The "Sticky Chain": A Kinetic Model for the Deformation of Biological Macromolecules

Ingomar L. Jäger
Department of Metal Physics, University of Leoben and Erich-Schmid-Institut of Materials Science, Austrian Academy of Sciences, A-8700 Leoben, Austria

ABSTRACT The deformation behavior of certain biologic macromolecules is modeled by the "sticky chain," a freely jointed chain with weak bonds between subsequent joints. Straining the chain leads to thermally assisted breaking of the weak bonds, yielding a characteristic shape of the force-elongation curve, usually with a pronounced plateau, but sometimes displaying a pseudo-Hookean behavior over a wide range of deformations. The number of individual links is assumed to be large, so the stochastic time evolution of the individual events can be approximated by a differential equation. The cases of individual and collective bond breaking are treated and formulae given for various measurable quantities. A threshold strain rate is found, below which the deformation force no longer depends on the deformation velocity. The method is applied to experimental results for the deformation of single molecules like titin or DNA and the results agree with the parameters deduced from the same experiments by the original authors using Monte Carlo (MC) calculations. Despite its intrinsic continuous character, the model, therefore, is applicable even for the deformation of macromolecules with only a few discrete unfolding elements, yielding physical quantities from experimental results using simple formulae instead of a host of MC computations.

INTRODUCTION

Macromolecules forming biologic materials (polypeptides, polysaccharides, and others) share a most interesting feature, a pronounced tendency toward forming second- and higher-order structures. Rarely if ever the secondary structure of such molecules is the standard configuration of large, multiply articulated polymer molecules, the random coil, but nearly always a regularly coiled-up or folded configuration. The most important of the possible configurations are the helical and the globular structures, whereas the configuration within a globule may consist of β-sheets. The folded structures are stabilized by weak bonds (hydrogen bridges, hydrophobic interactions, etc.), which, in spite of their individual weakness, are quite effective due to their number, but can be severed rather easily one by one. Folded structures may be coupled together by linker elements without a pronounced secondary structure, like, e.g., in the string-of-pearls structure found in titin (Kellermayer et al., 1998).

Now the fast-growing bulk of mechanical experiments on single molecules (see, e.g., reviews by Strick et al., 2000; Bustamante et al., 2000; Smith, 2000; Carrion-Vazquez et al., 2000; Zlatanova et al., 2000) shows a rich variety of force-elongation diagrams, many of them displaying quite distinctive plateaus—a constant or slightly rising force over a wide range of strain—resembling the transformation plasticity of, e.g., shape memory alloys. Likewise, the "yield" of a stressed macromolecule with a folded configuration can frequently be reversed: when the external stress is removed, the molecule, after a certain time, may return to its native state and the experiment can be repeated with identical results (Smith et al., 1999; Cluzel et al., 1996). This pseudo-reversible behavior may be reversible in the thermodynamic sense (at infinitely low speed of deformation, the "down" force-elongation curve is expected to follow the "up" curve), but, at finite speed, this is no longer necessarily true and a hysteresis develops. It is now generally explained by unfolding and refolding of the molecule (or parts of it, e.g., titin) under test and has been modeled using various methods: Molecular Dynamics (MD, Rohs et al., 1999; Lu and Schulten, 1999; Lu et al., 1998), diverse theories of phase transitions (Zimm and Bragg, 1959; Strick et al., 1999; Cluzel et al., 1996), and MC (Oberhauser et al., 1998; Rief et al., 1999). Each method has its own advantages and shortcomings. MD, e.g., can only tackle very short times, phase transition methods are obviously only applicable to collective deformation, and MC suffers from the problem that the input parameters are not known a priori. But no comprehensive theoretical treatment seems to exist yet.

Therefore, we introduce a compact model for a folded chain, the "sticky chain," which is basically a freely jointed chain with weak bonds between subsequent joints. In the following, complications arising from tertiary structures such as supercoiling (Strick et al., 1996, 1998; Marko and Siggia, 1995) are explicitly excluded. The fundamental equations governing our model are developed in The Basic Model, and some typical results of force-elongation curves are given in Results. Because there are good reasons to assume that the deformation of folded macromolecules may be a collective phenomenon, both basic equations and results are given for individual (random) breaking of the weak bonds and for collective behavior. An important result is the existence of a threshold strain rate, below which the defor-

Received for publication 15 February 2001 and in final form 3 July 2001.
Address reprint requests to Ingomar L. Jäger, University of Leoben, Dept. of Metal Physics, Jahnstr. 12, A-8700 Leoben, Austria. Tel.: 43-3842-45511-36; Fax: 43-3842-45512-60; E-mail: ingomar@unileoben.ac.at.

0006-3495/01/10/1897/10 $2.00

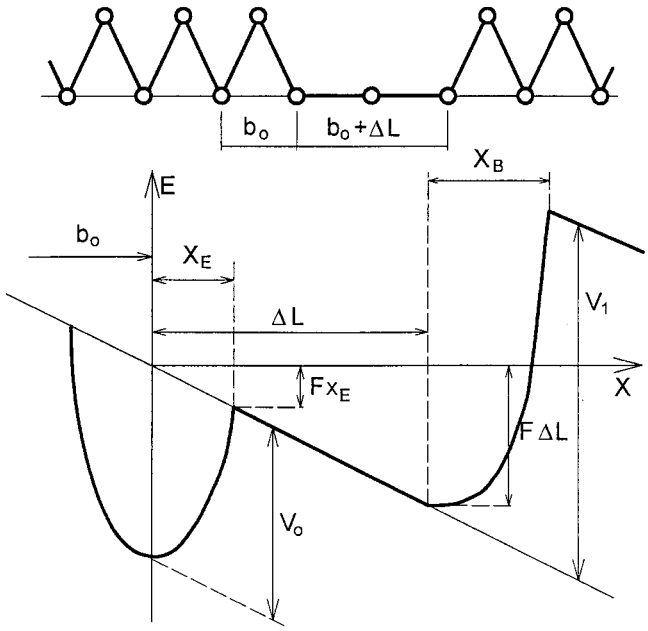

FIGURE 1 *Upper part*, symbolic representation of the Sticky Chain. *Lower part*, energy landscape associated with a single link defining some of the parameters used. For detailed description, see text.

mation force no longer depends on the deformation velocity. The section, Closed Form Solutions, shows how to extract formulae for various easily measurable quantities; they are compiled in the Appendix. In Application to Experiments these formulae are applied to experimental results for biological macromolecules such as titin, DNA, and others, and the physical parameters thus derived compared with results of MC calculations given by the original authors. A short summary is given in the Conclusion.

The model considerations quite deliberately start from a ground state of the chain, the straight configuration, which is, strictly speaking, physically impossible, because to remove entropy completely, infinite energy has to be spent. But the review articles cited above show that 1) the entropic kinking superimposed on the secondary structure of macromolecules can be perfectly modeled by the worm-like chain model, choosing appropriate values for the model parameters persistence length and contour length, and that 2) in many cases a very small stress is sufficient for the molecule to be stretched to near its contour length. Therefore, for practical purposes, the entropic elasticity can be easily subtracted by a suitable choice of the origin of the deformation scale. This view has been verified by numerical simulations including entropic elasticity (not shown).

