# Supplementary Information: DNA Calorimetric Force Spectroscopy at Single Base Pair Resolution

P. Rissone<sup>1</sup>, M. Rico-Pasto<sup>2,3</sup>, S. B. Smith<sup>4</sup> and F. Ritort<sup>1,5\*</sup>

<sup>1</sup>Small Biosystems Lab, Condensed Matter Physics Departement, Universitat de Barcelona, C/ Marti i Franques 1, Barcelona, 08028, Spain.

 <sup>2</sup>Unit of Biophysics and Bioengineering, Department of Biomedicine, School of Medicine and Health Sciences, Universitat de Barcelona, C/Casanoves 143, Barcelona, 08036, Spain.
 <sup>3</sup>Institute for Bioengineering of Catalonia (IBEC), The Barcelona Institute for Science and Technology (BIST), Barcelona, 08028, Spain.

<sup>4</sup>Steven B. Smith Engineering, Los Lunas, New Mexico, USA. <sup>5</sup>Institut de Nanociència i Nanotecnologia (IN2UB), Barcelona, Spain.

\*Corresponding author(s). E-mail(s): ritort@ub.edu; Contributing authors: rissone.paolo@gmail.com; m.rico.pasto@gmail.com; stevesmith@gmail.com;

# Contents

Supple	mentary Methods	<b>2</b>
1	Temperature Dependence of the DNA FDCs	2
2	Stochastic Gradient Descent in a Nutshell	2
3	Prediction of the DNA Unzipping Curve	3
	3.1 Computation of the Equilibrium FDC	4
	3.2 Equilibrium Free Energy	4
4	Fit of the NNBP parameters	5

#### **Extended Data: Figures**

#### **Extended Data: Tables**

# Supplementary Methods

#### 1 Temperature Dependence of the DNA FDCs

The elastic properties of ssDNA are strongly temperature dependent (see Fig. 1B, main text). Accurately measuring these properties requires modeling all contributions to the trap-pipette distance,  $\lambda$ , 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  $(x_{\rm b} \text{ and } x_{\rm h})$  can be evaluated by using the *effective stiffness* method <sup>48</sup>. According to it, these terms are approximated by an effective stiffness,  $k_{\rm eff}^{-1} \approx k_{\rm h}^{-1} + k_{\rm b}^{-1}$ . The use of short handles (29bp) makes the evaluation of the stretching terms easier as their stiffness is much larger as compared to the trap stiffness  $(k_{\rm h} \gg k_{\rm b})$ , implying that  $k_{\rm eff} \approx k_{\rm b}$ . Moreover, if the force varies in a relatively narrow range  $(f_{\rm max} - f_{\rm min} \leq 10 \text{pN})$ , the trap stiffness can be considered nearly force-independent so  $k_{\rm eff}$  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. 1). This allows us to compute the (effective) contribution of the handles and optical trap,  $x_{\rm eff}$ , to the total trap-pipette distance,  $\lambda$ .

#### 2 Stochastic Gradient Descent in a Nutshell

The basic principle behind stochastic approximation can be backtracked to the Robbins–Monro algorithm of the  $1950s^{60}$ . Since then, stochastic gradient descent (SGD) methods have become one of the most widely used optimization methods<sup>61–65</sup>. SGD is an iterative method for optimizing an objective function, J(w), with suitable smoothness properties (e.g., differentiable or subdifferentiable). The set of parameters,  $w^*$ , minimizing J(w), is iteratively approximated according to an update algorithm proportional to the antigradient of the objective function,  $-\nabla_w J(w)$ . Starting from an initial guess of w, at each step of the algorithm, the parameters are updated according to

$$\begin{cases} w_{t+1} = w_t + v_{t+1} \\ v_{t+1} = \beta v_t - \eta \nabla_{w_t} J(w) , \end{cases}$$
(1)

where  $v_t$  is the velocity of the optimization and  $\eta \geq 0$  is the step size (called *learing rate*). The parameter  $\beta$  (the so-called *momentum coefficient*) accounts for a fraction of the previous step in the current update. The critical difference between SGD and standard gradient descent algorithms is that information (total entropy and coefficients) from only one FEC segment ( $\Delta x_k$ ) at a time is used to calculate the step, and the segment is picked randomly at each step.

6

13

The SGD convergence rate can be improved by considering Nesterov's Accelerated Gradient (NAG), introduced in 1983<sup>66,67</sup>. According to NAG, the update equations are

$$\begin{cases} w_{t+1} = w_t + v_{t+1} \\ v_{t+1} = \beta v_t - \eta \nabla_{w_t + \beta v_t} J(w) \,. \end{cases}$$
(2)

While the classic momentum (CM) algorithm updates the velocity vector by computing the gradient at  $w_t$ , the NAG algorithm computes the gradient at  $w_t + \beta v_t$ . To make an analogy, while CM faithfully trusts the gradient at the current iteration, NAG puts less faith in it and looks ahead in the direction suggested by the velocity vector; it then moves in the direction of the gradient at the look-ahead point. If  $\nabla_{w_t+\beta v_t} J(w) \approx \nabla_{w_t} J(w)$ , then the two updates are similar. The advantage of using NAG is that it converges at a rate of  $\mathcal{O}(1/t^2)$ , while CM converges at a rate of  $\mathcal{O}(1/t)$ .

