ELSEVIER

Biophysical Chemistry 101–102 (2002) 513–533

Biophysical Chemistry

www.elsevier.com/locate/bpc

The effect of force on thermodynamics and kinetics of single molecule reactions

Ignacio Tinoco Jr.[a],*, Carlos Bustamante[b]

[a]*Department of Chemistry, University of California, Berkeley, CA 94720-1460, USA*
[b]*Departments of Physics, and Molecular and Cell Biology, Howard Hughes Medical Institute, University of California, Berkeley, CA 94720, USA*

Abstract

The usual variables chemists use to affect a chemical reaction are temperature and pressure. We consider here an additional variable: force, *F*. By attaching a molecule to the tip of a cantilever of an atomic force microscope, or to a bead in a laser light trap, we can control the force on a single molecule. This mechanical force can drive a reaction to completion, or stabilize the reactants. Force changes the thermodynamic stability of a molecule; it can thus increase or decrease the free energy change for the reaction. Force can also speed or slow rates of reactions; it changes the free energy of activation of the reaction.

Keywords: Single molecule reactions; Force; Transition state

---

1. Introduction

This paper is dedicated with respect, and admiration to our friend John Schellman. His papers in the 1950s on the thermodynamic stability of peptides showed how to apply basic physical chemistry and physics to complex biochemical systems. He has been a model for us ever since. Following in his steps, here we analyze the fundamental physical chemistry of single biological macromolecules under force.

Since the classical studies of the Dutch scientist J.D. van der Waals, published in 1873, physical scientists have been aware that many properties of matter could be rationalized by the strength and direction of the forces that molecules exert on each other. The existence of these forces makes it possible to rationalize phenomena as diverse as the elasticities and the melting points of solids, the viscosities and the boiling points of liquids, and the compressibilities of gases, as macroscopic manifestations of the myriads of small interactions between the constituent molecules. These fundamental ideas were soon extended to explain the reactivity and affinity between chemical species by Svante Arrhenius. To formulate the first quantitative description of chemical kinetics, Arrhenius proposed that only molecules with an excess of energy over the average energy of the population could participate in chemical reactions. Accordingly, reactions between molecular species follow pathways that involve the formation of a high-energy species whose accessibility along the reac-

---

*Corresponding author. Tel.: +1 510 642 1440; fax: +1 510 643 6232.
E-mail address: intinoco@lbl.gov (I. Tinoco Jr.).

0301-4622/02/$ - see front matter © 2002 Elsevier Science B.V. All rights reserved.
PII: S0301-4622(02)00177-1

tion coordinate ultimately controls the rate of the reaction. This species, called the transition state was postulated to be an unstable molecular structure, a result of the internal stresses and forces developed in the course of the reaction. The advent of quantum mechanics, in the 1920s, made it possible to recognize the universal electrostatic nature of these intermolecular forces. This synthesis, led to formulation of a general recipe known as the *Hellman–Feynman theorem*, which states that all forces between molecules can be calculated from classical electrostatics [1] once the spatial distributions of the electron densities have been determined from a solution to Schrödinger's equation.

Despite these conceptual advances, until recently, scientists have been able to act on molecules and follow their reactions only if the molecules were present in large numbers. This approach made it difficult to investigate the detailed changes developed in the molecules in the course of the reactions. During the last decade, the advent of novel methods of single molecule manipulation has begun to offer researchers, for the first time, the opportunity to measure directly the forces generated in chemical reactions, and even to exert external forces to alter the extent and fate of these reactions; an area of enquiry that can be called *Mechanochemistry*. This approach is particularly useful in biochemistry, where mechanochemical effects often involve the modification of a large number of highly stereospecific, force-sensitive, weak interactions. Thus, protein and RNA folding, polymer elasticity, protein-induced bending of DNA, stress-induced catalysis of enzymes, or the behavior of molecular motors, are all processes in which stresses and strains develop along the reaction coordinate, and whose extent and fate are, therefore, susceptible to being altered by the application of external force. Beginning with the demonstration of direct mechanical manipulation of single DNA molecules [2,3], scientists have been using single molecule manipulation methods to investigate the folding of proteins and RNA [4–8], the separation of DNA strands [9,10], and the mechanical properties of various motor-like proteins [11–18].

The use of force as a variable under experimental control in chemical reactions gives researchers direct access to new physical parameters of the reaction. The analysis and interpretation of the data require a reformulation of the expressions of traditional thermodynamics and kinetics to arrive at the dependence of both thermodynamic and kinetic variables on the external force. Here we summarize these expressions and discuss their applications to a few special cases.

We do not consider instrumental aspects of single-molecule studies here, although it is clear that they are crucial to the design and analysis of experiments. To apply a force to a single molecule it must be connected to macroscopic objects such as micron-sized beads or an atomic force microscope cantilever. Measurements of piconewton forces and nanometer distances requires sensitive detection methods in instruments free from vibrations and thermal drift. Single-molecule force measurements are at present more difficult to do than bulk measurements, but instrumentation is improving rapidly and we look forward to the day when a force–unfolding curve will be as easy to obtain as a thermal melting curve.

2. Thermodynamics

Chemical thermodynamics is usually applied to systems containing many molecules. We will apply it to a system consisting of a single molecule in a solution of order 10^{19} solvent molecules. In order to apply the concepts of energy, entropy, free energy, length, etc. to the molecule we invoke the ergodic hypothesis—the properties of a single molecule averaged over time is equal to the average properties of many molecules at one time. The single molecule must be given enough time to sample its conformational space during an experiment, and we must repeat each experiment enough times to obtain an average property that remains constant to within the desired precision. Just as in bulk studies the time of a measurement determines which states of a system are accessible, and the variation of a property with time, or with repetition, identifies when a system is at equilibrium. Single-molecule thermodynamics is not very different from classical thermodynamics.

2.1. General thermodynamic relations

We first consider the effect of applying a force on a homogenous solid, for example a glass rod or a spring. The change in energy E is the sum of the reversible heat ($\mathrm{d}q_{\mathrm{rev}} = T\mathrm{d}S$) and reversible work ($\mathrm{d}w_{\mathrm{rev}} = -P\mathrm{d}V + \boldsymbol{F}\cdot\mathrm{d}\boldsymbol{x}$).

$$\mathrm{d}E = \mathrm{d}q_{\mathrm{rev}} + \mathrm{d}w_{\mathrm{rev}} = T\mathrm{d}S - P\mathrm{d}V + \boldsymbol{F}\cdot\mathrm{d}\boldsymbol{x} \tag{1}$$

$\boldsymbol{F}$ is the force and $\mathrm{d}\boldsymbol{x}$ is the differential change in length of the material. Both are vectors, but from now on we will write $F\mathrm{d}x$ as a scalar with $\mathrm{d}x$ the component of $\mathrm{d}\boldsymbol{x}$ along the direction of $\boldsymbol{F}$. The independent extensive variables characterizing the state of the material are entropy S, volume V, and extension or length x. Temperature T, pressure P and force F are the dependent intensive variables corresponding to S, V, x. As S, V, and x are not convenient independent variables, we make the usual Legendre transforms [19] to define new energy functions that have T or P, or both, as independent variables. The thermodynamic functions are thus enthalpy ($H = E + PV$), Helmholtz free energy ($A = E - TS$), and Gibbs free energy ($G = E - TS + PV$).

$$\mathrm{d}H = T\mathrm{d}S + V\mathrm{d}P + F\mathrm{d}x \tag{2}$$

$$\mathrm{d}A = -S\mathrm{d}T - P\mathrm{d}V + F\mathrm{d}x \tag{3}$$

$$\mathrm{d}G = -S\mathrm{d}T + V\mathrm{d}P + F\mathrm{d}x \tag{4}$$

These equations show that the force on the molecule can be determined from the change in energy function with respect to molecular length (end-to-end distance), x.

$$F = \left(\frac{\partial E}{\partial x}\right)_{S,V} = \left(\frac{\partial H}{\partial x}\right)_{S,P} = \left(\frac{\partial A}{\partial x}\right)_{T,V} = \left(\frac{\partial G}{\partial x}\right)_{T,P} \tag{5}$$

It is sometimes more convenient to have the intensive variable force instead of length as an independent variable; we subtract Fx from each energy function to accomplish this. As T, P, and F are the most convenient independent variables to control; we apply a Legendre transform to the Gibbs free energy ($G' = G - Fx$).

$$\mathrm{d}G' = -S\mathrm{d}T + V\mathrm{d}P - x\mathrm{d}F \tag{6}$$

Obviously, G and G' are equal at zero force.

By equating second derivatives of Eq. (4), or of Eq. (6), we obtain useful equations for the effect of T and P on the force and length.

$$\left(\frac{\partial F}{\partial T}\right)_{P,x} = -\left(\frac{\partial S}{\partial x}\right)_{T,P} \quad \text{and} \quad \left(\frac{\partial F}{\partial P}\right)_{T,x} = \left(\frac{\partial V}{\partial x}\right)_{T,P} \tag{7}$$

$$\left(\frac{\partial x}{\partial T}\right)_{P,F} = \left(\frac{\partial S}{\partial F}\right)_{T,P} \quad \text{and} \quad \left(\frac{\partial x}{\partial P}\right)_{T,F} = -\left(\frac{\partial V}{\partial F}\right)_{T,P} \tag{8}$$

We can hold either the length constant or the force constant experimentally, so from the temperature dependences we learn about the effect of length [Eq. (7)] or force [Eq. (8)] on the entropy of the material. The effect of pressure is much more difficult to measure experimentally; it tells us how the volume of the material changes with force or length.

From Eq. (4), we see that at constant T and P

$$\mathrm{d}G = F\mathrm{d}x \tag{9}$$

The change in free energy on applying a force at constant T, P is the reversible work done in stretching or compressing the material from length x_1 to x_2.

$$G(x_2) = G(x_1) + \int_{x_1}^{x_2} F\mathrm{d}x \tag{10}$$

The work is positive for stretching as well as for compression because F and x_2-x_1 are both positive for stretching, and both are negative for compression. Positive work means that the surroundings have done work on the system; the free energy of the material is increased.

