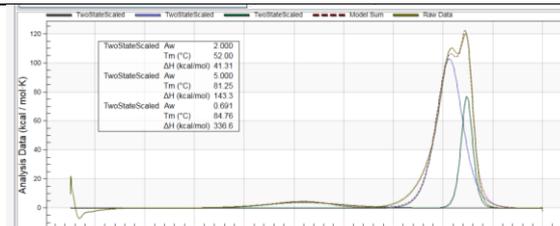
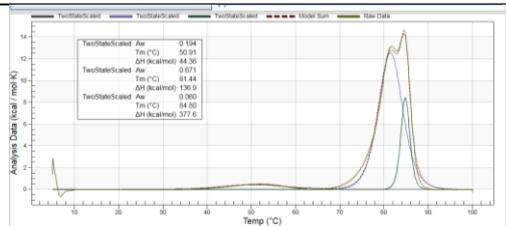
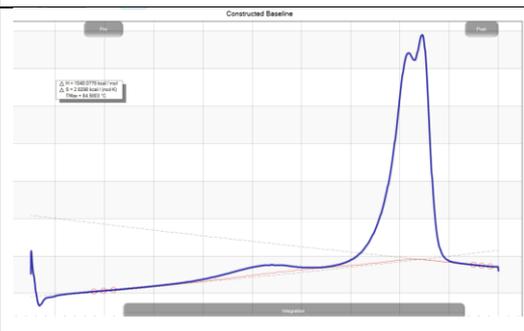
	$\alpha$	$B$	$\delta$
$\Delta H$ (kCal/mol)	38.63	143.3	338.8
$T_m$ (°C)	52.55	81.27	84.76

	$\alpha$	$B$	$\delta$
$\Delta H$ (kCal/mol)	53.99	143.1	319.0
$T_m$ (°C)	45.78	81.1	84.7

1<sup>st</sup> order/1<sup>st</sup> order



2<sup>nd</sup> order/2<sup>nd</sup> order



	$\alpha$	$B$	$\delta$
--	----------	-----	----------

	$\alpha$	$B$	$\delta$
--	----------	-----	----------

$\Delta H$ (kCal/mol)	44.36	136.9	337.6	$\Delta H$ (kCal/mol)	41.3	143.3	336.6
$T_m$ (°C)	50.91	81.44	84.8	$T_m$ (°C)	52.0	81.25	84.76

Table S4: Examples of DSC data analysis

## Section 5: Hybridization-Only system comparison

Below are the fitted  $\Delta H(44^\circ\text{C})/\Delta S(44^\circ\text{C})$  values for the **Hybridization-Only** controls in all  $L_p$  systems. As one would expect from the  $T_m$  values reported in the main text, the  $\Delta H/\Delta S$  values differ between the native/blocked and short/medium/long  $L_p$  samples. The DSC control is different in  $\Delta H/\Delta S$  although its  $T_m$  as reported in the main text, is consistent.

The fluorescence measurements of the DSC system required measuring a 0.75  $\mu\text{L}$  droplet in oil to simultaneously maintain constant concentration and avoid overexposing the qPCR optical detector. Another potential source of discrepancy is the high magnesium concentration used in DNA nanotechnology, for which the corrections to the nearest neighbor model may be insufficient. As such it is difficult to definitively say why the van't Hoff  $\Delta H/\Delta S$  values are different between these systems.

	<i>dH (kCal/mol)</i>	<i>error</i>	<i>dS (cal/mol-k)</i>	<i>error</i>
<i>Short</i>	69.7	0.9	193.4	2.9
<i>Medium</i>	72.3	1.2	201.7	3.9
<i>Long</i>	72.4	4.4	202.0	14.0
<i>Unblocked</i>	57.5	0.9	155.5	2.9
<i>Blocked</i>	55.2	1.7	147.9	5.3
<i>DSC</i>	103.7	1.6	298.3	4.9

## Section 6: Pipetting variation & Hybridization-Only systems

When  $\Delta H$ ,  $\Delta S$ , &  $\Delta C_p$  are fitted independently the  $\Delta S$  values for the no fold system are higher than anticipated, as discussed in the main text. One possible explanation for this is that the no-fold system is bimolecular more susceptible to variation in the concentration (or pipetting variation). While our experimental system was designed to minimize the effect of pipetting variation, initial error in concentration measurements is possible.

To examine the possible magnitude of this effect, the analysis was rerun 3 times assuming different pipetting variation on the scaffold and fold strands.

Percent Variation		Concentration		$\Delta H$	$\Delta S$
Scaff	Fold	Scaff	Fold		
0	0	45	225	$70.0 \pm 1.1$ KCal/mol	$195.0 \pm 3.5$ cal/mol-K
-5%	5%	42.75	236.25	$69.7 \pm 1.1$ KCal/mol	$194.2 \pm 3.5$ cal/mol-K
5%	-5%	47.25	213.75	$70.0 \pm 1.1$ KCal/mol	$196.0 \pm 3.5$ cal/mol-K

-10%	10%	40.5	247.5		69.5 ± 1.1 KCal/mol	193.4 ± 3.5 cal/mol-K
------	-----	------	-------	--	---------------------	-----------------------

While 5% error is more than would be anticipated for these systems, even 10% error in opposite directions for the scaffold and fold strands result in < 2 cal/mol-K variation in  $\Delta S$ . As such it is highly unlikely that the pipetting variation meaningfully contributes to differences in the extracted  $\Delta S$ .

## Section 7: $\Delta H^0$ vs. $\Delta H^0(T)$

As discussed in the main text, one can assume that  $\Delta H$  is temperature independent,  $\Delta H^0$ , or assume that it follows Sup. Eq. s2. As such, it is worth comparing  $\Delta H(T_m)$  to  $\Delta H^0$  as a measure of whether fitting the curvature in van't Hoff analysis is overfitting noise.

$$\Delta H(T) = \Delta H^0 + \Delta C_p \cdot \Delta T = \Delta H^0 + \Delta C_p \cdot (T_{\text{eval}} - T_m) \quad [s2]$$

Whether or not temperature independence was assumed, the same values for  $\Delta H^0$  were calculated, within error. Fig S4. shows an example plot comparing the two quantities.

If the  $\Delta H^0$  values from both fits considering and neglecting  $\Delta C_p$  diverged, it might imply that the system was over defined, and the fitting algorithm was using play between  $\Delta H^0$  and  $\Delta C_p$  to overfit noise in the data.

Phrased another way, this indicates that fit of the slope of the van't Hoff plot is the same whether or not curvature is accounted for.

