PHYSICAL REVIEW E **81**, 010902(R) (2010)

Unfolding times for proteins in a force clamp

Stefano Luccioli,[1,2] Alberto Imparato,[3] Simon Mitternacht,[4,5] Anders Irbäck,[4] and Alessandro Torcini[1,2]
[1]*Istituto dei Sistemi Complessi, CNR, via Madonna del Piano 10, I-50019 Sesto Fiorentino, Italy*
[2]*INFN, Sez. Firenze, CSDC, via Sansone, 1-I-50019 Sesto Fiorentino, Italy*
[3]*Department of Physics and Astronomy, University of Aarhus, Ny Munkegade, Building 1520, DK-8000 Aarhus C, Denmark*
[4]*Department of Theoretical Physics, Lund University, Sölvegatan 14A, SE-223 62 Lund, Sweden*
[5]*Department of Informatics, University of Bergen, PB 7803, N-5020 Bergen, Norway*
(Received 1 June 2009; revised manuscript received 9 November 2009; published 29 January 2010)

The escape process from the native valley for proteins subjected to a constant stretching force is examined using a model for a β barrel. For a wide range of forces, the unfolding dynamics can be treated as one-dimensional diffusion, parametrized in terms of the end-to-end distance. In particular, the escape times can be evaluated as first passage times for a Brownian particle moving on the protein free-energy landscape, using the Smoluchowski equation. At strong forces, the unfolding process can be viewed as a diffusive drift away from the native state, while at weak forces thermal activation is the relevant mechanism. An escape-time analysis within this approach reveals a crossover from an exponential to an inverse Gaussian escape-time distribution upon passing from weak to strong forces. Moreover, a single expression valid at weak and strong forces can be devised both for the average unfolding time as well as for the corresponding variance. The analysis offers a possible explanation of recent experimental findings for the proteins ddFLN4 and ubiquitin.

DOI: 10.1103/PhysRevE.81.010902 PACS number(s): 87.15.hm, 05.40.Jc, 87.15.A−, 87.15.La

Single-molecule pulling experiments have become an important and widely used tool for examining mechanical properties of proteins [1]. These experiments have stimulated a renewed interest in the escape processes from metastable potential wells in the presence of a biasing force [2]. Traditionally, the dependence of the escape rate k on the stretching force F has often been modeled using the phenomenological Bell formula $k(F)=k_0 e^{\beta F x_u}$ [3], where x_u is the distance from the native to the transition state and assumed constant ($\beta=1/k_B T$, with k_B being the Boltzmann constant and T being the temperature). The zero-force rate k_0 satisfies $k_0 \propto e^{-\beta \Delta G_0}$, where ΔG_0 is the escape free-energy barrier at zero force. There are, however, uncertainties about how to extract the zero-force properties k_0, x_u, and ΔG_0 from observed escape rates at nonzero force. One problem is the unknown constant of proportionality in the expression for k_0. Another difficulty is that the distance x_u between the native free-energy minimum and the unfolding barrier, which is assumed constant in the Bell formula, generally depends on the applied force.

To address these problems, several generalizations of the Bell formula have recently been proposed [4–7]. Most of the extensions are based on the same underlying picture as for the Bell formula; the protein is viewed as a Brownian particle moving in a tilted one-dimensional potential, $G(x)=G_0(x)-Fx$, where $G_0(x)$ is the zero force equilibrium free-energy profile. Using different approximations and parametrizations of $G_0(x)$, key properties of the escape process have been analyzed, such as the mean and variance of the rupture force at constant velocity pulling [4,6]. It was further shown [7] that the approach of Dudko, Hummer, and Szabo (DHS) [4] is able to describe experimentally observed deviations from the Bell formula for the fourth domain of Dictyostelium discoideum filamin (ddFLN4) [8].