If the material undergoes a phase change at a certain force, F_{tr}, a spontaneous change in length will occur. For a spontaneous increase in length during stretching (F is positive) the free energy increases.

$$G(x_3)=G(x_1)+\int_{x_1}^{x_2}F\mathrm{d}x+F_{\mathrm{tr}}\cdot(x_3-x_2) \tag{10a}$$

2.2. Single-molecule thermodynamics

We have been discussing a macroscopic solid object (the system) in the presence of an external force. Now we consider a system at constant T and P in which a force is applied to a single molecule in a solution. The thermodynamic properties now depend on the concentrations of all the molecules in the solution, as well as the force applied to one molecule. The molecule, solvent and solute species are all at constant T and P; only the chosen molecule is under the influence of force. The free energy of the system can be changed by doing work on the molecule, or by changing the number of moles n_i of a species in the solution. For the species present in bulk in the solution the change in free energy for a differential change in number of moles, n_i, is

$$\mathrm{d}G=\sum_i \mu_i \mathrm{d}n_i \tag{11}$$

The sum is over the species i present in bulk in the solution. The chemical potentials, μ_i, are defined by this equation; they are also called partial molar free energies.

$$\mu_i=\left(\frac{\partial G}{\partial n_i}\right)_{T,P,x_i,n_j\neq n_i} \tag{12}$$

Clearly, the chemical potential of each component i depends on the temperature and pressure, and the composition of the system—the values of the n_j. For a pure component the chemical potential is just its free energy per mole, but in a solution the chemical potential is the change in free energy of the solution when a mole of component is added (keeping the composition of the solution constant).

The conventional definition of activity a for a solute in a solution with $P=1$ atm is

$$\mu_i(T)=\mu_i^0(T)+RT\ln a_i \tag{13}$$

R is the gas constant (Avogadro's number times Boltzmann's constant k). The activity is the unitless effective concentration; it becomes numerically equal to the concentration of solute (by convention in molarity) as all *solute* concentrations approach zero. The chemical potential in the standard state $\mu_i^0(T)$ is the chemical potential—partial molar free energy—of the solute at activity $=1$ ($\ln a_i=0$) in the chosen solvent, such as water. This means we study a solution that is dilute in all solute species—so activities are equal to concentrations—and extrapolate to unit activity. However, we can define the solvent to include some of the solutes. For example, the solvent could be 200 mM NaCl, pH 7 phosphate buffer, and a solution would then be studied that was dilute in only a nucleic acid or protein species. Now the chemical potential in the standard state $\mu_i^0(T)$ is the chemical potential of the protein or nucleic acid at activity $=$ 1 in the 200 mM buffer solvent.

For the molecule subject to force, its chemical potential also depends on its extension x_i. For a single molecule with a force F applied to it we make a simple extension of Eq. (13). The chemical potential of the molecule depends on its activity α_i, as in Eq. (13), as well as its extension x_i under force.

$$\mu_i(T,x_i)=\mu_i^0(T,F=0)+kT\ln \alpha_i+\int_0^{x_i}F\mathrm{d}x'_i \tag{14}$$

The component of the length in the direction of the force is always zero at zero force, so the lower limit of the integral is zero. For a rigid molecule—one whose length does not change under force—

the work done in stretching it is zero. This is only a limiting case for an infinitely rigid molecule. In the presence of a force the chemical potential depends on the length or extension, x, of the molecule. The activity α_i can vary from 0 to 1 for the single molecule. Thus the chemical potential differs from that in the standard state depending on its activity and its free energy of stretching. The chemical potential in the standard state $\mu_i^0(T, F=0)$ is similar to that of the standard state in Eq. (13). It is the free energy of the molecule in the solvent at zero force; multiplication by Avogadro's number gives the partial molar free energy—chemical potential—for the standard state used in bulk solution.

2.3. Stretching a single molecule

One of the first reactions studied was the effect of force on stretching DNA from a coiled form with an end-to-end distance small compared to its contour length (length measured along the polynucleotide chain) to a conformation with an end-to-end distance equal or greater than its normal contour length [3,20]. The increase in chemical potential—partial molar free energy—on stretching the molecule (reversibly) to end-to-end distance x is just

$$G(x)_{\text{strectching}} = \int_0^x F \mathrm{d}x' \tag{15}$$

We remember that whatever the end-to-end molecular distance is, in the absence of a force its component along F is zero. The free energy of stretching a molecule is thus obtained from the integral of a force–extension curve. For a rigid molecule the molecule goes from zero length along F to its full length with essentially zero applied force. The free energy of stretching is negligible; it is a measure of the loss of rotational entropy, which is small compared to conformational entropies. The entropy and enthalpy of stretching are obtained from temperature derivatives at constant pressure.

$$\begin{aligned} S(x)_{\text{stretching}} &= -\left(\frac{\partial G(x)_{\text{stretching}}}{\partial T}\right)_P \quad H(x)_{\text{stretching}} \\ &= \left(\frac{\partial G(x)_{\text{stretching}}/T}{\partial(1/T)}\right)_P \end{aligned} \tag{16}$$

To calculate the stretching free energy to compare with experiments on DNA or RNA we need a model for the conformation of the polymer chain. For polynucleotides and polypeptides the models used most often include the freely-jointed-chain and the Porod–Kratky wormlike chain (see for example chapters 18 and 19 of [21]). For both double-stranded and single-stranded nucleic acids, the wormlike chain fits the experimental data adequately. A useful interpolation equation for the force–extension curve of a nucleic acid is [22]:

$$F = \left(\frac{kT}{P}\right)\left[\frac{1}{4(1-x/L)^2} + \frac{x}{L} - \frac{1}{4}\right] \tag{17}$$

here P is the persistence length, L is the contour length, and x is the end-to-end distance. The persistence length of a polymer is defined as the average projection of the end-to-end vector of the polymer on the direction of the first subunit of the polymer for an infinite number of subunits [21]. For a rigid rod the persistence length is the contour length; for a freely-jointed-chain the persistence length is one half the length of a statistical segment. For example, the persistence length of double-stranded DNA is approximately 50 nm (depending on the solvent) [22], and the persistence length of single-stranded RNA is approximately 1 nm [8]. Integration of Eq. (17) over extension x gives the free energy of stretching [Eq. (15)] the wormlike chain; the entropy and enthalpy come from Eq. (16).

$$\begin{aligned} G(x)_{\text{stretching}} &= -TS(x)_{\text{stretching}} \\ &= \left(\frac{kT}{P}\right)\left[\frac{L}{4(1-x/L)}\right][3(x/L)^2 \\ &\quad -2(x/L)^3] H(x)_{\text{stretching}} \\ &= 0 \end{aligned} \tag{18}$$

Note that the force, the free energy, and the entropy are all inversely proportional to the persistence length, so a flexible single strand like RNA ($P=1$ nm) requires 50 times the free energy needed to stretch a double-stranded DNA with $P=50$ nm. In this model the increase in free energy on stretching is all due to the loss of entropy. At zero force a polymer with a short persistence length has many more possible conformations than a polymer with a long persistence length, therefore it has much higher entropy. On stretching, the conformational entropy is lost, thus explaining the higher force and greater work needed to stretch. We can use Eq. (17) to calculate the force needed to stretch a polymer at 298 K with contour length $L=1000$ nm to 75% of its length ($x/L=0.75$). For single-stranded RNA (≈ 1700 nucleotides, $P=1$) it is 18.5 pN; for double-stranded DNA it is only 0.37 pN (≈ 2940 nucleotides, $P=50$). The corresponding free energies [Eq. (18)] are 2091 kJ/mol for the RNA and 41.8 kJ/mol for the DNA.

The wormlike chain is a good model for the conformations of a polynucleotide, but as the end-to-end distance approaches the contour length, effects occur other than the reduction of conformational entropy. Increasing force populates conformations with changes in bond angles, and other higher-energy states, causing both entropy and enthalpy to contribute to the free energy of stretching. The wormlike chain model is no longer valid. With increasing force DNA reaches its contour length then undergoes an overstretching transition that causes a 70% increase in length [20]. The double helix is thought to unwind to form a ladder-like structure.

2.4. Transforming one structure into another: $A \rightarrow B$

When a force is applied to a molecule that can undergo a reversible transition to a different conformation, the force will favor the longer species. For the reaction

$$\mathrm{A} \rightarrow \mathrm{B}$$

in the presence of a force, the change in free energy at constant T, P due to changing length at constant force F, is from Eq. (10a)

$$\Delta G(F) = \mu_{\mathrm{B}}(F, x_{\mathrm{B}}) - \mu_{\mathrm{A}}(F, x_{\mathrm{A}}) = F \cdot (x_{\mathrm{B}} - x_{\mathrm{A}}) \tag{19}$$

The difference in free energies of molecules A and B is equal to the reversible work $F \cdot (x_{\mathrm{B}} - x_{\mathrm{A}})$. As A converts to B, work is done by the system; when B reverts to A, the surroundings do the same amount of work on the system. The magnitudes of x_{B} and x_{A} depend on the force for most molecules, but a molecule can be rigid enough that its chemical potential is essentially independent of force.

