nature portfolio

Peer Review File

DNA Calorimetric Force Spectroscopy at Single Base Pair Resolution

Corresponding Author: Professor Felix Ritort

This file contains all reviewer reports in order by version, followed by all author rebuttals in order by version.

Version 0:

Reviewer comments:

Reviewer #1

(Remarks to the Author)
In this paper, Rissone et al. use a temperature jump optical trap to mechanically unzip single DNA molecules and employ theoretical analysis to derive the heat capacity change parameters of hybridization for the different DNA base pair motifs in the nearest-neighbor model. The heat capacity change ΔCp informs quantitatively about the temperature dependence of enthalpic and entropic contributions to the free energy hybridization, but experimental measurements through bulk and single-molecule experiments to date have yielded vastly different results, depending on the specific experimental conditions. This work seeks to resolve the challenges in determining ΔCp values through systematic mechanical pulling experiments over a range of temperatures and subsequent detailed theoretical analysis. The authors have laid the foundation for this work through a series of prior publications in mechanical pulling experiments using optical tweezers. The methodology put forward has been demonstrated on a DNA hairpin here and can be feasibly extended to other types of nucleic acids. I would recommend publication after the authors have addressed the following points:
1. The experiments were conducted at temperatures between 7°C and 42°C in increments of 2.5°C. How did the authors measure and ensure a constant temperature within the local environment in the microfluidic chamber? Can the authors also provide more detail on the design of the chamber and how it damps convection effects caused by the non-uniform temperature from the laser?
2. It is difficult to see the curves in Figure 1B, particularly with the error bars. I would suggest the authors remove the error bars from the figure and have an additional figure in the SI with the error bars, possibly plotted with some offset in the y-axis for better visualization of the error bars. The current error bars are plotted showing molecule-to-molecule variability, what about the variability for the same molecule across the different unzipping/rezipping cycles? The authors should also state what the error bars represent, e.g. confidence interval.
3. In computing the FEC from the experimental FDC, was keff estimated for each experimental run? If so, it would be interesting to report the variation in keff values across runs to validate the assumption of force-independent trap stiffness.
4. The errors reported in Figures 2 and 4, as well as the tables, should specify the confidence interval. It would also be helpful to include a brief overview of how the errors were calculated in the SI. I would imagine this was a nontrivial calculation involving error propagation and errors in fits to data.
5. Page 4, Line 148. I believe the authors are referring to Fig. 1B, not Fig. 2B.
6. Suggestion: add the loop size L to the schematic in Fig. 1A.
7. It is not very clear to me how the authors fit the WLC to the FEC to obtain the persistence length and inter-phosphate distance. There are two fitted parameters obtained from a fit of one equation to the data, this would add to the uncertainty of the fitted parameters and introduce some coupling between the uncertainties. As an aside, what is zss in the legend of Fig. 1, extended data?
8. For the decomposition of the FDC into different segments, how were the peaks corresponding to force rips identified? It is not very clear from Fig. 2C.
9. For the derivation of the free energies Δgi(T), it is not clear just from the explanation in the main text (Lines 238-243) how this is carried out. The authors can consider moving some of the details from the SI into the main text for this. In particular, more can be said about the theoretical FDCs and their role in obtaining the free energies.
10. Can the authors comment on the physicality of the obtained thermodynamic parameter values? For example, are the differences in values of Δsi(T), Δgi(T), Δhi(T), Δcp,i(T) for the different base pair motifs consistent with what one might expect, from a physical basis and also in comparison to findings in the literature?
11. How would the analysis need to be modified with a change in experimental conditions, for example a different salt condition?

(Remarks on code availability)
I had a look through the code and it is well-organized with enough detail for a reader to follow through with the calculations presented in the paper. I did not try to install and run the code, but there appears to be detailed instructions in a README file for one to do so.

Reviewer #2

(Remarks to the Author)
In this manuscript, Rissone et al. used single-molecule calorimetric force spectroscopy and statistical analysis to extract the thermodynamic parameters for different DNA nearest-neighbor base pair motifs. They showed the temperature dependence of the enthalpy and entropy of hybridization, thereby deriving non-zero heat capacity changes (ΔC_p). These are important results with implications for DNA metabolism, nanotechnology, and simulations. The experiments and analyses are well executed, and the manuscript is concisely written. However, the following points need to be addressed before publication.

(1) The temperature-jump optical trap is a critical feature of the study. But from the methods description it is not clear how to maintain a stable temperature below the room temperature.
(2) The full hairpin sequence used in this study should be made available. In addition, how the full unzipping curve is decomposed into segments needs to be described. It was mentioned that 400-600 segments are generated from the 3.6kbp-long hairpin, and Fig. 2C shows overlap between segments. How are the start and end points of each segment chosen?
(3) On several occasions, it was mentioned that the parameters for motifs GC/CG and TA/AT were obtained by applying the circular symmetry relations. The rationale and approach for this treatment need to be clarified.
(4) Fig. 2A: the persistence length values appear to form three clusters rather than follow a linear relationship with the temperature. The authors should examine whether any experimental artifact exists at high or low temperatures.
(5) Line 167: typo “intextensible”
(6) Line 230: Fig. 3D should be Fig. 2D.
(7) Line 343-344: not sure what “extensive” and “intensive” mean in this context.
(8) Line 457: xh (handle extension) is not expected to remain constant during the entire unzipping process.
(9) Line 462: this relation assumes ΔH to be T-independent?

(Remarks on code availability)
The code is clear and sufficient.

Reviewer #3

(Remarks to the Author)
This is a very interesting paper in which the authors apply single-molecule force measurements to determine the elusive heat capacity changes of DNA at the single base pair (or nearest neighbor) level. They mechanically unzip a long DNA hairpin (3593 bp) using an optical trap setup. The force-distance curve displays a sawtooth pattern whose analysis allows the measurement of entropies, enthalpies, and free energies for the different nearest neighbor base pairs. The results are interesting and important for everyone who studies or uses the biophysical properties of DNA.

Specific questions:

- It is not really explained how the hairpin is made in the first place. What is its sequence? Is it a plasmid that is opened and processed somehow to have a tetraloop? Do you need to anneal it to make sure it is double-stranded, etc.?

- It is not clear to this reviewer how the curve is decomposed into different segments. Apparently, you use K different segments – but what determines K? How does one arrive from 3593 bp at K = 400-600? Or are these just intervals between randomly chosen peaks in the FDC?

- How is the analysis performed then? Do you know the sequence for each given segment? How accurately is the sequence aligned with the measurement? Does that simply follow from the measured length?

- It would be great if the authors could explain better where the sawtooth pattern comes from. They explain how one of the force jumps occurs in the SI, but what determines the overall pattern? Are there preferred positions in the sequence where the force jumps?

- SI Fig. 4: lowest energy branch (n2, pink) → yellow?

- The choice of references for CRISPR and origami is unconventional. The first paper showing gene editing is not Ref. 1, but rather 10.1126/science.1225829, 10.1126/science.1231143.

Also, for origami, one should rather cite Rothemund (10.1038/nature04586) and Shih (10.1038/nature08016) and not, e.g., Ref. 3 in the manuscript.

(Remarks on code availability)

Version 1:

Reviewer comments:

Reviewer #1

(Remarks to the Author)
The authors have fully addressed my comments. I recommend publication of the manuscript.

(Remarks on code availability)

Reviewer #2

(Remarks to the Author)
The authors have done a good job addressing my comments, except one point (raised by all three reviewers) that still needs clarification:

Line 208-211: "Considering all possible combinations of starting and ending peaks, a set of K segments has been obtained for each T. Note that peaks are not necessarily consecutive, as different segments can overlap, making the analysis robust." The number of all possible combinations of starting and ending peaks must exceeds 400-600? Are the K segments randomly chosen from all possible combinations, or are there additional criteria?

(Remarks on code availability)

Reviewer #3

(Remarks to the Author)
The authors have responded well to all the reviewer comments and the paper may now be accepted for publication.

(Remarks on code availability)

NCOMMS-24-32591 - Reviewers' Response

We thank the Reviewers for their careful reading of our manuscript and are gratified that they all find the work valuable and interesting. The new revised version of the paper has been thoroughly revised and appropriately expanded to address the reviewers' comments.
In what follows, we present a point-by-point answer to all questions raised by the reviewers (shown in bold). Modifications to the main text are shown in red in this response document. Correspondingly, changes or adds are also shown in red in the resubmitted manuscript (with the custom-defined latex command \tcr{}).

Reviewer's Comments:

Reviewer #1 (Remarks to the Author)

In this paper, Rissone et al. use a temperature jump optical trap to mechanically unzip single DNA molecules and employ theoretical analysis to derive the heat capacity change parameters of hybridization for the different DNA base pair motifs in the nearest-neighbor model. The heat capacity change ΔC_p informs quantitatively about the temperature dependence of enthalpic and entropic contributions to the free energy hybridization, but experimental measurements through bulk and single-molecule experiments to date have yielded vastly different results, depending on the specific experimental conditions. This work seeks to resolve the challenges in determining ΔC_p values through systematic mechanical pulling experiments over a range of temperatures and subsequent detailed theoretical analysis. The authors have laid the foundation for this work through a series of prior publications in mechanical pulling experiments using optical tweezers. The methodology put forward has been demonstrated on a DNA hairpin here and can be feasibly extended to other types of nucleic acids. I would recommend publication after the authors have addressed the following points:

We thank the Reviewer for the positive appraisal of the manuscript and the constructive feedback that will contribute to improving our work. Below we address the points raised by the Reviewer.

1. The experiments were conducted at temperatures between 7°C and 42°C in increments of 2.5°C. How did the authors measure and ensure a constant temperature within the local environment in the microfluidic chamber? Can the authors also provide more detail on the design of the chamber and how it damps convection effects caused by the non-uniform temperature from the laser?

The Reviewer asks two important questions. These were already addressed in a previous work by some of us, where we described the temperature jump optical trap instrument. First, the temperature inside the microfluidic chamber is retrieved from the viscosity, η, of the buffer solution near the micropipette, i.e., the region where experiments are carried out. To measure η, we use Stokes' law for the force exerted on a sphere of radius R dragged at speed v in a viscous fluid, $f = -6\pi\eta Rv$. This is done by trapping an anti-digoxigenin (AD) bead with the optical tweezers and moving the fluidics chamber between two fixed positions along one direction (Ex., the x-axis). The T-dependence of η follows the well-known Vogel–Fulcher–Tammann-Hesse equation $\eta = \eta_0 e^{B/(T-T_0)}$, where η_0, B, and T_0 are empirical fitting parameters as explained in Ref.[2]. Inverting this equation gives the temperature T inside the chamber from the experimentally measured viscosity. This procedure is done upon incrementing the laser power output from zero to maximum and is repeated many times during the experiments to ensure the correct measurement of the temperature.

Second, the Reviewer asks about the convection induced by the heat flux inside the chamber. In fact, using a heating laser will, in principle, generate a non-uniform temperature profile and a convective current inside the chamber. This effect would yield a hydrodynamic flow between regions at different temperatures that increase with the heat supplied by the heating laser. This is inconvenient in single-molecule experiments because the flows might exert significant forces on the pulled molecule.

To minimize this effect and damp convection, our MiniTweezers setup uses Kohler's illumination to spread the heating beam out at the focus and change its shape (see Ref.[2] for details). Instead of focusing the heater to a point at the experimental plane, the heating laser is focused to a point located at the back-focal plane of the objective lens so that the heating beam emerges into the experimental area collimated as a cylinder about 66μm in diameter. Moreover, we use a microfluidic chamber thin enough (110μm) to avoid the creation of regions at different temperatures close to the micropipette. Experiments are performed by using this chamber design and the Cargille Labs Refractive Index matching Liquid, Series AAA, N=1.330. This liquid is placed between the objective lenses and both sides of the fluidics chamber as a replacement for water and ensures minimal heat absorption.

A full description of the temperature-jump optical trap instrument can be found in the following paper [2] (Ref. 20 of the manuscript).

We added the following paragraph to Methods, Sec 1:

The temperature inside the microfluidic chamber is measured from the viscosity, η, of the buffer solution in the proximity of the micropipette, i.e. the region where experiments are carried out. To measure η, we use Stokes' law for the force exerted on a sphere of radius R moving at speed v in a viscous fluid, $f = -6\pi\eta Rv$. This is done by trapping an anti-digoxigenin (AD) bead with the OT and moving it between two fixed positions along one direction (Ex., the x-axis). The T-dependence of η follows the well-known Vogel–Fulcher–Tammann-Hesse equation $\eta = \eta_0 e^{B/(T-T_0)}$, where η_0, B, and T_0 are empirical fitting parameters, see Ref.[2] for details. Inverting this equation gives the temperature T inside the chamber from the experimentally measured viscosity. Finally, convection effects were minimized using Kohler's illumination combined with a suitable design of the microfluidics chamber to reduce heat absorption. A full description of the temperature-jump optical trap instrument can be found in Ref.[2].

2. It is difficult to see the curves in Figure 1B, particularly with the error bars. I would suggest the authors remove the error bars from the figure and have an additional figure in the SI with the error bars, possibly plotted with some offset in the y-axis for better visualization of the error bars. The current error bars are plotted to show molecule-to-molecule variability, what about the variability for the same molecule across the different unzipping/rezipping cycles? The authors should also state what the error bars represent, e.g. confidence interval.

Unzipping experiments of DNA molecules, even for long hairpins, are characterized by reversibility among different cycles and among unfolding and refolding trajectories within the same cycle. Figure **??** shows the typical unzipping and rezipping of a pulling cycle (for one molecule) for the 3.6kbp DNA hairpin at each temperature. Remarkably, unzipping (red) and rezipping (blue) curves show small hysteresis and mostly in correspondence of the first rip upon refolding, which are due to the time needed to the unfolded ssDNA strands to align before hybridization. Raw data (light colored lines) are taken at 1kH and need to be processed to remove the statistical Brownian noise before analysis. To do this, we applied a Savitzky–Golay smoothing filter, i.e. a moving average with polinomial interpolation, over time windows of 150ms (darker lines). We added Fig. R1 to the Supp. Info as Extended Data Fig. 1

According to the Reviewer's suggestions, we removed the error bars from Fig.1B (see Fig. R2B). Error bars are now displayed in the new Extended Data Fig. 2 (here shown as Fig. R3). We used standard deviation for the error bars ($\pm\sigma$) rather than statistical errors or the confidence interval. The standard deviation quantifies thermal fluctuations along the FDC. Instead, statistical errors are roughly seven times smaller due to the large number of pulls $\sim$50, as indicated in the caption of Fig. R2.