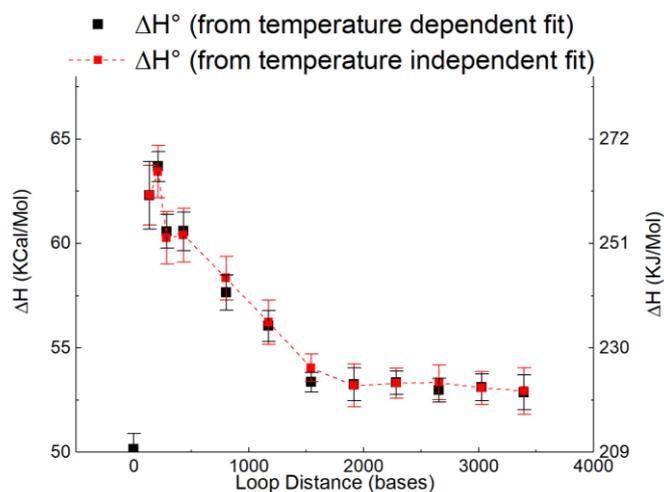


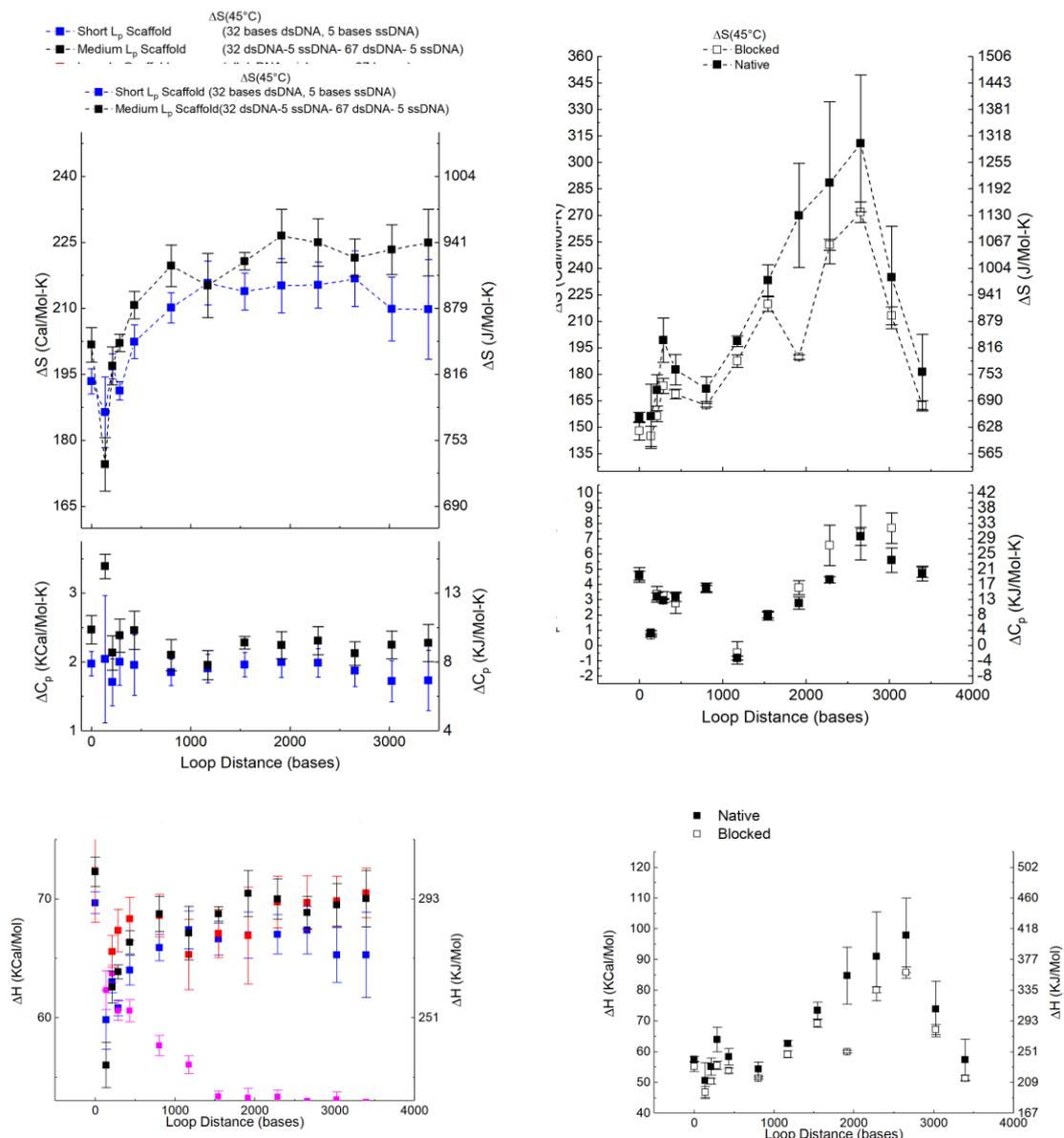
Figure s10.  $\Delta H^0$  values for fitted  $\Delta C_p$  and  $\Delta C_p = 0$ .

## Section 8: Data from All $L_p$ Systems

### Subsection 8a: $\Delta H(T)/\Delta S(T)$ from all $L_p$ systems

Below are the  $\Delta H$  and/or  $\Delta S$  for the various persistence length systems. Overall, the native/blocked systems to have too much variation in secondary structure to provide useful information. It is worth mentioning that some of these systems failed to achieve 5 % relative error between curves during affine transformation. Similarly, while it is possible to distinguish the small changes in  $\Delta S$  as a function of loop distance, when this is combined with simultaneous small changes in  $\Delta H$  associated with increased persistence length at short distances it becomes difficult to identify a trend in the data.