To derive the DNA NNBP entropies from unzipping experiments, we used an SGD minimization implementing NAG update equations. Let us rewrite Eq.(2) (main text) as  $\Delta S_0 = C \Delta s$ , where  $\Delta S_0$  is the vector of entropies measured with the Clausius-Clapeyron equation for each of the K FEC segments,  $\Delta s$  is the vector of the I = 8 NNBP entropy parameters, and C is the  $K \times I$ matrix of the coefficients,  $c_{k,i}$ .

Thus, for a given loss function (ex., least squares), the algorithm has to minimize

$$J(w) = \sum_{k=1}^{K} (\hat{w}_k - w_k)^2 = \sum_{k=1}^{K} (\Delta S_{0,k} - C_k \Delta \mathbf{s})^2.$$
(3)

By using this method, we measured the DNA entropies at the single base-pair level for each experimental temperature in the range [280, 315] K (see results in Fig. 3C, main text and Extended Data Table 3).

#### 3 Prediction of the DNA Unzipping Curve

In unzipping experiments, the total trap-pipette distance,  $\lambda$ , can be written as

$$\lambda(f) = x_{\rm b}(f) + x_{\rm h}(f) + x_{\rm ss}(f, n) + x_{\rm d}(f), \qquad (4)$$

where  $x_{\rm b}(f)$  is the displacement of the bead from the center of the optical trap,  $x_{\rm h}(f)$  and  $x_{\rm ss}(f,n)$  account for the extension of the two double-stranded handles and the ssDNA extension, respectively (described with the WLC model, Eq.(5), Methods), and  $x_{\rm d}(f)$  is the projection of the folded hairpin of diameter d (typically d = 2nm for DNA and RNA hairpins<sup>49</sup>) along the pulling axis<sup>68</sup>. It is modeled with the freely-jointed chain in Eq.(6), Methods. For a given  $\lambda$ , the total system free energy is given by

$$\Delta G_{\rm tot}(\lambda, n) = \Delta G_0(n) + \Delta G_{\rm b}(x_{\rm b}) + \Delta G_{\rm h}(x_{\rm h}) + + \Delta G_{\rm ss}(x_{\rm ss}, n) + \Delta G_{\rm d}(x_{\rm d}), \qquad (5)$$

where  $\Delta G_0(n) = \sum_{i}^{n} \Delta g_{0,i}$ , is the hairpin free-energy of hybridization according to the NN model and the other terms are the energy contributions of the corresponding elastic terms in Eq.(4)

#### 3.1 Computation of the Equilibrium FDC

Let us consider the case where thermal fluctuations are neglected in the FDC computation. Thus, at a given value of  $\lambda$ , the system is always in the state of minimum energy,  $\Delta G_{eq}(\lambda) = \Delta G_{tot}(\lambda, n^*)$ . To compute the equilibrium free energy of the system, let us first introduce the system partition function, Z. At each  $\lambda$ , this is defined as the sum over all the possible states, i.e., all the possible sequences of n open base pairs, which is

$$Z(\lambda) = \sum_{n=0}^{N} \exp\left(-\frac{\Delta G_{\text{tot}}(\lambda, n)}{k_{\text{B}}T}\right), \qquad (6)$$

where N is the total number of base pairs of the sequence. Finally, by recalling that  $\Delta G = -k_{\rm B}T \ln Z$ , the equilibrium force is given by:

$$f_{\rm eq}(\lambda) \equiv \frac{\partial \Delta G(x_{\rm eq})}{\partial \lambda} = -k_{\rm B}T \frac{\partial \ln Z(\lambda)}{\partial \lambda}.$$
 (7)

Computing Eq.(6) requires solving the transcendental Eq.(4) with respect to f (that can be performed numerically) and then computing Eq.(5) for all  $n \in [0, N-1]$ . For each  $\lambda$ , the value  $n^*$  minimizing the equilibrium freeenergy  $\Delta G_{\rm eq} = \Delta G_{\rm tot}(\lambda, n^*(\lambda))$  gives the most probable number of open basepairs. Eventually, the computation of the equilibrium force in Eq.(7) gives a theoretical prediction for the unzipping curve of a given sequence (Extended Data Fig. 5).

#### 3.2 Equilibrium Free Energy

The total free energy in Eq.(5) is the sum of two main contributions: the hybridization energy  $\Delta G_0(n)$ , which linearly depends on the number of hybridized NNBPs n, and the stretching energy  $\Delta G_{\rm el}(\lambda, n) = \Delta G_{\rm b}(x_{\rm b}) + \Delta G_{\rm h}(x_{\rm h}) + \Delta G_{\rm ss}(x_{\rm ss}, n) + \Delta G_{\rm d}(x_{\rm d})$  depending on both n and  $\lambda$ . For a given  $\lambda$ , the equilibrium configuration of the system is that with minimum  $\Delta G_{\rm el}(\lambda, n^*)$  and maximum  $\Delta G_0(n^*)$  among all possible values of n. Notice that for a hairpin of N bp, n ranges from 0 (native state) to N-1 NNBPs (totally unfolded), which gives N-1 possible system configurations for each value of  $\lambda$ .