These extensions based on Kramers theory [9] assume that the escape barrier is high compared to $k_B T$, leading to single-exponential kinetics. Very recently, Yew *et al.* analyzed deviations from single-exponential kinetics in unfolding simulations based on a C_α model [10]. By including the next-to-leading term in an eigenfunction expansion, they obtained an improved description of the unfolding dynamics at strong force. However, a comprehensive picture describing $k(F)$ and the full escape-time distribution at both weak and strong forces is still missing. A key parameter when describing the force dependence is the critical force F_c, at which the escape barrier disappears. In the DHS approach [4,7], one has $F_c=\Delta G_0/\nu x_u$, where ν is a model parameter ($\nu=1$ corresponds to the Bell formula). The above-mentioned ddFLN4 analysis [7] (with $\nu=1/2$ or 2/3) suggests that $F_c \sim 80-110$ pN for this protein. For the titin module I27, on the other hand, F_c appears to be significantly larger ($\Delta G_0/x_u \sim 640$ pN [11]). Due to different F_c, when analyzing experimental data, the strong-force regime $F>F_c$ may or may not be relevant, depending on the protein.

In this Rapid Communication we investigate the response of a model protein to a wide range of constant pulling forces. We show that, once the free-energy landscape is known with sufficient accuracy, the usual Smoluchowski equation [9] in one dimension is sufficient to obtain a good estimate of the average escape time from the native valley and the associated variance. Two force regimes, separated by the critical force F_c, are observed. For $F<F_c$, unfolding occurs through a thermally activated escape process. For $F>F_c$, the unfolding dynamics can instead be interpreted as pure diffusion with an external bias. The transition from the weak- to the strong-force regime is accompanied by a drastic change in the shape of the escape-time distribution, from exponential to inverse Gaussian. The applicability of this approach to real proteins, at forces studied experimentally, is addressed using recently reported data for ddFLN4 [7,8] and ubiquitin [12,13].

The protein model we consider is the three-dimensional off-lattice hydrophobic-polar-neutral (BPN) model [14–16], where each residue is represented by a single point and is of one of the following three types: hydrophobic (B), polar (P),

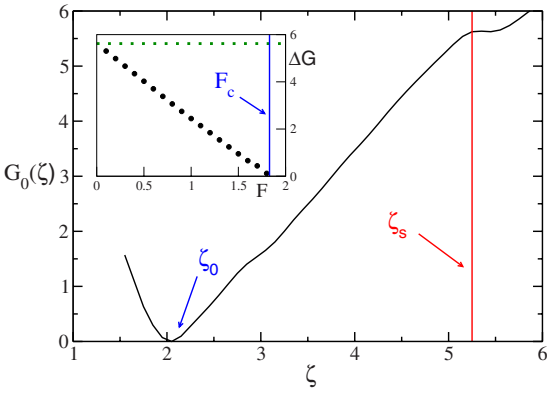

FIG. 1. (Color online) Free energy $G_0(\zeta)$ at zero force for the BPN protein, calculated as a function of the end-to-end distance ζ. The positions of the native state, $\zeta_0 \approx 2.0$, and the saddle, $\zeta_s \approx 5.25$, are indicated. The inset shows the escape barrier ΔG versus F. The vertical (blue) and horizontal (green) lines indicate F_c and the zero-force barrier, ΔG_0.

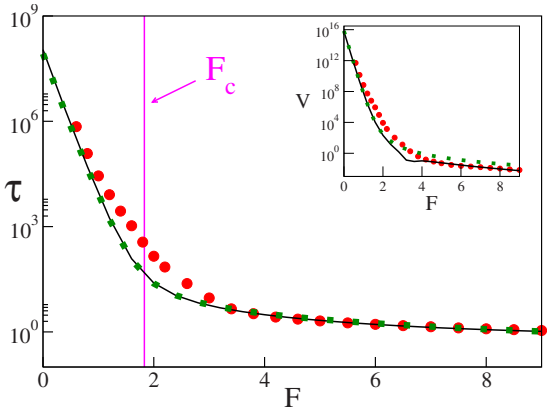

FIG. 2. (Color online) Average escape time against force for the BPN protein. Filled (red) circles are simulation results and the (black) curve is the prediction τ_S obtained from Eq. (1), with $\gamma=0.05$ and $M=46$. The vertical (magenta) line indicates F_c. The dotted (green) line is the estimate τ_L in Eq. (2), with M and γ as above and $a=3.25$. The inset shows the variance, V. Filled (red) circles are simulation results, whereas the (black) curve and the dotted (green) line represent the estimates V_S and V_L, respectively.

or neutral (N). We study a 46-residue sequence, which is known to form a four-stranded β barrel in its native state. The folding [14–17] and mechanical unfolding [18,19] of this sequence have been extensively studied. We analyze via Langevin dynamics the response of this model protein to external forces acting on the chain ends in the proximity of its folding temperature, namely, at $T=0.3$ [19]. Parameter values are as in Ref. [19] and all model quantities are dimensionless; for a comparison with physical units, see Ref. [16].

