BIOINFORMATICS APPLICATIONS NOTE

Vol. 27 no. 3 2011, pages 423–425
doi:10.1093/bioinformatics/btq663

*Structural bioinformatics*

Advance Access publication December 1, 2010

# Open source platform for the execution and analysis of mechanical refolding experiments

Daniel Aioanei*, Marco Brucale and Bruno Samorí

Department of Biochemistry 'G.Moruzzi', University of Bologna, Via Irnerio 48, 40126 Bologna, Italy

Associate Editor: Anna Tramontano

**ABSTRACT**

**Motivation:** Single-molecule force spectroscopy has facilitated the experimental investigation of biomolecular force-coupled kinetics, from which the kinetics at zero force can be extrapolated via explicit theoretical models. The atomic force microscope (AFM) in particular is routinely used to study protein unfolding kinetics, but only rarely protein folding kinetics. The discrepancy arises because mechanical protein refolding studies are more technically challenging.

**Results:** We developed software that can drive and analyse mechanical refolding experiments when used with the commercial AFM setup 'Picoforce AFM', Bruker (previously Digital Instruments). We expect the software to be easily adaptable to other AFM setups. We also developed an improved method for the statistical characterization of protein folding kinetics, and implemented it into an AFM-independent software module.

**Availability:** Software and documentation are available at http://code.google.com/p/refolding under Apache License 2.0.

**Contact:** aioaneid@gmail.com

**Supplementary information:** Supplementary data are available at *Bioinformatics* online.

Received on September 13, 2010; revised on November 3, 2010; accepted on November 27, 2010

## 1 INTRODUCTION

Biochemical reactions commonly proceed via large conformational changes, resulting in a well-defined mechanical reaction coordinate on which they can be monitored. Since force is a determinant factor in the rate of such reactions, single-molecule force spectroscopy (SMFS) emerged as an invaluable tool in their investigation under mechanical tension (Bustamante *et al.*, 2004; Kumar and Li, 2010). Thanks to its remarkable ability to stretch and monitor one molecule at a time, SMFS seeks to achieve the long-standing goal of mapping the energy landscape of biomolecules (e.g. proteins, RNA) on a well-defined reaction coordinate, even for proteins which show irreversible thermal or chemical unfolding (Jollymore and Li, 2010).

In particular, protein folding kinetics can be studied at the single molecule level using the atomic force microscope (AFM). To this end, protein modules can be first unfolded and subsequently allowed to refold while subjected to a 'force-clamp' (Fernandez and Li, 2004; Garcia-Manyes *et al.*, 2007, 2009b), or they can be directly observed to refold at fixed extension via 'lock-in force spectroscopy' (Garcia-Manyes *et al.*, 2009a). Such techniques depend on recent technological advances implemented in custom-built AFMs with very limited availability. Alternatively, the AFM can be used in the more traditional 'velocity-clamp' mode to drive protein modules to fold under mechanical tension via the 'double-pulse protocol' (see Suplementary Material). Shortly, the distance between the base of the cantilever and the surface, rather than the stretching force, is maintained constant for some amount of time, allowing previously unfolded protein modules to refold (Bullard *et al.*, 2006; Cao and Li, 2007; Carrion-Vazquez *et al.*, 1999).

Despite the widespread availability of AFM instrumentation supporting the velocity-clamp mode of operation, single molecule folding kinetics studies remain rather scarce in the scientific literature, likely due to the unavailability of mandatory software technology. We fill this gap by making such software freely available. To validate our software, we studied the folding kinetics of protein GB1 (Cao *et al.*, 2006) and obtained a kinetics characterization similar to the previously published one (Cao and Li, 2007).

## 2 APPROACH

Our software contains three main components:

(1) An automated procedure for driving refolding experiments through Nanoscope v6 software, in conjunction with Picoforce AFM and Nanoscope IIIa controller, Bruker. For each execution of the double-pulse protocol, our software instructs the Nanoscope software to execute a 'Nanoscope script' and capture a 'Nanoscope strip chart'. Importantly, the actual bending of the cantilever is detected from the strip chart file and the starting position is adjusted accordingly for the next double-pulse so as to counterbalance accumulating drift (Oberhauser *et al.*, 2001). Section 1 in Supplementary Material contains more information on the implemented double-pulse protocol.

(2) Offline tools for automated peak identification, force measurements, Worm-like chain (Bustamante *et al.*, 1994) fits and data filtering. Tools are also included for manually improving the results of some of the automated tasks such as zero-force baseline and contact point identification.