Let us suppose that the system starts at equilibrium, with  $n_1$  open bp. Upon increasing  $\lambda$ , the elastic term in Eq.(5) also increases. The number of open bp,  $n_1$ , remains constant until a value of  $n = n_2 > n_1$  is found so that  $\Delta G_{\text{tot}}(\lambda, n_1) \equiv \Delta G_{\text{tot}}(\lambda, n_2)$  (Extended Data Fig. 4A, top): even though the total energy of these two states is the same, the energetic internal balance is different (Extended Data Fig. 4A, bottom). The system minimizes the elastic free energy and switches to state  $n_2$  by releasing  $\Delta n = n_2 - n_1$  bp. Notice that, despite opening  $\Delta n$  bp increases the system's energy, the released ssDNA causes the elastic contribution to decrease. In general,  $\Delta G_{\rm el} \gg \Delta G_0$  so the global balance of the state  $n_2$  is lower than the one of  $n_1$ . Therefore, the equilibrium free energy of hybridization,  $\Delta G_0(n^*)$ , is a step function increasing with  $\lambda$  (Extended Data Fig. 4B) with each discontinuity corresponding to a rip along the equilibrium FDC.

#### 4 Fit of the NNBP parameters

The *T*-dependent NNBP entropies and enthalpies permit us to derive the heat capacity changes  $\Delta c_{p,i}$  for each motif from the relations,

$$\Delta s_i = \Delta s_{m,i} + \Delta c_{p,i} \log(T/T_{m,i}) \tag{8a}$$

$$\Delta h_i = \Delta h_{m,i} + \Delta c_{p,i} (T - T_{m,i}), \qquad (8b)$$

where  $T_{m,i}$  is the melting temperature of motif *i*, and  $\Delta s_{m,i}$  and  $\Delta h_{m,i}$  are the entropy and enthalpy at  $T = T_{m,i}$ , respectively. The extraction of the NNBP thermodynamics parameters  $(\Delta c_{p,i}, \Delta s_i, \Delta h_i, T_{m,i})$  has to be carried out carefully as the results are susceptible to experimental errors and parameters initialization. In particular,  $\Delta s_{m,i}$ ,  $\Delta h_{m,i}$ , and  $T_{r,i}$  strongly depend on their initialization values when directly fitted from Eqs.(8) as an error in  $\Delta s_{m,i}$  $(\Delta h_{m,i})$  get compensated by  $T_{m,i}$  and vice versa.

To derive the  $\Delta c_{p,i}$ , we fit the NNBP entropies to the equation  $\Delta s_i(T) = A_i + \Delta c_{p,i} \log(T)$ , being  $A_i = \Delta s_{m,i} - \Delta c_{p,i} \log(T_{m,i})$ . Notice that we derive  $\Delta c_{p,i}$  from the NNBP entropies as they are obtained from the experimental data, in contrast to enthalpies that are computed from the free energies. Given  $\Delta c_{p,i}$ , we fit the NNBP free energies,  $\Delta g_i(T)$ , to the equation

$$\Delta g_i(T) = \Delta h_i(T) - T\Delta s_i(T) =$$

$$= \Delta h_{m,i} + \Delta c_{p,i}(T - T_{m,i}) - T\left(\Delta s_{m,i} + \Delta c_{p,i}\log\left(\frac{T}{T_{m,i}}\right)\right) \quad (9)$$

$$= B_i + \Delta c_{p,i}T - T\left(A_i + \Delta c_{p,i}\log\left(T\right)\right).$$

obtained by combining Eqs.(8) (blue dashed lines in Fig. 3B, main text). Notice that  $B_i = \Delta h_{m,i} - \Delta c_{p,i} T_{m,i}$ . By definition,  $T_{m,i}$  is the high temperature value where  $\Delta g_i(T_{m,i}) = 0$ . Finally, a new fit to Eqs.(8a) and (8b) by using the previously derived values of  $\Delta c_{p,i}$  and  $T_{m,i}$  (red and blue dashed lines in Fig. 2D, main text), gives  $\Delta s_{m,i}$  and  $\Delta h_{m,i}$ . The results are shown in Fig. 4 and Table 1 of the main text.





Fig. 1: Computation of the FEC from the experimental FDC. The force versus hairpin extension,  $x_H$ , (black line) is computed by subtracting to the trap position,  $\lambda$ , (grey line) the elastic contribution of the optically trapped bead,  $x_b$ , and DNA handles,  $2x_h$ , (green dashed line). To measure the *T*-dependent ssDNA elasticity, we fit the FEC after the last rip to the WLC model (orange dashed line). Notice that the average unzipping force,  $f_m$ , (red line) remains constant upon computing the FEC. Data are shown at  $T = 25^{\circ}$ C.



Fig. 2: T-dependence of the measured force versus hairpin extension. At each T, average unzipping forces are shown by dashed lines. The extension change over the studied temperature range is ~ 500nm (grey vertical lines).