We want to know the free energy change at zero force to compare with bulk measurements. Therefore, we use the chemical potentials [Eq. (14)] that relate the free energy at any extension x (and therefore force F) to the free energy at $F=0$. At equilibrium the free energy difference is equal to the difference in chemical potentials.

$$F \cdot (x_{\mathrm{B}} - x_{\mathrm{A}}) = \mu_{\mathrm{B}}^{0} - \mu_{\mathrm{A}}^{0} + kT \ln \frac{[\alpha_{\mathrm{B}}(F)]}{[\alpha_{\mathrm{A}}(F)]} + \int_{0}^{x_{\mathrm{B}}} F_{\mathrm{B}} \mathrm{d}x'_{\mathrm{B}} - \int_{0}^{x_{\mathrm{A}}} F_{\mathrm{A}} \mathrm{d}x'_{\mathrm{A}} \tag{20}$$

where $[\alpha_{\mathrm{B}}(F)]$ and $[\alpha_{\mathrm{A}}(F)]$ are the equilibrium activities of A and B at force F. The difference in standard free energy *at zero force* in a chosen solvent at temperature T and pressure $P=1$ atm is thus

$$\begin{aligned} \mu_{\mathrm{B}}^{0} - \mu_{A}^{0} &= F \cdot (x_{\mathrm{B}} - x_{\mathrm{A}}) - kT \ln K_{\mathrm{eq}}(F) - \Delta G(x)_{\mathrm{stretch}} \\ K_{\mathrm{eq}}(F) &= \frac{[\alpha_{\mathrm{B}}(F)]}{[\alpha_{\mathrm{A}}(F)]} \\ \Delta G(x)_{\mathrm{stretch}} &= \int_{0}^{x_{\mathrm{B}}} F_{\mathrm{B}} \mathrm{d}x'_{\mathrm{B}} - \int_{0}^{x_{\mathrm{A}}} F_{\mathrm{A}} \mathrm{d}x'_{\mathrm{A}} \end{aligned} \tag{21}$$

The standard free energy of reaction at zero force, $\mu_{\mathrm{B}}^{0} - \mu_{\mathrm{A}}^{0}$, is thus the reversible work at force F minus the effect of the equilibrium constant at that force, and minus the difference in free energies of stretching product and reactant to that F. We note that Eq. (21) reduces to the correct expression at zero force.

$$\mu_B^0 - \mu_A^0 = -kT \ln K_{eq}(F=0) \tag{22}$$

Eq. (21) relating the standard chemical potentials at zero force to a measured mechanical work at any force is our most important result. We can also derive it from the thermodynamic function [Eq. (6)] $G' = G - Fx$. The equation analogous to Eq. (14) is

$$\mu'_i(T,x_i) = \mu_i^0(T,F=0) + kT \ln \alpha_i - \int_0^F x_i \mathrm{d}F' \tag{23}$$

For a reversible transition at constant T, P, and F, $\Delta G'$ is zero, so the equivalent of Eq. (19) is

$$\Delta G'(F) = \mu'_B(F,x_B) - \mu'_A(F,x_A) = 0 \tag{24}$$

which leads to the equivalent of Eq. (21).

$$\mu_B^0 - \mu_A^0 = -kT \ln K_{eq}(F) + \Delta G'(F)_{\mathrm{stretch}}$$

$$K_{eq}(F) = \frac{[\alpha_B(F)]}{[\alpha_A(F)]}$$

$$\Delta G'(F)_{\mathrm{stretch}} = \int_0^F x_B \mathrm{d}F' - \int_0^F x_A \mathrm{d}F' \tag{25}$$

We note that Eq. (21) and Eq. (25) are identical because of the identity

$$Fx = \int_0^x F \mathrm{d}x' + \int_0^F x \mathrm{d}F' \tag{26}$$

Eq. (21) is directly related to experiment, but Eq. (25) is helpful in calculations when an explicit function of extension vs. force is available, such as the wormlike chain.

The change of K_{eq} with F at constant T, P is obtained by taking the derivative with respect to F of Eq. (21) or Eq. (25)

$$\left(\frac{\partial \ln K_{eq}}{\partial F}\right)_{T,P} = \frac{x_B - x_A}{kT} \tag{27}$$

In Eq. (25), the derivative of $\Delta G'(F)_{\mathrm{stretch}}$ is $(x_B - x_A)$. In Eq. (21), we need to consider the derivatives of $F \cdot (x_B - x_A)$ and $\Delta G(x)_{\mathrm{stretch}}$; the derivative of $(x_B - x_A)$ is cancelled by the derivative of $\Delta G_{\mathrm{stretch}}$ thus giving Eq. (27). The change in length of the reaction $(x_B - x_A)$ will depend on force (unless both molecules are essentially rigid). However, as force approaches zero, $(x_B - x_A)$ approaches zero.

As an example of the effect of force on an equilibrium constant, Fig. 1 shows a calculated $\ln K_{eq}$ vs. F for the unfolding of an RNA hairpin. We choose $\mu_B^0 - \mu_A^0 = 150$ kJ/mol (35.8 kcal/mol) for the standard free energy at zero force. The hairpin is assumed rigid with an end-to-end distance of 2.0 nm; the single strand contains 50 nucleotides with a contour length of 29.5 nm (0.59 nm/nt). We use the wormlike chain Eq. (17), Eq. (18) to calculate the $F \cdot (x_B - x_A)$ and $\Delta G_{\mathrm{stretch}}$ terms in Eq. (21) to obtain $\ln K_{eq}$ vs. F. The logarithm of K_{eq} is approximately linear with force in the region where it can be measured ($\ln K_{eq} \cong 0$ at 17.5 pN), but it is obvious from the graph that a linear extrapolation from that region cannot be used to obtain the standard free energy at zero force. A correction must be made for the change of $(x_B - x_A)$ with force; this is done by using the wormlike chain to estimate the free energy of stretching the RNA to the force where the transition occurs.

Our studies on the unfolding of RNA hairpins [8] are consistent with the results derived in this section. RNA hairpins in solution at room temperature unfold reversibly to single strands; an experimental curve for a tetraloop with 22 bp is shown in Fig. 2a. The reversible work done is in reasonable agreement (after correction for the free energy of stretching) with the free energy change at zero force measured in bulk [Eq. (21)]. At constant force in the transition region the RNA jumped back and forth between hairpin and single strand (Fig. 2b), thus indicating a reversible reaction. Values of the equilibrium constant were measured from the fraction of time the molecule spent in each form; K_{eq} values ranged from 0.02 to 50 as force varied by 1 pN. The logarithm of K_{eq} was linear in force [Eq. (27)], and the value of $(x_B -$

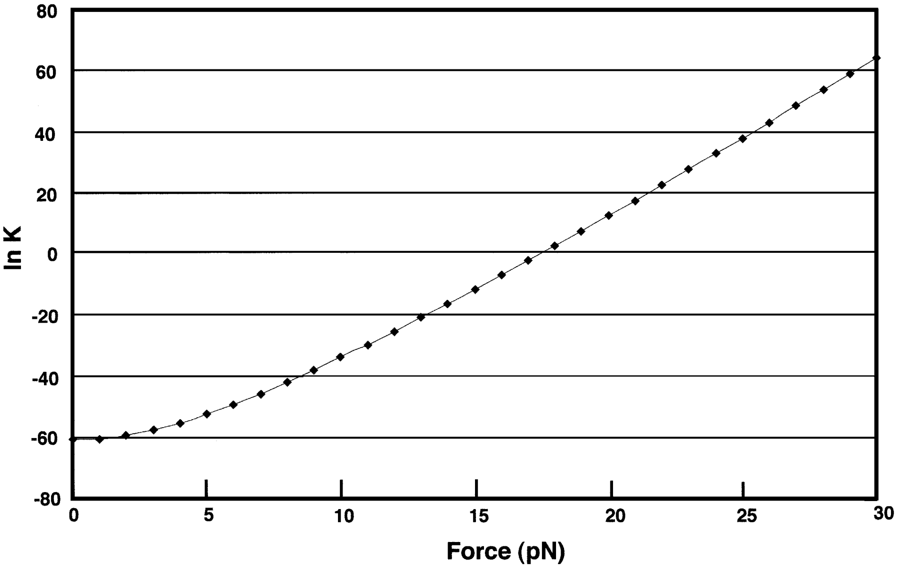

Fig. 1. A plot of the calculated logarithm of the equilibrium constant for the all-or-none unfolding of a 50-nucleotide RNA hairpin with a standard free energy at 25 °C of $\mu_B^0 - \mu_A^0 = 150$ kJ/mol (35.8 kcal/mol). We used Eq. (21) with the wormlike chain of persistence length 1 nm and a contour length of 0.59 nm/nucleotide. Experimentally we can only measure values of K near 1; ln K between −5 and +5.

x_A) obtained was consistent with the change in extension measured for the transition.

The change in $\ln K_{eq}(F)$ with T at constant F, P can be obtained from the temperature derivative of Eq. (21), but it is not particularly useful because all terms in Eq. (21) are temperature dependent. It is more informative—and experimentally easier—to consider the temperature derivative of the reversible work done at constant $K_{eq}(F)$. Any fixed value of $K_{eq}(F)$ can be used, but clearly $K_{eq}(F) = 1$ is the simplest; this choice defines $F_{1/2}$, the force at the midpoint of the transition.

$$F_{1/2}\cdot(x_B - x_A) = \mu_B^0 - \mu_A^0 + \Delta G(x)_{\text{stretch}} \tag{28}$$

From derivatives of Eq. (28) with respect to temperature we obtain the change in entropy and enthalpy of the reaction: $(\partial\Delta G/\partial T)_P = -\Delta S$; $(\partial\Delta G/T/\partial 1/T)_P = \Delta H$. The change of reversible work with temperature at the midpoint of the transition gives the standard entropy change of the reaction at zero force ΔS^0 minus $\Delta S_{\text{stretch}}$.

$$\frac{\partial[F_{1/2}\cdot(x_B - x_A)]}{\partial T} = -\Delta S^0 - \Delta S(x)_{\text{stretch}} \tag{29}$$

The derivative with respect to $1/T$ of reversible work divided by temperature, gives the standard enthalpy change of the reaction at zero force ΔH^0 (and a usually negligible $\Delta H_{\text{stretch}}$).

$$\frac{\partial[F_{1/2}\cdot(x_B - x_A)/T]}{\partial(1/T)} = \Delta H^0 + \Delta H(x)_{\text{stretch}} \tag{30}$$

Eq. (29), Eq. (30) show how to obtain the entropy and enthalpy of the reaction, however, we have not made any measurements of temperature dependence yet.

