Reconstructing the free-energy landscape of a polyprotein by single-molecule experiments

A. Imparato¹,²(a), F. Sbrana³ and M. Vassalli³,⁴

¹ Dipartimento di Fisica and CNISM, Politecnico di Torino - c. Duca degli Abruzzi 24, Torino, Italy, EU
² INFN, Sezione di Torino - Torino, Italy, EU
³ CSDC-Dipartimento di Fisica, Università di Firenze - via Sansone, 1, Sesto Fiorentino, Italy, EU
⁴ Istituto Sistemi Complessi, CNR - via Madonna del Piano 10, Sesto Fiorentino, Italy, EU

received 21 December 2007; accepted in final form 17 April 2008
published online 29 May 2008

PACS 82.37.Rs – Single molecule manipulation of proteins and other biological molecules
PACS 87.15.La – Mechanical properties
PACS 05.70.Ln – Nonequilibrium and irreversible thermodynamics

Abstract – The mechanical unfolding of an engineered protein composed of eight domains of Ig27 is investigated by using atomic force microscopy. Exploiting a fluctuation relation, the equilibrium free energy as a function of the molecule elongation is estimated from pulling experiments. Such a free energy exhibits a regular shape that sets a typical unfolding length at zero force of the order of 20 nm. This length scale turns out to be much larger than the kinetic-unfolding length that is also estimated by analyzing the typical rupture force of the molecule under dynamic loading.

Force spectroscopy techniques have enormously increased our knowledge on the structure of biopolymers such as proteins and nucleic acids. The possibility of controlling very precisely the force applied to the free ends of such molecules allowed the experimental evaluation of the typical lengths and energies of the bonds stabilizing the molecular structures [1–9]. The mechanical unfolding of a protein is typically an out-of-equilibrium process, where the unfolding occurs in time scales much shorter than the typical molecule relaxation time: this prevents the possibility of performing the experiments in quasiequilibrium conditions and thus of obtaining direct measurements of thermodynamic variables. However, this problem can be overcome by using the remarkable equality derived by Jarzynski [10], which allows one to measure the free-energy difference between the folded state and the unfolded one of a biomolecule [11]. By exploiting an extended version of the Jarzynski equality (JE) the free-energy landscape of model biopolymers as a function of the molecular elongation has been evaluated [12–18]. Furthermore in a recent experimental work [19], the free-energy landscape of a molecule similar to the one we use here has been reconstructed, in a range of the molecular elongation corresponding the unfolding of a single domain (see below for a discussion of the protein structure). Usually, the information concerning the landscape of a protein, obtained by single-molecule experiments, has been limited to the position and to the height of the energy barrier along the reaction coordinate, which is usually the molecule elongation. The actual shape of the landscape can be only guessed by adapting it to the information about the kinetics gathered during the unzipping experiments, see, e.g., [20]. However, in a recent paper [16] it has been argued that the unfolding length and the unfolding rate that are obtained by investigating the kinetics of protein unfolding, are not simply related to the position and height of free-energy barriers as given by the extended JE. In other words, the kinetic parameters governing the protein unfolding may not correspond to the equilibrium thermodynamical properties of the molecules.

The aim of this paper is thus to show that it is possible to reconstruct the free-energy landscape (FEL) of a protein for any value of its elongation. Furthermore, we aim to compare the typical length scale of this landscape with the unfolding length measured in kinetic experiments, and to discuss the meaning of this parameter in the context of the equilibrium properties of proteins.

Unfolding of large proteins is a very rare and slow event, but by using single-molecule techniques it is possible to drive mechanical unfolding and make it feasible to study
the unfolding kinetics of these molecules. The molecule investigated in this work is a recombinant polyprotein composed of eight repeats of the Ig27 domain. The typical distance between two consecutive peaks ranges between 22 and 26 nm.

the AFM. Bottom panel: force extension curve of the molecule force spectroscopy experiments exploiting an atomic force microscope (AFM), the molecule deposited onto the substrate is grabbed by the AFM probe. Mechanical unfolding of the protein is thus induced by moving the probe away from the substrate with a constant velocity \( v \) (linear protocol). As the probe is retracted, the force exerted on the molecule increases until the molecule (linear protocol). As the probe is retracted, the force exerted on the molecule increases until the molecule

\[
\begin{align*}
\ell &= \text{the elongation of the molecule and } \\
z(t) &= \text{the actual position of the probe with respect to the substrate (see fig. 1). In the present work, mechanical-unfolding}
\end{align*}
\]