3. In computing the FEC from the experimental FDC, was k_{eff} estimated for each experimental run? If so, it would be interesting to report the variation in k_{eff} values across runs to validate the assumption of force-independent trap stiffness.

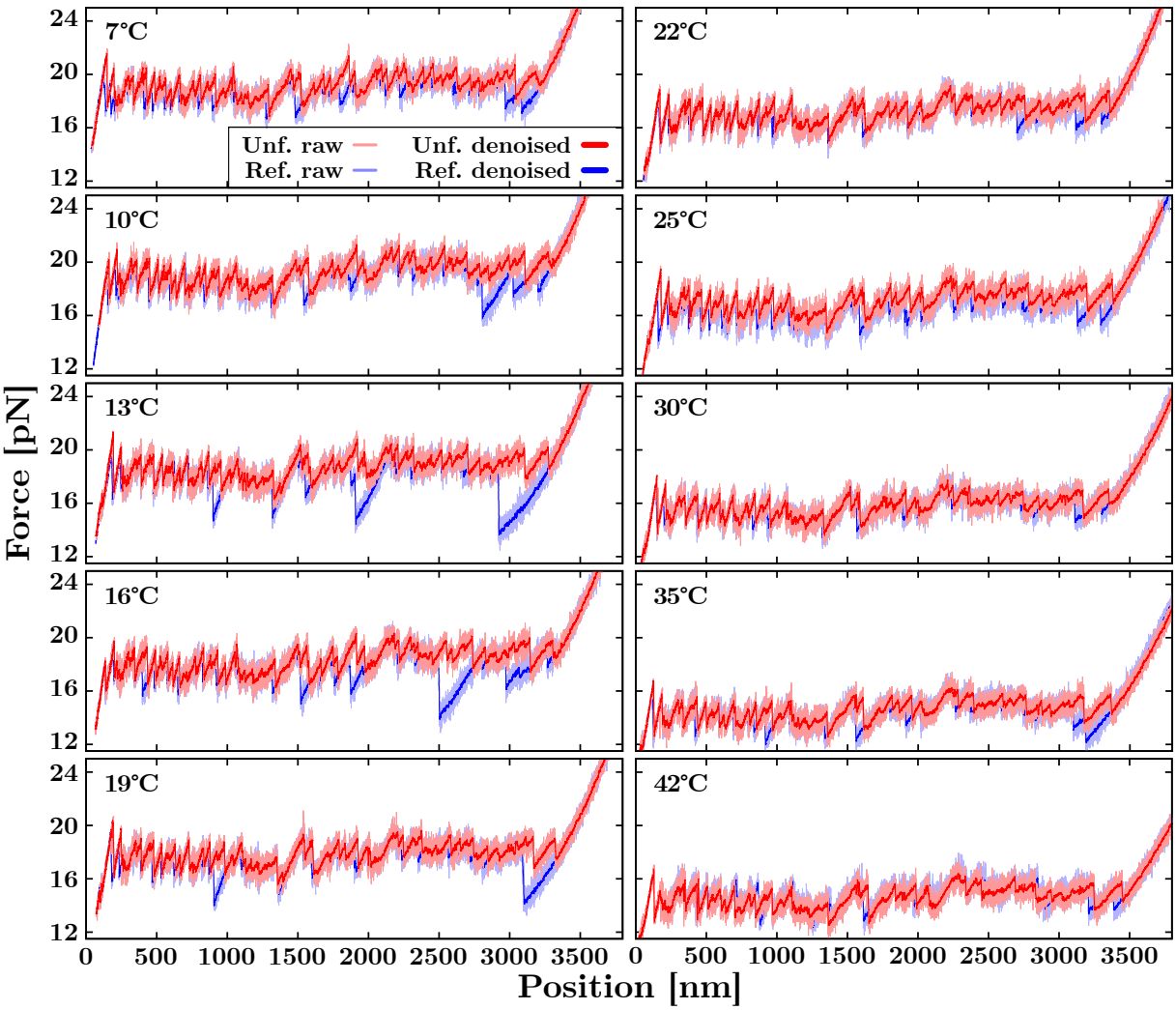

Figure R1: **Raw unzipping data.** Each panel shows one unfolding/refolding cycle for one molecule at a given T. Unzipping (red) and rezipping (blue) curves show a small hysteresis except for sporadic refolding events and particularly in correspondence of the first force jump observed during the refolding upon initiation of stem formation from the ssDNA. The light-colored lines are the raw data at 1kHz, while darker lines are raw data after denoising. This is done by applying a Savitzky–Golay filter over time windows of 150ms.

The effective stiffness, $k_{\rm eff}$, accounts for the contributions of the optically trapped bead and the dsDNA short handles, which are treated as serially connected springs so that $k_{\rm eff}^{-1} = k_{\rm b}^{-1} + k_{\rm h}^{-1}$. Both terms can be considered force-independent in the explored range of forces ($\sim 0-25$pN). In fact, the trapped bead is subject to a harmonic potential $\propto k_b x^2$ so that it has a linear response with force. In general, deviations from the linearity of the optical trap are minor, with k_b increasing by $\sim 5\%$ from 0 to 20pN [3], within statistical errors. The contribution from the dsDNA short handles (29bp each), $k_{\rm h}$, can be neglected as they become fully stretched at low forces (2-3pN), behaving as rigid rods with $k_h \gg k_{\rm b}$. Thus, for all practical purposes $k_{\rm eff}^{-1} \approx k_{\rm b}^{-1}$ or $k_{\rm eff} \approx k_{\rm b}$. For each temperature, we have measured $k_{\rm eff}$ by fitting the slope preceding the first FDC rip to a straight line. To better capture the variability of this parameter, we extracted it independently for each unzipping/rezipping trajectory and finally averaged the results over all molecules at a given T. In Fig. R4A, B, we show the stiffness variability among trajectories and molecules at 25°C, respectively. Results are compatible with the standard errors, proving that these measurements are sample-independent. Finally, in Fig. R4C, we show $k_{\rm eff}$ averaged over all molecules at each studied temperature. It is apparent that the stiffness does not exhibit any temperature dependence.

By following the above discussion, we have rewritten Sec. 2 of the Supp. Info as follows:

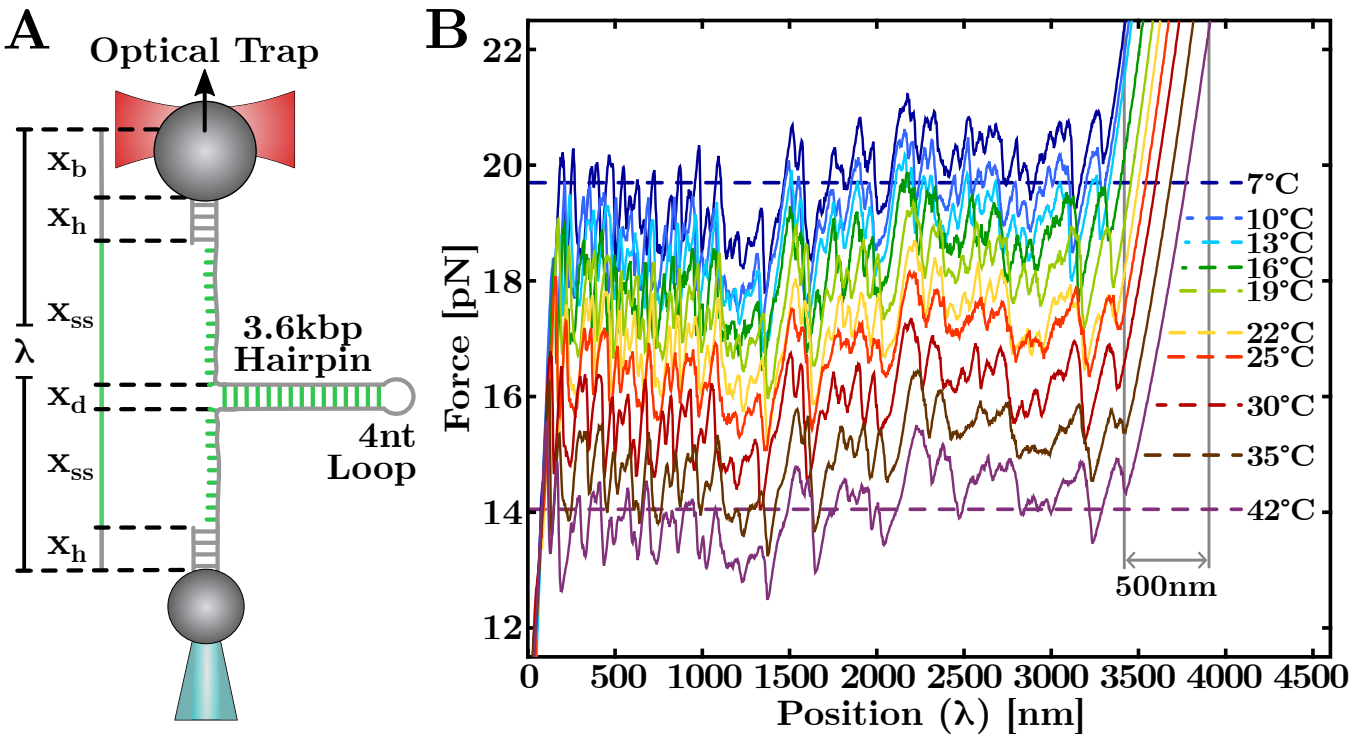

Figure R2: **Single-molecule calorimetric force-spectroscopy.** (**A**) Experimental setup (Sec. 1, Methods) showing each component of the (measured) total trap-pipette distance, λ. The hairpin is made of a 3.6kbp stem ending in a 4 nucleotides (nt) loop. (**B**) Temperature dependence of FDCs obtained by pulling the 3.6kbp DNA hairpin. At each T, the FDC results from averaging over $5-6$ molecules for $40-50$ unzipping/rezipping cycles (see also Extended Data, Fig, 1).

The elastic properties of ssDNA are strongly temperature dependent (see Fig. 1B, main text). Accurately measuring the elastic parameters requires modeling all contributions to the trap-pipette distance, λ, which includes the optical trap displacement ($x_{\rm b}$), the dsDNA handles ($x_{\rm h}$), the ssDNA ($x_{\rm ss}$), and the molecular diameter (d_0). The setup contributions ($x_{\rm b}$ and $x_{\rm h}$) have been evaluated according to the *effective stiffness* method [48], by which are treated as serially connected springs so that $k_{\rm eff}^{-1} = k_{\rm b}^{-1} + k_{\rm h}^{-1}$. Both terms can be considered force-independent in the explored range of forces ($\sim 0-25$pN). In fact, the trapped bead is subject to a harmonic potential $\propto k_b x^2$ so that it has a linear response with force. In general, deviations from the linearity of the optical trap are minor, with k_b increasing by $\sim 5\%$ from 0 to 20pN [3], within statistical errors. Moreover, as the force varies in a relatively narrow range ($f_{\rm max} - f_{\rm min} \lesssim 10$pN), the trap stiffness can be considered nearly force-independent so $k_{\rm eff}$ is constant along the folded branch of the FDC. The contribution from the dsDNA short handles (29bp each), $k_{\rm h}$, can be neglected as they become fully stretched at low forces (2-3pN), behaving as rigid rods with $k_h \gg k_{\rm b}$. Thus, for all practical purposes $k_{\rm eff}^{-1} \approx k_{\rm b}^{-1}$ or $k_{\rm eff} \approx k_{\rm b}$. We computed $k_{\rm eff}$ by fitting the slope preceding the first force rip in the FDC to the linear equation $f = k_{\rm eff} x$ (orange dashed-line in Extended Data Fig. 3). For each temperature, we have measured $k_{\rm eff}$ by fitting the slope preceding the first FDC rip to a straight line. To better capture the variability of this parameter, we extracted it independently for each unzipping/rezipping trajectory and finally averaged the results over all molecules at a given T. In Extended Data Fig. 4A, B, we show the stiffness variability among trajectories and molecules at 25°C, respectively. Results are compatible with the standard errors, proving that these measurements are sample-independent. Finally, in Extended Data Fig. 4C, we show $k_{\rm eff}$ averaged over all molecules at each studied temperature. It is apparent that the stiffness does not exhibit any temperature dependence.

Fig. R4 has also been added to the Supp. Info (Extended Data Fig. 4).