2.5. The overstretching transition of DNA

When a DNA molecule is subjected to forces of 65 pN or more, it suddenly yields, stretching up to 70% beyond its canonical B-form contour length. This overstretching transition occurs within

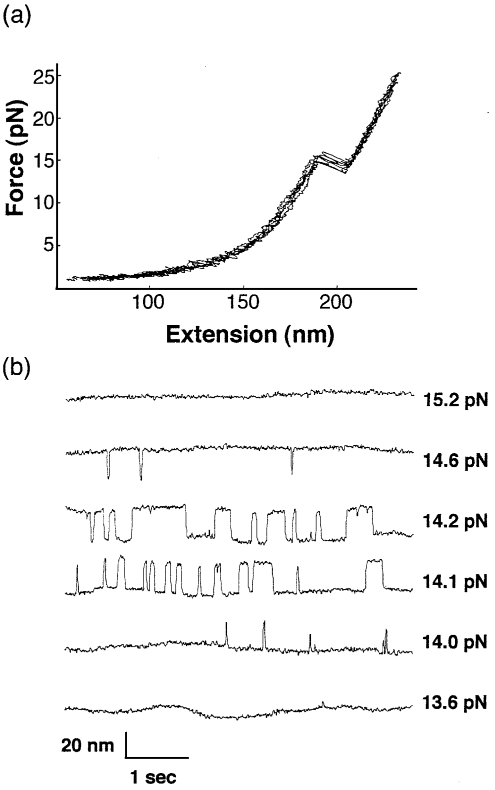

Fig. 2. (a) Pulling (force–extension) curves for the unfolding of a 22 bp RNA hairpin [8]. The stretching and relaxing curves are superimposed; the lack of hysteresis indicates that the unfolding is reversible. The proof of reversibility is shown in (b). When the force is held constant at the transition, the RNA switches back and forth with time from folded hairpin to unfolded single strand. The equilibrium constant for the transition is obtained from the total time spent in each conformation; K is close to 1 at 14.1 and 14.2 pN. The rate constants for the forward and back reactions are obtained from the average times spent in each conformation.

a narrow range of forces, resulting in the flat plateau seen in Fig. 3. We can treat this transition as a cooperative transition in which a number of base pairs of the DNA switch from B-form to stretched-form as a cooperative unit. We use Eq. (21) with the $\Delta G(x)_{\text{stretch}}$ term set to zero.

$$\Delta G^0 = F\Delta_c - kT\ln K$$

$$K = \frac{f}{1-f} \tag{31}$$

here ΔG^0 is the standard free energy of the transition at zero force, Δ_c is the length of the cooperatively stretching unit of the DNA, K is the equilibrium constant, and f is the fraction of base pairs that have undergone the transition. Solving for f we obtain

$$f = \frac{e^{-(\Delta G^0 - F\Delta_c)/kT}}{1 + e^{-(\Delta G^0 - F\Delta_c)/kT}} \tag{32}$$

The breadth of the transition, ΔF, can be written in terms of the slope at the mid point of the transition for which $f = 0.5$.

$$\Delta F = \frac{1}{(\partial f/\partial F)_{f=0.5}} \tag{33}$$

From Eq. (32), we see that $f = 0.5$ when $F = \Delta G^0/\Delta_c$, thus substituting for F in Eq. (33), we obtain the size of the cooperative unit.

$$\Delta_c = \frac{4k_BT}{\Delta F} \tag{34}$$

The degree of cooperativity of the system can then be expressed at any given temperature as the ratio between the size of the cooperative unit and the physical units, Δ_c/δ_0, where δ_0 is the spacing between adjacent base pairs in an overstretched DNA molecule. A ratio of 1 implies that the cooperative unit is the size of the physical unit and is the signature of a non-cooperative process. Using [20] $\Delta F \sim 3$ pN and $\delta_0 = 0.58$ nm, Eq. (34) yields a ratio $\Delta_c/\delta_0 \sim 10$ bp, i.e. about one whole turn of the molecule collectively overstretches during the transition. Little is known about the structure of the molecule in the overstretched state, but various models have been suggested [23].

In the discussion of the reaction $A \rightarrow B$, we assumed an all-or-none reaction. In single-molecule studies, we can easily tell if this assumption is valid because we are not averaging properties over many molecules, as is done in bulk studies. Sometimes the unfolding is obviously not two-state, as in the overstretching of DNA or when pulling apart strands of DNA [10] with thousands of base pairs. However, shorter regions of the DNA can react cooperatively to give phenomena similar to that seen in the RNA hairpins.

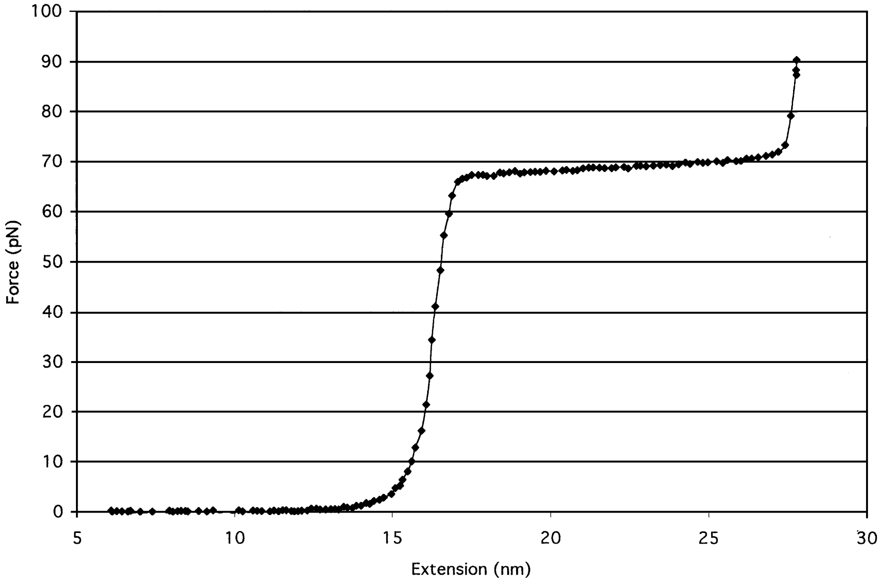

Fig. 3. The overstretching of λ-DNA in 500 mM NaCl, 20 mM Tris, pH 8 at room temperature. The DNA is first unfolded to its B-form contour length. The force then rises rapidly until near 65 pN there occurs a flat plateau where the DNA length increases over a narrow force range. The DNA undergoes a cooperative transition to a stretched, presumably unwound, form.

2.6. Ligand binding to reactants and products

In general the different species in a reaction will bind ligands, such as metal ions M, with different affinities. We can then write an equilibrium reaction as

$$A \rightarrow B + \Delta n M$$

with a net number, Δn, of metal ions released or bound, where Δn can be positive (if A binds M more strongly than B) or negative (if B binds M more strongly than A). The equilibrium constant now depends on the concentration—activity—of the metal ions.

$$K_{\mathrm{eq}}(F) = \frac{[\alpha_{\mathrm{B}}][M]^{\Delta n}}{[\alpha_A]} = K_{\mathrm{obs}}(F)[M]^{\Delta n} \tag{35}$$

We have defined $K_{\mathrm{obs}}(F)$ as the experimentally observable ratio of species A and B. As $K_{\mathrm{eq}}(F)$ is a constant, the magnitude of $K_{\mathrm{obs}}(F)$ will depend on the metal ion concentration (or other specific ligand concentration) at constant force.

$$\frac{\partial \ln K_{\mathrm{obs}}(F)}{\partial \ln [M]} = -\Delta n \tag{36}$$

Eq. (36) is true at any force including zero force; it is traditionally used in bulk experiments to measure the number of ligands involved in a reaction.

2.7. Probabilistic nature of single molecule reactions

Within a collection of many molecules at any temperature, a few molecules will have sufficient energy to react. For example, a stable hairpin that undergoes thermal unfolding at 75 °C can unfold at room temperature, although the probability is very low. Similarly, each time a force vs. extension curve is obtained for a molecule that can change from A to B, the force that causes the reaction can be different. The transition forces will cluster around the mean, with a decreasing probability of occurrence at larger or smaller forces. In bulk measurements we think of the fraction of species in each form, where the fraction in form B, X_{B} is

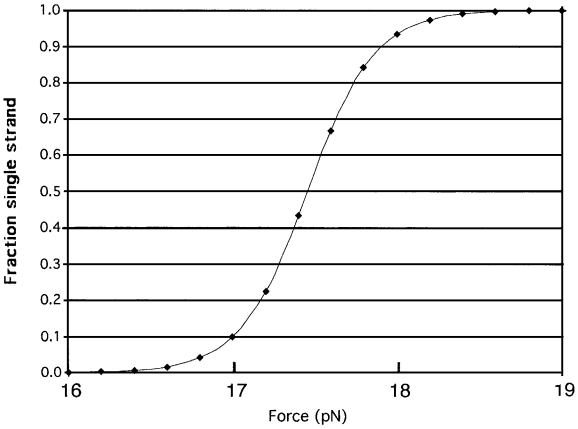

Fig. 4. A plot of the calculated fraction single strand vs. force for the RNA hairpin described in Fig. 1. The curve can be interpreted as giving the fraction of experiments that will result in a transition from a hairpin to a single strand at a force equal to, or less than, that of the curve. Experimentally, the curve can be measured either from the fraction of the time the molecule spends as a single strand when the force is held constant, or from summing a histogram of the number of transitions that occur at each force vs. force.

equal to $K/(1+K)$ with K the equilibrium constant. For the example presented in Fig. 1, the equilibrium constant at 25 °C is approximately 10^{-26} at zero force and becomes 10^{+26} at 30 pN. The fraction single strand thus varies from 10^{-26} to 1. The change from 0 to 1 occurs in a narrow force region near the transition force (17.5 pN in Fig. 1). This region is shown in Fig. 4, where the fraction single strand is plotted vs. force. For single molecule reactions the fraction $X_B(F)$ is interpreted as the probability that the reaction has occurred. In bulk studies of unfolding, a single-strand fraction of 0.10 means that there is 10% single strand and 90% folded form in the solution at that force. In single-molecule experiments, a single-strand fraction of 0.10 means that after many experiments we find that only 10% of molecules have unfolded to single strands at that force, or at lower forces. In our example, 10% of the reactions occur below 17 pN and 93% of the reactions occur below 18 pN. In single-molecule experiments we can obtain a force–melting curve by counting the number of transitions that occur in each force range, such as 17.0–17.1, etc. The sum of the number of transitions that occur below a force plotted vs. that force is the melting curve. This method is unique to single-molecule methods; it cannot be done in bulk. We can also measure a force–melting curve by holding the force constant and measuring the ratio of species: $K=[B]/[A]$. This is analogous to what is done in bulk for thermal melting curves. The two independent methods (directly measuring K at each force, and counting the number of times a transition occurs below a force) provide a useful check on the results [8].

3. Statistical thermodynamics

Thermodynamics is traditionally applied to macroscopic systems, although the formalism for applying it to small systems is well established [24]. However, the natural method to use with single molecules is statistical thermodynamics, which relates average measured properties to molecular properties. The variations—fluctuations—in the properties can also be calculated.