Fig. 3: Clausius-Clapeyron equation applied over the full FDC. (A) T-dependence of the entropy change per base,  $\Delta s_{ss}$ . It accounts for the work to stretch the ssDNA and orient the folded molecule along the direction of the external force, from f = 0pN to  $f_m(T)$  (integral term in Eq.(2), main text). The results are reported in Extended Data Table 1. A fit to data according to  $\Delta s_{ss}(T) = \Delta s_{ss,0} + \Delta c_p^{ss} \log(T/T_m)$  (orange dashed line), gives the ssDNA heat capacity change per base at zero force,  $\Delta c_p^{ss} = -11.2 \pm 0.2$  cal mol<sup>-1</sup>K<sup>-1</sup>. (B) T-dependence of the total entropy change,  $\Delta S_0(T)$ , upon unzipping the 3.6kbp DNA hairpin measured using the Clausius-Clapeyron equation (see Eq.(2), main text).



Fig. 4: Derivation of the theoretical FDC. (A) Schematics of the stretching and hybridization energy contributions. Upon unzipping, the molecule has  $n_1$  open bp before the force rip (left) and  $n_2 > n_1$  open bp after the rip (right). At the force rip (black dots), the total free energy of the system is the same in both states, and the system changes from the highest free energy branch  $(n_1, pink)$  to the lowest energy branch  $(n_2, pink)$ . (B) The free energy of hybridization upon unzipping the hairpin is a monotonically increasing step function, with each discontinuity corresponding to a rip along the equilibrium FDC.



Position [nm]

Fig. 5: T-dependent theoretical FDC predictions. Experimental FDCs (dark-colored solid lines and points) and theoretical predictions (light-colored lines) obtained with the free energy parameters derived at each T. Error bars indicate the variability of the experimental FDCs.



Fig. 6: Prediction of the DNA duplexes melting temperatures. (A) Comparison of the melting temperatures for the set of 92 DNA oligos studied by Owczarzy *et al.* in Ref.<sup>55</sup> (horizontal axis) and the values predicted with the unzipping energy parameters (vertical axis). Perfect agreement between the two data sets would imply all points falling on the dashed grey line x = y. Predictions obtained with Eq.(10) of Sec. 6, Methods (blue squares) show a systematic discrepancy with respect to the experimental values (dashed red line). By accounting for the entropic correction (Eq.(11) of Sec. 6, Methods), predictions agree with the experimental measurements within errors (orange triangles). Results are reported in Extended Data Table 5. (B) Derivation of the entropic correction,  $\delta \Delta s$ . To do this, we subtracted the inverse of the measured,  $T_m^{Bulk}$ , and predicted,  $T_m^{Unz}$ , melting temperatures (blue squares). This equals the difference between the inverse of Eq.(11) and Eq.(10) (see Eq.(12) in Sec. 6, Methods). A linear fit to data (dashed red line) yields  $\delta \Delta s =$  $6(1) \operatorname{cal mol}^{-1} \mathrm{K}^{-1} \sim 4R \log 2$ , where  $R = 1.987 \operatorname{cal mol}^{-1} \mathrm{K}^{-1}$  is the ideal gas constant. The orange triangles show the theoretical correction to  $T_m$  per DNA duplex predicted by assuming  $\delta \Delta s \equiv 4R \log 2$ .



Fig. 7: Prediction of DNA cold denaturation temperatures. Histograms of the melting (red) cold denaturation (blue) temperatures predicted using the ten NNBP thermodynamics parameters (Table 1, main) for all possible DNA sequences of 3, 6, and 8 bp ending with a GAAA tetraloop.

# **Extended Data: Tables**

Table 1: T-dependence of the DNA FDCs

$T [^{\circ}C]$	T[K]	$\mathbf{f}_{\mathrm{m}}$ [pN]	$l_p \ [nm]$	$d_b \ [nm]$	$\Delta \mathrm{s_{ss}} \; [\mathrm{cal} \; \mathrm{mol}^{-1} \mathrm{K}^{-1}]$
7	280	19.72 (3)	0.74(7)	0.647(3)	3.33(2)
10	283	19.08(4)	0.68(2)	0.631(9)	3.18(2)
13	286	18.71 (4)	0.73(3)	0.662(1)	3.07(2)
16	289	18.26 (7)	0.67(2)	0.672(1)	2.95 (2)
19	292	17.87 (4)	0.78(2)	0.655(1)	2.84(2)
22	295	17.12(2)	0.79(3)	0.657(1)	2.67(2)
25	298	16.75(3)	0.77(2)	0.647(1)	2.58(2)
30	303	15.86(2)	0.75(2)	0.665(1)	2.39(2)
35	308	14.96(3)	0.88(2)	0.639(1)	2.21(1)
42	315	14.06(4)	0.88(4)	0.641(1)	2.02(1)

FDC average unzipping force,  $f_{\rm m}$ , persistence length,  $l_p$ , interphosphate distance,  $d_b$ , and ssDNA stretching entropy per base,  $\Delta s_{ss}$  in the studied temperature range (in Celsius and Kelvin degrees). The errors (in brackets) refer to the last digit. The error in temperature is  $\pm 1^{\circ}$ C (K).