Let us now consider the case where a molecular bond, or a group of molecular bonds, is subject to a force increasing linearly with the time \( f = rt \), where \( r \) is the force increase rate. It can be shown that the rupture force is distributed according to the probability distribution function [21]

\[
P(f) = \frac{1}{\tau_0} f e^{\beta f x_u} \exp \left[ - \frac{k_B T}{x_u \tau_0} \left( e^{\beta f x_u} - 1 \right) \right],
\]

and the typical rupture force, defined as the maximum of \( P(f) \), is given by

\[
f^* = \frac{k_B T}{x_u} \ln[\beta r x_u \tau_0],
\]

where \( x_u \) is the unfolding length, usually interpreted as the deformation after which the molecule unfolds, \( \tau_0 \) is the typical unfolding time at zero force, and \( \beta = 1/(k_B T) \). Since we use a rather soft spring \((k = 0.04 \text{N/m})\), we take the rate of increase of the force in eqs. (1), (2) to be \( r = kv \). By plotting the force as a function of the time, we verified that this is a good approximation for the force rate (data not shown). Because the AFM tip can bind any module of the protein at random, one is not able to control the site from which the protein is picked up [6,7]. Therefore, we obtained a random sample of single-molecule unfolding trajectories containing from one to eight unfolding events. In fig. 2 we plot the typical rupture force \( f^* \) of a single Ig27 module as a function of the AFM probe velocity \( v \), where trajectories with at least three repeats (i.e., rupture events) are considered. By fitting the data to eq. (2) we obtain an estimate of the unfolding length \( x_u = 0.30 \pm 0.07 \text{nm} \), which is in good agreement with the value \( x_u = 0.25 \text{nm} \) found in [2].
We now focus on the evaluation of the free-energy landscape $F_0(\ell)$, which is a function of the molecular elongation $\ell$. As discussed in ref. [12], such an energy landscape can be obtained by exploiting an extended version of the Jarzynski equality
\[
\langle e^{\beta W} \rangle_t e^{\beta U(\ell, z(t))} = \exp \left[ -\beta F_0(\ell) \right] / Z_0, \tag{3}
\]
where $W$ is the work done on the molecule during the unfolding process, $Z_0$ is the partition function of the unperturbed system (folded molecule with $f = 0$), and $U(\ell, z(t)) = k/2(z(t) - \ell)^2$ is the external potential associated to the tip of the AFM. Note that when one takes the logarithm of the rhs of eq. (3), the partition function appears in the additive constant $-\log Z_0$, and thus it plays no role in the setting of $F_0(\ell)$. In the following, we do not consider all the unfolding trajectories that we used to obtain the data in fig. 2, but we take only those trajectories with at least six rupture events. For each of these trajectories, we compute the work $W$ done on the molecule by the external force $f(t) = k/z(t) - \ell(t))$. Practical procedures for obtaining the optimal estimate of $F_0(\ell)$ from eq. (3) are discussed in refs. [12–14]. However, as the value of the work $W$ exceeds few hundreds of $k_BT$, eq. (3) gives unreliable results, since in that case the average is estimated by considering very small numbers ($\exp(-\beta W) \ll 1$).

In order to avoid this numerical problem, we calculate here the average of the quantity $\exp(-\beta W + \Delta)$, where $\Delta$ is a fixed quantity, so as eq. (3) provides an estimate of $\exp(-\beta F_0(L + \Delta)]$. By taking different values for $\Delta$ one selects ranges of values of the work $W$ such that $W + \Delta$ is not larger than a few hundreds of $k_BT$, and thus one can reconstruct the free-energy landscape $F_0(\ell)$ piecewise. In particular we take $\Delta = 0, 1000, 2000, 3000, 4000 k_BT$.

In refs. [14,15] it was argued that eq. (3) supplies a reliable estimate of the molecule free energy, provided that, as the pulling rate is decreased, the estimated curves $F_0(\ell)$ collapse onto a single curve. In fig. 3, the free-energy landscape $F_0(\ell)$ as reconstructed from eq. (3) is plotted for different values of the AFM tip velocity $v$. We actually observe a collapse of the curves for the two smallest velocities, while the agreement with the curve obtained with the largest value of $v = 2000 \text{ nm/s}$ worsens as the coordinate $\ell$ increases.

The energy landscape $F_0(\ell)$ exhibits a single minimum at $\ell = 0$, corresponding to the molecule in the folded native state. However, the curve $F_0(\ell)$ exhibits cusps which are equispaced at a distance $\Delta \ell \approx 20 \text{ nm}$. As a constant force $f$ is applied, the free-energy landscape is tilted according to $F(\ell, f) = F_0(\ell) - f\ell$. In fig. 4 we plot the FEL, as obtained with the smallest velocity here considered, for different values of the force $f$. As the external force $f$ is applied, the landscape $F(\ell, f)$ becomes more intricate: the cusps become local maxima, delimiting equispaced minima. The overcoming of each such cusp corresponds thus to the breaking of a single Ig27 domain. For each value of the force, a well-defined global minimum appears in the function $F(\ell, f)$, therefore $F(\ell, f)$ predicts how many domains out of the eight will be unfolded at equilibrium for that value of the force. As an example, we find that the second minimum becomes deeper than the first one when $f \geq 50 \text{ pN}$, see inset of fig. 4. Thus, we expect that in a force clamp experiment, where a feedback system allows to apply a constant force to the molecule free ends [7,8], this would be the constant force required to unfold a single Ig27 domain. In a different experimental set-up, where force-ramp pulling was used, the first Ig27 domain was found to unfold at a
typical force $f \simeq 75\,\text{pN}$ (see fig. 5b of ref. [5]): such a value is consistent with $f \simeq 50\,\text{pN}$ predicted by our results.