A typical unfolding trajectory begins with a waiting phase, where the end-to-end distance ζ stays close to its native value. This phase is followed by a sudden increase in ζ. A fundamental question is whether the escape from the native valley can be effectively described as one-dimensional diffusion, parametrized in terms of ζ. Based on this assumption the unfolding process is commonly described as the motion of a pointlike Brownian particle in the potential $G(\zeta)=G_0(\zeta)-F\zeta$, where $G_0(\zeta)$ is the equilibrium free-energy profile. The average first passage time $\tau(x)$ at a threshold ζ_s for a particle with initial position $x \in [\zeta_0, \zeta_s]$ can be obtained by solving the Smoluchowski equation. One finds that [9]

$$\tau(x) = \beta M \gamma \int_x^{\zeta_s} dy\, e^{\beta G(y)} \int_{\zeta_0}^{y} dz\, e^{-\beta G(z)}, \qquad (1)$$

where M is the particle mass and γ is the damping constant. The boundaries at ζ_0 and ζ_s are reflecting and absorbing, respectively. When using Eq. (1) to calculate the escape time from the native valley, ζ_0 is the native ζ and ζ_s is that of the saddle, or the barrier, to be crossed. The escape time is obtained as $\tau_S \equiv \tau(\zeta_0)$. In our simulations, escape times are measured using a threshold slightly larger than ζ_s to avoid saddle recrossing [9].

We begin by testing the escape-time prediction τ_S directly against simulation results for the BPN protein, without making any further assumption on the form of $G(\zeta)$. For this purpose, we determine $G(\zeta)$ numerically, using methods described in Ref. [19]. Figure 1 shows the calculated free-energy profile at zero force, $G_0(\zeta)$, which exhibits a pronounced native minimum at $\zeta_0 \approx 2.0$ and a barrier at $\zeta_s \approx 5.25$. The height of the barrier is $\Delta G_0 = G_0(\zeta_s) - G_0(\zeta_0) \approx 5.62$. The application of a stretching force F tilts the free-energy landscape to $G(\zeta)=G_0(\zeta)-F\zeta$ and reduces the barrier height ΔG. As shown in the inset of Fig. 1, ΔG decreases almost linearly with F. The barrier finally disappears at $F_c \approx 1.83$.

Knowing $G(\zeta)$, the escape-time prediction τ_S can be obtained by numerically evaluating the double integral in Eq. (1). In Fig. 2 we compare τ_S with simulated escape times. The agreement is very good for strong forces ($F \gtrsim 3$) as well as at weak forces ($F \lesssim 1.2$). Due to computational limitations, it was impossible to investigate forces <0.6. The regime in which the simulated escape times are most difficult to reproduce is around the critical force F_c, where there is no clear free-energy gradient either toward or away from the native state. In this regime, the details of the free-energy profile matter. It is remarkable, however, that this simple picture—without employing any fitting parameter—is able to describe the behavior at both strong and weak forces, despite escape-time differences of almost six orders of magnitude.

This analysis, based on the full profile $G(\zeta)$, addresses in a direct manner the question of whether or not the system can be described in terms of one-dimensional diffusion. In unfolding experiments, $G(\zeta)$ is unknown, and the challenge is to extract the main features of the free-energy landscape from measured escape times. This task is greatly facilitated if the free energy can be linearly approximated in the interval $[\zeta_0, \zeta_s]$, as $G(\zeta)=(F_c-F)(\zeta-\zeta_0)$ (up to an additive constant). With this approximation, the integrals in Eq. (1) can be evaluated analytically. The resulting expression, for the average escape time of a diffusive particle in one dimension in the presence of a bias (F in the present context), is [20]

$$\tau_L = \frac{M\gamma a}{F-F_c} - \frac{M\gamma k_B T}{(F-F_c)^2}\left[1 - e^{-\beta(F-F_c)a}\right], \qquad (2)$$

where $a=\zeta_s-\zeta_0$ is the distance between the reflecting and absorbing boundaries. Unlike the result reported in Ref. [20],