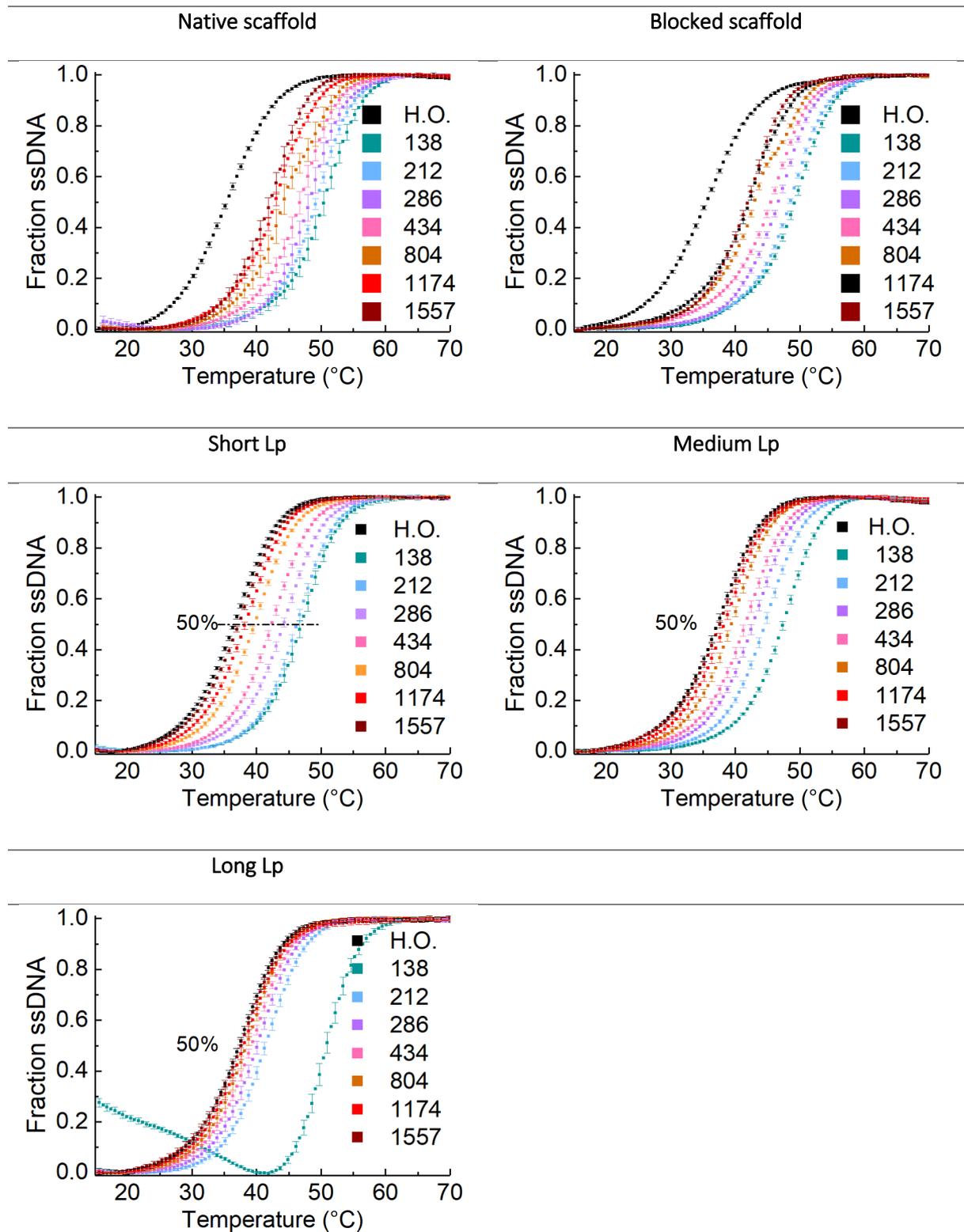
It is possible that with higher quality data the differences in  $\Delta H$  and/or  $\Delta S$  under these conditions could be resolved





### Subsection 8b: Average fluorescence from all systems

The table below shows the average melt traces for all  $L_p$  systems.

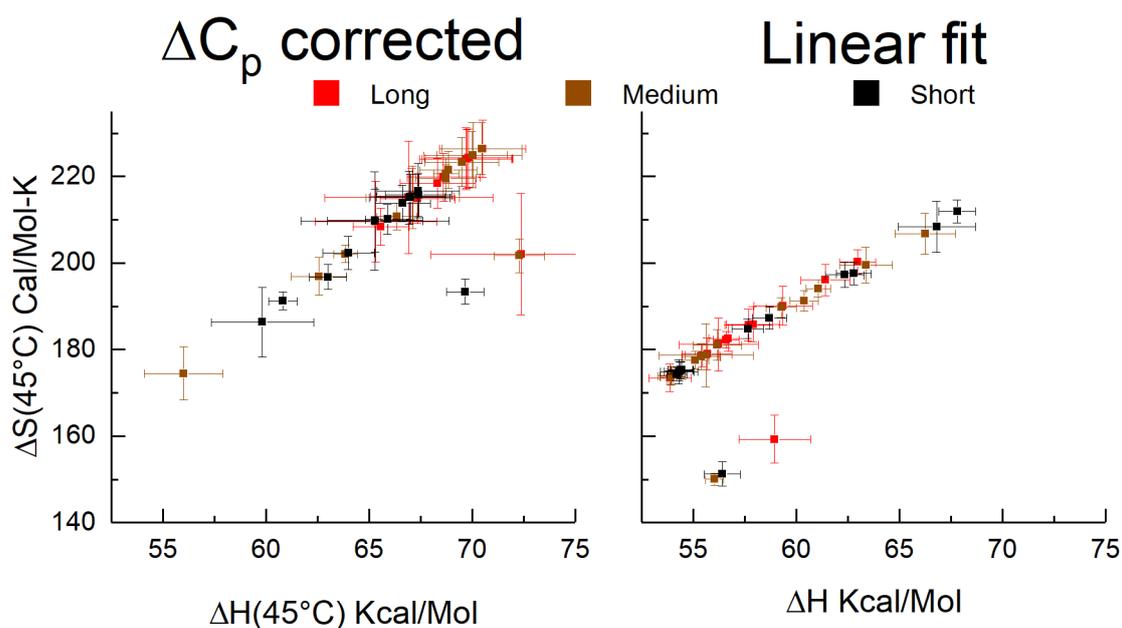


## Section 9: Entropy Enthalpy Compensation

For literature discussion of entropy/enthalpy compensation, we refer the reader to the following review. While this compensation can be real, it is often an artifact resulting from measurement uncertainty or neglect of  $\Delta C_p$ .<sup>5</sup>

The figures below show plotted entropy and enthalpy values for the linear fit, neglecting  $\Delta C_p$ , on the right, with the same plot is shown on the left for values fitted including  $\Delta C_p$ . It is worth noting that the error bars in these figures and in the main text are based on the standard deviation between replicates and don't include the quality of fit. While the error bars in the figure on the left are larger, the residuals on this fit are much lower. It is more accurate to consider the error bars on the left as an underrepresentation of uncertainty.

The important distinction between these figures is the extent to which the values in the linear fit are spread out evenly across compensating  $\Delta H$  and  $\Delta S$ , while the extracted values are much more clustered when  $\Delta C_p$  is fitted. This indicates that much of the observed  $\Delta H$ - $\Delta S$  compensation is statistical in nature.



## Section 10: uncertainty in van't Hoff analysis of optical melt curves

The propagation of uncertainty in van't Hoff analysis is worth consideration if one wishes to extract thermodynamic parameters from optical melt curves. Assume that the uncertainty on the true signal is  $\pm N$ , as below.

$$\sigma(T) = \sigma(T) \pm N$$

Given that van't Hoff analysis is performed on  $\ln(K)$  vs  $1/T$ , one calculates the equilibrium constant, shown below for unimolecular reactions.

$$K(T) \pm \delta = \frac{[ssDNA]}{[1 - ssDNA]} = \frac{[\sigma(T) \pm N]}{[1 - \sigma(T) \pm N]}$$

$$\delta = \sqrt{\left(\frac{N}{\sigma(T)}\right)^2 + \left(\frac{N}{1 - \sigma(T)}\right)^2}$$

As the melt curve,  $\sigma(T)$ , approaches either plateau, i.e., as  $[ssDNA]$  approaches 0, or  $[1 - ssDNA]$  approaches 0, the equilibrium constant is dominated by noise. As such, in fitting  $\ln(K)$  vs  $1/T$ , we only used the middle 9 points, or 5.5 °C, of the melt curve.