3.1. General relations

All the thermodynamic variables can be obtained from partition functions [25,26]. For a single molecule, the canonical partition function Q at constant T and V is

$$Q=\sum_{\text{states } i} e^{-E_i/kT} \tag{37}$$

where E_i is the energy in state i and k equals Boltzmann's constant. The probability P_i of finding the molecule in state i is

$$P_i=\frac{e^{-E_i/kT}}{Q} \tag{38}$$

The average energy $\langle E\rangle$ is

$$\langle E\rangle=\sum_{\text{states } i} P_iE_i=\frac{1}{Q}\sum_{\text{states } i} E_i e^{-E_i/kT} \tag{39}$$

From the above equation, it follows that the average energy (and all the thermodynamic variables) can be obtained from the logarithm of the partition function Q and its derivatives.

$$\langle E\rangle = kT^2\left(\frac{\partial \ln Q}{\partial T}\right)_V \tag{40}$$

If the molecule has a force applied to it, the average force is $\langle F\rangle$ when the end-to-end distance is x.

$$\langle F\rangle = -kT\left(\frac{\partial \ln Q}{\partial x}\right)_{T,V} \tag{41}$$

The energy of a molecule can be considered (approximately) as a sum of translational energy, rotational energy, and configurational energy. This means that the partition function is a product of partition functions for each type of energy. The well-known results are that the average translational and rotational energies each equal $(3/2)kT$; $\langle E(\text{translational})\rangle = \langle E(\text{rotational})\rangle = (3/2)kT$. Each degree of freedom (x, y, z translational motion and rotation about three perpendicular axes) has on average $(1/2)kT$ of kinetic energy. The configurational energy of the molecule includes vibrations, torsion angle rotations, and all the characteristics that distinguish one molecule from another.

We can also calculate the variations in the average energy $\langle E\rangle$. The root-mean-square deviation from the mean, $\sigma(E)$, is related to the heat capacity at constant volume, $(\partial E/\partial T)_V = C_V$.

$$\sigma(E)^2 = \langle E^2\rangle - \langle E\rangle^2 = kT^2 C_V \tag{42}$$

The simple relation between the average translational and rotational energies and temperature depends on the fact that the spacing between translational energy levels and between rotational energy levels is small compared to kT. For the configurational energy levels (such as vibrations and internal rotations) with spacing small compared to kT, the atoms involved will also have $(3/2)kT$ of thermal kinetic energy. This means that C_V per atom is of order k. The root-mean-square fluctuations in energy for a molecule are thus of order $N^{\frac{1}{2}}kT$, where N is the number of atoms. As the average energy is of order NkT, fluctuation in the energy divided by the energy is

$$\frac{\sigma(E)}{\langle E\rangle} = \frac{1}{N^{1/2}} \tag{43}$$

For a molecule with 2500 atoms (approx. 80 nucleotides), the fluctuations in energy are approximately 2% of the average.

The canonical partition function [Eq. (37)] corresponds to the thermodynamic equation that has S, V, and x (if force is applied) as variables; it corresponds to Eq. (1). It is more convenient for us to have T. P and F as variables instead, and to calculate the probability P_i of the molecule occurring as a partially folded species when T and F are specified for a system at 1 atm.

$$P_i(T,F) = \frac{e^{-(\Delta G_i(T,x_i) - Fx_i)/kT}}{Z} \tag{44}$$

The partition function then becomes (note sum over species not states).

$$Z = \sum_{\text{species } i} e^{-(\Delta G_i(T,x_i) - Fx_i)/kT} \tag{45}$$

The $\Delta G_i(T, x_i)$ are the free energies at constant T, pressure and average extension x_i of species i relative to an arbitrary reference species, such as the single strand or the completely folded species at zero force. Thus, each species is considered as an independent molecule under force F. The partition function [Eq. (45)] is analogous to Eq. (20) relating the equilibrium constant to the reversible work done during a two-state, all-or-none, transition. We can calculate the probability of each species as a function of temperature and force. That is, instead of assuming an all-or-none melting curve, we obtain melting curves for all partially folded species. Average thermodynamic and mechanical properties can also be calculated from these equations. For example, the mean end-to-end distance of all the species in the mixture is

$$\langle x\rangle = \sum_{\text{species } i} P_i x_i \tag{46}$$

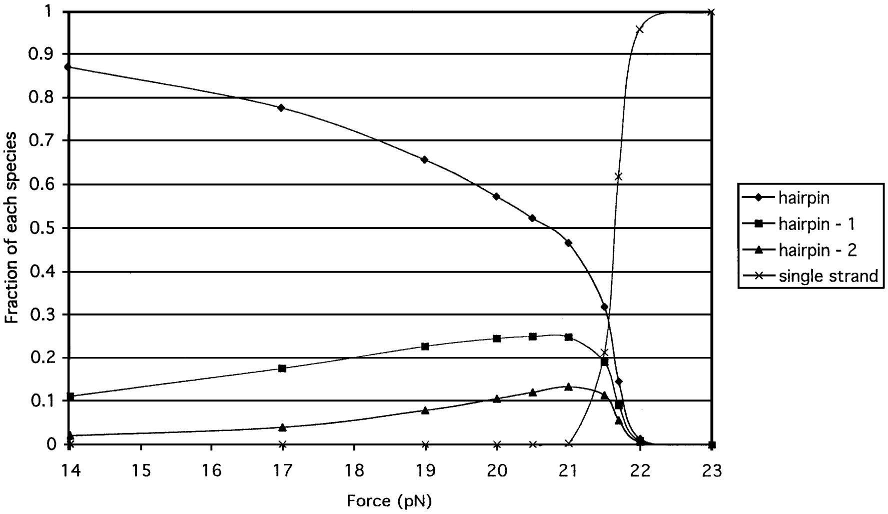

Fig. 5. A plot of the calculated fraction of each species vs. force in the unfolding of a 40 bp RNA hairpin. The hairpin, the hairpin species with 1 bp melted (hairpin-1), the hairpin species with 2 bp melted (hairpin-2), and the single strand are plotted. No other species is present in significant amounts. The single strand and the sum of the three species shown are essentially in a two-state, all-or-none equilibrium. The RNA hairpin has 84 nucleotides with a nearest-neighbor free energy of −10.3 kJ/mol, and a tetraloop free energy of 22.13 kJ/mol; the temperature is 25 °C. The wormlike chain parameters are the same as Fig. 1. We used Eq. (44) to calculate the fraction of each species.

Of course the above equations also apply at zero force, so thermal melting curves as well as force–melting curves can be calculated.

3.2. Force–melting curves for a 40-bp hairpin

For thermal melting curves of hairpins, all-or-none melting is an increasingly poorer approximation as the number of base pairs passes 20. Melting starts at the 5′ and 3′ ends and at the loop closing base pair; melted bubbles also occur in the middle of long stems. There are significant fractions of partially unfolded species in the thermal transition range, so it is essential to use the partition function methods rather than all-or-none equations. In contrast, for force–melting curves of hairpins, the melting is always from the 5′ and 3′ ends where the force is applied. The partition function shows that the only partially unfolded species present are those with two or three base pairs melted from the 5′ and 3′ ends. The single-strand fraction is nearly the same as an all-or-none treatment.

Fig. 5 shows the melting of a 40 bp hairpin with an average nearest-neighbor free energy of −10.30 kJ/mol per bp at 25 °C. The hairpin is closed by a tetraloop with a free energy of 22.13 kJ/mol, so the total free energy relative to the single strand is $39\cdot(-10.30)+22.13=-379.6$ kJ/mol. We used Eq. (44) to calculate the fraction of each species as a function of force. The only species that occur in visible amounts are hairpin, single strand, and the hairpin with one or two base pairs melted. At 25 °C the base pairs are very stable and will only melt when a force is directly applied. As the transition region is approached the first base pair melts, then the second one melts. Once the melting force is reached all the base pairs break and only the single strand is left. This pattern is nearly the same for a 30-bp hairpin, or for a much larger hairpin. We are calculating equilibrium fractions of each species, so the results mean that at the melting force the base pairs will keep melting until the single strand is obtained. We chose all the base pairs to be identical. If there are regions of different stability in the stem, intermediate species can occur. Of course, more complicated structures with bulges, internal loops and tertiary interactions will have many intermediate species. However, a general conclusion is that force–melting curves can be better approximated by two-state, all-or-none, transitions than thermal melting curves. Unfolding molecules with force is simpler to understand and interpret than thermal, or denaturant unfolding, because the force is applied locally rather than globally and randomly.

Force unfolding of nucleic acids or proteins is probably more relevant to what happens biogically than is thermal unfolding. When a messenger RNA is being translated its folded structure must be unfolded; it has effective forces acting on each end as it is translocated through the ribosome. Unfolding caused by helicases is sometimes described as a snow plow wedging apart the RNA strands; a very local force is applied. Clearly, all molecular interactions and reactions involve local forces plus random thermal effects. Single-molecule studies allow us to investigate the effect of both.

4. Single molecule kinetic studies

Single-molecule studies are ideal to investigate the dynamic behavior of complex biological processes. In fact, kinetic studies in bulk are often frustrated by the presence of many different intermediates and of multiple species in solution. Moreover, appropriate time resolution in bulk kinetic studies impose stringent requirements on population synchronization, which only lasts the amount of time that it takes the molecular ensemble to dephase and loose coherence. All these problems are avoided in single-molecule studies where the trajectory of an individual molecule can be followed in time and where intermediates appear as longer lasting states along the trajectory. All these considerations are valid regardless of whether the dynamical process is followed using single molecule manipulation or single molecule detection (fluorescence) methods.

Here, we will concentrate on the use of single molecule manipulation methods to follow dynamical processes. This mechanical approach offers one additional advantage: the use of force as a variable under experimental control. As mentioned in the introduction, if strains and stresses develop in the course of chemical reactions, it is possible to modify the rate and even the fate of processes by the application of force. Moreover, the application of force imposes a well-defined reaction coordinate for the process in the direction along which the force is being applied. Thus, one can follow precisely in time the extent of the reaction, the sensitivity of its rate to the applied force and, from the latter, the location of the transition state along this coordinate.

5. The kinetic regime

Just as is the case in bulk studies, kinetic effects are seen whenever the system under consideration is away from equilibrium. In pulling experiments, these conditions occur whenever the rate at which the force is being applied to the molecular system is faster than its slowest relaxation time. For example, the mechanical extension of a DNA molecule will display kinetic effects if the rate at which the molecule is extended is large compared to the rate at which the molecule can sample its accessible configurations for a given end-to-end distance constraint.