4. The errors reported in Figures 2 and 4, as well as the tables, should specify the confidence

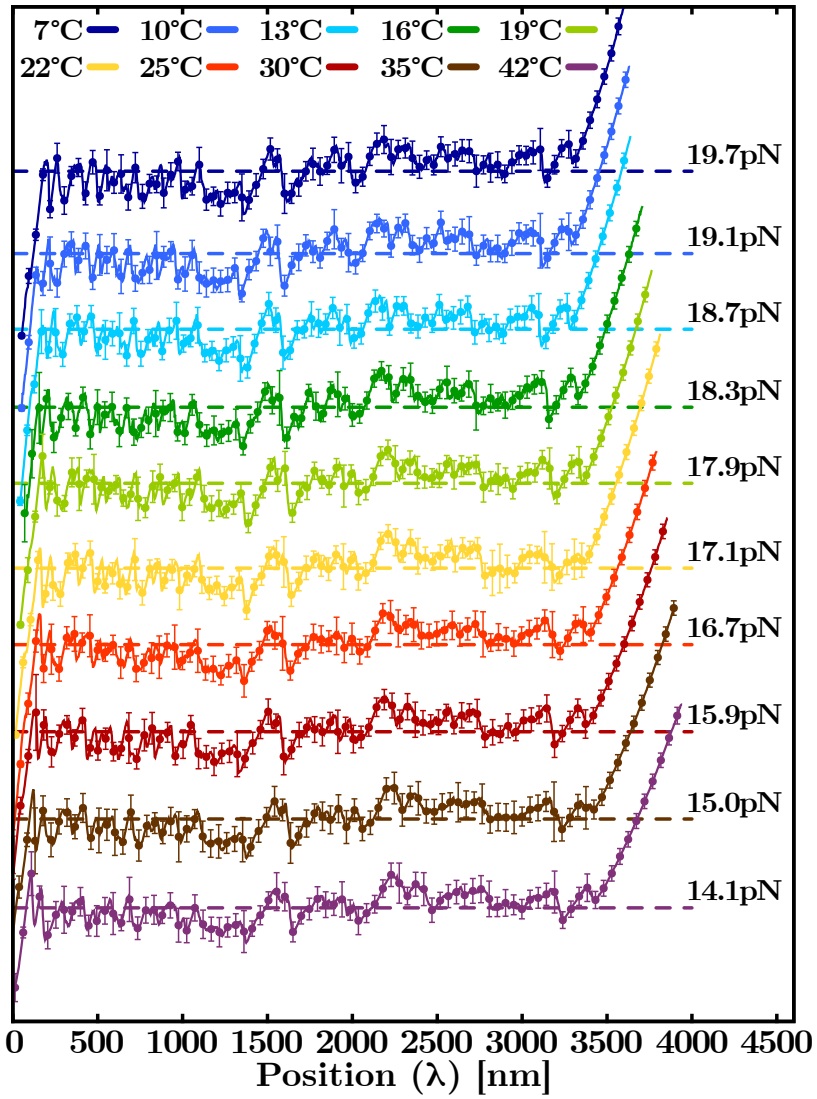

Figure R3: **Experimental variability of the T-dependent FDCs.** Temperature dependence of FDCs obtained by pulling the 3.6kbp DNA hairpin. At each T, the FDC results from averaging over $5-6$ molecules for $40-50$ unzipping/rezipping cycles. The error bars are the standard deviation in force (plotted for a fraction of the total data points) and show the molecule-to-molecule variability.

interval. It would also be helpful to include a brief overview of how the errors were calculated in the SI. I would imagine this was a nontrivial calculation involving error propagation and errors in fits to data.

This is an important point that we have revisited in the resubmission of the paper. Statistical error estimation is not straightforward in our analysis, where the measured thermodynamic parameters depend on several variables, each with its uncertainty. The stochastic nature of the gradient descent algorithm allowed us to derive the NNBP entropies and compute the standard deviation of the parameters with the bootstrap method. Bootstrapping is a statistical method that involves repeatedly sampling from a dataset to estimate the distribution of a set of parameters. It permits us to estimate the variability and confidence intervals of the parameters without making strong assumptions about the underlying distributions. For example, the standard deviations reported for the entropy values in the figures and tables have been derived by performing 10^3 bootstrap resampling for the data at each temperature. An analogous argument can be made for the Monte Carlo minimization algorithm to determine the NNBP free energies.
In the first version of the paper, statistical errors were derived from bootstrapping only, thereby underestimating the errors. However, the main uncertainty of energy measurements in force spectroscopy comes

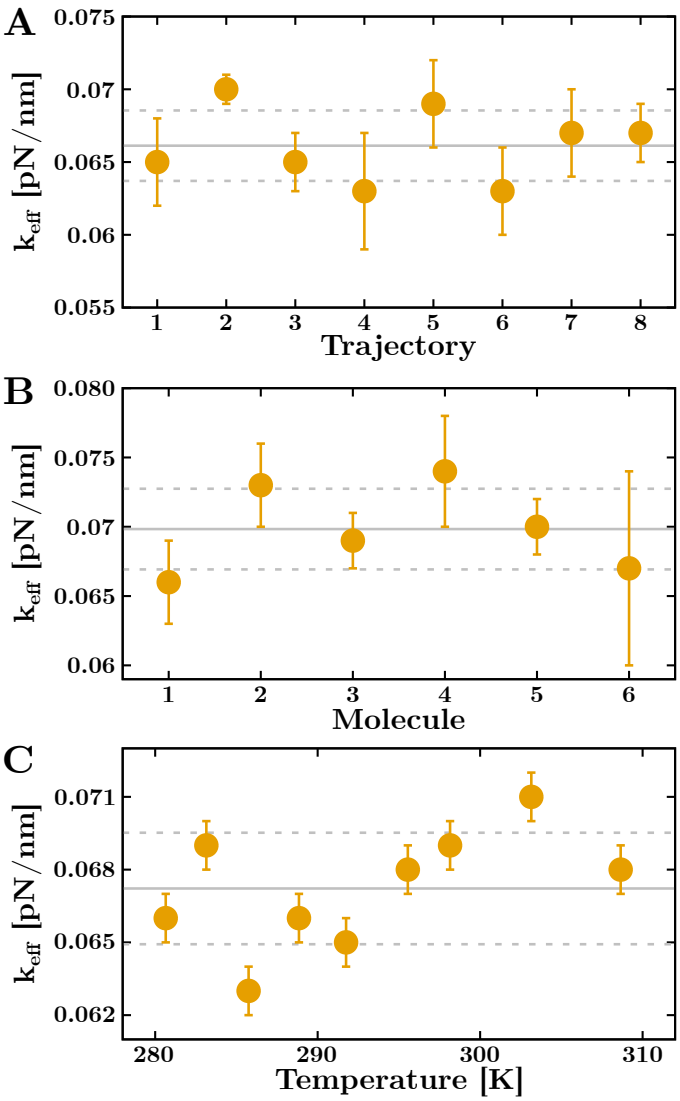

Figure R4: **Experimental measurement of the effective stiffness.** The effective stiffness, k_{eff}, has been obtained by fitting the first slope of the FDC between 3pN and the first rip to the linear equation $f = k_{\text{eff}}x$ (see Sec. 2, Supp. Info). Panels (**A**) and (**B**) show the k_{eff} variability among different trajectories (for one selected molecule) and molecules (after having averaged over cycles for each molecule) at 25°C. Panel (**C**) reports k_{eff} at each experimental temperature. The continuous line indicates the average value and the dashed lines show the corresponding error.

from the incomplete knowledge of the ssDNA elastic response. It has been shown [6] that the total error in determining the energy parameters is dominated by the uncertainty of the ssDNA elastic parameters: a $5-10\%$ error in the estimation of the persistence length of ssDNA translates into a similar error in the NNBP energies. This yields an accuracy of ≈ 0.1kcal/mol in estimating the NNBP free energies. To determine the statistical errors, we have propagated the values of the elastic parameters at the limits of their confidence interval to determine the corresponding free energies, enthalpies, and entropies at all temperatures. Minimum and maximum values for $\Delta g_i(T)$, $\Delta s_i(T)$, and $\Delta h_i(T)$ set the confidence intervals on each quantity with the corresponding mean.

Once $\Delta s_i(T)$ and $\Delta h_i(T)$ were determined with their statistical errors, the standard errors of the thermodynamic parameters in Table 1 and Fig. 4 ($T_{m,i}$, $\Delta s_{m,i}(T)$ and $\Delta h_{m,i}(T)$ and $\Delta c_{p,i}$) have been derived from the fit to Eqs.5 as $s(\mathbf{x}) = \sqrt{(\text{diag}[C(\mathbf{x})]}$, where $C(\mathbf{x})$ is the covariance matrix of the parameters vector including all errors of the fitted datasets.

All error bars in every quantity in the plots and tables show a confidence interval equal to 2σ centered at the mean, with σ the statistical error. For instance, $\Delta s_{GA/CT} = -23.6(7)\text{cal}\,\text{mol}^{-1}\text{K}^{-1}$ in Table 4

(main text) has $\sigma = 0.7\text{cal}\,\text{mol}^{-1}\text{K}^{-1}$ meaning that the confidence interval goes from $-22.9\text{cal}\,\text{mol}^{-1}\text{K}^{-1}$ to $-24.3\text{cal}\,\text{mol}^{-1}\text{K}^{-1}$.

The above discussion has been added as Sec. 5 of the Supp. Info:

Statistical error estimation is crucial to our analysis, where the measured thermodynamic parameters depend on several variables, each with its uncertainty. The stochastic nature of the gradient descent algorithm allowed us to derive the NNBP entropies and compute the standard deviation of the parameters with the bootstrap method. Bootstrapping is a statistical method that involves repeatedly sampling from a dataset to estimate the distribution of a set of parameters. It permits us to estimate the variability and confidence intervals of the parameters without making strong assumptions about the underlying distributions. For example, the standard deviations reported for the entropy values in the figures and tables have been derived by performing 10^3 bootstrap resampling for the data at each temperature. An analogous argument can be made for the Monte Carlo minimization algorithm to determine the NNBP free energies. However, the main uncertainty in the NNBP parameters derivation using force spectroscopy comes from the measurement of the ssDNA elastic response. It has been shown [29] that the total error in determining the energy parameters is dominated by the uncertainty of the ssDNA elastic parameters: a $5-10\%$ error in the estimation of the persistence length of ssDNA translates into a similar error in the NNBP energies. This yields an accuracy of ≈ 0.1kcal/mol in estimating the NNBP free energies.
To determine the statistical errors, we have propagated the values of the elastic parameters at the limits of their confidence interval to determine the corresponding free energies, enthalpies, and entropies at all temperatures. Minimum and maximum values for $\Delta g_i(T)$, $\Delta s_i(T)$, and $\Delta h_i(T)$ set the confidence intervals on each quantity with the corresponding mean (see Extended Data Tables 4, 5, and 6). Once $\Delta s_i(T)$ and $\Delta h_i(T)$ were determined with their statistical errors, the standard errors of the thermodynamic parameters in Table 1 and Fig. 4 of the main text ($T_{m,i}$, $\Delta s_{m,i}(T)$ and $\Delta h_{m,i}(T)$ and $\Delta c_{p,i}$) have been derived from the fit to Eqs.5 as $s(\mathbf{x}) = \sqrt{(\text{diag}[C(\mathbf{x})]}$, where $C(\mathbf{x})$ is the covariance matrix of the parameters vector including all errors of the fitted datasets.
All error bars in every quantity in the plots and tables show a confidence interval equal to 2σ centered at the mean, with σ the statistical error. For instance, $\Delta s_{\text{GA/CT}} = -23.6(7)\text{cal}\,\text{mol}^{-1}\text{K}^{-1}$ in Table 4 (main text) has $\sigma = 0.7\text{cal}\,\text{mol}^{-1}\text{K}^{-1}$ meaning that the confidence interval goes from $-22.9\text{cal}\,\text{mol}^{-1}\text{K}^{-1}$ to $-24.3\text{cal}\,\text{mol}^{-1}\text{K}^{-1}$.

Moreover, we added the following sentence in Sec. "Derivation of the NNBP entropies" of the main text:

Notice that to determine the statistical errors of $\Delta s_i(T)$, we have propagated the values of the elastic parameters at the limits of their confidence interval. Minimum and maximum values for $\Delta s_i(T)$ set the confidence intervals with the corresponding mean. The same method has been applied to all NNBP thermodynamic parameters.

All figures and tables have been updated to account for the revised error estimation.

5. Page 4, Line 148. I believe the authors are referring to Fig. 1B, not Fig. 2B.

We thank the Reviewer for noticing this typo. It was indeed a reference to Fig. 1B, not 2B. This has now been corrected.

6. Suggestion: add the loop size L to the schematic in Fig. 1A.

The illustration of the DNA molecule in Fig. R2A now explicitly indicates that the hairpin ends with a loop of 4 nucleotides.

7. It is not very clear to me how the authors fit the WLC to the FEC to obtain the persistence length and inter-phosphate distance. There are two fitted parameters obtained from a

fit of one equation to the data, this would add to the uncertainty of the fitted parameters and introduce some coupling between the uncertainties. As an aside, what is z_{ss} in the legend of Fig. 1, extended data?

The Reviewer is asking about the derivation of the ssDNA elasticity parameters. Indeed, we have fitted the elastic response part of the FEC to the inextensible WLC with two fitting parameters: the persistence length l_p and the inter-phosphate distance d_b. A larger extension can be obtained at a given force and temperature by increasing l_p or d_b. Therefore, both parameters are anticorrelated in the fits: increasing l_p will reduce db to compensate for the fit, and vice-versa. Much progress has been made in studying the temperature-dependent elastic properties of nucleic acids with single-molecule pulling experiments. In particular, the persistence length for semi-flexible polymers can be written as $l_p = l_p^0 + l_p^{el}$, where l_p^0 is the intrinsic persistence length and l_p^{el} is the electrostatic correction. For ssDNA, it has been shown [12] that the overall increase of l_p with T is due to the dominant term l_p^{el}, which increases due to a screening collective effect by water and ions in the surrounding media. In that reference, the inter-phosphate distance d_b exhibits a weak temperature dependence for ssDNA in the explored temperature range. Our fits confirm the expected temperature dependencies of l_p and d_b: while l_p changes at the rate 10^{-3} nm/K, d_b remains almost constant in the studied temperature range. A detailed procedure description can be found in the above-quoted reference [12].

To clarify this, we edited the last paragraph of section "T-dependence of the ssDNA elasticity" of the manuscript:,

Similar behavior has been observed for shorter ssDNAs of $20-40$ nucleotides [12] and polypeptide chains [13]. The observed increase in $x_{\rm ss}$ with T is predicted in Debye-Huckel theory due to the entropy of the cloud of counterions. Upon increasing temperature, the screening of the phosphates repulsion is reduced, and l_p increases. Our results confirm this trend with l_p increasing with temperature tenfold relative to d_b, which remains almost constant in the studied temperature range.

Finally, we thank the Reviewer for pointing out a typo in the legend of the Extended Data Fig. 1 (now Extended Data Fig. 3). z_{ss} was the name chosen for the ssDNA single base-pair extension in a previous manuscript version. The error has been corrected.

8. For the decomposition of the FDC into different segments, how were the peaks corresponding to force rips identified? It is not very clear from Fig. 2C.