## Section 11: fold strand sequences

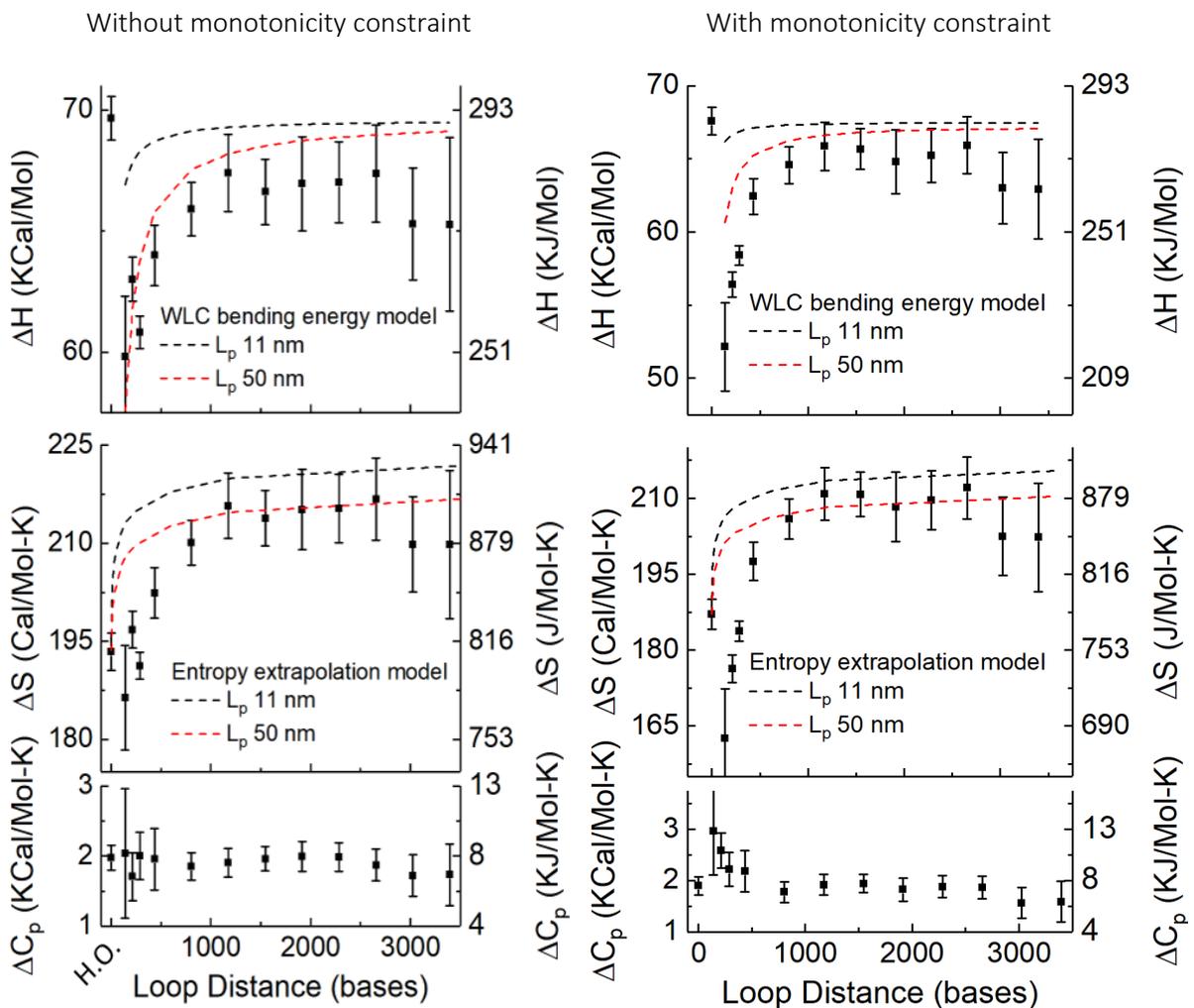
The subsequences listed below match those described ( $\alpha, \beta', \beta, \gamma, \delta$ ) in Fig. 1c of the main text. As such the full strand sequences contain combinations of these with different anchor sequences,  $\gamma$ , to enforce different fold distances.

### SubSequences

Fluorophore ( $\delta$ )	GTC CAC TAT TAA AGA ACG TGG ACT CCA ACG TCA AAG GGC GAA - 3'JOE
Quench Complement ( $\beta'$ )	C GCT GAT GGT AAC TGT CGG GTA TAA CGT CGG TAG CAG GCG
Quench ( $\beta$ )	<b>Iowa Black Quencher 5'</b> - CGC CTG CTA CCG ACG TTA TAC CCG ACA GTT ACC ATC AGC G
Melt ( $\alpha$ )	AAA CCG TCT A
Truncated Scaffold	GGCACCTGATTGGGTGATGGTTCACGTAGTGGGCCATCGCCCTGATAGACGGTTTTTCGCCCTTGACGTTGGAGTCCACGTTCTTAATAGTGGACT C
	<b>Full Strand Sequence (anchor subsequence bolded)</b>
Hybridization Only	CGC TGA TGG TAA CTG TCG GGT ATA ACG TCG GTA GCA GGC GAA ACC GTC TA
Fold - 138	CCA CGC TGG TTT GCC CCA GCA GGC GAA AAC GCT GAT GGT AAC TGT CGG GTA TAA CGT CGG TAG CAG GCG AAA CCG TCT A
Fold - 212	TCT TTT CAC CAG CGA GAC GGG CAA CAG CTC GCT GAT GGT AAC TGT CGG GTA TAA CGT CGG TAG CAG GCG AAA CCG TCT A
Fold - 286	CGT GCC AGC TGC ATT AAT GAA TCG GCC AAC GCT GAT GGT AAC TGT CGG GTA TAA CGT CGG TAG CAG GCG AAA CCG TCT A
Fold - 434	TTT CCT GTG TGA AAT TGT TAT CCG CTC ACC GCT GAT GGT AAC TGT CGG GTA TAA CGT CGG TAG CAG GCG AAA CCG TCT A
Fold - 804	ACG TTG GTG TAG ATG GGC GCA TCG TAA CCC GCT GAT GGT AAC TGT CGG GTA TAA CGT CGG TAG CAG GCG AAA CCG TCT A