What are the factors that define this kinetic regime? In particular, what pulling rates are required to visit this regime and how do these rates depend on the nature of the molecular system, and on the instrumental characteristics (such as the stiffness of the force transducer used)? Here, it is useful to define the *loading rate*, or the rate at which the force is applied to the molecular system. First introduced by Evans et al. [27,28] the force loading rate, ρ (force/s), is equal to the product of the pulling speed, v (distance/s), and the stiffness (force/distance) of the transducer κ_{trans}. Let a molecule of DNA be attached by one of its ends to a fixed position while we pull on this molecule from its other end by means of a linear force transducer, such as a bead in an optical trap or the tip of an AFM cantilever. The mean quadratic spatial fluctuation Δx^2 experienced by the position of the bead or cantilever tip is then given by the equipartition theorem [26]:

$$\Delta x^2 = \frac{kT}{\kappa_{\text{net}}} = \frac{kT}{\kappa_{\text{trans}} + \kappa_{\text{mol}}(F)} \tag{47}$$

where, the net stiffness of the system is given by the sum of the stiffness of the transducer and the molecule (κ_{mol}). Notice that the latter has been written explicitly as a function of the force since, in general, the elastic response of the molecule will not be linear. Nonetheless, the equipartition result is still applicable to this composite system, since its behavior can be approximated as linear for small fluctuations away from the position of mechanical equilibrium between the transducer and the molecule. The root-mean-square fluctuation in the force is then:

$$\Delta F_{\text{RMS}} = \sqrt{[\kappa_{\text{trans}} + \kappa_{\text{mol}}(F)]kT} \tag{48}$$

We expect that at any given force, the system will remain in a quasi-static equilibrium as long as the force applied to the molecule during its slowest relaxation time, τ, is smaller or equal to the root-mean-square fluctuation in force experienced by the molecule, i.e,

$$\rho\tau = \kappa_{\text{trans}} v\tau \leqslant \Delta F_{\text{RMS}} v \leqslant \frac{\sqrt{[\kappa_{\text{trans}} + \kappa_{\text{mol}}(F)]\kappa_{\text{B}}T}}{\kappa_{\text{trans}}\tau} \tag{49}$$

Eq. (49) defines the boundary between the quasi-static and the kinetic regimes. It states that kinetics become important depending on the pulling rate, v, the slowest relaxation time of the molecular system, τ, and the force constants, κ of the transducer and molecule. If at any given force, the molecule is stressed at rates larger than those given in Eq. (49), the molecule moves from the equilibrium to the kinetic regime. The relaxation time τ is the time necessary for the molecule to return to equilibrium after a newly imposed end-to-end extension.

A measure of the relaxation rate of a polymer whose ends are subjected to these boundary conditions is provided by the Rouse-Zimm model in which the molecule behaves as a non-self avoiding chain made up of masses connected by linear springs which are coupled hydrodynamically through the solvent (discussed in [29]). Quake et al. [30] have recently investigated the dynamics of a single DNA molecule attached by its two ends to polystyrene beads and showed its behavior to be in excellent agreement with the predictions of the Rouse-Zimm model for the first few modes. The relaxation time associated with the νth ($\nu = 1, 2,\ldots$) dynamic mode in the Rouse-Zimm model for an unattached molecule is [31]:

$$\tau_\nu = 2\sqrt{\frac{6}{\pi}}\frac{(LP)^{3/2}\eta_s}{kT\nu^{3/2}} \tag{50}$$

where P is, as before, the persistence length of the molecule, L its contour length, and η_s the viscosity of the solvent. Eq. (50) shows that higher modes, characterized by larger ν-values, attenuate more rapidly than lower modes. Attachment speeds up the relaxation times associated with the various modes of the polymer [30], but for the present purpose we can use the values predicted by Eq. (50) to obtain an upper-bound estimate of the maximum speed at which we can pull a polymer at equilibrium.

Notice that the relaxation time of the polymer increases as the 3/2 power of its length. For example, the slowest relaxation time, ($\nu = 1$) for lambda-DNA ($L = 16.4$ μm, $P = 53$ nm), with $\eta_s = 10^{-3}$ N s/m^2, $kT = 4.1 \times 10^{-21}$ J, is $\tau_1 = 0.5$ s, whereas, $\tau_1 = 1.25$ ms for a 300-nm long DNA molecule. Therefore, to insure that the force measurements on a λ-DNA are quasi-static, each new extension should be allowed to equilibrate for times on the order of seconds. In the study of Smith et al. [20], a molecule of λ-DNA was extended in increments of half a micron. Each measurement consisted of the extension step followed by a 2-s wait (for the force signal to settle) and then a 2-s period of averaging the force readings. Thus, the rate of extension and the relative size of the fluctuations insure quasi-static conditions, and it is valid to identify the area under the force–extension curve with the maximum reversible work done on the molecule, w_{max}, which is equal to the change in the Gibbs free energy of the molecule, $\Delta G(x)$.

In the study of Essevaz-Roulet et al. [10] for example, the DNA strand separation was performed continuously and therefore Eq. (49) directly applies. Assuming the 'average' length of the single strand chain $L=10$ μm, and using the persistence length, P, of single strand DNA as 0.7 nm [20], the slowest ($\nu=1$) relaxation time of the chain is [Eq. (50)], $\tau_1=3.8\times10^{-4}$ s. In these experiments, $\kappa_{\text{trans}}=1.7\times10^{-3}$ pN/nm, and $\upsilon=40$ nm/s [28]. Moreover, since the force is always $<$ 20 pN in these experiments, we take [32] $\kappa_{\text{mol,ssDNA}}(F=20\ \text{pN})=20\times10^{-3}$ pN/nm, as an upper-bound value. Eq. (49) can be re-written:

$$\frac{\upsilon\kappa_{\text{trans}}\tau_1}{\sqrt{[\kappa_{\text{trans}}+\kappa_{\text{mol}}(F)]k_BT}}=8.7\times10^{-5}\ll1 \tag{51}$$

Thus, again, these experiments are performed quasi-statically and the results are directly interpretable in thermodynamics terms.

In contrast to the experiments on DNA, force hysteresis has been described during the stretching of protein molecules [4–7,33]. The forces associated with unfolding of proteins in these experiments vary widely, (20–300 pN), depending on the nature of the protein, the rate of pulling, and the stiffness of the force transducer. These experimentally determined forces are far in excess of the equilibrium force predicted previously for the unfolding of globular domains. In particular for titin, this force has been estimated between 3.5 and 5 pN [34]. Unlike the case of the random polymer, unfolding and refolding of a protein globular domain such as those present in titin must involve the formation of re-entrant structures in which the polypeptide chain folds onto itself. The application of an external force acting perpendicular to the bonds do not always lower the free energy of the transition state for unfolding. Similarly, during refolding, the external force will increase the activation energy of the transition states required to reform the folded structure thus slowing down the process. Consequently, both the equilibrium rates of unfolding and refolding are significantly lower than the rate at which the molecule is experimentally stretched or released. As a result, at the experimental rates of approximately 60 pN/s used in these experiments [4,33], the force at which the molecule unfolds is larger and the force at which the molecule refolds is smaller than the equilibrium forces for those two processes. These discrepancies are the origin of the observed hysteresis. To summarize, the observed hysteresis indicates that the rate at which the molecule is stretched or released exceeds the equilibrium rate of domain unfolding and refolding, respectively. Reversible unfolding of these molecules would necessitate rates of pulling several orders of magnitude slower than have been used to date. Such experiments are difficult because of limitations imposed by instrumental drift and other instabilities.

Many single molecule manipulation experiments have been carried out in the kinetic regime, such as the mechanical rupture of single bonds and the unfolding of macromolecules, as described below.

5.1. The effect of force on the rate coefficients of chemical reactions

Our starting point here is an elementary unimolecular reaction.

$$\text{A}\rightarrow\text{B}$$

We follow the convention of transition state theory [35] and assume that for this reaction to proceed, reactant molecule A forms a high energy transition state $\text{A}^{\neq}$ that is in rapid equilibrium with reactant:

$$\text{A}\underset{k_{-1}^{\neq}}{\overset{k_1^{\neq}}{\leftrightarrow}}\text{A}^{\neq}\overset{k_2^{\neq}}{\rightarrow}\text{B}$$

The transition state equilibrium with reactants is characterized by an equilibrium constant, $K^{\neq}$, and thus by the standard free energy to the transition state, $\Delta G^{\neq}$. Thus the initial rate of the reaction can be written as:

$$\begin{aligned}-\frac{\text{dA}}{\text{d}t}&=k_2^{\neq}(\text{A}^{\neq})=k_2^{\neq}K^{\neq}(\text{A})\\&=k_2^{\neq}\text{e}^{-\frac{\Delta G^{\neq}}{k_BT}}(\text{A})=\frac{k_BT}{h}\text{e}^{-\frac{\Delta G^{\neq}}{k_BT}}(\text{A})\end{aligned} \tag{52}$$

where, h is Planck's constant, k_B is Boltzmann's constant, and $\Delta G^{\neq}$ is the free energy difference between reactant and the transition state. The experimental rate constant for the reaction is thus

$$k=\frac{k_BT}{h}\mathrm{e}^{\frac{-\Delta G^{\neq}}{k_BT}} \tag{53}$$

Let us now assume that the formation of the transition state involves the generation of strain in the reacting species A. If this strain can be induced in the molecule by the direct application of external force, then such force will favor the reaction and its conversion into B. The effect of the externally applied force is two-fold. First it introduces a privileged axis that becomes the natural reaction coordinate along which the reaction will proceed under the effect of the force. The new reaction coordinate need not coincide with, and in general will be different from, the one that the reaction follows in the absence of the applied force. However, because of the force, the energy diagram of the molecule can now be described as a function of the position x along this reaction coordinate. In particular, the energy of the transition state can now be explicitly written as $\Delta G^{\neq}(X_{ts})$. Thus, the transition state lies at a free energy $\Delta G^{\neq}$ above the reactants at a specific distance, X_{ts}, along the reaction coordinate. Second, the application of force deforms and tilts the energy profile at every point along the reaction coordinate. As a first approximation we assume that the application of a force F lowers the free energy of the reaction by an amount proportional to the position along the coordinate

$$k_{\mathrm{forward}}=\frac{k_BT}{h}\mathrm{e}^{-\frac{\Delta G^{\neq}_{\mathrm{forward}}(X_{\mathrm{ts,f}})-FX_{\mathrm{ts,f}}}{k_BT}} \tag{54}$$

