

# Technical Brief:

## The I27<sup>RS8</sup> Protein as a Reference for Force Extension

Sheldon E. Broedel, Jr., Ph. D.

Athena Environmental Sciences, Inc., Baltimore, MD



Athena Enzyme Systems™

1450 South Rolling Road

Baltimore, MD 21227

www.athenaes.com

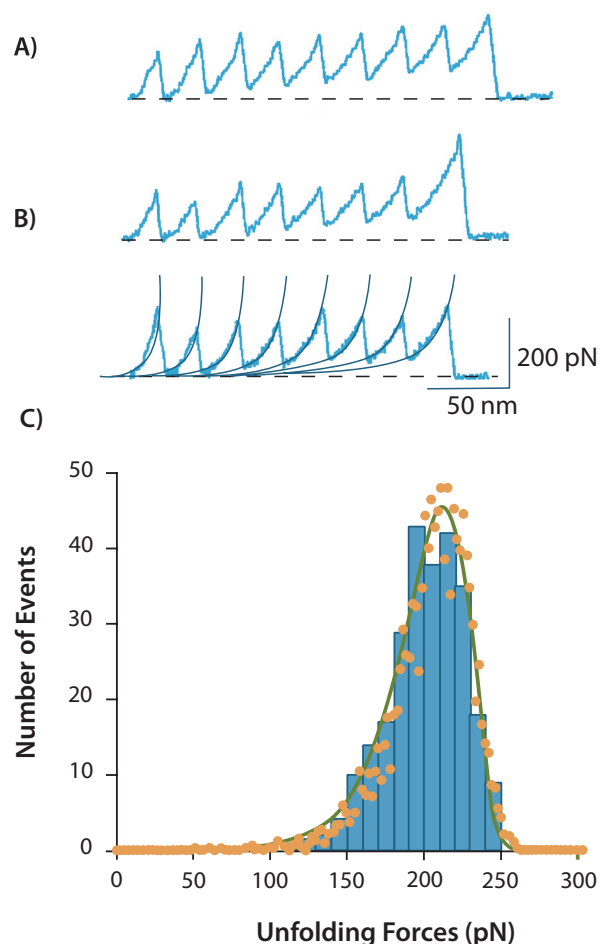
### The I27<sup>RS8</sup> Protein as a Reference for Force Extension

August 10, 2004

Atomic force microscopy (AFM) is a remarkably simple instrument. Originally used for measuring surface contours of individual molecules, it has provided important information about the structure of proteins. In the force measuring mode, AFM is capable of measuring forces down to piconewtons and can resolve force changes caused by the displacement of its probe by a fraction of a nanometer. Single molecule measurements are routinely done (for reviews see references 2, 3). Since measurements can be made on many individual molecules in a short period of time, a statistical evaluation of the measurements can be performed. Perhaps the most powerful aspect of AFM is that the measurements can be done in an aqueous environment, allowing for the study of biological material under conditions that resemble those *in vivo*. The folding and unfolding of the protein can be studied in the presence of substrate, product, and co-factors. The temperature can be controlled as can the buffer conditions, *i.e.*, ionic strength, pH, solvents, etc.

The mechanical unraveling of proteins using atomic force microscopy is a powerful tool for the study of protein structure. Single molecules stretched between the AFM probe and stage result in force-extension recording which provide information about the stability and kinetics of protein and protein domain unfolding. AFM in a force extension mode offers several advantages over thermal or chemical denaturation to the study of protein structure. Until recently, the only way to measure the stability of a protein was to change its physical (*e.g.*, with heat or high pressure) or chemical environment (*e.g.*, with guanidinium chloride or urea, acidic or alkaline pH) and monitor the loss of protein conformation, using spectroscopic techniques from which folding free energies could be obtained. AFM allows the measurement of folding reactions at the single-molecule level and therefore can uncover rare unfolding events that are not observed by using traditional bulk methods, in which measured parameters are averaged over many molecules. It allows a direct measurement of the mechanical stability of the fold. The process of mechanical unfolding is therefore defined as a function of extension (here the reaction coordinate is well defined, *i.e.* length), which in the case of proteins exposed to stress *in vivo*, may be a more physiological perturbation. In contrast to other structural techniques (*i.e.*, X-ray crystallography), force spectroscopy is dynamic and it works in real time. AFM works in aqueous solution, so crystal formation is not required. Thus, force extension using AFM promises to elucidate the dynamic

structural properties of a wide variety of proteins at the single molecule level and provide an important complement to other structural and dynamic techniques (*e.g.*, X-ray crystallography, NMR spectroscopy).