Fold - 1174 CTA TTT TTG AGA GAT CTA CAA AGG CTA TCC GCT GAT GGT AAC TGT CGG GTA TAA CGT CGG TAG CAG GCG AAA CCG TCT A  
 Fold - 1544 TGT TTA GCT ATA TTT TCA TTT GGG GCG CGC GCT GAT GGT AAC TGT CGG GTA TAA CGT CGG TAG CAG GCG AAA CCG TCT A  
 Fold - 1914 AAA AAT CAG GTC TTT ACC CTG ACT ATT ATC GCT GAT GGT AAC TGT CGG GTA TAA CGT CGG TAG CAG GCG AAA CCG TCT A  
 Fold - 2284 GCT CAT TAT ACC AGT CAG GAC GTT GGG AAC GCT GAT GGT AAC TGT CGG GTA TAA CGT CGG TAG CAG GCG AAA CCG TCT A  
 Fold - 2655 AAG CGC GAA ACA AAG TAC AAC GGA GAT TTC GCT GAT GGT AAC TGT CGG GTA TAA CGT CGG TAG CAG GCG AAA CCG TCT A  
 Fold - 3025 CCA AAA GGA GCC TTT AAT TGT ATC GGT TTC GCT GAT GGT AAC TGT CGG GTA TAA CGT CGG TAG CAG GCG AAA CCG TCT A  
 Fold - 3394 ATA GGT GTA TCA CCG TAC TCA GGA GGT TTC GCT GAT GGT AAC TGT CGG GTA TAA CGT CGG TAG CAG GCG AAA CCG TCT A

## Section 12: Extracted thermodynamics with and without low temperature monotonicity constraint



The monotonicity constraint primarily affected short fold distances as shown in the figure above. There were only approximately 6 data points below the melt curve to which this constraint could reasonably be applied. As such, it is likely that the improvement in analysis quality when removing the constraint are due to insufficient data points for its proper application, rather than some fundamental failure of the constraint.

## Section 13: Discussion of heat capacity, enthalpy and entropy conventions

To be brief, we here use conventions for entropy and enthalpy already existing in the literature<sup>5-7</sup> with the understanding that they may or may not continue to be the primary convention.

By this definition  $\Delta H(T)$  and  $\Delta S(T)$  are the slope and intercept of the derivative of the  $\ln(K)$  vs  $1/T$  curve. As such, both  $\Delta H(T)$  and  $\Delta S(T)$  vary as a function of temperature, where the temperature dependence of both (as well as the curvature of the plot) are determined by the  $\Delta C_p$ .

The benefit of this convention is that it describes the change in entropy and enthalpy of a transition, at a given temperature, while accounting for the difference in heat capacity between the two states.



Fold	Position from B-Side which would block the anchor for that fold..
138	A1,A2
212	A3,A4
286	A5,A6
434	A9,A10
804	B7,B8
1174	C5,C6
1544	D3,D4
1914	E1,E2
2284	E11,E12
2655	F9,F10
3025	G7,G8
3394	H5,H6

#### DNA Sequences

##### A-Side, 32bp oligos

##### Sequence

##### Pos

tccagtaagcgtcatacatggctttgatgat	3690
aagccagaatggaaagcgagctctgaattt	3727
ttggccttgatattcacaacgaatggatctt	3764
cgccagcattgacaggaggtgaggcaggtca	3801
ccctcagagccgccaccagaaccaccaccaga	3838
gccgccaccctcagaaccgccaccctcagagc	3875
cggaaaccagagccaccaccggaaccgcctccc	3912
ttattagcgtttgccatctttcataatcaaa	3949
tagcgcgtttcatcggcattttcggcatag	3986
tagcgacagaatcaagttgcctttagcgtca	4023
tcaccaatgaaaccatcgatagcagcaccgta	4060
accagtagcaccattaccattagcaaggccgg	4097
cgacttgagccatttgggaattagagccagca	4134
cggaaattattcattaaaggtgaattatcacc	4171

##### A-Side, 37bp oligos Sequence

##### Pos

tccagtaagcgtcatacatggctttgatgat	3690
aagccagaatggaaagcgagctctgaattaccgt	3727
ttggccttgatattcacaacgaatggatcttcatta	3764
cgccagcattgacaggaggtgaggcaggtcagacga	3801
ccctcagagccgccaccagaaccaccaccagaccgc	3838
gccgccaccctcagaaccgccaccctcagagccacca	3875
cggaaaccagagccaccaccggaaccgcctcccaga	3912
ttattagcgtttgccatctttcataatcaaaatcac	3949
tagcgcgtttcatcggcattttcggcatagcccc	3986
tagcgacagaatcaagttgcctttagcgtcagactg	4023
tcaccaatgaaaccatcgatagcagcaccgtaacg	4060
accagtagcaccattaccattagcaaggccggaacg	4097
cgacttgagccatttgggaattagagccagcaaaatc	4134
cggaaattattcattaaaggtgaattatcaccgtcac	4171

acattcaaccgattgagggaggggaaggtaaat	4208	acattcaaccgattgagggaggggaaggtaaatattga	4208
attcatatggtttaccagcgctaaagacaaaa	4245	attcatatggtttaccagcgctaaagacaaaagggcg	4245
ccacggaataagtttattttgtcacaatcaat	4282	ccacggaataagtttattttgtcacaatcaatagaaa	4282
cataaaggtggcaacatataaaagaaacgcaa	4319	cataaaggtggcaacatataaaagaaacgcaaagaca	4319
tattacgcagtatgtagcaaacgtagaaaat	4356	tattacgcagtatgtagcaaacgtagaaaatacata	4356
cggaatacccaaaagaactggcatgattaaga	4393	cggaatacccaaaagaactggcatgattaagactcct	4393
aagttaccagaaggaaaccgaggaaacgcaat	4430	aagttaccagaaggaaaccgaggaaacgcaataataa	4430
aagcccttttaagaaaagtaagcagatagcc	4467	aagcccttttaagaaaagtaagcagatagccgaaca	4467
agcaagaaacaatgaaatagcaatagctatct	4504	agcaagaaacaatgaaatagcaatagctatcttaccg	4504
gataaccacaagaattgagttaagccaataa	4541	gataaccacaagaattgagttaagccaataataag	4541
acaaagtacagaggtaattgagcgctaataatc	4578	acaaagtacagaggtaattgagcgctaataatcagaga	4578
aagcgcattagacgggagaattaactgaacac	4615	aagcgcattagacgggagaattaactgaacaccctga	4615
tagcagcctttacagagagaataacataaaaa	4652	tagcagcctttacagagagaataacataaaaaacaggg	4652
taagaaacgattttttgtttaacgtcaaaaat	4689	taagaaacgattttttgtttaacgtcaaaaatgaaaa	4689
acaaaataaacagccatattttatcccaat	4726	acaaaataaacagccatattttatcccaatccaaa	4726
gtaacgagcgtctttccagagcctaatttgc	4763	gtaacgagcgtctttccagagcctaatttgcagtt	4763
tgcaccagctacaattttatcctgaatctta	4800	tgcaccagctacaattttatcctgaatcttaccac	4800
aggttttgaagccttaaatcaagattagttgc	4837	aggttttgaagccttaaatcaagattagttgctattt	4837
aacgcgaggcgttttagcgaacctcccgactt	4874	aacgcgaggcgttttagcgaacctcccgacttgcggg	4874
gcaaatcagatatagaaggcttatccggtatt	4911	gcaaatcagatatagaaggcttatccggtattctaag	4911
ttttcatcgtaggaatcattaccgcccgaat	4948	ttttcatcgtaggaatcattaccgcccgaatagcaa	4948
agtaccgcactcatcgagaacaagcaagccgt	4985	agtaccgcactcatcgagaacaagcaagccgtttta	4985
gtctttccttatcattccaagaacgggtatta	5022	gtctttccttatcattccaagaacgggtattaacca	5022
aatttacgagcatgtagaaaccaatcaataat	5059	aatttacgagcatgtagaaaccaatcaataatcggct	5059
agataagtcctgaacaagaaaaataatccc	5096	agataagtcctgaacaagaaaaataatcccacct	5096
gttcagctaatacgagaacgcgcctgtttatca	5133	gttcagctaatacgagaacgcgcctgtttatcaacaat	5133
aaagtaattctgtccagacgacgacaataaac	5170	aaagtaattctgtccagacgacgacaataaacat	5170