Notice that this approximation is equivalent to tilting the free energy profile of the reaction with a pivot point at $X=0$, without deforming the free energy landscape of the reaction. A similar expression holds for the effect of the force on the back reaction

$$k_{\mathrm{back}}=\frac{k_BT}{h}\mathrm{e}^{-\frac{\Delta G^{\neq}_{\mathrm{back}}(X_{\mathrm{ts,b}})+FX_{\mathrm{ts,b}}}{k_BT}} \tag{55}$$

The equilibrium constant for the reaction is equal to the ratio of the forward and back rate constants.

$$K=\frac{k_{\mathrm{forward}}}{k_{\mathrm{back}}}$$
$$=\exp\left\{-\frac{[\Delta G^{\neq}_{\mathrm{forward}}(X_{\mathrm{ts,f}})-\Delta G^{\neq}_{\mathrm{back}}(X_{\mathrm{ts,b}})]}{k_BT}\right\}\exp\left\{\frac{F(X_{\mathrm{ts,f}}+X_{\mathrm{ts,b}})}{k_BT}\right\} \tag{56}$$

As the standard free energy of the reaction, $\Delta G^0=\Delta G^{\neq}_{\mathrm{forward}}(X_{\mathrm{ts,f}})-\Delta G^{\neq}_{\mathrm{back}}(X_{\mathrm{ts,b}})$, and the change in length for the reaction is $\Delta X=X_{\mathrm{ts,f}}+X_{\mathrm{ts,b}}$, we obtain

$$K=\mathrm{e}^{-\Delta G^0/k_BT}\mathrm{e}^{+F(\Delta X)/k_BT} \tag{57}$$

consistent with Eq. (21).

The temperature dependence of the rate constant is traditionally given in terms of its logarithm vs. $1/T$. At constant force

$$\frac{\mathrm{d}\ln k}{\mathrm{d}(1/T)}=\frac{-(\Delta H^{\neq}(X_{ts})-FX_{ts}+k_BT)}{k_B} \tag{58}$$

Here, we have used the fact that $\Delta G^{\neq}=\Delta H^{\neq}-T\Delta S^{\neq}$. Nearly always the k_BT term in the numerator is negligible compared to the other terms, so that the slope of $\ln k$ vs. $1/T$ depends on the transition enthalpy, which is approximately equal to the Arrhenius activation energy, and the force applied.

The dependence of each rate constant on the applied external force is:

$$\frac{\mathrm{d}\ln k}{\mathrm{d}F}=\frac{X_{ts}}{k_BT} \tag{59}$$

thus, the response of the reaction rate to the application of force can be used to locate the position of the transition state X_{ts}, along the reaction coordinate. The rates of reactions whose transition states are close to the reacting species along the reaction coordinate, will be less sensitive to the applied force, as the force favors the formation of the products lying farther away along the coordinate. Conversely, rates of reactions for which the transition state lie close to the products, will be very sensitive to the application of the force. In the absence of the force, molecules of reactant undergo a transformation to product due to thermal crossing of the potential energy barrier. In the presence of the force, the deformation and/or tilt of the potential energy profile along the reaction coordinate affects the free energy of the products of the reaction the most, the free energy of the reactants the least and that of the transition state by an amount in between. Thus, if the applied force lowers the free energy of the products below that of the reactants, the force can favor product formation and speed up the reaction even if in the absence of force the reaction is not spontaneous. However, even in the presence of force the crossing of the barrier is still a thermally activated process.

6. Bond rupture

In a typical bond rupture experiment, a molecule (ligand) bound to a mechanical transducer, is pulled away at a constant velocity from its binding partner (receptor) until the weak interaction holding the two together breaks apart. By its design, such mechanically induced dissociation is a far from equilibrium kinetic process in which the final state of the system corresponds to the state of infinite dilution of the ligand and the receptor with no likelihood for the binding pair to re-associate [27]. At a molecular level the simplest ligand–receptor interaction can be visualized in terms of a potential energy diagram with two local minima separated by an energy barrier. When no force is applied, the associated state is the global minimum of the system. The effect of the force is to tilt and deform the energy diagram until the energy of the dissociated state becomes the favored state. The rupture process is, nonetheless, a thermally induced process, i.e. the crossing of the barrier can occur spontaneously at any given force due to thermal fluctuations. If we are willing to wait a long enough time, the bond can break at any arbitrarily low or even zero force. The result is that bond rupture is a stochastic phenomenon: if a given bond is pulled 100 times at a given loading rate and applied force duration, a distribution of break forces will be observed. The bond strength is the force that produces the most frequent failure in these repeated experiments and corresponds to the force at which the distribution has a maximum.

We can describe this repeated breaking experiment as the simple reaction:

$$N_{\text{intact}} \rightarrow N_{\text{broken}} \tag{60}$$

where N is the number of pullings. Then, the fraction of bonds broken between time t and $t+\mathrm{d}t$ is:

$$\frac{\mathrm{d}N_i}{N_i} = -k(t)\mathrm{d}t \tag{61}$$

here, $k(t)$ is the rate coefficient which is written explicitly as a function of time, through the force applied to the bond which is itself a function of time. The time dependence of the force is dictated by the force–extension behavior of the molecular pair. In many cases, however, it is possible to approximate the force as varying linearly with time, i.e.

$$F(t) = \rho t = \kappa v t \tag{62}$$

We can now integrate Eq. (61)

$$\int_{N_i(0)}^{N_i(t)} \frac{\mathrm{d}N_i}{N_i} = -\int_0^t k(t)\mathrm{d}t \tag{63}$$

where the rate coefficient is given by

$$k(t) = k_0 \mathrm{e}^{-\frac{\Delta G^{\neq} - F(t)X_{\text{ts}}}{k_{\text{B}}T}} = k_0 \mathrm{e}^{-\frac{\Delta G^{\neq} - \kappa v t X_{\text{ts}}}{k_{\text{B}}T}} \tag{64}$$

where $\Delta G^{\neq}$ is the transition state free energy of the process, X_{ts} is the distance to the transition state along the reaction coordinate, and where Eq. (62) has been used. Integration of Eq. (64) leads to

$$\frac{N_i(t)}{N_i(0)} = \exp\left\{ -k_0 \frac{k_B T}{\kappa v X_{\text{ts}}} \mathrm{e}^{-\frac{\Delta G^{\neq}}{k_{\text{B}}T}} \left(\mathrm{e}^{\frac{\kappa v X_{\text{ts}}}{k_{\text{B}}T} t} - 1 \right) \right\} \tag{65}$$

This expression gives the fraction of bonds still intact at force $F(t)=\kappa \upsilon t$ or, equivalently, after time t; i.e. the probability that a bond will not have been broken in the interval between 0 and t. Now, the probability of observing a bond rupture in the interval between t and $t+\mathrm{d}t$ is equal to the product of the probability of the bond breaking in the interval between t and $t+\mathrm{d}t$, times the probability that the bond will have not broken in the interval between 0 and t. The probability of bond rupture is thus given by the product of Eq. (64) and Eq. (65) [8]:

$$P_{\mathrm{rupture}}(t)=k_0 \mathrm{e}^{-\frac{\Delta G^{\neq}-\kappa \upsilon X_{\mathrm{ts}} t}{k_{\mathrm{B}}T}} \times \exp\left\{-k_0 \frac{k_{\mathrm{B}}T}{\kappa \upsilon X_{\mathrm{ts}}} \mathrm{e}^{-\frac{\Delta G^{\neq}}{k_{\mathrm{B}}T}} \left(\mathrm{e}^{\frac{\kappa \upsilon X_{\mathrm{ts}}}{k_{\mathrm{B}}T} t}-1\right)\right\} \quad (66)$$

The first term in this expression is a fast increasing function of the force $(=\kappa \upsilon t)$, whereas the second term is an even faster decreasing function of the applied force, since the life-time of an intact bond decreases rapidly with time for a given loading rate. Note also that both terms depend explicitly on the experimental parameters through κ and υ. Thus, increasing the pulling rate or the stiffness of the transducer, will increase the likelihood of bond survival to time t or (for a given loading rate) to force F. Similarly, decreasing the temperature will also increase the likelihood of bond survival.

The rupture force F^* corresponds then to the maximum of the distribution. This rupture force can be obtained by equating the derivative (with respect to F) of Eq. (66) to zero, to yield

$$F^*=\frac{\Delta G^{\neq}}{X_{\mathrm{ts}}}+\frac{k_{\mathrm{B}}T}{X_{\mathrm{ts}}} \ln\left(\frac{\kappa \upsilon X_{\mathrm{ts}}}{k_0 k_{\mathrm{B}}T}\right) \quad (67)$$

giving the rupture force in terms of all the significant parameters of the system. The second term vanishes whenever:

$$\kappa \upsilon X_{\mathrm{ts}}=k_0 k_{\mathrm{B}}T$$

The term on the right hand side is the rate of energy exchange of the system with the thermal bath, whereas the term on the left is the rate of energy delivery into the system by the pulling process. When these two terms equal each other, the rupture force equals the ratio of the activation energy of the barrier to the distance to the transition state along the reaction coordinate. Under any other conditions, the second term in Eq. (67) can be positive or negative. If $\kappa \upsilon X_{\mathrm{ts}} \ll k_0 k_{\mathrm{B}}T$, then the process is largely thermally activated and $F^*<\Delta G^{\neq}/X_{\mathrm{ts}}$. If $\kappa \upsilon X_{\mathrm{ts}} \gg k_0 k_{\mathrm{B}}T$, a significant amount of energy is dissipated during the pulling process and this energy cannot be exchanged sufficiently fast by the system and the bath. Values of the rupture force larger than $\Delta G^{\neq}/X_{\mathrm{ts}}$ reflect this imbalance. In principle a plot of the rupture force of the bond with the logarithm of the loading rate or as a function of temperature, can be used to extract the position of the transition state and the activation energy of the dissociation reaction.