**Figure 1.** Force-extension relationships for recombinant poly-I27 measured with AFM techniques. A) Sawtooth pattern of unfolding observed with the I2711 polyprotein. B) Stretching of single I27<sup>RS8</sup> polyproteins gave force-extension curves with a sawtooth pattern with equally spaced force peaks. The sawtooth pattern is well described by the WLC equation (B, bottom trace, continuous lines). C) Unfolding force frequency histogram for I27<sup>RS8</sup>. The filled circles correspond to a Monte Carlo simulation of the mean unfolding forces ( $n=1,000$ ) of eight domains placed in series using an unfolding rate constant,  $k_u^\circ$ , of  $3.3 \times 10^{-4} \text{ s}^{-1}$ , an unfolding distance,  $\Delta x_u$ , of 0.25 nm and a  $V = 0.6 \text{ nm/ms}$ .

When working with a protein with unknown unfolding properties, it is helpful to employ a reference protein as a means of calibrating the AFM instrumentation. The I27<sup>RS8</sup> engineered polyprotein is ideal for this purpose. I27<sup>RS8</sup> is composed of eight direct repeats of the Ig module 27 of the human titin protein. The construction and characterization of the polyprotein has been previously described by Dr. Julio Fernandez and his colleagues.<sup>4</sup> In this paper, the properties of the protein are reviewed in the context of using it as a reference in AFM analyses.

## Force Extension Properties of I27<sup>RS8</sup>

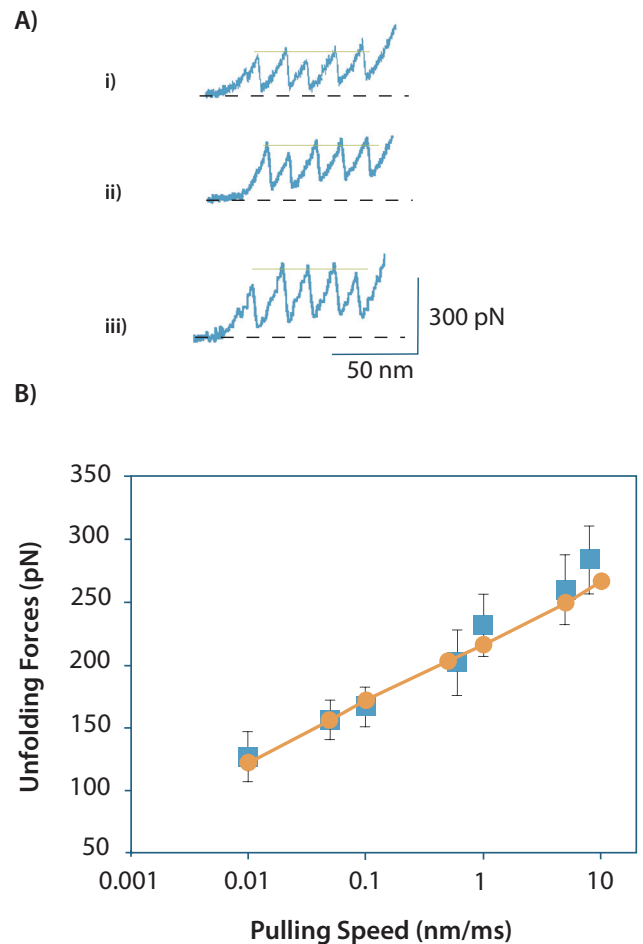
### Forced unfolding of the polyprotein

The force-extension relationships for the I27<sup>RS8</sup> polyprotein were measured using a single axis atomic force microscope. Single proteins were fully extended by retracting the sample holder away from the AFM tip. Stretching the I27<sup>RS8</sup> polyprotein resulted in a force-extension curve with peaks that varied randomly in amplitude with an average value of ~200 pN (Fig. 1). A histogram of force peaks (unfolding force) measured from 266 unfolding events revealed an asymmetrical distribution of events that had a maximum at ~200 pN (Fig. 1C). Hence, the only variations in the force required to unfold were stochastic, consistent with the idea that during stretching, the tandem I27 modules unfold independently from each other.