THE BASIC MODEL

The basis of the model proposed is shown in Fig. 1, *upper part*. Bold lines denote the backbone of the chain and thin lines the weak bonds between subsequent joints. It is assumed that the joints themselves impose no restrictions to motion, so the rupture of a weak bond immediately leads to stretching of the respective Λ-shaped link; effects of inertia are neglected. Please note that the (angled) chain links are only meant symbolically, i.e., to visualize a system of elements in tandem that can individually exist in one of two states, a shorter one (energetically favored), and an extended one. In particular, the angled chain links are not representations of any configuration found in nature.

A schematic drawing of the energy landscape associated with a single chain link under a force F is shown in Fig. 1, *lower part*, where a number of important parameters are defined. In the undisturbed (folded) state the link rests in the leftmost potential minimum with a local depth of $-V_0$. This position is furthermore considered as the origin of the x-(deformation)coordinate. The potential around this minimum is assumed to be harmonic with an elastic constant k connected with V_0 and the potential width x_E by

$$V_0 = \tfrac{1}{2} k x_E^2. \tag{1}$$

If the link manages to get out of this potential minimum, either by applied stress or by thermal excitation, the link elongates by an amount of ΔL into the second minimum. From this position, it may return into the original minimum (refolding), or, again by applied stress or thermal excitation, exit into a position $x \geq \Delta L + x_B$, which means irreversible fracture, a case we will not consider in this paper. Because the second potential is due to deformation of the strong backbone bonds of the chain, it is characterized by a well depth of $V_1 \gg V_0$ (not drawn to scale!) and width x_B, connected by an elastic constant κ,

$$V_1 = \tfrac{1}{2} \kappa x_B^2. \tag{2}$$

The applied force, F, is the inclination of the x axis with respect to the horizontal. It lowers the exit point of the ground state potential by $F \cdot x_E$, the second minimum by $F \cdot \Delta L$.

Let us consider a chain of length L_0, composed of a total of N units of length b_0 each, out of which N_Λ units are in the folded state (in the following, all quantities referring to the folded state will be denoted by the index Λ). Then the rate equation for N_Λ is, according to the model of strain-assisted bond breaking (Bell, 1978),

$$\frac{dN_\Lambda}{dt} = -v_0 e^{Fx_E/RT}\{N_\Lambda e^{-V_0/RT} - (N - N_\Lambda)e^{-F\Delta L/RT}\}, \tag{3}$$

where v_0 is the usual attempt rate. We assume N and N_Λ to be large numbers so that the intrinsically stochastic character of unfolding and refolding of single links is smeared out and a differential equation can be used.

If ε denotes the dimensionless strain, $\delta L/L_0$, then the applied force F is connected to ε by

$$F = \frac{\varepsilon N b_0 - (N_0 - N_\Lambda)\Delta L}{N_\Lambda/k + (N - N_\Lambda)/\kappa}. \tag{4}$$

Here, N_0 denotes the equilibrium number of folded links at $F = 0$, which is derived from Eq. 3 with $dN_\Lambda/dt = 0$,

$$N_0 e^{-V_0/RT} - (N - N_0) = 0. \quad (5)$$

Eq. 4 is simply Hooke's law for N_Λ elements with elastic constant k in tandem with $(N - N_\Lambda)$ with κ. If the strain ε is time-dependent, so is, of course, F, so is N_Λ. For the rest of this paper, we will assume that the time-dependence of ε is linear, i.e., that $\varepsilon = \dot{\varepsilon} \cdot t$, or, for more complicated cases, that the time-dependence of ε can be composed from parts with individually constant $\dot{\varepsilon}$. This corresponds to the usual experimental conditions, where the testing device runs at constant speed.

Because the model has no intrinsic length scale, we can reduce it to dimensionless parameters: $\alpha = V_0/RT$, $\alpha_0 = e^{-\alpha}$, $\beta = b_0/x_E$, $\gamma = \Delta L/b_0$, $\mu = \kappa/k$, the normalized strain rate $\eta = \dot{\varepsilon}/v_0$, and the dimensionless time $\tau = v_0 t$. The dimensionless force, φ, is given by $F = kb_0\varphi = (2RT/x_E)\alpha\beta\varphi$, using Eq. 1. To avoid confusion with the elastic constants k and κ, the gas constant is denoted by R. The fraction of folded links with and without stress is $f_\Lambda = N_\Lambda/N$ and $f_0 = N_0/N$, respectively. For the application of the model to experimental results, another parameter turns out to be important, the rate parameter $\alpha_R = v_0 e^{-\alpha}$. Its reciprocal value is the mean life time of a folded link without a force acting on the chain.

Using these parameters, Eqs. 3 and 4 yield

$$\frac{df_\Lambda}{d\tau} = -e^{2\alpha\beta\varphi}\{f_\Lambda\alpha_0 - (1 - f_\Lambda)e^{-2\alpha\beta^2\gamma\varphi}\}, \quad (6)$$

$$\varphi = \frac{\eta\tau - \gamma(f_0 - f_\Lambda)}{f_\Lambda + (1 - f_\Lambda)/\mu}. \quad (7)$$

Now these basic equations for the sticky chain model are derived under the tacit assumption that the potential governing the unfolding of any link is independent of the state of the neighboring links. Let us call this "individual" bond breaking. This is probably true for string-of-pearls structures, but, especially for helical configurations, it is not a good description. In an α helix, e.g., the stabilizing H bonds are not between adjacent monomers, but rather from monomer n to $n + 3$, so, to start a helix, it takes at least three monomers in the correct position before the fourth one experiences a weak bond stabilizing the first turn of the helix. From then on, each monomer adjacent to the already completed helical portion of the molecule can add one weak bond and, therefore, is subject to a potential according to Fig. 1. This is a typical collective behavior and has therefore been modeled in the literature with typical collective models, e.g., the Zimm–Bragg model (Zimm and Bragg, 1959). Now, collective behavior usually introduces a host of problems, but, fortunately, not here. Because the unfolding of a single link in the interior of an otherwise well-ordered helix involves the fracture of a number $n > 1$ (e.g., $n = 4$) weak bonds, its probability is $p_n = \exp(-nV_0/RT) = (p_1)^n$, which can be safely neglected compared to the probability of unfolding a link adjacent to one or more links already unfolded, $p_1 = \exp(-V_0/RT)$, which may well be of the order of 0.1. We shall therefore assume that helices only unfold in positions adjacent to already unfolded links, beginning at the obvious starting points, the ends of the helix. An MD simulation confirming this view is given by Rohs et al. (1999). Therefore, Eq. 6 need only be slightly altered: instead of the fractions of links that can be unfolded or refolded, f_Λ and $(1 - f_\Lambda)$, respectively, the constant fraction of links adjacent to already unfolded links, $2/N$, enters, and the rate equation for collective behavior thus reads

$$\frac{df_\Lambda}{d\tau} = -e^{2\alpha\beta\varphi}\frac{2}{N}\{\alpha_0 - e^{-2\alpha\beta^2\gamma\varphi}\}. \quad (8)$$

For the sake of completeness, if the macromolecule considered consists of several folded subunits linked together by linker elements without a secondary structure, the factor 2 in Eq. 8 has to be replaced by $2u$, with u the number of subunits.