gccagtaataagagaatataaagtaccgacaa	5207	gccagtaataagagaatataaagtaccgacaaaaggt	5207
caacgccaacatgtaatttaggcagaggcatt	5244	caacgccaacatgtaatttaggcagaggcatttcga	5244
tcaacagtagggcttaattgagaatcgccata	5281	tcaacagtagggcttaattgagaatcgccatattta	5281
tgcgttatacaaattcttaccagtataagcc	5318	tgcgttatacaaattcttaccagtataagccaacgc	5318
aatcataattactagaaaaagcctgttagta	5355	aatcataattactagaaaaagcctgttagtatcata	5355
tgtgataaataaggcgtaaataagaataaac	5392	tgtgataaataaggcgtaaataagaataaacaccgg	5392
cttctgacctaaatthaatggttgaaatacc	5429	cttctgacctaaatthaatggttgaaataccgaccg	5429
gagaaaacttttcaaataatatttagttaat	5466	gagaaaacttttcaaataatatttagttaatcat	5466
aatgctgatgcaaatccaatcgcaagacaaag	5503	aatgctgatgcaaatccaatcgcaagacaaagaacgc	5503
aacctccggcttaggttgggttatataactat	5540	aacctccggcttaggttgggttatataactatatgta	5540
attatcaaaatcataggtctgagagactacc	5577	attatcaaaatcataggtctgagagactacctttt	5577
tagcttagattaagacgctgagaagagtcaat	5614	tagcttagattaagacgctgagaagagtcaatagtgta	5614
aattaattttcccttagaatccttgaaaacat	5651	aattaattttcccttagaatccttgaaaacatagcga	5651
tgagtgaataaccttgcttctgtaaatcgctc	5688	tgagtgaataaccttgcttctgtaaatcgctcatt	5688
cctttttaatggaaacagtacataaatcaat	5725	cctttttaatggaaacagtacataaatcaatatg	5725
acaaaattaattacatttaacaatttcattg	5762	acaaaattaattacatttaacaatttcattgaatta	5762
ctgagcaaaagaagatgatgaaacaaacatca	5799	ctgagcaaaagaagatgatgaaacaaacatcaagaaa	5799
caaaatcgcgagaggcgaattattcatttca	5836	caaaatcgcgagaggcgaattattcatttcaattac	5836
ataacggattcgctgattgcttgaatacca	5873	ataacggattcgctgattgcttgaataccaagtta	5873
atatacagtaacagtacctttacatcgggag	5910	atatacagtaacagtacctttacatcgggagaaaca	5910
agaaattgcgtagattttcaggttaacgtca	5947	agaaattgcgtagattttcaggttaacgtcagatga	5947
accatatcaaaattattagcacgtaaaacaga	5984	accatatcaaaattattagcacgtaaaacagaataa	5984
tggattatacttctgaattatggaaggaattg	6021	tggattatacttctgaattatggaaggaattgaacca	6021
cagatgatggcaattcatcaatataatcctga	6058	cagatgatggcaattcatcaatataatcctgattgtt	6058
ccagaaggagcgggaattatcatcatattcctg	6095	ccagaaggagcgggaattatcatcatattcctgattat	6095
ttgagtaacattatcattttgcggaacaaaga	6132	ttgagtaacattatcattttgcggaacaaagaacca	6132
ttaaatcctttgccgaacgttattaatttta	6169	ttaaatcctttgccgaacgttattaattttaaaagt	6169

gaagtattagactttacaacaattcgacaac	6206	gaagtattagactttacaacaattcgacaactcgta	6206
attagagccgtcaatagataatacatttgagg	6243	attagagccgtcaatagataatacatttgaggattta	6243
tatctaaaatatcttttaggtgcactaacaact	6280	tatctaaaatatcttttaggtgcactaacaactaatag	6280
gttggcaaatcaacagtagaaaggaattgagg	6317	gttggcaaatcaacagtagaaaggaattgaggaggt	6317
aacctcaaatatcaaaccctcaatcaatatct	6354	aacctcaaatatcaaaccctcaatcaatatctggtca	6354
cagcagcaaatgaaaaatctaaagcatcacct	6391	cagcagcaaatgaaaaatctaaagcatcaccttgctg	6391
gtattaacaccgctgcaacagtgccacgctg	6428	gtattaacaccgctgcaacagtgccacgctgagagc	6428
accaccagcagaagataaaaacagaggtgaggc	6465	accaccagcagaagataaaaacagaggtgaggcggtca	6465
atagccctaaaacatcgccattaaaaataccg	6502	atagccctaaaacatcgccattaaaaataccgaacga	6502
atTTTTgaatggctattagtctttaatgcgcg	6539	atTTTTgaatggctattagtctttaatgcggaactg	6539
ctgacctgaaagcgtaagaatactggcacag	6576	ctgacctgaaagcgtaagaatactggcacagacaat	6576
taaaaggacattctggccaacagagatagaa	6613	taaaaggacattctggccaacagagatagaaccctt	6613
tttaccattggcagattcaccagtcacacgacc	6650	tttaccattggcagattcaccagtcacacgaccagtaa	6650
acctacatttgacgctcaatcgtctgaaatg	6687	acctacatttgacgctcaatcgtctgaaatggatta	6687
cgccagccattgcaacaggaaaaacgctcatg	6724	cgccagccattgcaacaggaaaaacgctcatggaat	6724
ctatcggccttgctggaatatccagaacaat	6761	ctatcggccttgctggaatatccagaacaatattac	6761
agtaataacatcacttgctgagtagaagaac	6798	agtaataacatcacttgctgagtagaagaactcaaa	6798
cacgcaaattaaccgtttagcaatacttctt	6835	cacgcaaattaaccgtttagcaatacttctttgatt	6835
taatcagtgaggccaccgagtaaaagagtctg	6872	taatcagtgaggccaccgagtaaaagagtctgtccat	6872
caggaacggtacgccagaatcttgagaagtgt	6909	caggaacggtacgccagaatcttgagaagtgtttta	6909
gggagctaaacaggaggccgattaaaggatt	6946	gggagctaaacaggaggccgattaaaggattttaga	6946
agcacgtataacgtgcttctcgttgaatc	6983	agcacgtataacgtgcttctcgttgaatcagagc	6983
gcgccgctacagggcgctactatggttgctt	7020	gcgccgctacagggcgctactatggttgcttgacg	7020
cacgctgcgcgtaaccaccacaccgcccgcg	7057	cacgctgcgcgtaaccaccacaccgcccgcgctaat	7057
aaggagcggcgctagggcgctggcaagtgt	7094	aaggagcggcgctagggcgctggcaagtgtagcggt	7094
ccggcgaacgtggcgagaaaggaaggaagaa	7131	ccggcgaacgtggcgagaaaggaaggaagaaagcga	7131
taaaggagccccgatttagagcttgacggg	7168	taaaggagccccgatttagagcttgacgggaaag	7168