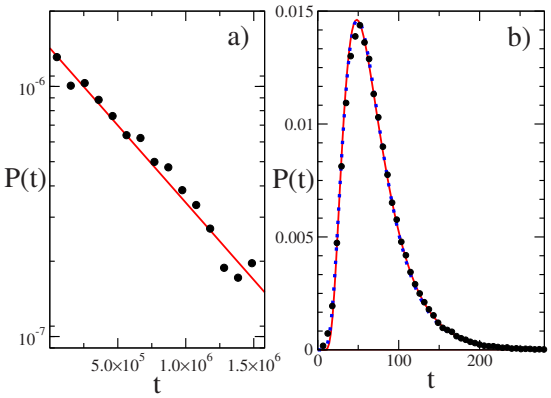

FIG. 3. (Color online) Escape-time distribution $P(t)$ for the BPN protein at two different forces. Large (black) dots are simulation results. (a) $F=0.6$ ($<F_c$). The (red) curve is an exponential ($\tau=7.02\times10^5$). (b) $F=2.2$ ($>F_c$). The (red) curve is an inverse Gaussian ($\tau=71.1$, $V=1.35\times10^3$), whereas small (blue) dots represent a log-normal fit to data.

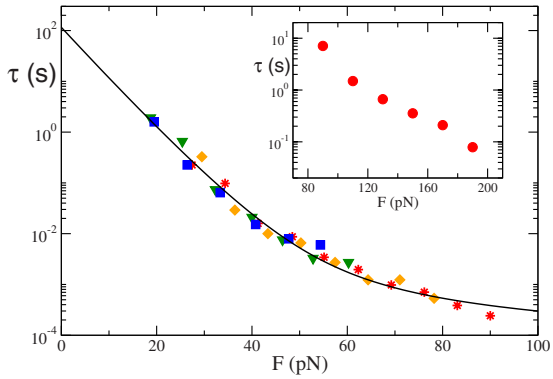

FIG. 4. (Color online) Average unfolding time versus force for ddFLN4. The symbols (explained in detail in Ref. [23]) represent experimental data, originally obtained at constant velocity [8] and transformed to constant-force conditions in Ref. [7]. The (black) curve is a fit of Eq. (2) ($M\gamma=0.012$ pN s/nm, $a=1.1$ nm, and $F_c=60$ pN). The inset shows ubiquitin data from force-clamp experiments reported in Ref. [13].

Eq. (2) includes the effect of a zero-force barrier, represented by the term $F_c a$. The singular terms at $F=F_c$ in Eq. (2) cancel out, as they should.

The assumption that $G(\zeta)$ is linear between the native state and the saddle is quite well satisfied for the BPN protein (see Fig. 1). Actually, the escape times obtained using this approximation, τ_L, essentially coincide with the estimated τ_S obtained using the full $G(\zeta)$, as can be seen from Fig. 2. Note that Eq. (2), like Eq. (1), has no parameter that needs to be fitted because we can use the value of F_c previously determined. While Eq. (2) well describes the escape time down to the lowest forces that could be studied, one should still be cautious in using this expression to extrapolate to zero force, because a "turnover" to a force-independent process is likely to occur at weak force [21]. The extent of this weak-force regime might be non-negligible if the temperature is high [21].

The variance of the escape time is in the Smoluchowski approach given by $V_S=\tau_{2,S}-\tau_S^2$, where the second moment $\tau_{2,S}$ reads [22]

$$\tau_{2,S}=2\left(\frac{M\gamma}{k_BT}\right)\int_{\zeta_0}^{\zeta_s}dy\,e^{\beta G(y)}\int_{\zeta_0}^{y}dx\,e^{-\beta G(x)}\tau(x). \qquad (3)$$

Like the mean, the variance can be obtained analytically if $G(\zeta)$ depends linearly on ζ. This estimate of the variance, V_L, can be found in Eq. (S1) [23]. The inset of Fig. 2 shows our simulation results for the variance of the escape time for the BPN protein, along with the estimates V_S and V_L. For $F<F_c$, V_S and V_L are almost identical, while for $F>F_c$, V_L is slightly larger than both V_S and the simulation results, although the corresponding three average times are very similar in this regime. Overall, both V_S and V_L agree well with the simulation results.