Eq. 7 remains untouched, only the meaning of f_0 has to be reconsidered: in the case of individual bond breaking, it is the equilibrium fraction of folded links without stress, given by Eq. 5. Because the term in parenthesis in Eq. 8 now does not contain f_Λ, another approach has to be taken. Consider a perfect helix of N links: unfolding the first or last link increases the energy of the system by V_0, unfolding a second link (either adjacent to the first one or starting at the other end of the helix) again increases the energy by V_0, and so on. Unfolding of an isolated link is ruled out. The probability of a helix with n (out of N) unfolded links is $p_n = \exp(-n \cdot V_0/RT)$ and its folded length $N - n$. A little algebra then yields the mean length of a chain in equilibrium without acting force to be $\langle N\rangle = N - 1/(1 - p_1)$. If $p_1 = \exp(-V_0/RT) \ll 1$, then $f_0 = \langle N\rangle/N = 1 - 1/N$, for the long chains mainly considered here, therefore, $f_0 \approx 1$.

RESULTS

The equations implicitly defining the force-elongation diagram of the model, Eqs. 6 and 7 for individual, or 8 and 7 for collective bond breaking, can easily be solved numerically, and some typical diagrams are given in Figs. 2 and 3. Because, as stated above, the real-world force, F, is connected to the model force, φ, by $F = (2RT/x_E)\alpha\beta\varphi$, where $2RT/x_E$ is the scale factor connecting the model with the real world, model results are displayed as $\alpha\beta\varphi$ versus ε.

Model parameters used for the diagrams are: For the bond energy of weak bonds, values between a few hundredths up to a few tenths of an electron volt are given in textbooks, depending, among other factors, on the chemical environment. Therefore, the range of $\alpha = 2$–10 was considered. Lower values seem unrealistic because then the fraction of

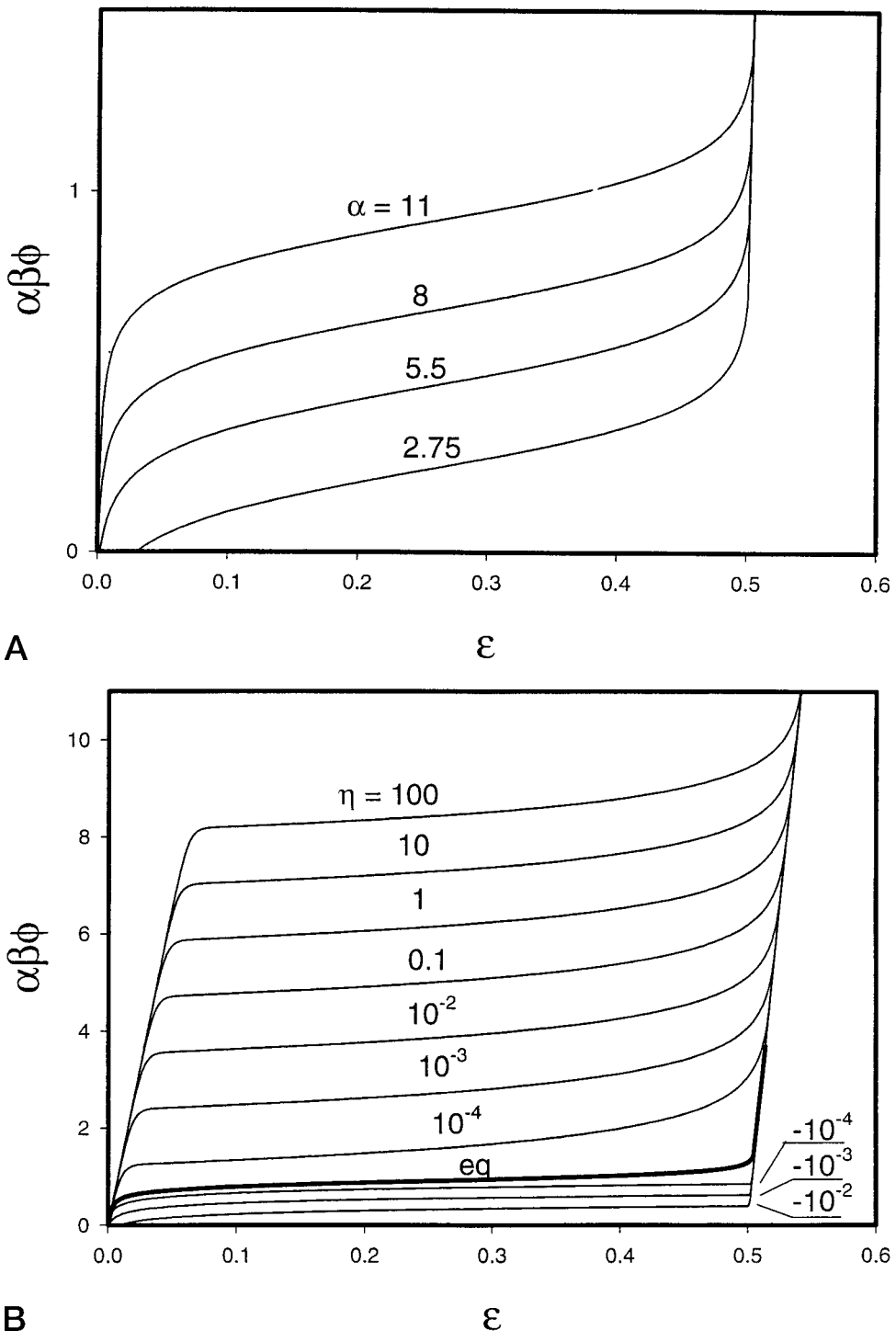

FIGURE 2 (*A*) Calculated model force $\alpha\beta\varphi$ versus strain ε for the case of individual bond breaking at equilibrium. Parameter $\alpha = V_0/RT$. (*B*) Calculated model force $\alpha\beta\varphi$ versus strain ε. Rate dependence of the model force for individual bond breaking and $\alpha = 11$. The parameter is the model strain rate, η.