(3) A standalone, AFM-independent offline procedure for the statistical characterization of protein folding kinetics from mechanical refolding experiments with homomeric polyproteins, based on the analytical model of Section 3.1.

*To whom correspondence should be addressed.

## 3 METHODS

### 3.1 Maximum likelihood estimation of folding kinetic parameters

We adopt Bell's approximation (Bell, 1978; Walcott, 2008) to Kramers' reaction-rate theory (Hänggi *et al.*, 1990), which describes the force-dependent folding rate as $k_f(F) = k_f^0 \exp[-F\Delta x_f/(K_b T)]$, where $K_b$ is the Boltzmann constant, $T$ is the temperature in Kelvin, $k_f^0$ is the spontaneous folding rate and $\Delta x_f$ is the folding distance. We aim to extract the last two mentioned parameters from refolding experiments.

Traditionally, data were collected in a few fixed configurations, where a configuration is defined by the amount of time allowed for refolding and the inferred force at the start of the waiting period. It was therefore assumed that the same configuration can be reproduced exactly multiple times, ignoring any variation between double-pulse protocol executions. It was also assumed that the residual force remains constant during the waiting time (Cao and Li, 2007; Carrion-Vazquez *et al.*, 1999), ignoring the fact that it increases after each folding event. Furthermore, it was assumed that the total number of modules that could refold is limited by the extension during the waiting period (Cao and Li, 2007), breaking the assumption of the ideal spring cantilever. The mean and SD of the refolding ratio would then be computed for each configuration and then fitted to an exponential formula based on Bell's equation or to Monte Carlo simulations based on it, ignoring the fact that such summary statistics are not *sufficient* (Lehmann and Casella, 1998), i.e. they do not capture all possible information about the parameters.

We overcome all the above limitations by introducing a Maximum Likelihood estimation procedure. Shortly, the stretching force is computed by solving the WLC cubic equation (Aioanei *et al.*, 2009), and the likelihood function is computed as the product of the probability to observe the actual number of folding events for each double-pulse protocol execution. The maximum likelihood is then located over a grid of $(k_f^0, \Delta x_f)$ values, and estimation errors are computed by case resampling (see Section 2 in Supplementary Material).

### 3.2 Folding kinetics of protein GB1

We estimated the kinetic parameters of protein GB1 in buffer Tris–HCl (10 mM, pH 7.5) by performing mechanical refolding experiments with homomeric polyproteins $(GB1)_8$ and $(GB1)_{16}$ (see Section 3 in Supplementary Material for experimental data statistics). A sample trace can be seen in Figure 1.

We obtained the kinetic parameters $\Delta x_f = 2.53 \pm 0.12$ nm and $k_f^0 = 500 \pm 85 s^{-1}$ errors representing one SD. Our kinetics characterization is roughly compatible with previously published values of $\Delta x_f = 2.1$ nm and $k_f^0 = 720 \pm 120 s^{-1}$ (Cao and Li, 2007).

## 4 DISCUSSION

Mechanical refolding experiments can be performed with typical commercial velocity-clamp AFM instrumentation, and we provide an out-of-the-box software solution for performing and analysing such experiments in conjunction with Picoforce AFM, Bruker. We expect our software to be easily adaptable to other AFM setups. In fact, since the analytical model of Section 3.1 is not specific to a particular AFM, its implementation can already be used with refolding data obtained with any other AFM. Furthermore, we developed all the software in the Java and Python programming languages to ensure its portability across all major operating systems.