tggggtcgaggtgccgtaaagcactaaatcg  
7205

tggggtcgaggtgccgtaaagcactaaatcggaaccc 7205

**B-side Oligos, 32bp**

**s**

**Sequence**

**Pos**

gatggtggttccgaaatcggaatccctta	143
ccacgctggttgcagcagggcaaatcc	180
ttcaccgctggcctgagagagttgagcaa	217
tctttcaccagcgagacgggcaacagctgat	254
gagaggcggttgcgtattggcgccagggtg	291
cgtagcagctgattaatgaatcgccaacgc	328
ttgcgctcactgcccgtttcagtcgggaaa	365
gggtgcctaataagtgagtaactcacattaa	402
acaacatacagccggaagcataaagtgtaaa	439
tttctgtgtgaaattgtatccgctcacaat	476
ggtaccgagctcgaattcgaatcatggtcat	513
gcttgcatgcctgcaggtcgactctagaggat	550
cccagtcacgacgttgtaaacgacggccagtc	587
gctgcaaggcgattaagttgggtaacccaggg	624
ctcttcgctattacgccagctggcgaaagggg	661
ggctgcgcaactgttgggaagggcgatcggtg	698
ggtgccggaaccaggcaaagcgcattcgccat	735
aagatcgactccagccagctttccggcaccg	772
gccagtttgaggggacgacgaccgtatcgcc	809
acgttggtgtagatgggcatcgtaaccgtg	846
gggaacaaacggcggttaccgtaatgggat	883
taaatgtgagcgagtaacaaccgctcggttc	920
cgctctggccttctgtagccagctttcatc	957

**B-side Oligos, 37bp**

**Sequence**

**Pos**

gatggtggttccgaaatcggaatccctataaat	143
ccacgctggttgcagcagggcaaatccctggtt	180
ttcaccgctggcctgagagagttgagcaagcggt	217
tctttcaccagcgagacgggcaacagctgattgcc	254
gagaggcggttgcgtattggcgccagggtggttt	291
cgtagcagctgattaatgaatcgccaacgcgcggg	328
ttgcgctcactgcccgtttcagtcgggaaacctgt	365
gggtgcctaataagtgagtaactcacattaaatgcg	402
acaacatacagccggaagcataaagtgtaaagcctg	439
tttctgtgtgaaattgtatccgctcacaattccac	476
ggtaccgagctcgaattcgaatcatggtcatagctg	513
gcttgcatgcctgcaggtcgactctagaggatccccg	550
cccagtcacgacgttgtaaacgacggccagtcgcaa	587
gctgcaaggcgattaagttgggtaacccaggggttt	624
ctcttcgctattacgccagctggcgaaaggggatgt	661
ggctgcgcaactgttgggaagggcgatcggtgcgggc	698
ggtgccggaaccaggcaaagcgcattcgccattca	735
aagatcgactccagccagctttccggcaccgcttct	772
gccagtttgaggggacgacgaccgtatcgccctcagg	809
acgttggtgtagatgggcatcgtaaccgtgcatct	846
gggaacaaacggcggttaccgtaatgggataggtt	883
taaatgtgagcgagtaacaaccgctcggttctccgt	920
cgctctggccttctgtagccagctttcatcaacat	957

catttttaaccaataggaacgcatcaaaaa	994	catttttaaccaataggaacgcatcaaaaaataatt	994
ttgttaaattcgcgttaaattttgttaa	1031	ttgttaaattcgcgttaaattttgttaa	1031
tgataagcaaatatttaaattgtaa	1068	tgataagcaaatatttaaattgtaa	1068
cggttgataatcagaaaagccccaaaa	1105	cggttgataatcagaaaagccccaaaa	1105
ggtaatcgtaaaactagcatgtcaat	1142	ggtaatcgtaaaactagcatgtcaat	1142
gcctgagagtctggagcaacaagaga	1179	gcctgagagtctggagcaacaagaga	1179
ctattttgagagatctacaaaggct	1216	ctattttgagagatctacaaaggct	1216
accgttctagctgataaattaatgcc	1253	accgttctagctgataaattaatgcc	1253
gccggagacagtcaaatcacatcaat	1290	gccggagacagtcaaatcacatcaat	1290
agtaatgtgtaggtaaagattcaaa	1327	agtaatgtgtaggtaaagattcaaa	1327
ttttagaaccctcatatattttaa	1364	ttttagaaccctcatatattttaa	1364
tcggggagaagcctttatttcaacg	1401	tcggggagaagcctttatttcaacg	1401
cggttgtaacaaaaacattatgacc	1438	cggttgtaacaaaaacattatgacc	1438
aaattaagcaataaagcctcagagc	1475	aaattaagcaataaagcctcagagc	1475
atccaataaatcatacaggcaaggca	1512	atccaataaatcatacaggcaaggca	1512
aggtagcatcaatttactaataagta	1549	aggtagcatcaatttactaataagta	1549
tgtttagctatatttcatttggggcg	1586	tgtttagctatatttcatttggggcg	1586
tgaccattagatacatttcgcaaat	1623	tgaccattagatacatttcgcaaat	1623
cagttgattccaattctcgaacgag	1660	cagttgattccaattctcgaacgag	1660
actaaagtacggtgtctggaagttc	1697	actaaagtacggtgtctggaagttc	1697
aatctggtgctgtagctcaacatgt	1734	aatctggtgctgtagctcaacatgt	1734
aggtcattttgcggatggcttagag	1771	aggtcattttgcggatggcttagag	1771
ggattagagagtaccttaattgctc	1808	ggattagagagtaccttaattgctc	1808
aaagcgaaccagaccggaagcaaac	1845	aaagcgaaccagaccggaagcaaac	1845
cgaagacttcaaatatcgcgtttta	1882	cgaagacttcaaatatcgcgtttta	1882
gcaaagcggattgcatcaaaaagatt	1919	gcaaagcggattgcatcaaaaagatt	1919
aaaaatcaggtctttaccctgactat	1956	aaaaatcaggtctttaccctgactat	1956