spontaneously broken links is improbably high, values larger than 10 yield no qualitatively new information. For the various ratios, on which the results do not depend strongly, $\beta = 12$, $\gamma = 0.5$, and $\mu = 2$ were chosen more or less arbitrarily, and $N = 1000$ for the collective case. The parameter defining the time scale, ν_0, cannot, at the moment, be estimated properly, values given in the literature for, e.g., the time constants for the coiling of polypeptides, vary over many orders of magnitude, depending on a host of factors, e.g., the surrounding medium (Klimov and Thirumalai, 1997). Therefore, the parameter in the diagrams drawn is the model deformation rate, η. Results for the case of individual bond breaking, using Eq. 6, are given in Fig. 2 for various values of α and η; the corresponding diagrams for collective behavior, using Eq. 8, are shown in Fig. 3. The strong dependence of the force on the potential depth, α, is

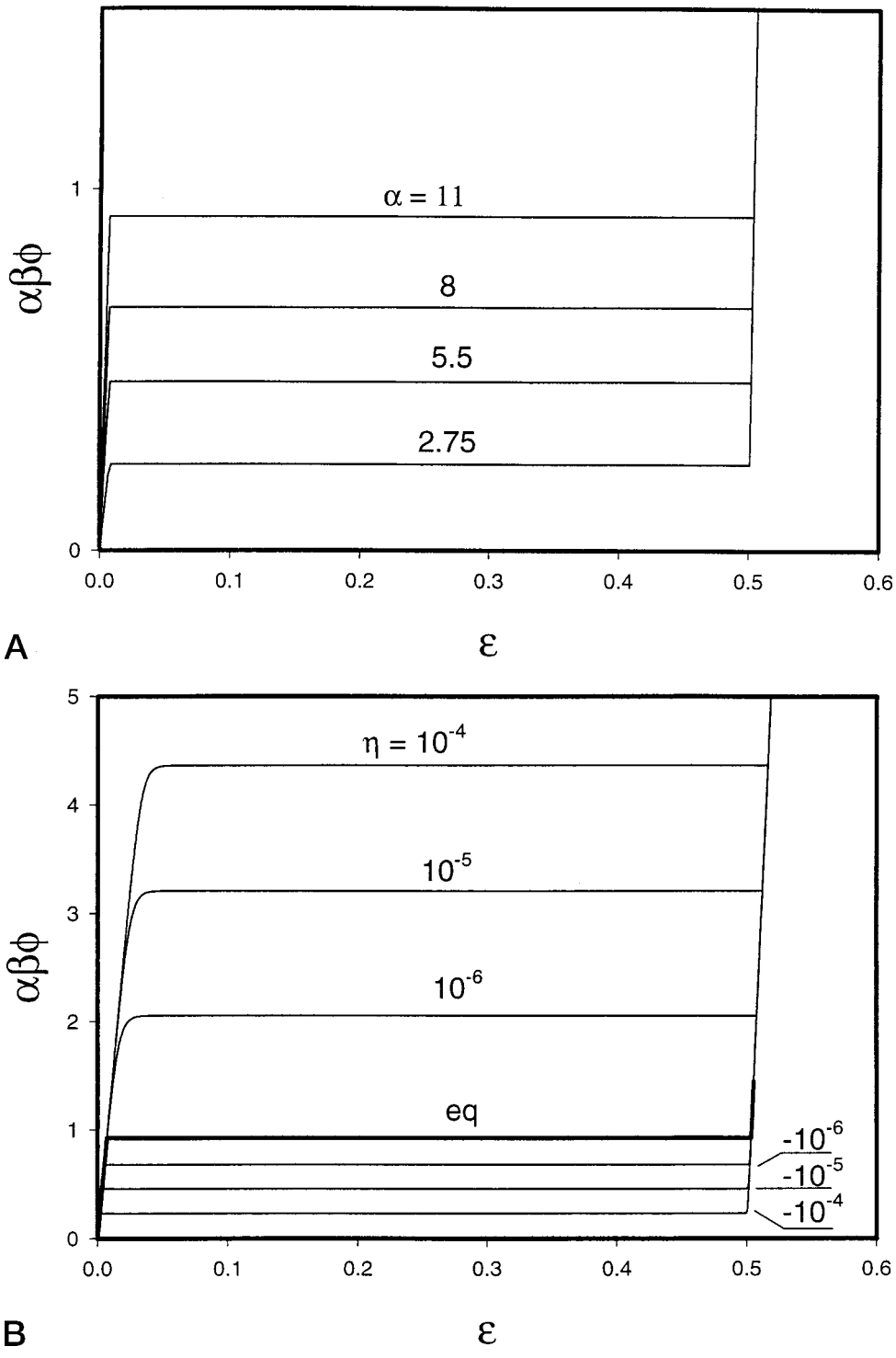

FIGURE 3 (*A*) Calculated model force $\alpha\beta\varphi$ versus strain ε for the case of collective bond breaking at equilibrium. Parameter $\alpha = V_0/RT$. (*B*) Calculated model force $\alpha\beta\varphi$ versus strain ε. Rate dependence of the model force for collective bond breaking and $\alpha = 11$. The parameter is the model strain rate, η.

as expected for a thermally activated process, also the logarithmic dependence on the strain rate.

However, interestingly, there is a threshold strain rate, η_{th}, at which the strain-rate dependence of the deformation stress ceases abruptly and is replaced by a universal curve, labeled eq for equilibrium. This is a little counterintuitive, because, from a thoughtless application of thermally assisted bond breaking, one would expect the logarithmic dependence of the force on the deformation speed to hold down to arbitrary low speeds. But this would only be true if refolding could be neglected totally. Therefore, the equilibrium force-elongation curve may be considered to represent the "correct" or intrinsic tensile strength of the molecule.

All curves, except for very low bond strengths α, show a pronounced plateau, where the stress is slowly rising with strain for individual (Fig. 2) or constant for collective bond breaking (Fig. 3). This is the proper bond-breaking region after the initial Hookean incline. Only for very low α and individual bond breaking, this plateau vanishes and is replaced by a more or less featureless pseudo-Hookean almost straight line. Finally, at the rightmost of the diagram, all curves show a very steep incline, caused by the Hookean stretching of the backbone bonds after (nearly) all weak bonds have been broken and the molecule has adopted the most stretched-out configuration possible. In this region, the gradient of the model force-elongation curve is $\alpha\beta\mu$, whereas it is $\alpha\beta$ at the beginning, according to the different elastic constants of backbone and weak bonds, respectively.

CLOSED FORM SOLUTIONS

Some of the main features of Figs. 2 and 3 can be calculated or estimated analytically. From these calculations, a number of formulae are derived, which are compiled in the Appendix, permitting a comfortable derivation of physical parameters from experimental data. Because the dependence of the stress on f_Λ is weak for individual and nonexistent for collective bond breaking, we shall retain this dependence only in a few selected cases and confine ourselves generally to the center point of the bond-breaking region, $f_\Lambda = ½$. In this case, Eqs. 6 and 8 and all derived from them can be written in a common, generalized form as

$$\frac{df_\Lambda}{d\tau} = -e^{2\alpha\beta\varphi}\frac{1}{G}\{\alpha_0 - e^{-2\alpha\beta^2\gamma\varphi}\}, \tag{9}$$

with $G := 2$ for individual bond breaking and $f_\Lambda = ½$, whereas $G := N/2$ for the collective case.