The raw unzipping data in a pulling cycle have been averaged over 150ms time windows using the Savitzky–Golay filter (see Fig. R1). The resulting FDCs were first averaged over pulling cycles for a given molecule and next averaged over different molecules, and statistical errors were determined (see Figure R3). Such an averaging procedure was repeated to generate one thermodynamic FDC at each temperature. Each FDC shows the characteristic sawtooth pattern with its sequence of force slopes and rips. Therefore, each peak along the FDC corresponds to a force rip concurrent with the unfolding of a group of base pairs. In practice, identifying the FDC peaks is done with an algorithm that finds the maxima of a given function $f(x)$, corresponding to the FDC. To do so, the function derivatives of the FDC are computed numerically from the data, and the maxima are determined by imposing $df(x)/dx = 0$. We notice that the magnitude of the force rip is proportional to the number of released bases in that unfolding event. The larger the rip, the more reproducible the number of bases is among repeated unzipping-rezipping cycles. In contrast, small force rips are related to releasing a small number of bases (10-40), becoming indistinguishable within the statistical errors along the FDC (see Figure R3). Therefore, we set $\Delta f \geq 2pN$ as the force-rip threshold for a force rip to be considered in the analysis. This procedure can be achieved with the built-in analysis tools of the most common programming languages (Python, in our case) and is automatically done in the code we shared on the public repository, so we avoided a detailed description in the manuscript.
Once the peaks (maxima) have been selected, we defined a segment, Δx_k, as the FDC region delimited by two peaks. Note that peaks are not necessarily consecutive, as different segments can overlap, permitting

us to analyze many segments and making the analysis robust. A set of segments has been obtained at each temperature by considering all possible combinations of starting and ending peaks along the FDC (see Fig. 2C, main text). This gives $K = 400 - 600$ segments per temperature. For each segment, we fit the slopes preceding the starting and ending peaks to the WLC elastic model to extract the number of released bases at the force rip after the peak. Determining the number of released bases in all force rips along the FDC permits us to identify the segment's location along the DNA sequence.

We also added the following sentence to the section "Derivation of the NNBP entropies" in the main:

To derive the entropies of the different NNBPs, we have decomposed the full unzipping curve into segments of variable length encompassing different regions along the FEC. Each segment is delimited by two peaks corresponding to force-rips along the FEC. The peaks have been selected by looking for local maxima along the experimental signal. To do so, the FEC function derivatives have been numerically computed, and the maxima were determined by imposing $df(x)/dx = 0$. We notice that the magnitude of the force rip is proportional to the number of released bases in that unfolding event. To ensure the maximal signal-to-noise ratio, we discarded events within the FEC statistical errors by only accounting for force-rips with $\Delta f \leq 2$pN. Figure 2C shows examples of segments starting and ending at a peak (colored circles). By considering all possible combinations of starting and ending peaks, a set of K segments has been obtained for each T. Note that peaks are not necessarily consecutive, as different segments can overlap, making the analysis robust.

9. For the derivation of the free energies $\Delta g_i(T)$, it is not clear just from the explanation in the main text (Lines 238-243) how this is carried out. The authors can consider moving some of the details from the SI into the main text for this. In particular, more can be said about the theoretical FDCs and their role in obtaining the free energies.

The derivation of the theoretical FDC is thoroughly described in detail in Refs. [6, 8]. It is based on calculating the equilibrium FDC at different values of the trap position given a set of energies and elastic parameters for the dsDNA handles and the ssDNA released during the unzipping. At each step, the total trap-pipette distance $\lambda(f) = x_b(f) + 2x_h(f) + x_{ss}(n, f) + x_d(f)$ (Eq.1, main text) is increased by a fixed $\Delta\lambda = 3$nm. The contribution of each element of the molecular construct (optical trap, dsDNA handles, ssDNA, and hairpin diameter) is computed. For a given λ and the number of unzipped base pairs n, the corresponding applied force, f, is retrieved by inverting Eq.(1) (main text). This calculation is repeated for different values of n from $n = 0$bp to $n = 3593$bp. The equilibrium value $n = n^*$ and the applied force f are determined by finding the absolute minimum over n of $\Delta G(n, f)$ for that λ. The integer function of n^* versus λ is monotonically increasing with jumps of variable size Δn^* at specific values of λ. While Δn^* determines the force rip at the jump, the specific λ gives the force values along the FDC before and after the jump. We notice that for a given λ, the total energy of the molecular system has two competing contributions: the positive stretching free energy, $\Delta G_{el}(n, f)$, acting as an externally applied work to unfold the hairpin, and the negative hybridization free energy, $\Delta G_0(n)$, which keeps the hairpin folded. At each force rip, it is fulfilled that $\Delta G_{el}(\Delta n^*, f) = \Delta G_0(\Delta n^*)$ rendering the equilibrium FDC a sawtooth pattern made of a sequence of gentle slopes separated by force rips.
The derivation of the equilibrium FDC for each set of energies is the step in an optimization Monte Carlo algorithm used to minimize the error between the theoretical FDC and the experimental one. A steepest gradient procedure combined with a simulated annealing protocol permits the determination of the best energy set, $\Delta g_i(T)$. The complete description of the theoretical FDC prediction and the Monte Carlo fit of the NNBP free energies can be found in [6, 8] (Refs., 25, 26 of the main text).

By following the Reviewer's suggestion, we expanded the description of the FDC theoretical prediction in the main text as follows:

The $\Delta g_i(T)$ values are derived via a Monte Carlo optimization based on the prediction of the theoretical FDC according to the NN model [7, 6, 8]. The equilibrium FDC is obtained by calculating the equilib-

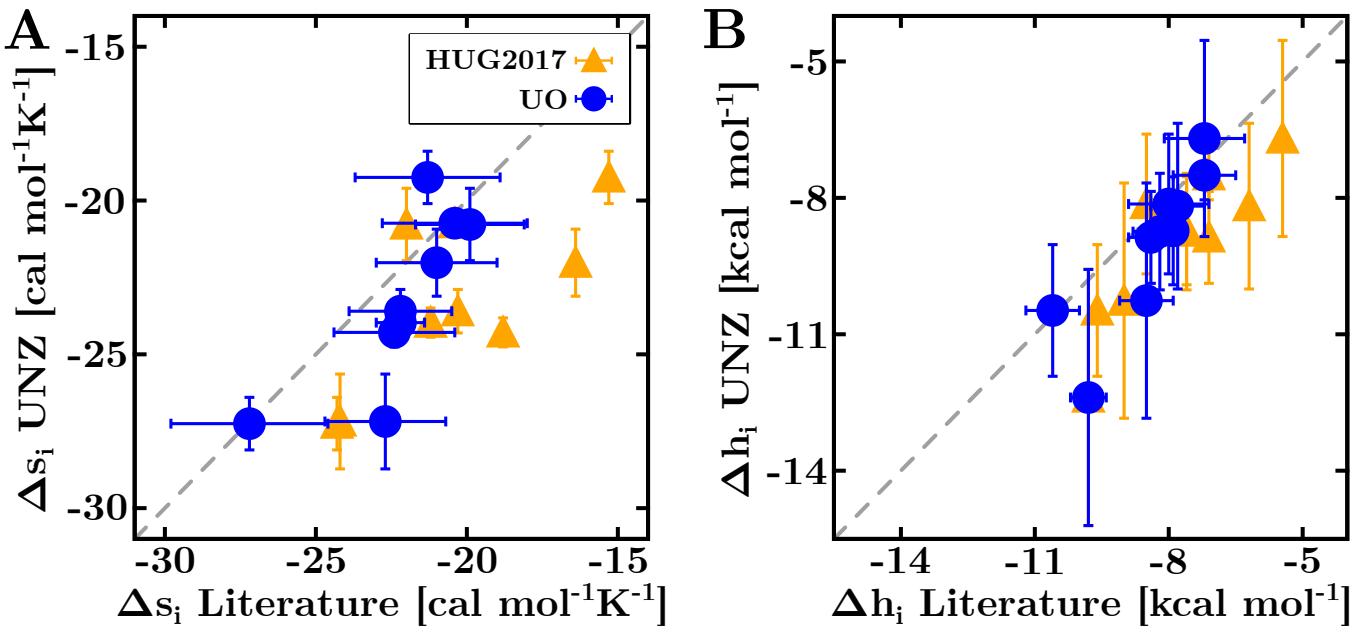

Figure R5: **Comparison of the NNBP parameters with the literature.** Correlation of NNBP unzipping (UNZ) parameters (y-axis) with the results (x-axis) by Huguet *et al.* [8] (orange) and Santalucia *et al.* [15] (blue) for the NNBP entropies, Δs_i (**A**), and enthalpies, Δh_i (**B**). Each point corresponds to a NNBP motif, a perfect correlation implying $x = y$ (dashed grey line). Notice that, since both the HUG2017 and UO datasets were obtained at different experimental conditions and $\Delta c_{p,i} = 0$, we compared these values to $\Delta s_{m,i}$ and $\Delta h_{m,i}$ extrapolated at T_m (main text, Fig. 4B at the bottom).

rium force at different values of the trap position given a set of energies and elastic parameters. At each step, λ (Eq.(1)) is increased by a fixed $\Delta\lambda = 3$nm. The contribution of each element of the molecular construct (optical trap, dsDNA handles, ssDNA, and hairpin diameter) is computed. For a given λ and number of unzipped base pairs n, the corresponding applied force, f, is retrieved by inverting Eq.(1). This calculation is repeated for all values of n. For a given λ, the total energy of the molecular system has two competing contributions: the positive stretching free energy, $\Delta G_{el}(n, f)$, acting as an externally applied work to unfold the hairpin, and the negative hybridization free energy, $\Delta G_0(n)$, which keeps the hairpin folded. The equilibrium values of n^* and f^* are determined by finding the absolute minimum over n of $\Delta G(n, f)$ for that λ. At each force rip, it is fulfilled that $\Delta G_{el}(\Delta n^*, f) = \Delta G_0(\Delta n^*)$ rendering the equilibrium FDC a sawtooth pattern made of a sequence of gentle slopes separated by force rips (Sec. 3, Supp. Info). The derivation of the equilibrium FDC for each set of energies is the step in an optimization Monte Carlo algorithm used to minimize the error between the theoretical and the experimental FDCs in Fig. R2B (Sec. 5, Methods). This procedure gives the eight NNBP independent free energies. The other two parameters (GC/CG and TA/AT) are obtained from the circular symmetry relations [4, 9, 8].

10. Can the authors comment on the physicality of the obtained thermodynamic parameter values? For example, are the differences in values of $\Delta s_i(T)$, $\Delta g_i(T)$, $\Delta h_i(T)$, $\Delta c_{p,i}(T)$ for the different base pair motifs consistent with what one might expect, from a physical basis and also in comparison to findings in the literature?

In Fig. R5, we show the correlation between the results reported in the literature shown in the x-axis versus our results (UNZ) shown in the y-axis and extrapolated to the NNBP melting temperatures, $T_{m,i}$ for each motif i. Two panels are shown, for the entropies (left panel) and the enthalpies (right panel). We compare our results with two sets of values from the literature: unzipping experiments from Huguet *et al.* [8] (HUG2017, yellow triangles) obtained at room temperature; and melting data unified by Santalucia *et al.* [15] (UO, blue circles) obtained by averaging over datasets of duplexes of varying sequences measured by different labs. Notice that despite the differences in the experimental conditions and techniques, an overall agreement is observed within the experimental errors.

Notice that a direct comparison of Δs_i and Δh_i values between our experiments and those in the existing literature is misleading. In fact, our Δs_i and Δh_i are temperature dependent, while those in the literature are temperature independent. Which temperature should be chosen to make a comparison? In Fig. R5, we chose T_m as the temperature for the comparison. However, one might also choose the physiological temperature $T = 37°\text{C}$ instead, as several web-server predicting tools do. A comparison between our results and those in the literature can also be made at the level of the NNBP free energies. Such a comparison was already shown in Fig. 3B of the main text. The strong compensation between the temperature-dependent enthalpies and entropies masks the finite Δc_p^i's. Our free energies shown in Fig. 3B of the main text (blue circles) agree with the UO dataset [15] (black line) and the energy parameters obtained by Huguet *et al.* [8] (grey line). Overall agreement is observed, except for motifs AC/TG and GA/CT where the UO free energies are lower.
Ultimately, to compare the $\Delta c_{p,i}$ values with the literature, we computed the average over all motifs that gives $\overline{\Delta c_p} = -30(10)$ cal mol^{-1}K^{-1}bp^{-1}. This should be compared with the highly dispersed set of values obtained from bulk experiments ranging between -20 and -160 cal mol^{-1}K^{-1}bp^{-1}, depending on the experimental technique, setup, and DNA sequence [14, 11]. In contrast, recent molecular dynamic simulations estimated an average $\Delta c_p \sim -30$ cal mol^{-1}K^{-1}bp^{-1} [10, 5], in agreement with our results.
Finally, the correlation between the NNBP energy values and the sequence motif (GC vs AT content, purine-purine stacking e.g. GG/CC versus GC/CG) is not straightforward. However, there have been attempts to find correlation patterns between the NNBP type and structural parameters informative of the relative displacements of base pairs at each base pair step (twist, slide, roll,..) [1]. The correlation of the NNBP motifs with the energies proves difficult due to the complexity of the stacking interactions within base pairs that are highly sensitive to the interactions between electric charges in the aromatic rings.

By following this discussion, we added the following sentence to the "Discussion" section of the main text:

Notice that a direct comparison of Δs_i and Δh_i values between our experiments and those in the existing literature can appear misleading. Our Δs_i and Δh_i are temperature dependent, while those in the literature are temperature independent. Which temperature should be chosen to make a comparison? In Extended Data Fig. 9, we chose T_m as the temperature for the comparison. However, one might also choose the physiological temperature $T = 37°\text{C}$ instead, as several web-server predicting tools do. As we can see from Fig. 3B, an overall agreement is observed between the free energies of the different motifs measured from single-molecule unzipping and bulk experiments despite the differences in the experimental conditions and techniques.
The correlation between the NNBP energy values and the sequence motif (GC vs AT content, purine-purine stacking, e.g., GG/CC versus GC/CG) is not straightforward. Despite attempts to find correlation patterns between the NNBP type and structural parameters informative of the relative displacements of base pairs at each base pair step (twist, slide, roll,...) [1], correlations at the NNBP energy level prove difficult due to the complexity of the stacking interactions within base pairs that are highly sensitive to the interactions between electric charges in the aromatic rings.

Figure R5 has been added to the Supp. Info as Extended Data Fig. 9.

11. How would the analysis need to be modified with a change in experimental conditions, for example a different salt condition?