Temperature $\pm 1$ [K]	280	283	285	288	291	295	298	303	308	315
$\mathbf{AA}/\mathbf{TT}$	-12.4 (5)	-12.6(4)	-15.7 (7)	-12.8 (3)	-15.0(5)	-16.0(5)	-16.3(6)	-16.9(5)	-16.0(3)	-18.3(4)
AC/TG	-15.4(2)	-15.6(1)	-16.4(2)	-16.4(1)	-16.9(2)	-17.5 (1)	-17.7 (1)	-18.3(1)	-18.6(1)	-19.6(2)
AG/TC	-12.0 (3)	-13.1(2)	-11.8 (4)	-14.3(2)	-13.3(4)	-13.8(4)	-14.5(3)	-15.0(4)	-16.2(3)	-16.6(3)
$\mathbf{AT}/\mathbf{TA}$	-15.1(1)	-15.3(1)	-16.1(1)	-15.6(1)	-16.4(1)	-17.1 (2)	-16.7(1)	-17.3 (1)	-17.3(1)	-17.9(1)
CA/GT	-12.8 (2)	-13.7(2)	-13.8(2)	-14.9(2)	-15.1(2)	-16.6(4)	-16.0(2)	-17.3 (3)	-19.4(7)	-18.4(2)
CC/GG	-12.4 (3)	-13.3(3)	-12.2(4)	-14.6(2)	-13.5(4)	-13.7(4)	-14.8(3)	-15.2(3)	-16.0(3)	-16.4(3)
CG/GC	-22.2(3)	-22.4(4)	-21.4(5)	-23.8(2)	-22.8 (4)	-22.6(4)	-23.7(3)	-24.1(3)	-24.2(4)	-24.6(3)
GA/CT	-12.4 (4)	-13.4(3)	-12.4(5)	-14.6(3)	-14.1 (4)	-14.5(4)	-14.7(5)	-15.2(6)	-16.6(4)	-17.3(3)
GC/CG	-20.8 (3)	-21.3(3)	-19.7(5)	-22.8 (2)	-21.5(4)	-21.8 (3)	-22.7(3)	-23.5(2)	-24.4(3)	-23.7 (3)
$\mathbf{TA}/\mathbf{AT}$	-16.1(1)	-16.1(1)	-17.1(2)	-16.2(1)	-16.9(1)	-17.2(1)	-17.5(2)	-17.7(2)	-16.7(3)	-18.2(1)
The 10 DNA entronies	paniseam	from unzin	ning a 3.6l	thn hairnir	in the ter	nnerature	ran <i>o</i> e [980	315] K (s	ad text) T	ne entrony

The 10 DNA entropies measured from unzipping a 3.6kbp hairpin in the temperature range [280, 315] K (see text). The entropy of the last two motifs (GC/CG and TA/AT) has been computed by applying the circular symmetry relations. The error (in brackets) refers to the last digit.

**Table 2**: NNBP  $\Delta s_{0,i}$  [cal mol<sup>-1</sup>K<sup>-1</sup>] at different temperatures.

## SUPPLEMENTARY INFORMATION

Temperature $\pm 1$ [K]	280	283	285	288	291	295	298	303	308	315
AA/TT	-1.57(2)	-1.49(2)	-1.55(2)	-1.47(2)	-1.43(1)	-1.43(1)	-1.30(1)	-1.27(1)	-1.19(1)	-1.15(1)
AC/TG	-1.52(2)	-1.38(1)	-1.53(2)	-1.35(1)	-1.47(2)	-1.34(1)	-1.43(1)	-1.36(1)	-1.17(1)	-1.21(1)
AG/TC	-1.73(2)	-1.66(2)	-1.49(2)	-1.60(2)	-1.54(2)	-1.42(1)	-1.41(1)	-1.25(1)	-1.35(1)	-1.25(1)
AT/TA	-1.37(1)	-1.43(1)	-1.27(1)	-1.32(1)	-1.25(1)	-1.28(1)	-1.17(1)	-1.00(1)	-1.19(1)	-1.00(1)
CA/GT	-1.86(2)	-1.92(2)	-1.82(2)	-1.88(2)	-1.82(2)	-1.91(2)	-1.65(2)	-1.68(2)	-1.70(2)	-1.52(2)
CC/GG	-2.03(2)	-1.86(2)	-2.04(2)	-1.91(2)	-2.00(2)	-1.88(2)	-1.91(2)	-1.86(2)	-1.56(2)	-1.70(2)
CG/GC	-2.51(3)	-2.52(3)	-2.39(2)	-2.46(3)	-2.35(2)	-2.24(2)	-2.43(2)	-2.30(2)	-2.03(2)	-1.93(2)
GA/CT	-1.52(2)	-1.59(2)	-1.55(2)	-1.48(2)	-1.49(2)	-1.42(1)	-1.52(2)	-1.46(2)	-1.26(1)	-1.20(1)
GC/CG	-2.79(3)	-2.83(3)	-2.51(3)	-2.78(3)	-2.55(3)	-2.53(3)	-2.49(3)	-2.36(2)	-2.34(2)	-2.11(2)
$\mathbf{TA}/\mathbf{AT}$	-1.31(1)	-1.19(1)	-1.10(1)	-1.11(1)	-1.10(1)	-1.00(1)	-1.00(1)	-0.74(1)	-0.97(1)	-0.88(1)
The 10 DNA free-ene entropy of the last two (in bundloth) meters to d	rgies meas motifs (GC	ured from 2/CG and '	unzipping TA/AT) hi	a 3.6kbp ] as been cor	hairpin in mputed by	the tempe the applyi	rature ran ng circula	ge [280, 31 r symmetry	[5] K (see y relations.	text). The The error