Let us first consider the case of infinitely slow deformation, $\eta \to 0$. In this case, there is equilibrium at any time, $df_\Lambda/d\tau = 0$, meaning that the term in parenthesis in Eqs. 6 or 8, respectively, equals zero,

$$f_\Lambda\alpha_0 - (1 - f_\Lambda)e^{-2\alpha\beta^2\gamma\varphi_{eq}} = 0 \tag{10}$$

for individual bond breaking at arbitrary f_Λ, and

$$\alpha_0 - e^{-2\alpha\beta^2\gamma\varphi_{eq}} = 0 \tag{11}$$

for individual bond breaking at $f_\Lambda = ½$ and for the collective case.

From Eqs. 10 and 11, the equilibrium (model) force, φ_{eq}, can be determined. For individual bond breaking (Eq. 10) φ_{eq} is a function of f_Λ, and insertion of φ_{eq} and f_Λ into Eq. 7 yields ε. Thus the curves labeled eq are derived. For collective bond breaking (Eq. 11), f_Λ does not enter the formula, therefore the equation only determines the plateau value of φ_{eq}. From these considerations formulae A1–A3 are derived.

Once bond breaking has started, it is the dominant mechanism. Therefore, in this region, $df_\Lambda/d\tau \approx -\eta/\gamma$. Inserting this into Eq. 9 yields

$$\frac{\eta}{\gamma} = \frac{1}{G}e^{2\alpha\beta\varphi}\{\alpha_0 - e^{-2\alpha\beta^2\gamma\varphi}\}. \tag{12}$$

To calculate the stress in the region of high deformation rate, we split φ into two contributions: the equilibrium stress plus the rate-dependent part,

$$\varphi = \varphi_{eq} + \delta. \tag{13}$$

Using the fact that φ_{eq} is the solution of Eq. 11 yields an equation for δ,

$$\frac{\eta}{\gamma} = \frac{1}{G}e^{\alpha/\beta\gamma}e^{2\alpha\beta\delta}\alpha_0[1 - e^{-2\alpha\beta^2\gamma\delta}]. \tag{14}$$

For unfolding $\varphi > \varphi_{eq}$, i.e., $\delta > 0$, so if only $\alpha\beta^2\gamma\delta > 1$, we can neglect the second term in the parenthesis, meaning that the refolding rate is negligible compared to the unfolding rate. This yields, after some algebra, formulae A4–A4c.

Now let us consider the lower part of Figs. 2 and 3, corresponding to refolding. Suppose we have unfolded a chain up to a certain point on the right-hand Hookean line and then reverse the testing device. If we do so with a model strain rate, $|\eta|$, low enough, we just follow the line eq down to the origin. If, however, we use a larger value of $|\eta|$, the stress is lower than its equilibrium value, according to the rate equations, Eqs. 6 and 8. Formally, this may even lead to negative forces if the refolding rate of the chain cannot keep pace with the unloading rate, but this case has to be ruled out because one cannot push a chain, it just goes slack.

Again we can calculate the strain-rate dependence of the stress in this regime, starting from Eq. 13. In contrast to unfolding, now $\delta < 0$. Therefore, if $|\alpha\beta^2\gamma\delta| > 1$ the second term in the parenthesis of Eq. 14 now dominates, leading to formulae A5–A5c. The condition above, that one cannot push a chain, is reflected in Eqs. A6 and A6a. Given the fulfillment of condition A6a, the difference of stresses between unfolding and refolding at identical $|\dot\varepsilon|$ is given by Eq. A7. From this difference and the logarithmic dependences A4c and A5c, the threshold strain rate, $\dot\varepsilon_{th}$, the upper limit of $\dot\varepsilon$ for equilibrium, can be derived. It is given by A8. Formulae A1–A9, or their counterparts in terms of dimensionless parameters, have been checked against numerical solutions and found to agree perfectly.

APPLICATION TO EXPERIMENTS

Extension experiments on single polymer molecules are done nowadays rather routinely, a selected few of them shall be evaluated using the formulae of the Appendix.

DNA

Measured force-elongation curves of Fig. 2 A of Cluzel et al. (1996) agree perfectly with calculated curves of this paper after mental subtraction of the entropic elasticity. They show a marked horizontal plateau, indicating collective unfolding. The strain rates given by the authors, 1 and 10 μm/sec, equivalent to $\dot{\varepsilon} = 0.066$ and 0.66 with their $L_0 = 15.1$ μm, show no effect whatsoever on the stress, so even the higher deformation rate lies within the equilibrium range. Thus A3 can be applied yielding $V_0 = 0.122$ eV, therefore $\alpha = 4.7$. This is in reasonable agreement with the parameter used by Cluzel et al., 8.4 kJ/mol per bp, equivalent to $\alpha = 3.37$. Furthermore, Strick et al. (1999) give an elastic stiffness of the DNA molecule derived from the deviation of the deformation from the worm-like chain model near the onset of bond breaking, $S = 1$ nN. This value is corroborated by Smith et al. (1996) and Noy et al. (1997). From it, and using the textbook value for b_0, 0.34 nm (Kreutzig, 2000; Smith et al., 1996), we can derive an elastic $k(S = k \cdot b_0)$ and, with Eq. 1, estimate $x_E = 0.115$ nm. Because $\dot{\varepsilon} = 0.66$ obviously is an equilibrium deformation rate, let us assume the minimum threshold value for the rate dependence of the force to be $\dot{\varepsilon}_{th} = 1$. We use b_0 above to convert Cluzel et al.'s L_0 into the number of basepairs (links), $N = 44.4 \times 10^3$, and then insert this N (via G), and the textbook value $\Delta L = 0.27$ nm (Kreutzig, 2000; Smith et al., 1996) into A8, to obtain a minimum value of the rate parameter, $\alpha_R^{min} = 3785$. Now this value, suggesting a lifetime for an intact link of only 0.26 ms, seems impossibly high, but this conclusion would be erroneous. Because the deformation is collective, α_R only describes the fluctuation of the first (or last) link of the chain whereas links well within the chain feel the collective potential. Cluzel et al., e.g., use in their calculation a collective unfolding potential about three times V_0, this would result in a lifetime for a middle link of a few seconds, still very low, but then DNA is not a protein designed for mechanical stability but to store and reproduce information, and for the reproduction of the stored information it is necessary to straighten the Watson–Crick helix so it can be split and the sequence read. Thus a low mechanical stability of the helix might even be useful, as suggested, e.g., by Leger et al. (1998).