In this work, we report the complete characterization of the DNA thermodynamics at the nearest-neighbors base-pairs (NNBPs) level, measuring the entropies, free energies, enthalpies, and heat capacities from unzipping experiments. To do so, we present a general method that can be straightforwardly extended to other experimental conditions or setups as long as the experiments are carried out using force-spectroscopy techniques. Therefore, changing the salt concentration or ionic strength (divalent instead of monovalent), would not influence the analysis. We chose to carry out the pulling experiments at the reference salt condition 1M NaCl to compare with the most common reference in the literature. Ultimately, we notice that the effects of salt concentration in single-molecule unzipping of long DNA hairpins are well-known

and have already been studied by and have already been studied by Huguet *et al.* in [6, 8]. These works evidence that increasing the salt concentration causes an increase in the FDC unzipping force, stabilizing hydrogen bonding and then increasing the hybridization free energy.

To clarify the generality of our method, we added the following sentence to the "Discussion" section of the main text:

We notice that the method presented in this work is general and can be straightforwardly extended to other experimental conditions (e.g., different salt concentration or ionic strength).

Reviewer #1 (Remarks on code availability)

I had a look through the code and it is well-organized with enough detail for a reader to follow through with the calculations presented in the paper. I did not try to install and run the code, but there appears to be detailed instructions in a README file for one to do so.

Reviewer #2 (Remarks to the Author)

In this manuscript, Rissone et al. used single-molecule calorimetric force spectroscopy and statistical analysis to extract the thermodynamic parameters for different DNA nearest-neighbor base pair motifs. They showed the temperature dependence of the enthalpy and entropy of hybridization, thereby deriving non-zero heat capacity changes (ΔC_p). These are important results with implications for DNA metabolism, nanotechnology, and simulations. The experiments and analyses are well executed, and the manuscript is concisely written. However, the following points need to be addressed before publication.

We are happy that the Reviewer finds our work valuable and of broad interest to the scientific community and for the positive recommendation and assessment.

(1) The temperature-jump optical trap is a critical feature of the study. But from the methods description it is not clear how to maintain a stable temperature below the room temperature.

The Reviewer is asking for information about the experimental setup used to maintain the optical trap (OT) at a constant low temperature while performing the experiments. To do this, we use an insulated box integrated with a refrigeration system. Before starting the experiments and with no OT inside, the ice box is cooled down to subzero temperatures. A water reservoir inside the box stabilizes the temperature at 5°C after the refrigeration is turned off. This is necessary as vibrations from the refrigeration system would affect the pulling experiments. The OT is then placed in the ice box with approximately one day needed for the instrument to thermalize with the surroundings. This setup allows fir carring out experiments at a constant temperature of 5°C for 3-4 days.
A thermocouple thermometer monitors the temperature inside the box, while two other thermocouples within the OT give the temperature close to the microfluidics chamber. Finally, The temperature inside the microfluidic chamber is measured from the viscosity, η, of the buffer solution in the proximity of the micropipette, i.e. the region where experiments are carried out.
To measure η, we use Stokes' law for the force exerted on a sphere of radius R dragged at speed v in a viscous fluid, $f = -6\pi\eta Rv$. This is done by trapping an anti-digoxigenin (AD) bead with the optical tweezers and moving the fluidics chamber between two fixed positions along one direction (Ex., the x-axis). The T-dependence of η follows the well-known Vogel–Fulcher–Tammann-Hesse equation $\eta = \eta_0 e^{B/(T-T_0)}$, where η_0, B, and T_0 are empirical fitting parameters as explained in Ref.[2]. Inverting this equation gives the temperature T inside the chamber from the experimentally measured viscosity. This procedure is done upon incrementing the laser power output from zero to maximum and is repeated many times during the experiments to ensure the correct measurement of the temperature.

Following the Reviewer's comment, we added the following sentences to the Methods section of the main text:

To carry out experiments at low temperatures, the OT is placed into an insulated box integrated with a refrigeration system (ice-box) and containing a water reservoir. The ice-box is cooled down to subzero temperatures in order for the water to freeze. After the refrigeration is then turned, the temperature stabilizes at 5°C.

and later,

The temperature inside the microfluidic chamber is measured from the viscosity, η, of the buffer solution in the proximity of the micropipette, i.e. the region where experiments are carried out. To measure η, we use Stokes' law for the force exerted on a sphere of radius R moving at speed v in a viscous fluid, $f = -6\pi\eta Rv$. This is done by trapping an AD bead with the OT and moving it between two fixed positions along one direction (Ex., the x-axis).The T-dependence of η follows the well-known Vogel–Fulcher–Tammann-Hesse

equation $\eta = \eta_0 e^{B/(T-T_0)}$, where η_0, B, and T_0 are empirical fitting parameters. Inverting this equation gives the temperature T inside the chamber from the experimentally measured viscosity.

Table R1: **Hairpin sequence** ($5' \rightarrow 3'$)

ATAGAGACACATATATAATAGATCTTAGGTCGCCGCCCCGTAACCTGTCGGATCACCGGAAAGGACCCGTAAA
GTGATAATGATTATCATCTACATATCACAACGTGCGTGGAGGCCATCAAACCACGTCAAATAATCAATTATGA
CGCAGGTATCGTATTAATTGATCTGCATCAACTTAACGTAAAAACAACTTCAGACAATACAAATCAGCGACAC
TGAATACGGGGCAACCTCATGTCAACGAAGAACAGAACCCGCAGAACAACAACCCGCAACATCCGCTTTCCTA
ACCAAATGATTGAACAAATTAACATCGCTCTTGAGCAAAAAGGGTCCGGGAATTTCTCAGCCTGGGTCATTGA
AGCCTGCCGTCGGAGACTAACGTCAGAAAAGAGAGCATATACATCAATTAAAAGTGATGAAGAATGAACATCC
CGCGTTCTTCCCTCCGAACAGGACGATATTGTAAATTCACTTAATTACGAGGGCATTGCAGTAATTGAGTTGC
AGTTTTACCACTTTCCTGACAGTGACAGACTGCGTGTTGGCTCTGTCACAGACTAAATAGTTTGAATGATTAG
CAGTTATGGTGATCAGTCAACCACCAGGGAATAATCCTTCATATTATTATCGTGCTTCACCAACGCTGCCTCA
ATTGCTCTGAATGCTTCCAGAGACACCTTATGTTCTATACATGCAATTACAACATCAGGGTAACTCATAGAAA
TGGTGCTATTAAGCATATTTTTTACACGAATCAGATCCACGGAGGGATCATCAGCAGATTGTTCTTTATTCAT
TTTGTCGCTCCATGCGCTTGCTCTTCATCTAGCGGTTAAAATATTACTTCAAATCTTTCTGTATGAAGATTTGA
GCACGTTGGCCTTACATACATCTGTCGGTTGTATTTCCCTCCAGAATGCCAGCAGGACCGCACTTTGTTACGC
AACCAATACTATTAAGTGAAAACATTCCTAATATTTGACATAAATCATCAACAAAACACAAGGAGGTCAGACC
AGATTGAAACGATAAAAACGATAATGCAAACTACGCGCCCTCGTATCACATGGAAGGTTTTACCAATGGCTCA
GGTTGCCATTTTTAAAGAAATATTCGATCAAGTGCGAAAAGATTTAGACTGTGAATTGTTTTATTCTGAACTA
AAACGTCACAACGTCTCACATTATATTTACTATCTAGCCACAGATAATATTCACATCGTGTTAGAAAACGATA
ACACCGTGTTAATAAAAGGACTTAAAAAGGTTGTAAATGTTAAATTCTCAAGAAACACGCATCTTATAGAAAC
GTCCTATGATAGGTTGAAATCAAGAGAAATCACATTTCAGCAATACAGGGAAAATCTTGCTAAAGCAGGAGTT
TTCCGATGGGTTACAAATATCCATGAACATAAAAGATATTACTATACCTTTGATAATTCATTACTATTTACTGA
GAGCATTCAGAACACTACACAAATCTTTCCACGCTAAATCATAACGTCCGGTTTCTTCCGTGTCAGCACCGGG
GCGTTGGCATAATGCAATACGTGTACGCGCTAAACCCTGTGTGCATCGTTTTAATTATTCCCGGACACTCCCG
CAGAGAAGTTCCCCGTCAGGGCTGTGGACATAGTTAATCCGGGAATACAATGACGATTCATCGCACCTGACAT
ACATTAATAAATATTAACAATATGAAATTTCAACTCATTGTTTAGGGTTTGTTTAATTTTCTACACATACGATT
CTGCGAACTTCAAAAAGCATCGGGAATAACACCATGAAAAAAATGCTACTCGCTACTGCGCTGGCCCTGCTTA
TTACAGGATGTGCTCAACAGACGTTTACTGTTCAAAACAAACCGGCAGCAGTAGCACCAAAGGAAACCATCAC
CCATCATTTCTTCGTTTCTGGAATTGGGCAGAAGAAAACTGTCGATGCAGCCAAAATTTGTGGCGGCGCAGAA
AATGTTGTTAAAACAGAAACCCAGCAAACATTCGTAAATGGATTGCTCGGTTTTATTACTTTAGGCATTTATA
CTCCGCTGGAAGCGCGTGTGTATTGCTCACAATAATTGCATGAGTTGCCCATCGATATGGGCAACTCTATCTG
CACTGCTCATTAATATACTTCTGGGTTCCTTCCAGTTGTTTTTGCATAGTGATCAGCCTCTCTCTGAGGGTGA
AATAATCCCGTTCAGCGGTGTCTGCCAGTCGGGGGGAGGCTGCATTATCCACGCCGGAGGCGGTGGTGGCTT
CACGCACTGACTGACAGACTGCTTTGATGTGCAACCGACGACGACCAGCGGCAACATCATCACGCAGAGCATC
ATTTTCAGCTTTAGCATCAGCTAACTCCTTCGTGTATTTTGCATCGAGCGCAGCAACATCACGCTGACGCATC
TGCATGTCAGTAATTGCCGCGTTCGCCAGCTTCAGTTCTCTGGCATTTTTGTCGCGCTGGGCTTTGTAGGTAA
TGGCGTTATCACGGTAATGATTAACAGCCCATGACAGGCAGACGATGATGCAGATAACCAGAGCGGAGATAAT
CGCGGTGACTCTGCTCATACATCAATCTCTCTGACCGTTCCGCCCGCTTCTTTGAATTTTGCAATCAGGCTGTC
AGCCTTATGCTCGAACTGACCATAACCAGCGCCCGGCAGTGAAGCCCAGATATTGCTGCAACGGTCGATTGCC
TGACGGATATCACCACGATCAATCATAGGTAAAGCGCCACGCTCCTTAATCTGCTGCAATGCCACAGCGTCCT
GACTTTTCGGAGAGAAGTCTTTCAGGCCAAGCTGCTTGCGGTAGGCATCCCACCAACGGGAAAGAAGCTGGTA
GCGTCCGGCGCCTGTTGATTTGAGTTTTGGGTTTAGCGTGACAAGTTTGCGAGGGTGATCGGAGTAATCAGT
AAATAGCTCTCCGCCTACAATGACGTCATAACCATGATTTCTGGTTTTCTGACGTCCGTTATCAGTTCCCTCCG
ACCACGCCAGCATATCGAGGAACGCCTTACGTTGATTATTGATTTCTACCATCTTCTACTCCGGCTTTTTTAGC
AGCGAAGCGTTTGATAAGCGAACCAATCGAGTCAGTACCGATGTAGCCGATAAACACGCTCGTTATATAAGCG
AGATTGCTACTTAGTCCGGCGAAGTCGAGAAGGTCACGAATGAACCAGGCGATAATGGCGCACATCGTTGCGT
CGATTACTGTTTTTGTAAACGCACCGCCATTATATCTGCCGCGAAGGTACGCCATTGCAAACGCAAGGATTGC
CCCGATGCCTTGTTCCTTTGCCGCGAGAATGGCGGCCAACAGGTCATGTTTTTCTGGCATCTTCATGTCTTAC
CCCCAATAAGGGGATTTGCTCTATTTAATTAGGAATAAGGTCGATTACTGATAGAACAAATCCAGGCTACTGT
GTTTAGTAATCAGATTTGTTCGTGACCGATATGCACGGGCAAAACGGCAGGAGGTTGTTAGCGCGACCTCCTG
CCACCCGCTTTCACGAAGGTCATGTGTAAAAGGCCGCAGCGTAACTATTACTAATGAATTGCCAGTTCGCGTT
CGCCAGCATCCG

Sequence of the 3593bp Watson-Crick fully complementary DNA hairpin.

(2) The full hairpin sequence used in this study should be made available. In addition, how the full unzipping curve is decomposed into segments needs to be described. It was mentioned that 400-600 segments are generated from the 3.6kbp-long hairpin, and Fig. 2C shows overlap between segments. How are the start and end points of each segment chosen?