15

Temperature $\pm 1$ [K]	280	283	285	288	291	295	298	303	308	315
AA/TT	-5.05(15)	-5.04(12)	-6.04(21)	-5.16 (08)	-5.83 (16)	-6.17 (15)	-6.16 (17)	-6.39(15)	-6.13(09)	-6.93 (14)
AC/TG	-5.84(05)	-5.80(03)	-6.21(05)	-6.10(04)	-6.41(05)	-6.51(04)	-6.71(05)	-6.91(04)	-6.90(03)	-7.38 (06)
AG/TC	-5.09(10)	-5.36(07)	-4.86(13)	-5.74(06)	-5.41(12)	-5.51(12)	-5.73(10)	-5.80(12)	-6.34(09)	-6.49(09)
AT/TA	-5.60(03)	-5.76(03)	-5.88(04)	-5.82(03)	-6.03(04)	-6.32(05)	-6.15(03)	-6.26(04)	-6.54(04)	-6.65(03)
CA/GT	-5.45(05)	-5.81(05)	-5.76(06)	-6.19(06)	-6.22(07)	-6.81(11)	-6.41(05)	-6.93(09)	-7.67 (20)	-7.31(06)
CC/GG	-5.51(10)	-5.62(08)	-5.52(12)	-6.13(06)	-5.93(11)	-5.93(12)	-6.32(08)	-6.46(10)	-6.51(10)	-6.88 (08)
CG/GC	-8.75 (09)	-8.86 (11)	-8.49 (14)	-9.32(07)	-9.01(11)	-8.92(13)	-9.51(09)	(60) 09.6-	-9.50(12)	-9.68(10)
GA/CT	-4.99(11)	-5.37(08)	-5.09(14)	-5.71 (08)	-5.59(11)	-5.70(13)	-5.91(14)	-6.07(17)	-6.38(12)	-6.63(11)
GC/CG	-8.61 (10)	-8.85 (10)	-8.15 (14)	-9.38 (06)	-8.82 (11)	-8.98 (11)	-9.27 (09)	-9.48 (07)	-9.87 (09)	-9.57 (11)
$\mathbf{TA}/\mathbf{AT}$	-5.84(04)	-5.75(03)	-5.99(05)	-5.79(03)	-6.04(04)	-6.07(04)	-6.21(05)	-6.11(05)	-6.14(08)	-6.61(05)
The 10 DNA enthal entropy of the last two	pies measu motifs (G	rred from C/CG and	unzipping   TA / AT ) ]	a 3.6kbp ] ras been co	hairpin in monted br	the tempe v the apply	rature ran ing circula	ge [280, 31 r symmetr	5] K (see v relations.	text). The The error

16

### SUPPLEMENTARY INFORMATION

(in brackets) refers to the last digit.