Finally, let us recheck the validity of the initial assumptions: 1) If the deformation of DNA were not collective, then the plateau would not be horizontal but inclined, and from the inclination b_0 could be estimated using Eq. A2. Now, indeed the plateau of the measurements of Cluzel et al. (1996) is not exactly horizontal, but the height difference between start and end of the plateau is approximately 3–10 pN, depending on how the line interpolating the experimental data is drawn, this yields $b_0 = 1.6–6.4$ nm, far beyond the textbook value cited above. So this possibility is ruled out. 2) If the strain rates used were not in the equilibrium range, we should expect a strain rate dependence of the force according to Eq. A4c. With the data given, this would yield a difference $\Delta F(\dot{\varepsilon}, 10\,\dot{\varepsilon}) = 70$ pN, far too much to hide in experimental scatter. So this possibility is ruled out, too. However, there remains an unexplained detail of the measured force-extension curves. Bustamante et al. (2000) show an extension curve of DNA where the plateau seems to be absolutely horizontal, whereas the measurements of Smith et al. (1996), in accordance to Cluzel et al. (1996), show a slightly inclined plateau, but, again, the b_0 derived from it is unrealistic. In addition, under the influence of a changed chemical environment, the plateau is markedly smeared out at the beginning and at the end. Now, the change in plateau heights under the influence of varying chemistry can be understood as a chemical influence on V_0, but the pronounced smearing-out is beyond our model.

Titin, tenascin

Rief et al. (1997) report the reversible unfolding of titin immunoglobulin (IG) domains. Their force-elongation curve shows discrete serrations corresponding to discrete events of IG "barrels" unfolding one by one superimposed on a gradually rising background. Similar phenomena are shown by Oberhauser et al. (1998) for tenascin.

Both these molecules are of the string-of-pearls type, consisting of pearls of IG or fibronectin III (FNIII), respectively, and linkers in between. Although here, obviously, only a rather small number of large structures is unfolded, we shall try and apply the formulae derived for continuous deformation. Before doing so, however, we have to consider the role of the linkers. The strain rate $\dot{\varepsilon} = (1/L_0)(dL/dt)$, which plays such a dominant role in all formulae outside the equilibrium regime, denotes the deformation rate of the actually unfolding elements, so we have to determine, or at least estimate, the part of the deformation (rate) stemming from the linkers and, if necessary, subtract it. Now both the IG and FN III structures are tertiary structures held together by hydrogen bridges (Carrion-Vazquez, 2000; Lu and Schulten, 1999), commonly considered as weak and soft, whereas the linkers are covalently bound backbones considered strong and stiff. This view is corroborated, e.g., by Noy et al. (1997) reporting a seven-fold increase of the elastic modulus of DNA after the unfolding transition, compared to the modulus before. In the following, we shall therefore assume that the macroscopic deformation (rate) is mainly caused by deformation of the IG or FN III domains and neglect the contribution of the linkers. That means we take $L_0 = N \cdot b_0$, with b_0 being the distance of the points where the linkers couple to the IG or FN III domains, i.e., the N–C distance of Lu and Schulten (1999) or Paci and Karplus (2000). This b_0 is given for tenascin FN III domains as 3.6 nm (Leahy et al., 1992) and assumed approximately the same for titin IG domains. L_0 therefore differs markedly from any values of contour lengths derived from the application of a worm-like chain model.

Titin

Most laudably, Rief et al. (1997) measured the stretching force of titin over almost three decades of strain rate, from their linear interpolation we deduce, using Eq. A4c, x_E = 0.27 nm. This is in perfect agreement with the data used by the authors for their MC simulation, which gave the best fit, x_E(Rief) = 0.30 nm. Additional data used are: b_0 = 3.6 nm (see above), ΔL = 28 nm (Rief et al., 1997), $L_0 \approx 10b_0$ = 36 nm, yielding: deformation rate 1 μm/sec equivalent to $\dot{\varepsilon}$ = 27.8. With these data, Rief et al.'s interpolation line and Eq. A4b, we arrive at $\alpha_R = 3.7 \times 10^{-5}$, again agreeing with Rief et al.'s data and also with Carrion-Vazquez et al. (2000). The continuum model seems to yield perfect results even for this case of only a few unfolding elements, if only mean force values are inserted. Because nothing is known about v_0, α cannot be determined. But if we insert $v_0 = 5 \times 10^{12}$, which may be safely assumed to be an absolute maximum value, we arrive at α = 39.5, which, again, is in perfect agreement with the value $k = 10\ RT/\text{Å}^2$, equivalent to α = 39.2, chosen for the MD simulation of Lu and Schulten (1999), albeit without any reasons given. Finally, we estimate $\dot{\varepsilon}_{th}$. Using Eq. A8 yields $\dot{\varepsilon}_{th} = 2.09 \times 10^{-4}$ equivalent to 7.5×10^{-6} μm/sec. As can be seen, this $\dot{\varepsilon}_{th}$ depends weakly ($\beta\gamma$ = 100) on α, of which only an upper limit is known (and used), therefore it is an upper limit. The corresponding maximum value of the equilibrium force is F_{eq} = 5.8 pN. Both the values of $\dot{\varepsilon}_{th}$ and F_{eq} are rather low and therefore not easily accessible experimentally, but show that even Rief et al.'s lowest deformation rate was almost three decades above the maximum equilibrium strain rate, so all their experiments were done far in the strain-rate dependent force region.

Last, here again is an interesting contradiction. The MD simulation of Lu and Schulten (1999), done for a much higher deformation rate, not only yields much higher forces, but also a completely different rate dependence. Applying Eq. A4c to their simulation yields x_E = 0.0097 nm instead of 0.27 nm, which is clearly inacceptable.

Tenascin

Oberhauser et al. (1998) measured the unfolding force of tenascin, also over more than two decades of strain rate. From their interpolation x_E = 0.28 nm can be derived, identical to titin in accordance to the similarity of the FN III blocks with IG. Also the data for b_0 and ΔL are identical. Unfortunately, in this case, the atomic force microscope (AFM) picked up random segments of the molecule, so it is not possible to correlate pulling speeds and contour lengths exactly to calculate $\dot{\varepsilon}$, this may at least partly account for the scatter in the F versus $\dot{\varepsilon}$ diagram given. Let us assume that the repeated unfolding diagrams of Fig. 3 of Oberhauser et al. (1998) describe seven domains of FN III (the number of unfolding events) deformed at 0.01 μm/sec, i.e. L_0 = 25.2 nm and $\dot{\varepsilon}$ = 0.40. Then the mean unfolding force, F = 68.5 pN, with Eq. A4b yields $\alpha_R = 9.9 \times 10^{-4}$, while Oberhauser et al. assume $\alpha_R = 4.6 \times 10^{-4}$. With this α_R we estimate $\dot{\varepsilon}_{th} = 5.86 \times 10^{-3}$; with this estimate, even the lowest deformation speed used, 0.01 μm/sec, is well above the threshold, in accordance with the marked rate dependence of the force.