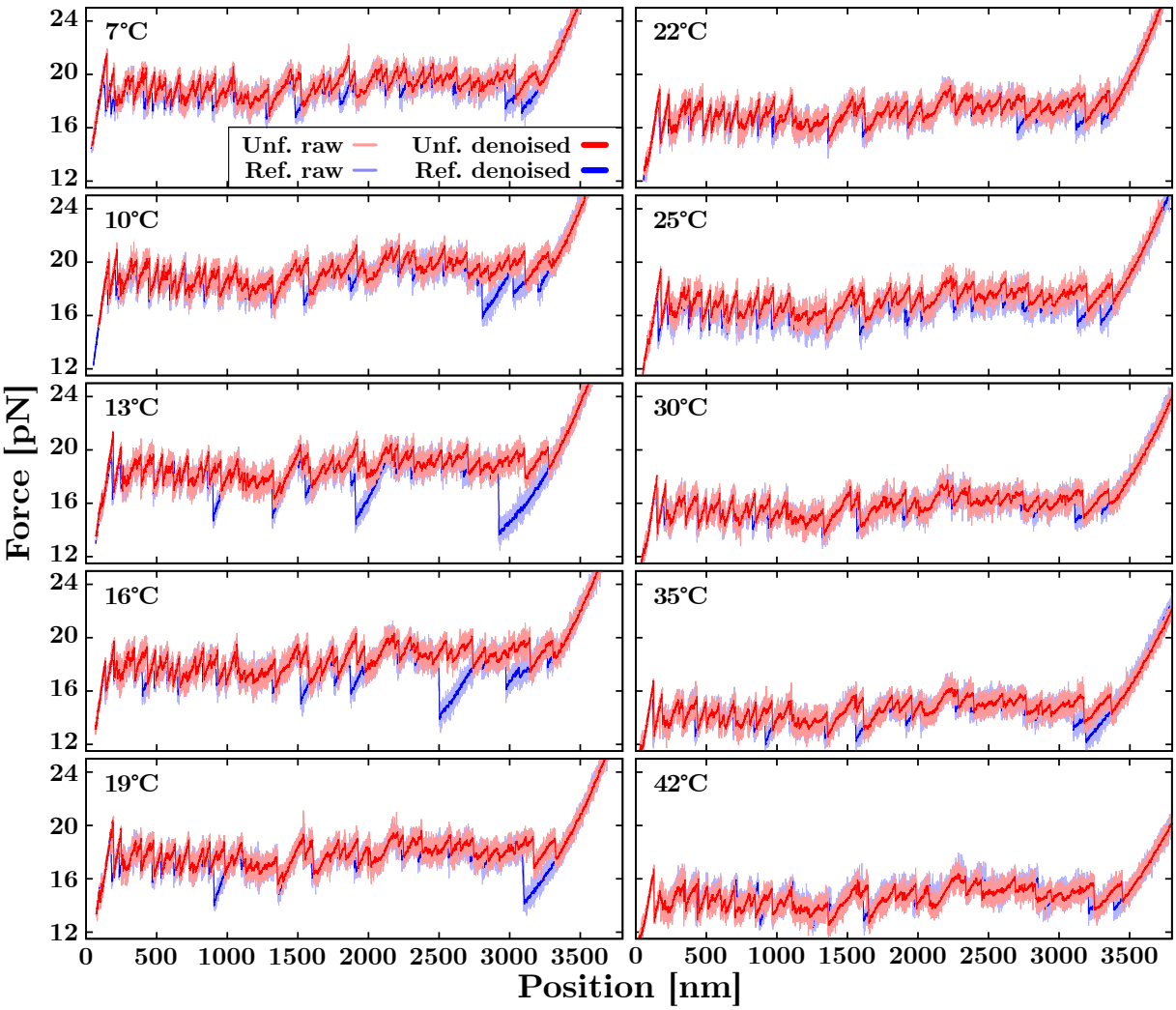

Figure R6: **Raw unzipping data.** Each panel shows one unfolding/refolding cycle for one molecule at a given T. Unzipping (red) and rezipping (blue) curves show a small hysteresis except for sporadic refolding events and particularly in correspondence of the first force jump observed during the refolding upon initiation of stem formation from the ssDNA. The light-colored lines are the raw data at 1kHz, while darker lines are raw data after denoising. This is done by applying a Savitzky–Golay filter over time windows of 150ms.

The Reviewer is asking about the procedure we developed to decompose the experimental FDC into segments. The raw unzipping data in a pulling cycle have been averaged over 150ms time windows using the Savitzky–Golay filter (see Fig. R6). The resulting FDCs were first averaged over pulling cycles for a given molecule and next averaged over different molecules, and statistical errors were determined (see Figure R7). Such an averaging procedure was repeated to generate one thermodynamic FDC at each temperature. Each FDC shows the characteristic sawtooth pattern with its sequence of force slopes and rips. Therefore, each peak along the FDC corresponds to a force rip concurrent with the unfolding of a group of base pairs. In practice, identifying the FDC peaks is done with an algorithm that finds the maxima of a given function $f(x)$, corresponding to the FDC. To do so, the function derivatives of the FDC are computed numerically from the data, and the maxima are determined by imposing $df(x)/dx = 0$. We notice that the magnitude of the force rip is proportional to the number of released bases in that unfolding event. The larger the rip, the more reproducible the number of bases is among repeated unzipping-rezipping cycles. In contrast, small force rips are related to releasing a small number of bases (10-40), becoming indistinguishable within the statistical errors along the FDC (see Figure R7). Therefore, we set $\Delta f \geq 2pN$ as the force-rip threshold for a force rip to be considered in the analysis. This procedure can be achieved with the built-in analysis tools of the most common programming languages (Python, in our case) and is

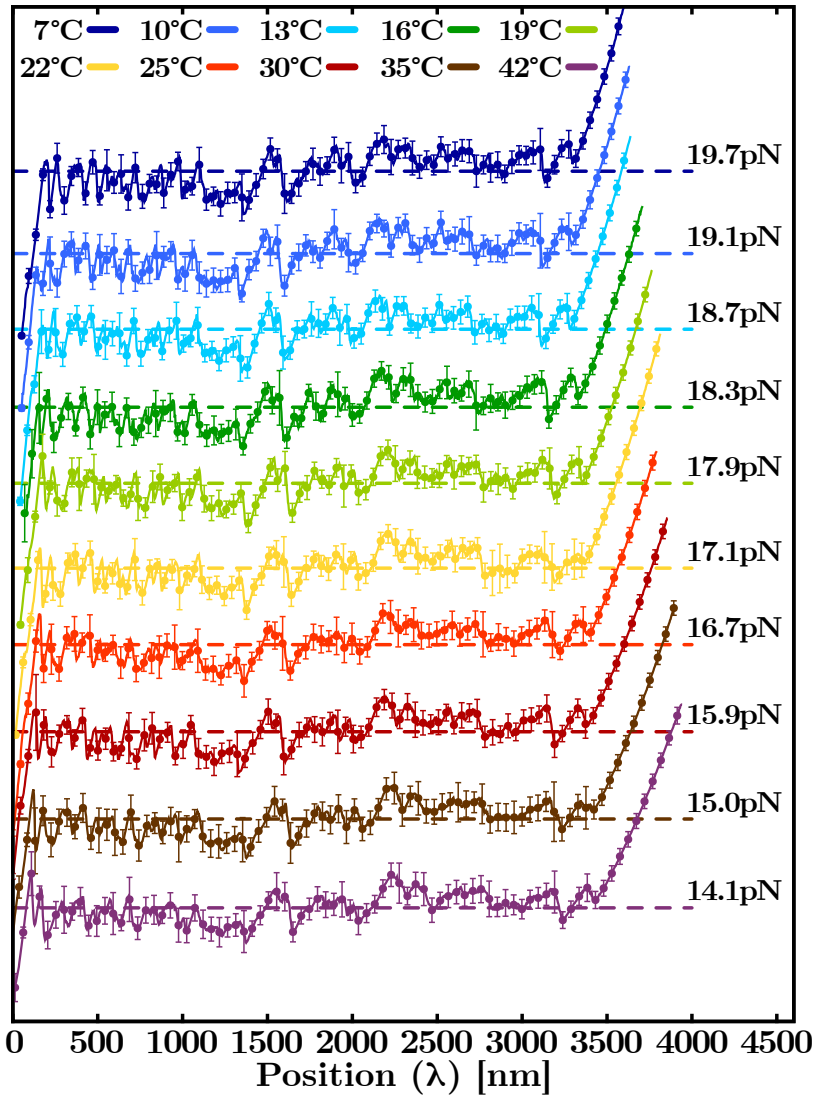

Figure R7: **Experimental variability of the T-dependent FDCs.** Temperature dependence of FDCs obtained by pulling the 3.6kbp DNA hairpin. At each T, the FDC results from averaging over $5-6$ molecules for $40-50$ unzipping/rezipping cycles. The error bars are the standard deviation in force (plotted for a fraction of the total data points) and show the molecule-to-molecule variability.

automatically done in the code we shared on the public repository, so we avoided a detailed description in the manuscript.
Once the peaks (maxima) have been selected, we defined a segment, Δx_k, as the FDC region delimited by two peaks. Note that peaks are not necessarily consecutive, as different segments can overlap, permitting us to analyze many segments and making the analysis robust. A set of segments has been obtained at each temperature by considering all possible combinations of starting and ending peaks along the FDC (see Fig. 2C, main text). This gives $K = 400 - 600$ segments per temperature. For each segment, we fit the slopes preceding the starting and ending peaks to the WLC elastic model to extract the number of released bases at the force rip after the peak. Determining the number of released bases in all force rips along the FDC permits us to identify the segment's location along the DNA sequence.

To better explain the method we use, we expanded the section "Derivation of the NNBP entropies" in the main text as follows:

To derive the entropies of the different NNBPs, we have decomposed the full unzipping curve into segments of variable length encompassing different regions along the FEC. Each segment is delimited by two peaks

corresponding to force-rips along the FEC. The peaks have been selected by looking for local maxima along the experimental signal. To do so, the FEC function derivatives have been numerically computed, and the maxima were determined by imposing $df(x)/dx = 0$. We notice that the magnitude of the force rip is proportional to the number of released bases in that unfolding event. To ensure the maximal signal-to-noise ratio, we discarded events within the FEC statistical errors by only accounting for force-rips with $\Delta f \leq 2$pN. Figure 2C shows examples of segments starting and ending at a peak (colored circles). By considering all possible combinations of starting and ending peaks, a set of K segments has been obtained for each T. Note that peaks are not necessarily consecutive, as different segments can overlap, making the analysis robust.

Ultimately, by following the Reviewer's suggestion, the full DNA hairpin sequence (Figure R1) is now reported in Extended Data Table S2.

(3) On several occasions, it was mentioned that the parameters for motifs GC/CG and TA/AT were obtained by applying the circular symmetry relations. The rationale and approach for this treatment need to be clarified.

The nearest-neighbour model predicts the free-energy of duplex formation (at zero force) by considering the stack of adjacent nucleotides along the hairpin backbone. Thus, the total formation free energy of an hairpin can be written as

$$E = \sum_{i,j=\{A,U,C,G\}} n_{ij} \Delta g_{ij} \,, \tag{1}$$

where n_{ij} is the occupancy of that dimer along the sequence and Δg_{ij} is the NNBP free energy between adjacent stacks or dimers (i, j). Considering all possible combination of Watson-Crick pairing, one gets 16 nearest-neighbour base-pair (NNBP) free energy values. However, 6 NNBPs must have the same formation energy due to complementary symmetry 5'-3', which reduces the number of free parameters to 10. Moreover, as Gray and Tinoco showed a long time ago [4], one can take advantage of the circular symmetry property to further reduce this number from 10 to 8. In fact, by neglecting end effects, the occupation numbers are proven to satisfy additional closure relations (see Appendix D in Ref. [8] for a derivation),

$$\begin{aligned} 2\Delta g_{CG} + \Delta g_{CA} + \Delta g_{AG} &= 2\Delta g_{GC} + \Delta g_{GA} + \Delta g_{AC} \\ 2\Delta g_{AU} + \Delta g_{AG} + \Delta g_{AC} &= 2\Delta g_{UA} + \Delta g_{GA} + \Delta g_{CA} \end{aligned} \tag{2}$$

which give,

$$\begin{aligned} \Delta g_{GC} &= \frac{1}{2}\left(2\Delta g_{CG} + \Delta g_{CA} + \Delta g_{AG} - \Delta g_{GA} - \Delta g_{AC}\right) \\ \Delta g_{UA} &= \frac{1}{2}\left(2\Delta g_{AU} + \Delta g_{AG} + \Delta g_{AC} - \Delta g_{GA} - \Delta g_{CA}\right) . \end{aligned} \tag{3}$$

As circular symmetry relations are commonly used in the thermodynamics of nucleic acids, we have omitted their detailed explanation in the present paper. An analogous argument can be done for the NNBP entropies (Δs_{ij}) and enthalpies (Δh_{ij}).

(4) Fig. 2A: the persistence length values appear to form three clusters rather than follow a linear relationship with the temperature. The authors should examine whether any experimental artifact exists at high or low temperatures.

The Reviewer points out that the measured temperature dependence for the persistence length (l_p) and the interphosphate distance (d_b) exhibits deviations from a linear behavior as some of the experimental points appear grouped in clusters. First, the elastic parameters' linear behavior as a function of temperature has already been observed in previous studies of ssDNA and dsDNA. However, corrections are of different sign [12]. We do not attach a particular significance to the apparent clustering observed in the l_p data, which

is compatible with the statistical errors originating from different causes. On the one hand, unzipping experiments are affected by an experimental drift, which is an uncontrolled variation of the relative distance between the optical trap and the micropipette. The direction and magnitude of the drift changes over time (during the same cycle and among different cycles) and is temperature-dependent. This effect is usually negligible for short hairpins but affects the derivation of the elastic properties of long ssDNA molecules, as in our unzipping experiments. Therefore, the issue of questionable data clustering is not related to the emergence of the complex temperature dependence of the elastic parameters but rather to the intrinsic statistical variability of the experimental measurements.

(5) Line 167: typo "intextensible".
(6) Line 230: Fig. 3D should be Fig. 2D.

We thank the Reviewer for pointing out these typos.

(7) Line 343-344: not sure what "extensive" and "intensive" mean in this context.

In physics, intensive properties do not vary with the amount of substance or size of the system, such as temperature, density, or hardness. By contrast, an extensive quantity is additive for subsystems. This includes mass, volume, and, in our case, the thermodynamics variables molecular extension, free energy, entropy, enthalpy, and heat capacity.

(8) Line 457: x_h (handle extension) is not expected to remain constant during the entire unzipping process.

The Reviewer is correct as, in general, this is not true. The contributions from the dsDNA handles will change upon varying the force according to the WLC dependence through its persistence length. However, the persistence length of dsDNA is roughly 20 times larger than that of ssDNA ($l_p^{ds} = 12.5$nm and $l_p^{ss} = 0.77$nm at 25°C). In addition, we are using very short dsDNA handles (29bp) compared to the length of the stem (3.6kbp). Therefore, the stiffness of the dsDNA handles is at least ten times larger than that of the released ssDNA, which is long and soft. In these conditions, the change in molecular extension of the dsDNA handles between the maximum and minimum force values along the unzipping curve is negligible. Therefore, the dsDNA handles contribution ($2x_h$) to the change in the total extension $\Delta\lambda$ is very small and can be safely taken as constant at the average unzipping force, f_m.
In light of the Reviewer's comment, we amended the corresponding sentence of the main text as follows:

The comparably high stiffness of the optically trapped bead and the short extension and large persistence length of the dsDNA handles, as compared to the released ssDNA ($l_p^{ds} = 12.5$nm and $l_p^{ss} = 0.77$nm at 25°C), permit us to approximate the contributions x_b and $2x_h$ to λ as constant near the average unzipping force. Therefore, $\Delta\lambda \approx \Delta x$, i.e., equals the ssDNA's extension change only.