The force extension curves of I27<sup>RS8</sup> are well described by the worm-like chain model (WLC) which predicts the entropic restoring force generated upon the extension of a polymer (Fig. 1B, thin lines). Fits of the WLC to the force-extension curves of I27<sup>RS8</sup> gave a persistence length of  $p = 0.39 \pm 0.07 \text{ nm}$  ( $n=10$ ). The persistence length is a measure of the distance over which the polymer retains memory of a direction. It is significant that the persistence length is of the size of a single amino acid. Consecutive peaks were fitted by the WLC with the same persistence length and a contour length increment of  $\Delta l_c = 28.4 \pm 0.3 \text{ nm}$  (Fig 1B, thin lines;  $n=16$ ). The increase in contour length of a module upon unfolding is 28.4nm, which is approximately equal to the expected length of 89 fully extended amino acids minus the folded length of the module ( $[89 \times 0.38 \text{ nm}] - 4 \text{ nm} = 29.8 \text{ nm}$ ).

### AFM Measurements of the Unfolding Rate Constant

In order to simulate a chain of modules placed under force we have used Monte Carlo techniques and compiled the force required to unfold a domain during the simulated extensions. The force applied to the modules was calculated from the WLC model of elasticity and varied depending on the contour length and the total extension of the protein.<sup>6,7</sup> The unfolding and refolding of the modules was assumed to be a simple two state reaction with rate constants that depend exponentially on the applied force.<sup>6,7</sup> We found that an unfolding rate at zero force of  $k_u^0 = 3.3 \times 10^{-4} \text{ s}^{-1}$  and an unfolding distance  $\Delta x_u = 0.25 \text{ nm}$  was required in order to fit the distribution of unfolding forces measured from the I27 polyprotein (Fig. 1B). An independent estimate is obtained by fitting a Monte Carlo simulation to the dependency of the average force of unfolding on the rate of pulling by the



**Figure 2.** Stretching single I27<sup>RS8</sup> proteins at different pulling speeds. (A): i) 0.05nm/ms, ii) 0.5 nm/ms and iii) 5 nm/ms. The average unfolding forces (continuous lines) are 116 pN, 224 pN, and 272 pN, respectively. B) Dependence of the unfolding forces on the pulling speed. Each symbol (squares) is the average of (from left to right) 16,19,16,266,21,21,9 data points. The solid line corresponds to the result of a Monte Carlo simulation. The best fit was obtained with  $k_u^0 = 3.3 \times 10^{-4} \text{ s}^{-1}$  and a  $\Delta x_u = 0.25 \text{ nm}$ . Estimation of the folding rate constant

AFM. Experiments where the polyprotein was pulled at rates varying over the range 0.01 - 10nm/ms are shown in Figure 2. The average force of unfolding was shown to depend on the pulling rate. This dependency was reproduced by a Monte Carlo simulation using  $k_u^0 = 3.3 \times 10^{-4} \text{ s}^{-1}$  and  $\Delta x_u = 0.25 \text{ nm}$ . Thus, by two independent AFM experiments, we found that the transition state for unfolding was reached after an extension of only 2.5 Å and at an unfolding rate constant at 0 force of  $3.3 \times 10^{-4} \text{ s}^{-1}$ .