Unfortunately, it is at the moment not possible to use the information contained in the reported refolding speeds. Reasons for this are that forces in the refolding range are very low and hard to distinguish from zero plus fluctuations, and that refolding may be slowed down by a host of geometric effects, misalignment, etc., described in the literature, see, e.g., Kurzynski (2000), Zhou and Karplus (1999) or Klimov and Thirumalai (1997).

Spectrin

Unfolding forces of spectrin were measured by Rief et al. (1999) and again show a marked rate dependence. Without any knowledge of geometrical parameters, the application of Eq. A4c to the measured forces yields x_E = 1.59 nm, ≈5–6 times the value of titin and tenascin. This interesting discrepancy is in perfect agreement with Rief et al.'s result and its physical significance is discussed by the authors. Assuming b_0 = 10 nm (Yan et al., 1993; Paci and Karplus, 2000), ΔL = 31.7 nm (Rief et al., 1999) and L_0 = 12 units of b_0 each, the deformation speeds result in $\dot{\varepsilon}$ = 6.67 and 0.667 respectively. Then the deformation force at $\dot{\varepsilon}$ = 6.67, F = 32.5 pN, yields, with Eq. A4b, $\alpha_R = 1.44 \times 10^{-5}$, similar to titin. However, due to the lack of information on α, and the low $\beta\gamma$ = 20, here we cannot apply Eq. A8 to extract the threshold strain rate. Also the information contained in the refolding times cannot, at the moment, be evaluated for the reasons given above.

CONCLUSION

A simple model for the deformation of macromolecules on the basis of the unfolding of certain domains is presented. In addition to reproducing some characteristic features of experimental data, such as the plateau of the force-elongation curve or the logarithmic dependence of the pulling force on the pulling speed, it discloses the fact that there is a rather well-defined strain rate, below which the force does not depend on the deformation rate any more. Closed form solutions are given that permit a simple derivation of physical parameters from measured quantities. With the aid of these formulae, force measurements on macromolecules are evaluated, using a pocket calculator, and the results thus derived agree perfectly with parameters derived by the original authors using lengthy and time-consuming MC calculations plus parameter fitting. Most surprisingly, the model formulae yield good results even for the case of only

few unfolding domains, e.g., for titin, in spite of the intrinsic continuum character of the model, provided that mean deformation forces are input. Moreover, in some cases, additional information can be derived, or, at least, estimated. Thus the model presented is a valuable aid for the simple evaluation of experiments involving the stretching and unfolding of certain macromolecules. Because the model lacks any built-in length scale, it may also be applied to deformation phenomena at higher hierarchical levels than the molecular one.

APPENDIX: FORMULAE

Please note that all formulae containing f_Λ explicitly are valid for individual bond breaking at that value of f_Λ. For these cases, formulae for the collective case do not exist. Formulae not containing f_Λ are valid both for individual bond breaking at $f_\Lambda = ½$ and collective bond breaking. If the formulae for the two cases differ, the parameter G is used $G := 2$ for individual bond breaking at $f_\Lambda = ½$, and $G := N/2$ for the collective case.

Equilibrium

$$F_{eq}(f_\Lambda) = \frac{V_0}{\Delta L} + \frac{RT}{b_0} \ln \frac{1 - f_\Lambda}{f_\Lambda} \tag{A1}$$

$$\Delta F_{eq}(f_\Lambda^1, f_\Lambda^2) = \frac{RT}{b_0} \ln \frac{(1 - f_\Lambda^1) f_\Lambda^2}{f_\Lambda^1 (1 - f_\Lambda^2)} \tag{A2}$$

$$F_{eq} = \frac{V_0}{\Delta L} \tag{A3}$$

Nonequilibrium

Unfolding

$$F(\dot\varepsilon, f_\Lambda) = \frac{RT}{x_E} \{\ln \dot\varepsilon - \ln \gamma - \ln \alpha_R - \ln f_\Lambda\} \tag{A4}$$

$$\Delta F(f_\Lambda^1, f_\Lambda^2) = \frac{RT}{x_E} \ln \frac{f_\Lambda^2}{f_\Lambda^1} \tag{A4a}$$

$$F(\dot\varepsilon) = \frac{RT}{x_E} \left\{\ln \dot\varepsilon + \ln \frac{G}{\gamma} - \ln \alpha_R\right\} \tag{A4b}$$

$$\Delta F(\dot\varepsilon_1, \dot\varepsilon_2) = \frac{RT}{x_E} \ln \frac{\dot\varepsilon_1}{\dot\varepsilon_2} \tag{A4c}$$

Refolding

$$F(|\dot\varepsilon|, f_\Lambda) = \frac{RT}{x_E} \frac{1}{1 - \beta\gamma} \{\ln|\dot\varepsilon| - \ln \gamma - \ln \alpha_R - \alpha - \ln(1 - f_\Lambda)\} \tag{A5}$$

The maximum refolding strain rate (at which the chain goes slack, $F = 0$), is

$$\ln|\dot\varepsilon|_{max} = \alpha + \ln \alpha_R + \ln(1 - f_\Lambda) + \ln \gamma \tag{A6}$$

for arbitrary f_Λ, or

$$\ln|\dot\varepsilon|_{max} = \alpha + \ln \alpha_R - \ln \frac{G}{\gamma}. \tag{A6a}$$

Subject to condition A6,

$$\Delta F(f_\Lambda^1, f_\Lambda^2) = \frac{RT}{x_E} \frac{1}{1 - \beta\gamma} \ln \frac{1 - f_\Lambda^1}{1 - f_\Lambda^2}, \tag{A5a}$$

subject to A6a,

$$F(|\dot\varepsilon|) = \frac{RT}{x_E} \frac{1}{1 - \beta\gamma} \left\{\ln|\dot\varepsilon| + \ln \frac{G}{\gamma} - \ln \alpha_R - \alpha\right\}, \tag{A5b}$$

and thus

$$\Delta F(\dot\varepsilon_1, \dot\varepsilon_2) = \frac{RT}{x_E} \frac{1}{1 - \beta\gamma} \ln \frac{\dot\varepsilon_1}{\dot\varepsilon_2}. \tag{A5c}$$

Unfolding–Refolding

$$\Delta F(|\dot\varepsilon|, \uparrow - \downarrow) = \frac{RT}{x_E} \frac{\beta\gamma}{\beta\gamma - 1} \left\{\ln|\dot\varepsilon| + \ln \frac{G}{\gamma} - \frac{\alpha}{\beta\gamma} - \ln \alpha_R\right\} \tag{A7}$$

Maximum strain rate for equilibrium

$$\ln|\dot\varepsilon|_{th} = \ln \alpha_R + \frac{\alpha}{\beta\gamma} - \ln \frac{G}{\gamma}. \tag{A8}$$

If a diagram of F versus $\ln(\dot\varepsilon)$ is extended beyond the physical regime, i.e., below $\dot\varepsilon_{th}$, down to $F = 0$, this purely hypothetical $\dot\varepsilon_0$ fulfills

$$\ln \dot\varepsilon_0 = \ln \alpha_R - \ln \frac{G}{\gamma}. \tag{A9}$$

The formulae given are by no means totally independent, so it is not possible, e.g., to calculate from $\dot\varepsilon_{th}$, via Eq. A4b, $F(\dot\varepsilon_{th})$, which is equal to F_{eq} and thus derive V_0 in an independent way. This would be a circle because the condition $F(\dot\varepsilon_{th}) = F_{eq}$ is an implicit definition of $\dot\varepsilon_{th}$.

