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.


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 Supplementary 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 (Ballard 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 homo- or polyproteins, based on the analytical model of Section 3.1.
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 \( k_\text{f}(F) = k_\text{f}^0 \exp\left(-F \Delta \chi / (k_\text{B} T)\right) \), where \( k_\text{f}^0 \) is the Boltzmann constant, \( T \) is the temperature in Kelvin, \( k_\text{f}^0 \) is the spontaneous folding rate and \( \Delta \chi \) 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 from 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 describe the following outlines by introducing a Maximum Likelihood estimation procedure. Shortly, the stretching force is computed from 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 describe the following outlines by introducing a Maximum Likelihood estimation procedure. Shortly, the stretching force is computed from 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.

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 homometric polypeptides (GB1(16) and (GB1)16 to (GB1)16 construct (see Section 3 in Supplementary Material). A sample trace can be seen in Figure 1.

We obtained the kinetic parameters \( \Delta \chi = 2.53 \pm 0.12 \text{nm} \) and \( k_\text{f}^0 = 500 \pm 85 \text{c}^{-1} \text{s}^{-1} \) errors representing one SD. Our kinetics characterization is roughly compatible with previously published values of \( \Delta \chi = 2.1 \text{nm} \) and \( k_\text{f}^0 \) is the spontaneous folding rate at 720 \( \pm 120 \text{c}^{-1} \text{s}^{-1} \) (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.
Software for protein refolding studies