We want to stress that at zero force the distance between the global minimum $\ell = 0$ and the first cusp is $\Delta \ell(f = 0) \simeq 20\,\text{nm}$, while at $f = 50\,\text{nm}$ the distance between the first minimum and the following maximum is $\Delta \ell(f = 50) \simeq 6\,\text{nm}$. These lengths are one order of magnitude larger than the kinetic-unfolding length $x_u = 0.25\,\text{nm}$ found in [2] and in the present work. Thus, as discussed in ref. [16], the quantity $x_u$ appearing in eq. (2) represents a kinetic parameter, and does not correspond to the typical length of the free-energy landscape: at zero force the thermal parameter, and does not correspond to the typical length scale of the protein. The typical length scale of this landscape is $20\,\text{nm}$, which is larger than the typical unfolding length $x_u$ as obtained by analyzing the typical rupture force. We want to remark that in ref. [2] the length $x_u = 0.25\,\text{nm}$ was identified as the distance of the free-energy barrier from the folded state. Our results suggest, on the contrary, that, at $f = 0$, the distance between the folded state and the first energy barrier is of the order of $20\,\text{nm}$. We attribute this difference to the fact that the quantity $x_u$ as obtained by eq. (2) is a kinetic parameter describing the unfolding kinetics of the Ig27 domain, while the extended JE (3) provides an equilibrium quantity, namely the equilibrium FEL.

It is worth noting that as our approach is successful to reconstruct the FEL of a protein made up of identical domains in the whole range of the molecule extension, we believe that the same approach can be successfully used to probe the FEL of heterogeneous proteins made up of different domains.

***

FS and MV acknowledge the support of the Strep European Project n. 12835 EMBIO (EMergent organization in complex BIOMolecular system).

REFERENCES

[1] Kellermayer M. S., Smith S. B., Granzier H. L. and Bustamante C., Science, 276 (1997) 1112.
[2] Carrion-Vasquez M. et al., Proc. Natl. Acad. Sci. U.S.A., 96 (1999) 3694.
[3] Carrion-Vasquez M. et al., Nat. Struct. Biol., 10 (2003) 738; Dietz H. and Rief M., Proc. Natl. Acad. Sci. U.S.A., 101 (2004) 16192.
[4] Dietz H., Berkemeier F., Bertz M. and Rief M., Proc. Natl. Acad. Sci. U.S.A., 103 (2006) 12724.
[5] Oberhauser A. F., Hansma P. K., Carrion-Vazquez M. and Fernandez J. M., Proc. Natl. Acad. Sci. U.S.A., 98 (2001) 468.
[6] Li H. et al., Proc. Natl. Acad. Sci. U.S.A., 97 (2000) 6527.
[7] Fernandez J. M. and Li H., Science, 303 (2004) 1674.
[8] Schlierf M., Li H. and Fernandez J. M., Proc. Natl. Acad. Sci. U.S.A., 101 (2004) 7299.
[9] Danilowicz C. et al., Proc. Natl. Acad. Sci. U.S.A., 100 (2003) 1694; Liphardt J. et al., Science, 292 (2001) 733; Onoza B. et al., Science, 299 (2003) 1892.
[10] Jarzynski C., Phys. Rev. Lett., 78 (1997) 2690; Phys. Rev. E, 56 (1997) 5018; Crooks G. E., J. Stat. Phys., 90 (1998) 1481; Phys. Rev. E, 61 (1999) 2361.
Reconstructing the free-energy landscape of a polyprotein by single-molecule experiments

[11] Liphardt J., Dumont S., Smith S., Tinoco I. and Bustamante C., *Science*, 296 (2002) 1832.
[12] Hummer G. and Szabo A., *Proc. Natl. Acad. Sci. U.S.A.*, 98 (2001) 3658.
[13] Braun O., Hanke A. and Seifert U., *Phys. Rev. Lett.*, 93 (2004) 158105.
[14] Imparato A. and Peliti L., *J. Stat. Mech.* (2006) P03005.
[15] Imparato A., Pelizzola A. and Zamparo M., *Phys. Rev. Lett.*, 98 (2007) 148102.
[16] Imparato A., Pelizzola A. and Zamparo M., *J. Chem. Phys.*, 127 (2007) 145105.
[17] Imparato A., Luccioli S. and Torcini A., *Phys. Rev. Lett.*, 99 (2007) 168101.
[18] Luccioli S., Imparato A. and Torcini A., preprint arXiv:0803.1811 (2008).
[19] Harris N. C., Song Y. and Kiang C.-H., *Phys. Rev. Lett.*, 99 (2007) 068101.
[20] Schierf M. and Rief M., *Biophys. J.*, 90 (2006) L33.
[21] Evans E., *Annu. Rev. Biophys. Biomol. Struct.*, 30 (2001) 105.
[22] Efron B., *The Jackknife, the Bootstrap, and Other Resampling Plans*, in *CBMS-NSF Reg. Conf. Ser. Appl. Math.*, Vol. 38 (SIAM) 1982.