ctttaaacagttcagaaaacgagaatgaccat	1993	ctttaaacagttcagaaaacgagaatgaccataaatc	1993
gaatcgtcataaatattcattgaatccccctc	2030	gaatcgtcataaatattcattgaatccccctcaaatg	2030
agtaaaatgtttagactggatagcgtccaata	2067	agtaaaatgtttagactggatagcgtccaatactgcg	2067
aggctttgcaaaagaagtttgccagagggg	2104	aggctttgcaaaagaagtttgccagaggggtaat	2104
tcgtttaccagacgacgataaaaacaaaata	2141	tcgtttaccagacgacgataaaaacaaaatagcgag	2141
attacgaggcatagtaagagcaacactatcat	2178	attacgaggcatagtaagagcaacactatcataccc	2178
cacattcaactaatgcagatacataacgcaa	2215	cacattcaactaatgcagatacataacgcaaaagga	2215
caggtagaaagattcatcagttgagatttagg	2252	caggtagaaagattcatcagttgagatttaggaatac	2252
tacgttaataaaaacgaactaacggaacaacat	2289	tacgttaataaaaacgaactaacggaacaacattatta	2289
gctcattataccagtcaggacgttggaagaa	2326	gctcattataccagtcaggacgttggaagaaaaatc	2326
aatcattgtgaattaccttatgcgattttaag	2363	aatcattgtgaattaccttatgcgattttaagaactg	2363
tagtaaattgggcttgagatggtttaattca	2400	tagtaaattgggcttgagatggtttaattcaacttt	2400
tgaataaggcttgcctgacgagaacaccaga	2437	tgaataaggcttgcctgacgagaacaccagaacgag	2437
cattacccaaatcaacgtaacaaagctgctca	2474	cattacccaaatcaacgtaacaaagctgctcattcag	2474
ccttcatcaagagtaatcttgacaagaaccgg	2511	ccttcatcaagagtaatcttgacaagaaccgatatt	2511
atgaacggtgtacagaccaggcgcataggctg	2548	atgaacggtgtacagaccaggcgcataggctgctga	2548
taagggaaccgaactgaccaacttgaaagag	2585	taagggaaccgaactgaccaacttgaaagaggacag	2585
tgttacttagccggaacgaggcgacagcggtc	2622	tgttacttagccggaacgaggcgacagcggtcaatca	2622
gcctgataaattgtgtcgaaatccgacactg	2659	gcctgataaattgtgtcgaaatccgacactgctcca	2659
aagcgcgaaacaaagtacaacggagatttgta	2696	aagcgcgaaacaaagtacaacggagatttgatcatc	2696
actaaaactcatctttgacccccagcgatt	2733	actaaaactcatctttgacccccagcgattatacc	2733
aaggcaccaacctaaaacgaaagaggcgaag	2770	aaggcaccaacctaaaacgaaagaggcgaagaatac	2770
tttcattaacgggtaaaatacgtaatgcca	2807	tttcattaacgggtaaaatacgtaatgcccactacg	2807
agaggctttgaggactaaagacttttcatga	2844	agaggctttgaggactaaagacttttcatgaggaag	2844
cgaaagacagcatcggaacgagggtagcaacg	2881	cgaaagacagcatcggaacgagggtagcaacggctac	2881
ttaaaggccgcttttgcgggatcgtcacctc	2918	ttaaaggccgcttttgcgggatcgtcacctcagcag	2918
ataaccgatattcggtcgctgaggcttgca	2955	ataaccgatattcggtcgctgaggcttgaggag	2955

tagttgcgccgacaatgacaacaaccatcgcccagc	2992	tagttgcgccgacaatgacaacaaccatcgcccagc	2992
gctttcgaggngaatttcttaaacagcttgat	3029	gctttcgaggngaatttcttaaacagcttgat	3029
ccaaaaggagccttaattgtatcggtttatc	3066	ccaaaaggagccttaattgtatcggtttatc	3066
aatttttcacgttgaaaatctccaaaaaaa	3103	aatttttcacgttgaaaatctccaaaaaaa	3103
agaatagaaaggaacaactaaaggaattgca	3140	agaatagaaaggaacaactaaaggaattgca	3140
gttttgctaaacaactttcaacagtttcagcggagtg	3177	gttttgctaaacaactttcaacagtttcagcggagtg	3177
tctttccagacgtagtaaatgaattttctgt	3214	tctttccagacgtagtaaatgaattttctgt	3214
accctcatagttagcgtaacgatctaaagttt	3251	accctcatagttagcgtaacgatctaaagttt	3251
cagtacaaactacaacgcctgtagcattccac	3288	cagtacaaactacaacgcctgtagcattccac	3288
ataggaaccatgtaccgtaacactgagtttc	3325	ataggaaccatgtaccgtaacactgagtttc	3325
agagccaccaccctcattttcagggatagcaa	3362	agagccaccaccctcattttcagggatagcaa	3362
caccctcagaaccgccaccctcagaaccgcca	3399	caccctcagaaccgccaccctcagaaccgcca	3399
ataggtgtatcaccgtactcaggaggtttagt	3436	ataggtgtatcaccgtactcaggaggtttagt	3436
agtgccgtcgagagggttgatataagtatagc	3473	agtgccgtcgagagggttgatataagtatagc	3473
taggattagcggggtttgctcagtaccaggc	3510	taggattagcggggtttgctcagtaccaggc	3510
aaagtattaagaggctgagactcctcaagaga	3547	aaagtattaagaggctgagactcctcaagaga	3547
cccctgcctatttcggaacctattattctgaa	3584	cccctgcctatttcggaacctattattctgaa	3584
tgcttgagtaacagtgccgtataaacagtt	3621	tgcttgagtaacagtgccgtataaacagtt	3621
acaggagtgtactggtaataagttttaacggg	3658	acaggagtgtactggtaataagttttaacggg	3658

## References

1. Schneider, F., Möritz, N. & Dietz, H. The sequence of events during folding of a DNA origami. *Sci. Adv.* **5**, (2019).
2. Turner, D. H. & Mathews, D. H. NNDB: The nearest neighbor parameter database for predicting stability of nucleic acid secondary structure. *Nucleic Acids Res.* **38**, 2009–2011 (2009).
3. Pedersen, R. *et al.* Properties of DNA. in *Handbook of Nanomaterials properties* 1125–1157 (2015). doi:10.1007/978-3-642-31107-9\_10

4. Measurements, F. Estimating the Persistence Length of a Worm-Like Chain Molecule from Force-Extension Measurements. **76**, 409–413 (1999).
5. Sharp, K. Entropy--enthalpy compensation: Fact or artifact? *Protein Sci.* **10**, 661–667 (2001).
6. Rouzina, I. & Bloomfield, V. A. Heat Capacity Effects on the Melting of DNA. 1. General Aspects. **77**, 3242–3251 (1999).
7. Mikulecky, P. J. & Feig, A. L. Heat Capacity Changes Associated with Nucleic Acid Folding. *Biopolymers* **81**, 38–58 (2006).