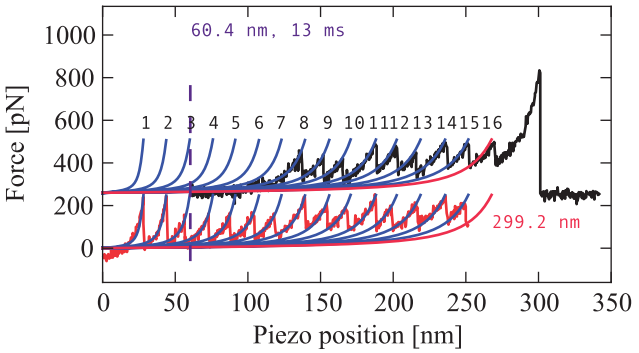


**Fig. 1.** (Colour online) A force-extension trace according to the double-pulse protocol with the homomeric polyprotein $(GB1)_{16}$. The lower curve represents the protein fetching phase, during which the polyprotein attached non-specifically to the cantilever and then a total of 15 modules have been subsequently unfolded. The higher curve is shifted by 260 pN just for display purposes, and it represents the phase where the same molecule is pulled for the second time. Note that only 15 out of the 16 modules could have refolded, since one module was not unfolded during the fetching phase. The vertical dashed line represents the piezo position during the waiting time lapse relative to the resting position of the cantilever tip, and its numerical value is shown together with the length of the time lapse at the top of the figure. Each WLC fit is shown redundantly shifted higher for display purposes. The contour length at the start of the waiting time lapse (before any refolding) is indicated in the bottom-right position.

## ACKNOWLEDGEMENTS

We thank Prof. Hongbin Li, University of British Columbia, Vancouver, Canada, for kindly providing the $(GB1)_8$ plasmid, Dr Isabella Tessari and Prof. Luigi Bubacco, University of Padova, Italy, for providing the $(GB1)_{16}$ construct.

*Funding*: Ministero dellUniversitá e della Ricerca Fondo per gli Investimenti della Ricerca di Base (MIUR FIRB) RBNE03PX83/001; MIUR FIRB Progetto NG-lab (G.U. 29/07/05 n.175); Progetti di Ricerca di Interesse Nazionale (PRIN) 2008 (Prot. 2008KZ3E5).

*Conflict of Interest*: none declared.

## REFERENCES

Aioanei,D. *et al.* (2009) Maximum likelihood estimation of protein kinetic parameters under weak assumptions from unfolding force spectroscopy experiments. *Phys. Rev. E*, **80**, 061916.

Bell,G. (1978) Models for the specific adhesion of cells to cells. *Science*, **200**, 618–627.

Bullard,B. (2006) The molecular elasticity of the insect flight muscle proteins projectin and kettin. *Proc. Natl Acad. Sci. USA*, **103**, 4451–4456.

Bustamante,C. *et al.* (1994) Entropic elasticity of lambda-phage DNA. *Science*, **265**, 1599–1600.

Bustamante,C. *et al.* (2004) Mechanical processes in biochemistry. *Annu. Rev. Biochem.*, **73**, 705–708.

Cao,Y. and Li,H. (2007) Polyprotein of GB1 is an ideal artificial elastomeric protein. *Nat. Mater.*, **6**, 109–114.

Cao,Y. *et al.* (2006) Nonmechanical protein can have significant mechanical stability. *Angew. Chem.*, **45**, 642–645.

Carrion-Vazquez,M. *et al.* (1999) Mechanical and chemical unfolding of a single protein: a comparison. *Proc. Natl Acad. Sci. USA*, **96**, 3694–3699.

Fernandez,J.M. and Li,H. (2004) Force-clamp spectroscopy monitors the folding trajectory of a single protein. *Science*, **303**, 1674–1678.

Garcia-Manyes,S. *et al*. (2007) Force-clamp spectroscopy of single-protein monomers reveals the individual unfolding and folding pathways of I27 and ubiquitin. *Biophys. J.*, **93**, 2436–2436.

Garcia-Manyes,S. *et al*. (2009a) Direct observation of an ensemble of stable collapsed states in the mechanical folding of ubiquitin. *Proc. Natl Acad. Sci. USA*, **106**, 10534–10539.

Garcia-Manyes,S. *et al*. (2009b) Osmolyte-induced separation of the mechanical folding phases of ubiquitin. *Proc. Natl Acad. Sci. USA*, **106**, 10540–10545.

Hänggi,P. *et al*. (1990) Reaction-rate theory: fifty years after kramers. *Rev. Mod. Phys.*, **62**, 251–341.

Jollymore,A. and Li,H. (2010) Measuring "unmeasurable" folding kinetics of proteins by single-molecule force spectroscopy. *J. Mol. Biol.*, **402**, 610–617.

Kumar,S. and Li,M.S. (2010) Biomolecules under mechanical force. *Phys. Rep.*, **486**, 1–74.

Lehmann,E.L. and Casella,G. (1998) *Theory of Point Estimation*. 2nd edn. Springer, New York, NY, USA.

Oberhauser,A.F. *et al*. (2001) Stepwise unfolding of titin under force-clamp atomic force microscopy. *Proc. Natl Acad. Sci. USA*, **98**, 468–472.

Walcott,S. (2008) The load dependence of rate constants. *J. Chem. Phys.*, **128**, 215101.