It is informative to go beyond the first and second moments and also study the full probability distribution of the escape time. For $F\lesssim F_c$, we find that the escape-time distribution of the BPN protein to a very good approximation is exponential, $P(t)=\tau^{-1}e^{-t/\tau}$, with τ being the mean [see Fig. 3(a)]. This observation confirms that at weak forces, where a free-energy barrier is still present, the main escape mechanism is thermal activation. At $F\sim F_c$, the escape process changes in character, from a thermally activated process to a diffusive process driven by an external bias (force). In the latter regime, it is known that first passage times follow a so-called inverse Gaussian distribution [24]. This distribution is given by

$$P(t)=\frac{\tau}{\sqrt{2\pi\Gamma t^3}}e^{-(t-\tau)^2/(2\Gamma t)}, \qquad (4)$$

where τ is the mean and $\Gamma=V/\tau$, with V being the variance. This expression indeed provides a very good description of our simulation results at strong forces, as illustrated in Fig. 3(b). It should be noticed that this comparison does not involve any parameter fitting because τ and V are determined directly from the simulations.

Previous studies have used a log-normal distribution, rather than the inverse Gaussian, to describe the escape-time distribution at strong forces [25,26]. While the log-normal distribution is similar to the inverse Gaussian [see Fig. 3(b)], there is no theoretical background to justify its use in the present context. The inverse Gaussian distribution is, by contrast, known to arise from the biased Brownian motion [24], which provides a simple physical picture of the unfolding dynamics at strong forces.

Having seen that our approach provides a good description of the unfolding dynamics of the BPN protein, we now turn to two real proteins: ddFLN4 and ubiquitin. Two results of the above analysis are particularly useful when comparing with experimental data. The first is Eq. (2), which provides an approximate closed-form expression for the average escape time $\tau(F)$ at both weak and strong forces. The second result is that the onset of the nonexponential strong-force behavior of $\tau(F)$ is accompanied by a change in the shape of the escape-time distribution, from exponential to inverse Gaussian.

Experimental unfolding times for ddFLN4 show, as mentioned earlier, clear deviations from the Bell formula [7,8]. It has been demonstrated [7] that the DHS approach [4] describes the data well. In Fig. 4, we show a fit of our Eq. (2) to the same data. The fit is good, and the fitted values $a=1.1$ nm (corresponding to x_u) and $\Delta G_0=F_c a=9.50$ kcal/mol are consistent with the results of Ref. [7]. Unlike the DHS approach, ours does not assume the escape barrier to be high. For ddFLN4, our fit to the $\tau(F)$ data indicates that the barrier disappears already at $F_c \sim 60$ pN. It would be very interesting to see whether the escape-time distribution is inverse Gaussian at, say, 100 pN, but this distribution has not been evaluated, as far as we know.

For ubiquitin, the escape-time distribution has been measured experimentally at 110 pN [12]. The data were found to be well described by a log-normal distribution [12], which is very similar to the inverse Gaussian one found above at strong forces. Our approach thus offers an explanation of the shape of the observed distribution. This explanation requires that $F_c < 110$ pN. Very recent experimental $\tau(F)$ data for ubiquitin [13] show signs of deviations from the Bell formula (see the inset of Fig. 4). However, it was found that the data could not discriminate between the Bell and DHS formulas [13]. Neither are the data sufficient to permit a stable fit of Eq. (2), which would have given us an independent estimate of F_c. The assumption that $F_c < 110$ pN seems, however, fully consistent with the experimental $\tau(F)$ data.

In this Rapid Communication we have shown for a model protein that the unfolding process from the native valley under force-clamp conditions can be modeled as a Brownian motion in a tilted one-dimensional free-energy landscape. Moreover, it turned out that this description could be further simplified with a surprisingly small loss of accuracy, by adopting a linear approximation for the free energy. This analysis links deviations from the Bell formula for $k(F)$ for $F > F_c$ to an altered shape of the escape-time distribution, from exponential to inverse Gaussian. Comparison with experiments indicates that the strong-force regime might set in at relatively weak force ($F_c \lesssim 100$ pN) for both ddFLN4 and ubiquitin.