This work is dedicated to Prof. H.-P. Stüwe, who taught me the merits of simple models.

The author is indebted to Dr. Udo Seifert, from the Max-Planck-Institut für Kolloid- und Grenzflächen forschung, Potsdam-Golm, Germany for pointing out the importance of Bell's work, and to Dr. Th. Schöberl for a critical reading of the manuscript.

REFERENCES

Bell, G. I. 1978. Models for the specific adhesion of cells to cells: a theoretical framework for adhesion mediated by reversible bonds between cell surface molecules. *Science.* 200:618–627.

Bustamante, C., St. B. Smith, J. Liphardt, and D. Smith. 2000. Single molecule studies of DNA mechanics. *Curr. Opin. Struct. Biol.* 10: 279–285.

Carrion-Vazquez, M., A. F. Oberhauser, T. E. Fisher, P. E. Marszalek, H. Li, and J. M. Fernandez. 2000. Mechanical design of proteins studied by single-molecule force spectroscopy and protein engineering. *Progr. Biophys. Mol. Biol.* 74:63–91.

Cluzel, P., A. Lebrun, C. Heller, R. Lavery, J.-L. Viovy, D. Chatenay, and F. Caron. 1996. DNA: an extensible molecule. *Science.* 271:792–794.

Kellermayer, M. S. Z., St. B. Smith, C. Bustamante, and H. L. Granzier. 1998. Complete unfolding of the Titin molecule under external force. *J. Struct. Biol.* 122:197–205.

Klimov, D. K., and D. Thirumalai. 1997. Viscosity dependence of the folding rates of proteins. *Phys. Rev. Lett.* 79:317–320.

Kreutzig, T. 2000. Biochemie. Urban & Fischer, München, Jena, Germany. 106–109.

Kurzynski, M. 2000. Internal dynamics of biomolecules and statistical theory of biochemical processes. *Physica A.* 285:29–47.

Leahy, D. J., W. A. Hendrickson, I. Aukhil, and H. P. Erickson. 1992. Structure of a fibronectin type III domain from tenascin phased by MAD analysis of the selenomethionylprotein. *Science.* 258:987–991.

Leger, J. F., J. Robert, L. Bourdieu, D. Chatenay, and J. F. Marko. 1998. RecA binding to a single double-stranded DNA molecule: a possible role of DNA conformational fluctuations. *Proc. Natl. Acad. Sci. U.S.A.* 95:12295–12299.

Lu, H., B. Isralewitz, A. Krammer, V. Vogel, and K. Schulten. 1998. Unfolding of Titin immunoglobulin domains by steered molecular dynamics simulation. *Biophys. J.* 75:662–671.

Lu, H., and K. Schulten. 1999. Steered molecular dynamics simulation of force-induced protein domain unfolding. *Proteins Struct. Funct. Genet.* 35:453–463.

Marko, J. F., and E. D. Siggia. 1995. Statistical mechanics of supercoiled DNA. *Phys. Rev. E.* 52:2912–2938.

Noy, A., D. V. Vezenov, J. F. Kayyem, T. J. Meade, and C. M. Lieber. 1997. Stretching and breaking duplex DNA by chemical force microscopy. *Chem. Biol.* 4:519–527.

Oberhauser, A. F., P. E. Marszalek, H. P. Erickson, and J. M. Fernandez. 1998. The molecular elasticity of the extracellular matrix protein tenascin. *Nature.* 393:181–185.

Paci, E., and M. Karplus. 2000. Unfolding proteins by external forces and temperature: the importance of topology and energetics. *Proc. Natl. Acad. Sci. U.S.A.* 97:6521–6526.

Rief, M., M. Gautel, F. Oesterhelt, J. M. Fernandez, and H. E. Gaub. 1997. Reversible unfolding of individual titin immunoglobulin domains by AFM. *Science.* 276:1109–1112.

Rief, M., J. Pascual, M. Saraste, and H. E. Gaub. 1999. Single molecule force spectroscopy of spectrin repeats: low unfolding forces in helix bundles. *J. Mol. Biol.* 286:553–561.

Rohs, R., C. Etchebest, and R. Lavery. 1999. Unraveling proteins: a molecular mechanics study. *Biophys. J.* 76:2760–2768.

Smith, B. L. 2000. The importance of molecular structure and conformation: learning with scanning probe microscopy. *Progr. Biophys. Mol. Biol.* 74:93–113.

Smith, St. B., Y. Cui, and C. Bustamante. 1996. Overstretching B-DNA. The elastic response of individual double-stranded and single-stranded DNA molecules. *Science.* 271:795–799.

Smith, B. L., T. E. Schäffer, M. Viani, J. B. Thompson, N. A. Frederick, J. Kindt, A. Belcher, G. D. Stucky, D. E. Morse, and P. K. Hansma. 1999. Molecular mechanistic origin of the toughness of natural adhesives, fibres and composites. *Nature.* 399:761–763.

Strick, T. R., J.-F. Allemand, D. Bensimon, A. Bensimon, and V. Croquette. 1996. The elasticity of a single supercoiled DNA molecule. *Science.* 271:1835–1837.

Strick, T. R., J. F. Allemand, D. Bensimon, and V. Croquette. 1998. Behavior of supercoiled DNA. *Biophys. J.* 74:2016–2028.

Strick, T. R., J. F. Allemand, D. Bensimon, R. Lavery, and V. Croquette. 1999. Phase coexistence in a single DNA molecule. *Physica A.* 263: 392–404.

Strick, T., J.-F. Allemand, V. Croquette, and D. Bensimon. 2000. Twisting and stretching single DNA molecules. *Progr. Biophys. Mol. Biol.* 74: 115–140.

Yan, Y., E. Winograd, A. Viel, T. Cronin, S. C. Harrison, and D. Branton. 1993. Crystal structure of the repetitive segments of spectrin. *Science.* 262:2027–2030.

Zhou, Y., and M. Karplus. 1999. Interpreting the folding kinetics of helical proteins. *Nature.* 401:400–403.

Zimm, B. H., and J. K. Bragg. 1959. Theory of the phase transition between helix and random coil in polypeptide chains. *J. Chem. Phys.* 31:526–535.

Zlatanova, J., St. M. Lindsay, and S. H. Leuba. 2000. Single molecule force spectroscopy in biology using the atomic force microscope. *Progr. Biophys. Mol. Biol.* 74:37–61.