(9) Line 462: this relation assumes ΔH to be T-independent?

The thermodynamic relation $\Delta S_0 = -\partial\Delta G_0/\partial T$ (as reported in the main text) defines the entropy ($\Delta S_0(T)$) as the partial derivative of the Gibbs free energy with respect to the temperature (T), $-\partial\Delta G_0/\partial T$. This relation is the standard definition of entropy in the constant P(pressure), T ensemble, $S = -\partial G/\partial T$. Notice that Gibbs's relation for the entropy is independent of the fact that ΔH is constant with T, in which case $\Delta C_p = 0$. In general, all thermodynamics quantities defined in the text are to be considered temperature dependent, even if the temperature argument has sometimes been dropped. In the case of the relation above, we did not explicitly make the T-dependence to lighten the notation, which might have confused the reviewer. However, in light of the Reviewer's comment, we make this dependency explicit to avoid misunderstanding.

Reviewer #2 (Remarks on code availability)

The code is clear and sufficient.

Reviewer #3 (Remarks to the Author)

This is a very interesting paper in which the authors apply single-molecule force measurements to determine the elusive heat capacity changes of DNA at the single base pair (or nearest neighbor) level. They mechanically unzip a long DNA hairpin (3593 bp) using an optical trap setup. The force-distance curve displays a sawtooth pattern whose analysis allows the measurement of entropies, enthalpies, and free energies for the different nearest neighbor base pairs. The results are interesting and important for everyone who studies or uses the biophysical properties of DNA.

We are pleased to learn that the Reviewer finds our work of broad interest to the DNA scientific community and supports the manuscript for publication. In the following, we reply point-by-point to the Reviewer's comments.

Specific questions:

- It is not really explained how the hairpin is made in the first place. What is its sequence? Is it a plasmid that is opened and processed somehow to have a tetraloop? Do you need to anneal it to make sure it is double-stranded, etc.?

To carry out the unzipping experiments with the temperature-jump optical trap, we designed and synthesized a 3593bp DNA hairpin ending with a GAAA tetraloop. This hairpin is flanked by two 29bp dsDNA handles tagged with biotin and digoxigenin at the other end to increase the tether lifetimes. The synthesis steps are described in red below.

We have added the synthesis description to the Supp. Info. and the sequence of all oligos used for the synthesis and the DNA sequence (Extended Data Tables 1 and 2 in the resubmitted version).

Restriction Digest of λ-DNA

We start by digesting the N6-methyladenine-free λ-DNA using the restriction enzyme EcoRI from New England Biolabs. The fragment spanning positions from 44972 to 48502 on the cosR end of the λ-DNA has been used to assemble the hairpin stem. The digestion reaction was prepared as follows:

Components. 64μl MilliQ water, 10μl 10X NEB reaction buffer, 1μl 100X BSA, 20μl λ-DNA (0.5μg/μl),

Table R2: **Oligo sequences**

Oligo name	**Sequence**
Loop	5′ – Pho – AAT T GC CAG TTC GCG TTC GCC AGC ATC CG A CTA CGG ATG CTG GCG AAC GCG AAC TGG C – 3′
Bio-cosRshort3	5′ – Bio – GAC TTC ACT AAT ACG ACT CAC TAT AGG GA A ATA GAG ACA CAT ATA TAA TAG ATC TT – 3′
cosRlong	5′ – Pho – GGG CGG CGA CCT AAG ATC TAT TAT ATA TGT GTC TCT ATT AGT TAG TGG TGG AAA CAC AGT GCC AGC GC – Dig – 3′
Split3	5′ – TCC CTA TAG TGA GTC GTA TTA GTG AAG TC – 3′
Inverted-splint	3′ – Dig – AAA AA – 5′ – 5′ – GCG CTG GCA CTG TGT TTC CAC CAC TAA C (SpC3)– 3′

List of the oligo sequences used for the synthesis of the DNA hairpin.

and 5μl EcoRI enzyme (20U/μl).

Procedure. The reaction was incubated for 3 hours at 37°C to ensure complete digestion. Next, EcoRI was thermally inactivated at 65°C for 20min.

The EcoRI restriction enzyme has been used to cut the λ-DNA into six fragments of lengths 21226bp, 7421bp, 5804bp, 5643bp, 4878bp, and 3530bp. Then, the digested DNA was purified (0.8% agarose gel, 1XTBE) using Sybr Safe staining (5μl) and blue illumination to minimize DNA damage and nicking. Finally, the 3530bp fragment was extracted using the Qiaquick gel extraction kit (Qiagen). The purified DNA was eluted in 50μl of 2mM Tris-Cl.

Phosphorylation of Digested λ-DNA

The digested λ-DNA fragment was phosphorylated using T4 polynucleotide kinase to prepare it for subsequent ligation. The phosphorylation reaction included:

Components. 19μl MilliQ water, 10μl 10X PNK buffer, 10μl, 10mM ATP, 50μl of digested λ-DNA ($\approx$180ng/μl) and 1μl of T4 polynucleotide kinase.

Procedure. The reaction was incubated for 30 minutes at 37°C. Next, polynucleotide kinase was thermally inactivated at 65°C for 20min.

Synthesis of dsDNA Handles and End-Loop

To create the hairpin structure, the phosphorylated λ-DNA fragment was annealed to oligonucleotides that form the hairpin loop and the dsDNA handles. The end-loop was synthesized by annealing the DNA fragment to an oligonucleotide (Oligonucleotide loop in Table R2), which self-assembled into a hairpin with a tetraloop at one end and an overhang complementary to the EcoRI restriction site. The dsDNA handles were formed using two partially complementary oligonucleotides (cosRlong and Bio-cosRshort3 in Table R2) that hybridize forming a protruding end complementary to the cosR end. The Bio-cosRshort3 oligonucleotide was purchased 5′-biotinylated. The molecular construction was doubly biotinylated in one handle and doubly Digoxigenin-tailed in the other handle.

Double Biotinylated Handle. Bio-cosRshort3, and the complementary oligonucleotide (Split3 in the Table R2) were tailed at the 3′ end with multiple biotins using terminal transferase and biotin-labeled dUTPs (Roche).

Double Digoxigenin-Tailed Handle. To create the digoxigenin dsDNA handle, we used cosR-long oligocleotide and a modified oligonucleotide (inverted-splint Table R2) complementary to the unpaired region of cosRlong. This oligonucleotide was modified with a polarity inversion at its 5′ end and a blocked 3′ end which allowed the oligonucleotide to be tailed with multiple digoxigenins at its 5′ end using digoxigenin-labeled dUTPs (Roche).

Oligonucleotide Tailing

The tailing reactions for both the biotin and digoxigenin-modified oligonucleotides were prepared using terminal transferase. The reaction mixture included:

Components. 8μl MilliQ water, 4μl 5X $CoCl_2$ solution, 4μl 5X reaction buffer, 1μl 10mM dATP, 1μl of either 1mM BIO-dUTP or DIG-dUTP (Roche), and 1μl 100μM oligonucleotide.

Procedure. The reaction was incubated for 15 minutes at 37°C, followed by the addition of 1μl of 0.5M EDTA to stop the reaction. The tailed oligonucleotides were purified using the QIAquick Nucleotide Removal Kit (Qiagen) and eluted in 50μl of 10mM Tris-Cl.

Annealing of the DNA Hairpin

Once the oligonucleotides for the handles and the end-loop were prepared, they were annealed to the phosphorylated λ-DNA fragment in a reaction containing:

Components. 29μl digested λ-DNA (330ng/μl), 5μl of each oligonucleotide (at varying concentrations for the loop and handles), 1.5μl 1M Tris pH 7.5, and 1.2μl 5M NaCl.

Procedure. The reaction was incubated at 70°C for 10 minutes, then at 55°C for 10 minutes, and gradually cooled down to room temperature in a thermal bath.

Ligation of the DNA Hairpin

Finally, the annealed hairpin construct was ligated overnight using T4 DNA ligase in the following ligation reaction:

Components. 55μl annealed product, 8μl 10X T4 DNA ligase buffer, 5μl 10mM ATP, and 3μl T4 DNA ligase (400U/μl).

Procedure. The reaction was incubated overnight at 16°C. After ligation, the reaction was thermally inactivated by incubating at 65°C for 10 minutes, and the final product was stored at −20°C.

This procedure resulted in the synthesis of a 3.6kbp DNA hairpin structure, ready for use in force-spectroscopy. The hairpin sequence is reported in Table R1.

- It is not clear to this reviewer how the curve is decomposed into different segments. Apparently, you use K different segments – but what determines K? How does one arrive from 3593 bp at K = 400-600? Or are these just intervals between randomly chosen peaks in the FDC?
- How is the analysis performed then? Do you know the sequence for each given segment? How accurately is the sequence aligned with the measurement? Does that simply follow from the measured length?

The raw unzipping data in a pulling cycle have been averaged over 150ms time windows using the Savitzky–Golay filter (see Fig. R8). The resulting FDCs were first averaged over pulling cycles for a given molecule and next averaged over different molecules, and statistical errors were determined (see Figure R9). Such an averaging procedure was repeated to generate one thermodynamic FDC at each temperature. Each FDC shows the characteristic sawtooth pattern with its sequence of force slopes and rips. Therefore, each peak along the FDC corresponds to a force rip concurrent with the unfolding of a group of base pairs. In practice, identifying the FDC peaks is done with an algorithm that finds the maxima of a given function $f(x)$, corresponding to the FDC. To do so, the function derivatives of the FDC are computed numerically from the data, and the maxima are determined by imposing $df(x)/dx = 0$. We notice that the magnitude of the force rip is proportional to the number of released bases in that unfolding event. The larger the rip, the more reproducible the number of bases is among repeated unzipping-rezipping cycles. In contrast, small force rips are related to releasing a small number of bases (10-40), becoming indistinguishable within the statistical errors along the FDC (see Figure R9). Therefore, we set $\Delta f \geq 2pN$ as the force-rip threshold for a force rip to be considered in the analysis. This procedure can be achieved with the built-in analysis tools of the most common programming languages (Python, in our case) and is automatically done in the code we shared on the public repository, so we avoided a detailed description in the manuscript.
Once the peaks (maxima) have been selected, we defined a segment, Δx_k, as the FDC region delimited by two peaks. Note that peaks are not necessarily consecutive, as different segments can overlap, permitting us to analyze many segments and making the analysis robust. A set of segments has been obtained at each temperature by considering all possible combinations of starting and ending peaks along the FDC (see Fig. 2C, main text). This gives $K = 400 - 600$ segments per temperature. For each segment, we fit the slopes

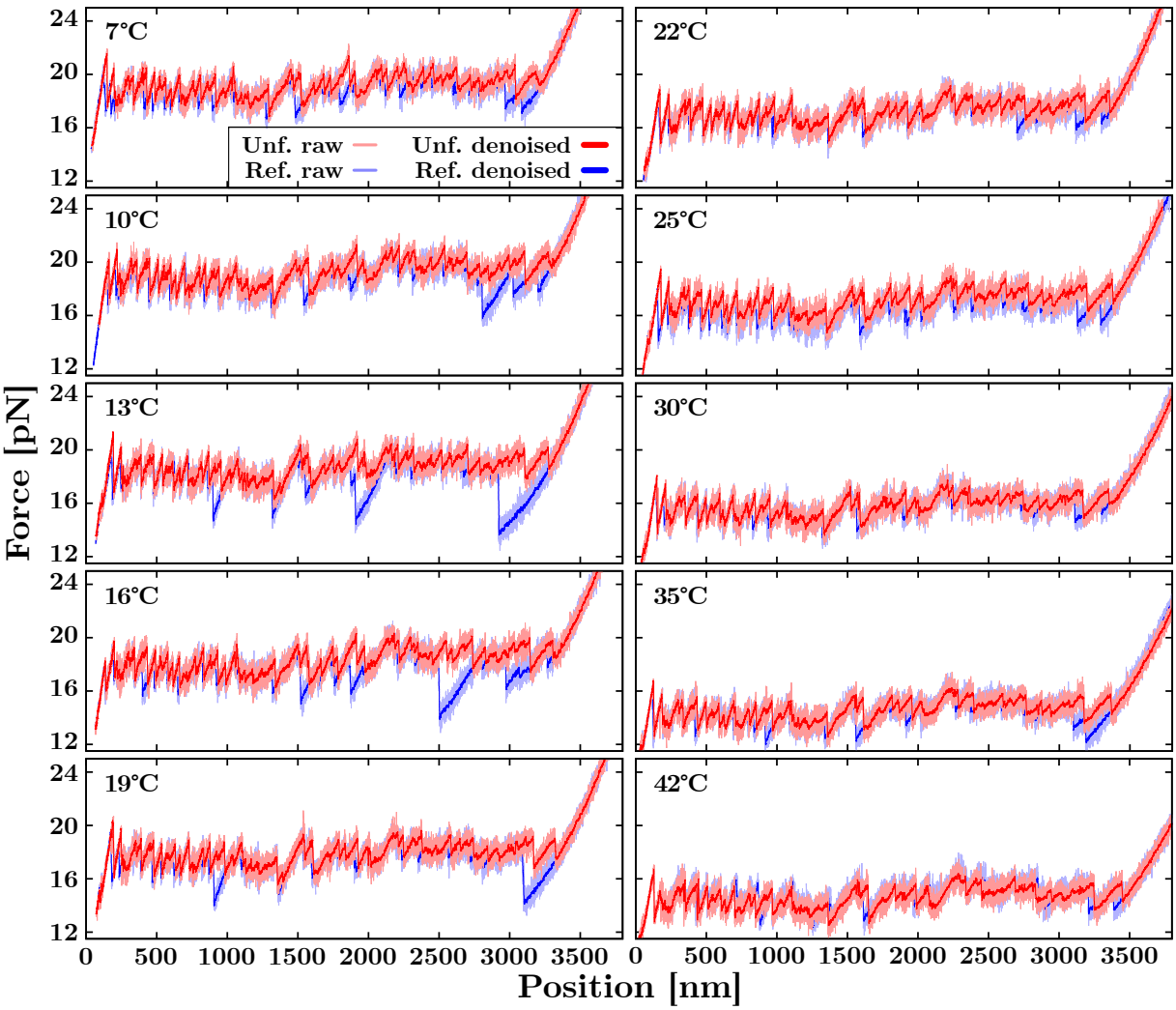

Figure R8: **Raw unzipping data.** Each panel shows one unfolding/refolding cycle for one molecule at a given T. Unzipping (red) and rezipping (blue) curves show a small hysteresis except for sporadic refolding events and particularly in correspondence of the first force jump observed during the refolding upon initiation of stem formation from the ssDNA. The light-colored lines are the raw data at 1kHz, while darker lines are raw data after denoising. This is done by applying a Savitzky–Golay filter over time windows of 150ms.

preceding the starting and ending peaks to the WLC elastic model to extract the number of released bases at the force rip after the peak. Determining the number of released bases in all force rips along the FDC permits us to identify the segment's location along the DNA sequence.