Sequence $(5' \rightarrow 3')$	T <sup>Exp</sup>	$\rm T_{Bi}^{Unz}$	$\rm T_{Uni}^{Unz}$	$\mathbf{T}^{\mathbf{UO}}$	$\mathrm{T}^{\mathrm{Hug}}$
ATCAATCATA	33.6	40.6	33.1	34.0	32.3
TTGTAGTCAT	36.0	42.4	35.0	36.7	36.2
GAAATGAAAG	34.4	44.6	37.3	34.6	33.8
CCAACTTCTT	40.6	47.9	40.6	40.4	38.9
ATCGTCTGGA	44.9	49.6	42.1	46.2	45.3
AGCGTAAGTC	40.3	48.2	40.9	45.1	43.4
CGATCTGCGA	49.1	54.5	47.3	50.5	50.2
TGGCGAGCAC	55.3	59.5	52.4	56.3	54.7
GATGCGCTCG	53.5	57.7	50.6	54.0	52.9
GGGACCGCCT	57.0	59.8	52.3	58.6	55.3
CGTACACATGC	49.9	53.8	47.3	51.2	50.6
CCATTGCTACC	48.9	55.5	48.7	49.6	48.2
	51.1	54.6	49.3	51.9	50.1
ATACTTACTGATTAG	51.5	56.0	50.6	49.7	50.4
GTACACTGTCTTATA	54.8	56.7	51.4	54.8	54.2
GTATGAGAGACTTTTA	55.4	58.5	53.2	54.8	54.2
	53.7	50.3	55.U	55.1	55.7 57.9
	57.1	09.0	04.0 EC E	50.9	57.2
	61.2	65.2	50.5 60.1	00.9 62.6	09.9 62.6
CTTTCATCCCCAT	62.8	68.0	62.0	63.0	62.6
TGGATGTGTGAACAC	60.4	64.8	59.8	62.3	62.0
ACCCCCCCAATACATC	62.9	68.0	63.8	65.6	64.5
GCAGTGGATGTGAGA	63.3	68.1	63.0	64 6	64.2
GGTCCTTACTTGGTG	60.3	65.2	60.0	62.0	61.7
CGCCTCATGCTCATC	65.8	70.9	65.8	66.5	65.9
AAATAGCCGGGCCGC	70.4	75.8	70.7	72.7	70.9
CCAGCCAGTCTCTCC	66.7	70.9	65.7	67.7	66.7
GACGACAAGACCGCG	68.6	69.7	64.7	69.7	70.3
CAGCCTCGTCGCAGC	72.0	74.8	69.8	73.0	72.7
CTCGCGGTCGAAGCG	70.7	73.7	68.7	72.9	73.6
GCGTCGGTCCGGGCT	74.1	76.5	71.4	77.8	76.1
TATGTATATTTTGTAATCAG	61.2	64.9	60.8	58.6	59.5
TTCAAGTTAAACATTCTATC	61.5	67.6	63.6	60.6	62.6
TGATTCTACCTATGTGATTT	64.4	69.5	65.4	63.7	64.8
GAGATTGTTTCCCTTTCAAA	65.3	72.8	68.8	66.3	67.1
ATGCAATGCTACATATTCGC	68.9	74.7	70.8	69.2	69.6
CCACTATACCATCTATGTAC	64.4	67.5	63.4	63.9	65.2
CCATCATTGTGTCTACCTCA	68.5	73.0	69.0	69.4	69.7
	68.5	74.5	70.5	70.3	70.5
TAGTGGCGATTAGATTCTGC	71.2	74.6	70.6	71.1	70.9
AGUTGUAGTGGATGTGAGAA	73.1	78.0	74.1	74.5	74.0
	73.0	76.0	72.1	70.0	74.0
	72.0	70.0	71.4	73.0	73.9
CCTACTACCCTTCCTCATCC	70.5	73.5	71.4	71.3	71.9
	76.2	74.0	70.7	77.9	74.0
ACCCACCACCCTCATCCCAT	77.3	78.3	74.4	78.7	78.6
AGCAGTCCGCCACACCCTGA	78.5	82.0	78.1	81.6	79.7
CAGCCTCGTTCGCACAGCCC	78.1	82.1	78.3	81.1	80.4
GTGGTGGGCCGTGCGCTCTG	81.0	83.2	79.4	83.6	82.0
GTCCACGCCCGGTGCGACGG	81.1	83.0	79.1	85.4	84.2
GATATAGCAAAATTCTAAGTTAATA	66.1	71.5	68.2	64.2	65.9
ATAACTTTACGTGTGTGACCTATTA	71.8	73.5	70.2	71.2	72.4
GTTCTATACTCTTGAAGTTGATTAC	67.7	72.2	68.9	67.3	69.9
CCCTGCACTTTAACTGAATTGTTTA	72.5	77.7	74.5	73.4	73.5
TAACCATACTGAATACCTTTTGACG	71.3	75.4	72.1	72.2	73.0
TCCACACGGTAGTAAAATTAGGCTT	73.8	78.2	74.9	74.6	75.8
TTCCAAAAGGAGTTATGAGTTGCGA	73.8	79.8	76.6	75.2	76.2
AATATCTCTCATGCGCCAAGCTACA	76.5	81.3	78.1	76.7	77.3
TAGTATATCGCAGCATCATACAGGC	75.0	78.8	75.5	75.5	75.8
TGGATTCTACTCAACCTTAGTCTGG	73.6	77.8	74.5	73.9	75.2
CGGAATCCATGTTACTTCGGCTATC	74.8	79.0	75.7	75.5	76.6

 Table 5: DNA Oligos Melting Temperatures [°C]