We acknowledge S. Lepri and A. T. Sanesi for useful discussions.

[1] J. R. Forman and J. Clarke, Curr. Opin. Struct. Biol. **17**, 58 (2007).
[2] E. Evans and K. Ritchie, Biophys. J. **72**, 1541 (1997).
[3] G. I. Bell, Science **200**, 618 (1978).
[4] O. K. Dudko, G. Hummer, and A. Szabo, Phys. Rev. Lett. **96**, 108101 (2006).
[5] H. J. Lin, H. Y. Chen, Y. J. Sheng, and H. K. Tsao, Phys. Rev. Lett. **98**, 088304 (2007).
[6] R. W. Friddle, Phys. Rev. Lett. **100**, 138302 (2008).
[7] O. K. Dudko, G. Hummer, and A. Szabo, Proc. Natl. Acad. Sci. U.S.A. **105**, 15755 (2008).
[8] M. Schlierf and M. Rief, Biophys. J. **90**, L33 (2006).
[9] P. Hänggi, P. Talkner, and M. Borkovec, Rev. Mod. Phys. **62**, 251 (1990).
[10] Z. T. Yew, T. McLeish, and E. Paci, Phys. Rev. Lett. **101**, 248104 (2008).
[11] L. Dougan, G. Feng, H. Lu, and J. M . Fernandez, Proc. Natl. Acad. Sci. U.S.A. **105**, 3185 (2008).
[12] S. Garcia-Manyes, J. Brujic, C. L. Badilla, and J. M. Fernandez, Biophys. J. **93**, 2436 (2007).
[13] S. Garcia-Manyes, L. Dougan, C. L. Badilla, J. Brujic, and J. M. Fernandez, Proc. Natl. Acad. Sci. U.S.A. **106**, 10534 (2009).
[14] J. D. Honeycutt and D. Thirumalai, Proc. Natl. Acad. Sci. U.S.A. **87**, 3526 (1990).
[15] R. S. Berry, N. Elmaci, J. P. Rose, and B. Vekhter, Proc. Natl. Acad. Sci. U.S.A. **94**, 9520 (1997).
[16] T. Veitshans, D. Klimov, and D. Thirumalai, Folding Des. **2**, 1 (1997).
[17] Z. Guo and C. L. Brooks III, Biopolymers **42**, 745 (1997); J. G. Kim, J. E. Straub, and T. Keyes, Phys. Rev. Lett. **97**, 050601 (2006).
[18] F.-Y. Li, J. M. Yuan, and C. Y. Mou, Phys. Rev. E **63**, 021905 (2001); D. J. Lacks, Biophys. J. **88**, 3494 (2005).
[19] A. Imparato, S. Luccioli, and A. Torcini, Phys. Rev. Lett. **99**, 168101 (2007); S. Luccioli, A. Imparato, and A. Torcini, Phys. Rev. E **78**, 031907 (2008).
[20] S. Redner, *A Guide to First Passage Processes* (Cambridge University Press, Cambridge, England, 2001).
[21] R. B. Best, E. Paci, G. Hummer, and O. K. Dudko, J. Phys. Chem. B **112**, 5968 (2008).
[22] C. W. Gardiner, *Handbook of Stochastic Methods* (Springer, Berlin, 2004).
[23] See supplementary material at http://link.aps.org/supplemental/10.1103/PhysRevE.81.010902 reporting the expressions for the variance V_L and for the average and the variance of the unfolding times in the strong force limit, as well as a description of the experimental data for ddFLN4 reported in Fig. 4.
[24] R. S. Chikara and J. L. Folks, *The Inverse Gaussian Distribution* (Marcel Dekker, New York, 1988).
[25] A. Imparato, A. Pelizzola, and M. Zamparo, Phys. Rev. Lett. **98**, 148102 (2007); J. Chem. Phys. **127**, 145105 (2007).
[26] P. Szymczak and M. Cieplak, J. Phys.: Condens. Matter **18**, L21 (2006).