To better explain the method we use, we expanded the section "Derivation of the NNBP entropies" in the main text as follows:

To derive the entropies of the different NNBPs, we have decomposed the full unzipping curve into segments of variable length encompassing different regions along the FEC. Each segment is delimited by two peaks corresponding to force-rips along the FEC. The peaks have been selected by looking for local maxima along the experimental signal. To do so, the FEC function derivatives have been numerically computed, and the maxima were determined by imposing $df(x)/dx = 0$. We notice that the magnitude of the force rip is proportional to the number of released bases in that unfolding event. To ensure the maximal signal-to-noise ratio, we discarded events within the FEC statistical errors by only accounting for force-rips with $\Delta f \leq 2$pN. Figure 2C shows examples of segments starting and ending at a peak (colored circles). Considering all possible combinations of starting and ending peaks, a set of K segments has been obtained for

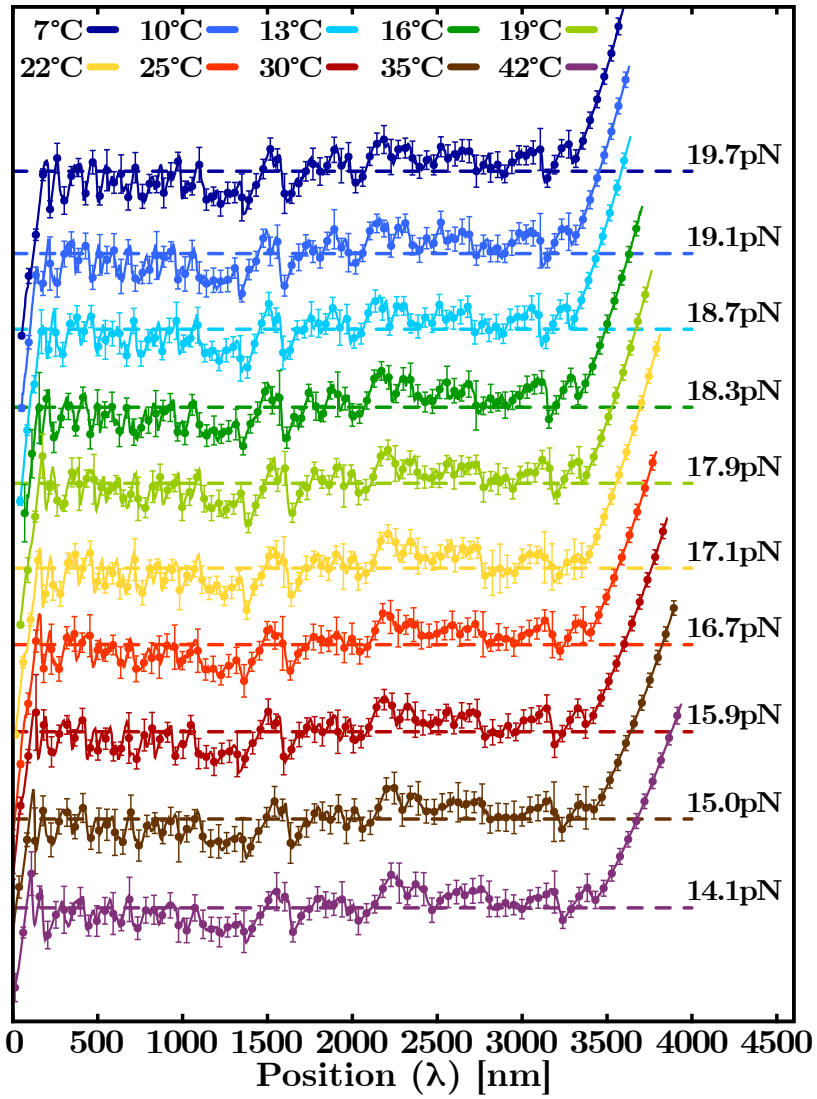

Figure R9: **Experimental variability of the T-dependent FDCs.** Temperature dependence of FDCs obtained by pulling the 3.6kbp DNA hairpin. At each T, the FDC results from averaging over $5-6$ molecules for $40-50$ unzipping/rezipping cycles. The error bars are the standard deviation in force (plotted for a fraction of the total data points) and show the molecule-to-molecule variability.

each T. Note that peaks are not necessarily consecutive, as different segments can overlap, making the analysis robust.

- It would be great if the authors could explain better where the sawtooth pattern comes from. They explain how one of the force jumps occurs in the SI, but what determines the overall pattern? Are there preferred positions in the sequence where the force jumps?

The force-distance curve (FDC) is the sequence of force slopes and rips describing the molecule unzipping. Therefore, all peaks along the FDC correspond to a force rip and the unfolding of a group of base pairs. In an unzipping experiment, the optical trap is moved at a constant speed with respect to the (fixed) micropipette. At the beginning of the protocol, the molecule is folded into its native double-stranded (ds) hairpin configuration (Fig. 1A, main text). Unzipping is a stick-slip process consisting of the succession of an elastic deformation (stick) followed by the release of groups of bases that collectively unfold cooperatively (slip), resulting in sudden force rips. As the optical trap moves away from the pipette, the force applied to the hairpin increases until the intramolecular bonds at the beginning of the stem break open, causing a sudden force drop in the FDC. Further moving away the optical trap from the micropipette causes

the force to increase again until the following group of bases opens, corresponding to another force rip in the FDC. The unfolding protocol proceeds until the hairpin is fully unzipped and the single-strand (ss) form is fully stretched. The reverse process starts (rezipping), and the molecule refolds from the loop until the native ds hairpin has been reformed. Upon rezipping, groups of bases are cooperatively absorbed into the stem, resulting in sudden force jumps. The force-distance curve (FDC) measured during unzipping and rezipping exhibits a saw-tooth pattern that depends on the sequence of the hairpin. The number of bp that collectively opens at each force rip is variable and depends on the specific sequence as the Watson-Crick base-pairing through the abundance of guanine(G)-cytosine(C) relative to the adenine(A)-thymine(T) base pairs. Therefore a larger external work, i.e. a higher force, is needed to unzip regions of the DNA sequence rich in G and C, causing higher force-rips in the FDC. In general, it has been observed in Ref. [7] that unzipping proceeds by cooperatively unzipping regions along the FDC that typically involve the release of 10-150 base-pairs per force rip.

Following the Reviewer's comment, we expanded Sec. 1, Methods as follows:

Each rip corresponds to the cooperative unfolding of a group of bp due to the applied (external) work exceeding the hybridization (internal) energy. When the molecule is completely unzipped, the rezipping protocol starts; λ is decreased, and the molecule folds back reversibly to the hairpin (native) state. We notice that the number of bp released in each unfolding event depends on the sequence composition so that the FDC can be considered as the DNA footprint, being unique for each sequence.

- SI Fig. 4: lowest energy branch (n2, pink) → yellow?

We thank the Reviewer for pointing out this typo.

- The choice of references for CRISPR and origami is unconventional. The first paper showing gene editing is not Ref. 1, but rather 10.1126/science.1225829, 10.1126/science.1231143. Also, for origami, one should rather cite Rothemund (10.1038/nature04586) and Shih (10.1038/nature08016) and not, e.g., Ref. 3 in the manuscript.

We thank the Reviewer for bringing these works to our attention. These references have been added according to their suggestion.

References

[1] Chris R Calladine and Horace Drew. *Understanding DNA: the molecule and how it works*. Academic press, 1997.

[2] Sara De Lorenzo, Marco Ribezzi-Crivellari, J Ricardo Arias-Gonzalez, Steven B Smith, and Felix Ritort. A temperature-jump optical trap for single-molecule manipulation. *Biophys. J.*, 108(12):2854–2864, 2015.

[3] Nuria Forns, Sara de Lorenzo, Maria Manosas, Kumiko Hayashi, Josep Maria Huguet, and Felix Ritort. Improving signal/noise resolution in single-molecule experiments using molecular constructs with short handles. *Biophys. J.*, 100(7):1765–1774, 2011.

[4] Donald M Gray and Ignacio Tinoco Jr. A new approach to the study of sequence-dependent properties of polynucleotides. *Biopolymers*, 9(2):223–244, 1970.

[5] S Hadži and J Lah. Origin of heat capacity increment in dna folding: The hydration effect. *Biochim. Biophys. Acta - Gen. Subj.*, 1865(1):129774, 2021.

[6] Josep M Huguet, Cristiano V Bizarro, Núria Forns, Steven B Smith, Carlos Bustamante, and Felix Ritort. Single-molecule derivation of salt dependent base-pair free energies in DNA. *Proc. Natl. Acad. Sci. U.S.A.*, 107(35):15431–15436, 2010.

[7] Josep Maria Huguet, Nuria Forns, and Felix Ritort. Statistical properties of metastable intermediates in DNA unzipping. *Phys. Rev. Lett.*, 103(24):248106, 2009.

[8] Josep Maria Huguet, Marco Ribezzi-Crivellari, Cristiano Valim Bizarro, and Felix Ritort. Derivation of nearest-neighbor DNA parameters in magnesium from single molecule experiments. *Nucleic Acids Res.*, 45(22):12921–12931, 2017.

[9] Pedro Licinio and João Carlos O Guerra. Irreducible representation for nucleotide sequence physical properties and self-consistency of nearest-neighbor dimer sets. *Biophys. J.*, 92(6):2000–2006, 2007.

[10] Alexander A Lomzov, Yury N Vorobjev, and Dmitrii V Pyshnyi. Evaluation of the gibbs free energy changes and melting temperatures of dna/dna duplexes using hybridization enthalpy calculated by molecular dynamics simulation. *J. Phys. Chem. B.*, 119(49):15221–15234, 2015.

[11] Peter J Mikulecky and Andrew L Feig. Heat capacity changes associated with nucleic acid folding. *Biopolymers*, 82(1):38–58, 2006.

[12] Marc Rico-Pasto and Felix Ritort. Temperature-dependent elastic properties of dna. *Biophys. Rep.*, 2(3):100067, 2022.

[13] Marc Rico-Pasto, Annamaria Zaltron, Sebastian J Davis, Silvia Frutos, and Felix Ritort. Molten globule–like transition state of protein barnase measured with calorimetric force spectroscopy. *Proc. Natl. Acad. Sci. U.S.A.*, 119(11):e2112382119, 2022.

[14] Ioulia Rouzina and Victor A Bloomfield. Heat capacity effects on the melting of dna. 1. general aspects. *Biophys. J.*, 77(6):3242–3251, 1999.

[15] John SantaLucia. A unified view of polymer, dumbbell, and oligonucleotide DNA nearest-neighbor thermodynamics. *Proc. Natl. Acad. Sci. U.S.A.*, 95(4):1460–1465, 1998.

NCOMMS-24-32591 - Reviewers' Response

We thank the Reviewers for carefully reading our response and are gratified that they all find the work valuable and recommend it for publication. The revised version of the manuscript addresses the second reviewer's final comment.

As for the previous round of review, comments by the reviewers are shown in bold text, and modifications to the main text are shown in red. Correspondingly, changes or additions are also shown in red in the resubmitted manuscript. Please notice that changes made during the first round of review are still marked in red in the resubmitted documents.

Reviewer's Comments:

Reviewer #1 (Remarks to the Author)

The authors have fully addressed my comments. I recommend publication of the manuscript.

We are pleased that the Reviewer found the answers to their comments exhaustive and recommended our work for publication.

Reviewer #2 (Remarks to the Author)

The authors have done a good job addressing my comments, except one point (raised by all three reviewers) that still needs clarification:

We thank the reviewer for the positive appraisal of our work. In what follows, we address the final comments by the Reviewer.

Line 208-211: "Considering all possible combinations of starting and ending peaks, a set of K segments has been obtained for each T. Note that peaks are not necessarily consecutive, as different segments can overlap, making the analysis robust." The number of all possible combinations of starting and ending peaks must exceeds 400-600? Are the K segments randomly chosen from all possible combinations, or are there additional criteria?

The selection of the FEC peaks is carried out by computing the numerical derivatives of the experimental signal to find the local maxima such that $df(x)/dx = 0$. By definition, each peak is followed by a force-rip proportional to the number of released bases in the unfolding event. To maximize the signal-to-noise ratio of rip detection, we only considered force-rips with $\Delta f \geq 2$pN and discarded the rest. This has been the only filtering process applied for selecting the peaks at which the rips occur. In a single unzipping curve, the typical number of selected peaks is $N_p \sim 30-35$ (see Fig. 2C, main text). Then, segments are built by considering all possible combinations of starting and ending peaks, resulting in a total number of segments, $K = N_p(N_p-1)/2 \sim 400-600$ at each temperature. We note that all K segments generated by this combinatorial approach are used for analysis.

We amended the text as follows:

Considering all possible combinations of starting and ending peaks, a set of K segments has been obtained for each T. Note that peaks are not necessarily consecutive, as different segments can overlap, making the analysis more robust. In a single unzipping curve, the typical number of peaks is $N_p \sim 30-35$ (see Fig. 2C). No further selection step is applied to the segments for analysis, resulting in a total number of segments $K = N_p(N_p-1)/2 \sim 400-600$ per temperature. Notice that such a large number of segments, each being a different sequence, permits measuring many sequences in a single unzipping experiment. In this regard, the unzipping approach is high throughout compared to oligo hybridization experiments, where the measurement of melting curves must be repeated for every oligo sequence.

Reviewer #3 (Remarks to the Author)

The authors have responded well to all the reviewer comments and the paper may now be accepted for publication.

We thank the Reviewer for positively evaluating our work on the manuscript based on the Reviewer's feedback and recommending it for publication.