Sequence $(5' \rightarrow 3')$	TExp	$\rm T_{Bi}^{Unz}$	$\mathbf{T}_{\mathrm{Uni}}^{\mathrm{Unz}}$	$\mathbf{T}^{\mathbf{UO}}$	$\mathrm{T}^{\mathrm{Hug}}$
CTGGTCTGGATCTGAGAACTTCAGG	75.6	80.1	76.8	76.6	77.3
ACAGCGAATGGACCTACGTGGCCTT	81.0	83.6	80.4	82.7	82.1
AGCAAGTCGAGCAGGGCCTACGTTT	81.5	84.5	81.3	82.8	82.8
GCGAGCGACAGGTTACTTGGCTGAT	80.1	83.1	79.9	81.3	81.7
AAAGGTGTCGCGGAGAGTCGTGCTG	82.4	83.5	80.4	83.0	83.6
ATGGGTGGGAGCCTCGGTAGCAGCC	83.4	87.4	84.1	86.6	84.6
CAGTGGGCTCCTGGGCGTGCTGGTC	83.4	87.6	84.4	87.5	85.7
GCCAACTCCGTCGCCGTTCGTGCGC	84.6	86.9	83.8	88.1	88.0
ACGGGTCCCCGCACCGCACCGCCAG	88.3	90.4	87.2	93.0	90.1
TTATGTATTAAGTTATATAGTAGTAGTAGT	66.6	71.4	68.5	65.8	69.7
ATTGATATCCTTTTCTATTCATCTTTCATT	70.4	78.0	75.2	70.3	71.8
AAAGTACATCAACATAGAGAATTGCATTTC	73.2	78.8	76.1	73.0	74.6
CTTAAGATATGAGAACTTCAACTAATGTGT	71.8	77.1	74.3	71.8	74.2
CTCAACTTGCGGTAAATAAATCGCTTAATC	75.5	80.5	77.8	75.2	77.3
TATTGAGAACAAGTGTCCGATTAGCAGAAA	76.4	81.2	78.5	77.5	78.4
GTCATACGACTGAGTGCAACATTGTTCAAA	76.9	80.8	78.2	78.0	79.3
AACCTGCAACATGGAGTTTTTGTCTCATGC	78.7	83.7	81.1	80.3	80.1
CCGTGCGGTGTGTACGTTTTATTCATCATA	77.6	81.2	78.5	80.0	80.5
GTTCACGTCCGAAAGCTCGAAAAAGGATAC	78.7	82.1	79.4	79.5	81.5
AGTCTGGTCTGGATCTGAGAACTTCAGGCT	80.6	84.7	81.9	82.2	82.5
TCGGAGAAATCACTGAGCTGCCTGAGAAGA	80.9	86.0	83.3	82.5	83.3
CTTCAACGGATCAGGTAGGACTGTGGTGGG	80.1	84.4	81.7	83.3	83.4
ACGCCCACAGGATTAGGCTGGCCCACATTG	84.0	88.9	86.2	87.5	85.5
GTTATTCCGCAGTCCGATGGCAGCAGGCTC	84.1	87.8	85.1	85.9	85.6
TCAGTAGGCGTGACGCAGAGCTGGCGATGG	84.6	88.8	86.1	88.1	88.2
CGCGCCACGTGTGATCTACAGCCGTTCGGC	84.5	88.2	85.6	89.0	89.3
GACCTGACGTGGACCGCTCCTGGGCGTGGT	86.4	89.3	86.6	91.2	89.9
GCCCCTCCACTGGCCGACGGCAGCAGGCTC	87.7	93.3	90.6	93.8	91.5
CGCCGCTGCCGACTGGAGGAGCGCGGGACG	88.6	93.4	90.8	94.8	93.9

**Table 5:** DNA Oligos Melting Temperatures [°C]

Melting temperatures of the 92 DNA duplexes studied by Owczarzy *et al.* in Ref. <sup>55</sup> at a concentration  $c = 2\mu$ M and 1020mM NaCl. The experimental values  $(T^{\text{Exp}})$  are compared with predictions obtained with the unzipping parameters by using Eq.(10)  $(T^{\text{Unz}}_{\text{Bi}})$  and Eq.(11)  $(T^{\text{Unz}}_{\text{Uni}})$  for bimolecular and unimolecular reactions, respectively (see main text and Sec. 6, Methods). Finally,  $T^{\text{UO}}$  and  $T^{\text{Hug}}$  are obtained with the unified oligonucleotide parameters (UO) in Ref. <sup>43</sup> and the Huguet *et al.* (2017) parameters in Ref. <sup>37</sup>. Results are reported with errors:  $T^{\text{Exp}} \pm 1.6^{\circ}$ C,  $T^{\text{Unz}} \pm 1.5^{\circ}$ C,  $T^{\text{UO}} \pm 1.5^{\circ}$ C, and  $T^{\text{Hug}} \pm 1.5^{\circ}$ C. Temperatures are given in Celsius degrees.

## Supplementary References

- [60] Robbins, H., Monro, S.: A stochastic approximation method. Ann. Math. Stat., 400–407 (1951)
- [61] Curry, H.B.: The method of steepest descent for non-linear minimization problems. Q. Appl. Math. 2(3), 258–261 (1944)
- [62] Polyak, B.T.: Introduction to Optimization. Optimization Software vol. 1, p. 32 (1987)
- [63] Qian, N.: On the momentum term in gradient descent learning algorithms. Neural Netw. 12(1), 145–151 (1999)
- [64] Bottou, L.: Large-scale machine learning with stochastic gradient descent. Proceedings of COMPSTAT'2010, 177–186 (2010)
- [65] Bottou, L.: Stochastic gradient descent tricks. Neural networks: Tricks of the trade, 421–436 (2012)
- [66] Nesterov, Y.: A method for unconstrained convex minimization problem with the rate of convergence o (1/k<sup>2</sup>). In: Doklady an Ussr, vol. 269, pp. 543–547 (1983)
- [67] Achlioptas, P.: Stochastic Gradient Descent in Theory and Practice. Stanford (2019)
- [68] Forns, N., de Lorenzo, S., Manosas, M., Hayashi, K., Huguet, J.M., Ritort, F.: Improving signal/noise resolution in single-molecule experiments using molecular constructs with short handles. Biophys. J. 100(7), 1765–1774 (2011)