6.1. Folding–unfolding kinetics

Most macromolecular unfolding studies have been done in the kinetic regime, i.e. in conditions in which the rate of pulling is faster than the slowest relaxation time of the molecule. In this regime, successive force–extension curves will yield a distribution of forces whose mean is always higher than the unfolding force at which the molecule spends 50% of the time in the folded and 50% of the time in the unfolded state. These distributions are still governed by Eq. (66). In particular, the fraction $f(F, \rho)$ of molecules still folded at force F or at time t [Eq. (65)], can be written as:

$$f(F,\rho)=\mathrm{e}^{(k_0/b\rho)(\mathrm{e}^{bF}-1)} \quad (68)$$

where k_0 is the unfolding rate coefficient at zero force, and $b=X_{\mathrm{ts}}/k_{\mathrm{B}}T$. This expression shows that it is possible to determine the position of the transition state along the reaction coordinate, if the molecule is pulled at different loading rates ρ. In a recent application of these ideas [8], the P5abc domain of the *T. thermophila* ribozyme which adopts secondary and, in the presence of Mg^{2+}, tertiary interactions, was mechanically unfolded.

The analysis showed that the transition state in this molecule is located only 1.6 nm from the folded state. This result is in direct contrast with the value of 11 nm for the transition state of the same molecule in the absence of Mg^{2+}, where the RNA can only make secondary interactions and where the mechanical unfolding can be carried out at equilibrium. The proximity of the transition state to the folded state means that the presence of tertiary interactions makes the unfolding rate of the molecule significantly less sensitive to the application of external force. Thus, in these conditions, thermal crossing of this early barrier becomes rate limiting for the unfolding of the molecule under force, and mechanical unfolding is observed at higher forces than its equilibrium counterpart (see below).

6.2. Folding–unfolding kinetics at equilibrium

As stated before, with single molecule methods one can follow the actual trajectories of molecules as they move on the potential energy surface of their reactions and thus, one can follow the kinetics of reactions even when these occur at equilibrium. Of course, we are limited by the limits of time resolution and stability of the pulling instrument. Even if the molecule dynamics falls within this window, the rates observed experimentally will be determined by a combination of molecular and instrumental factors. To a first approximation, the true rate constants of the forward and reverse reactions are modified by a multiplicative factor k_M, that depends on the size of the transducer (cantilever, bead, or bendable micropipet), the viscosity of the medium, and the tethering of the molecule to the transducer.

$$k_{\text{forward}} = k_M k_0 e^{-\frac{\Delta G^{\neq}_{\text{forward}} - F X_{\text{ts}}}{k_B T}} \tag{69}$$

with a similar expression holding true for the reverse reaction. The factor k_M is constant for any given experiment and can be determined to obtain the true rate constants.

Once again, this point is best illustrated by the recent observation of the mechanical unfolding–refolding reaction of a mutant of the P5abc ribozyme domain (P5ab) using optical tweezers [8]. In the absence of Mg^{2+}, as the force applied to the molecule approaches the transition force for the unfolding reaction (see Eq. (19) and Eq. (20)), the molecules display bi-stability. This means that in the vicinity of the transition force, the molecule can either exist in the totally folded or the totally unfolded configuration. To better display this molecular behavior, the experiment was performed using a constant feedback control in which the instrument adjusts the distance between the bead distance so as to keep the tension on the molecule invariant (Fig. 2b). Notice, that by varying the pre-set tension on the molecule, it is possible to tilt the equilibrium towards the folded or the unfolded forms. Moreover, because of the form of Eq. (69), the ratio of the rate constants at any give pre-set tension, gives the free energy of the unfolding reaction as a function of the applied force.

7. Conclusion

We have presented an introduction to the thermodynamics, statistical mechanics, and kinetics of physical processes and chemical reactions under the influence of force. Force is a very useful addition to the traditional variables of temperature and pressure in physical chemistry. The ability to apply a force along a chosen direction to a single molecule provides completely new ways to control reactions. The field of mechanochemistry is very young, but it already shows unlimited possibilities.

Acknowledgments

We wish to thank Ms Sophie Dumont, Dr Raven Hanna, Dr Chul-Hyun Kim, Dr Jan Liphardt, and Dr Bibiana Onoa for reading the paper and making helpful comments. Dr Steven Smith kindly provided the data for Fig. 3. Our research is supported by grants from NIH, NSF and DOE; we greatly appreciate their help.

References

[1] M. Daune, Molecular Biophysics, Oxford University Press, Oxford, 1999, p. 16.

[2] S.B. Smith, C. Heller, C. Bustamante, Model and computer simulations of the motion of DNA molecules

during pulse field gel electrophoresis, Biochemistry 30 (1991) 5264–5274.

[3] S.B. Smith, L. Finzi, C. Bustamante, Direct mechanical measurements of the elasticity of single DNA molecules using magnetic beads, Science 258 (1992) 1122–1126.

[4] M.Z. Kellermeyer, S.B. Smith, H.L. Granzier, C. Bustamante, Folding–unfolding transitions in single titin molecules characterized by force-measuring laser tweezers, Science 276 (1997) 1112–1116.

[5] L. Tskhovrebova, J. Trinick, J.A. Sleep, R.M. Simmons, Elasticity and unfolding of single molecules of the giant muscle protein titin, Nature 387 (1997) 308–312.

[6] M. Rief, M. Gautel, F. Oesterhelt, J.M. Fernandez, H.E. Gaub, Reversible unfolding of individual titin immunoglobulin domains by AFM, Science 276 (1997) 1109–1112.

[7] A.F. Oberhauser, P.E. Marszalek, H.P. Erickson, J.M. Fernandez, The molecular elasticity of the extracellular matrix protein tenascin, Nature 393 (1998) 181–185.

[8] J. Liphardt, B. Onoa, S.B. Smith, I. Tinoco Jr., C. Bustamante, Reversible unfolding of single RNA molecules by mechanical force, Science 292 (2001) 733–737.

[9] U. Bockelmann, B. EssevazRoulet, F. Heslot, DNA strand separation studied by single molecule force measurements, Phys. Rev. E 58 (1998) 2386–2394.

[10] B. Essevaz-Roulet, U. Bockelmann, F. Heslot, Mechanical separation of the complementary strands of DNA, Proc. Natl. Acad. Sci. USA 94 (1997) 11935–11940.

[11] M.J. Schnitzer, K. Visscher, S.M. Block, Force production by single kinesin motors, Nature Cell Biol. 2 (2000) 718–723.

[12] J.A. Spudich, How molecular motors work, Nature 372 (1994) 515–518.

[13] K. Svoboda, C.F. Schmidt, B.J. Schnapp, S.M. Block, Direct observation of kinesin stepping by optical trapping interferometry, Nature 365 (1993) 721–727.

[14] H. Yin, M.D. Wang, K. Svoboda, R. Landick, S.M. Block, J. Gelles, Transcription against an applied force, Science 270 (1995) 1653–1657.

[15] M.D. Wang, M.J. Schnitzer, H. Yin, R. Landick, J. Gelles, S.M. Block, Force and velocity measured for single molecules of RNA polymerase, Science 282 (1998) 902–907.

[16] R.J. Davenport, G.J.L. Wuite, R. Landick, Single-molecule study of transcriptional pausing and arrest by *E. coli* RNA polymerase, Science 287 (2000) 2497–2500.

[17] G.J. Wuite, S.B. Smith, M. Young, D. Keller, C. Bustamante, Single-molecule studies of the effect of template tension on T7 DNA polymerase activity, Nature 404 (2000) 103–106.

[18] D.E. Smith, S.J. Tans, S.B. Smith, S. Grimes, D.E. Anderson, C. Bustamante, The bacteriophage ϕ29 portal motor can package DNA against a large internal force, Nature 413 (2001) 748–752.

[19] H.B. Callen, Thermodynamics and an Introduction to Thermostatistics, Second ed., Wiley, New York, 1985, p. 493.

[20] S.B. Smith, Y. Cui, C. Bustamante, Overstretching B-DNA: the elastic response of individual double stranded and single stranded DNA molecules, Science 271 (1996) 795.

[21] C.R. Cantor, P.R. Schimmel, Biophysical Chemistry. Part III. The Behavior of Biological Macromolecules, WH Freeman, San Francisco, 1980.

[22] C. Bustamante, J.F. Marko, E.D. Siggia, S.B. Smith, Entropic elasticity of λ-phage DNA, Science 265 (1994) 1599–1600.

[23] A. Lebrun, R. Lavery, Modelling extreme stretching of DNA, Nucleic Acids Res. 24 (1996) 2260–2267.

[24] T.L. Hill, Thermodynamics of Small Systems, Dover, New York, 1963.

[25] T.L. Hill, Introduction to Statistical Thermodynamics, Addison-Wesley, Reading, MA, 1960.

[26] D. Chandler, Introduction to Modern Statistical Mechanics, Oxford University Press, New York, 1987.

[27] E. Evans, K. Ritchie, R. Merkel, Sensitive force technique to probe molecular adhesion and structural linkages at biological interfaces, Biophys. J. 68 (1995) 2580–2587.

[28] B. Samori, Stretching, tearing, and dissecting single molecules of DNA, Angewandte Chemie—Int. Ed. 37 (1998) 2198–2200.

[29] P.G. De Gennes, Scaling Concepts in Polymer Physics, Cornell U Press, Ithaca, NY, 1979.

[30] S.R. Quake, H. Babcock, S. Chu, The dynamics of partially extended single molecules of DNA, Nature 388 (1997) 151–154.

[31] A.Y. Grosberg, A.R. Khokhlov, Statistical Physics of Macromolecules, AIP Press, Woodbury, NY, 1994.

[32] S.B. Smith, Y. Cui, C. Bustamante, An optical-trap force transducer that operates by direct measurement of light momentum, Methods Enzymol. Biophotonics, vol. 361, in press..

[33] M.S. Kellermayer, S.B. Smith, C. Bustamante, H.L. Granzier, Complete unfolding of the titin molecule under external force, J. Struct. Biol. 121 (1998) 197–205.

[34] H.P. Erickson, Reversible unfolding of fibronectin type III and immunoglobulin domains provides the structural basis for stretch and elasticity of titin and fibronectin, Proc. Natl. Acad. Sci. USA 91 (1994) 10114–10118.

[35] I. Tinoco Jr., K. Sauer, J.C. Wang, J.D. Puglisi, Physical Chemistry: Applications to Biological Sciences, Prentice Hall, Upper Saddle River, NJ, 2001.