### Estimation of the Folding Rate Constant

A double pulse experiment designed to measure the folding rate of the I27<sup>RS8</sup> polyprotein is illustrated in Figure 3. A first extension of the polyprotein allows counting the available folded domains. After reaching the extended state, the protein was relaxed to its initial length. After a variable time period the protein was stretched again and a variable number of force peaks were observed. We interpret these results as an indication that some of the domains unfolded in the first extension of the protein had

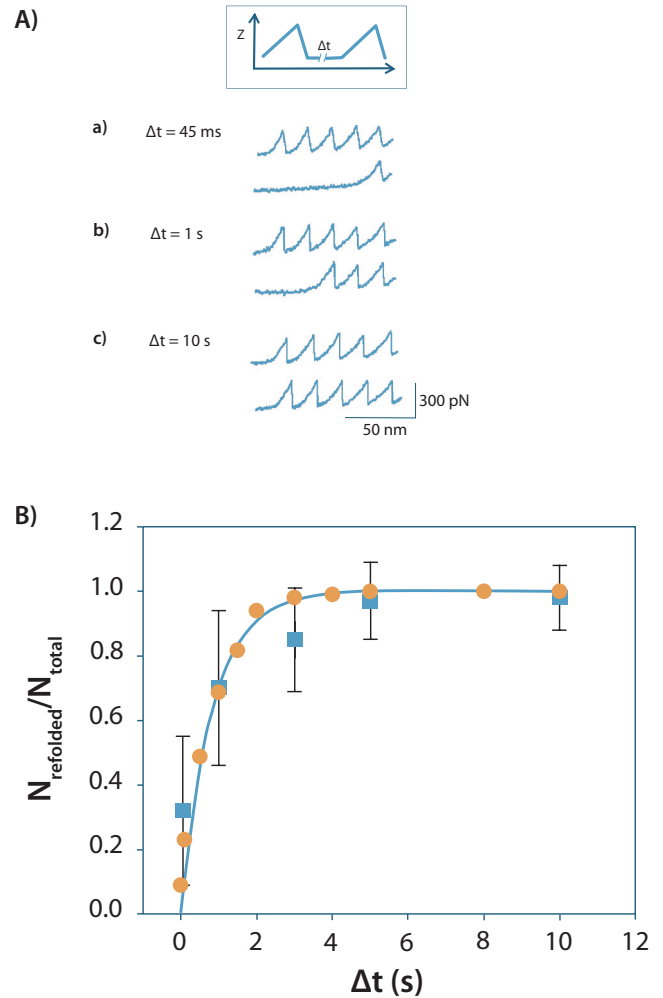
spontaneously refolded upon relaxation. The number of refolded domains counted this way was found to depend exponentially on the amount of time that the protein remained relaxed (Fig. 3). For a simple two-state model for the folding reaction, when a protein module is relaxed to its original length and there is no applied force, then  $k_f(F) = k_f^{\circ} \gg k_u(F) = k_u^{\circ} \sim 0$ , resulting in a simple first order folding reaction where the folding probability is given by  $P_f(t) = 1 - e^{-k_f^{\circ} t}$ . This function describes well the folding data of Figure 3B with  $k_f^{\circ} = 1.2 \text{ s}^{-1}$ .

### Using I27<sup>RS8</sup> as a Reference Protein in Force Extension Measurements

The consistency of measurements with the I27<sup>RS8</sup> polyprotein makes it an ideal reference protein for calibrating and controlling force extension experiments. Because the polyprotein is designed with a homogenous eight domain structure, it permits the measurement of extensions across a single molecule. Further, the polyprotein has proved to be robust allowing measurements to be made on multiple molecules in a single session as well as across multiple experiments. For example, the distance between the eight force peaks of 64 individual molecules was  $24.1 \pm 0.34 \text{ nm}$  (CV of 1.4%).<sup>4</sup> In eight separate experiments on a total of 266 molecules, the force peaks were measured at  $204 \pm 26 \text{ pN}$  (CV of 12.7%).<sup>4</sup>

The protein is currently available from Athena Enzyme Systems™ (Baltimore, MD, Cat. No. 0304) in a ready-to-use format under the trade name "I27O™". The protein is prepared for extension measurements by diluting the stock protein solution to a concentration of 100 μg/ml in PBS pH 7.3 buffer and applying the sample to freshly evaporated gold coverslips. The coverslips are then attached to the piezoelectric positioner stage. I27<sup>RS8</sup> was engineered with an amino-terminal hex-His peptide sequence and carboxy-terminal di-Cys. This configuration is hypothesized to allow the protein molecules to attach to the coverslip at the carboxy-terminal leaving the amino-terminus free to attach to the silicon nitride cantilever tip, when the tip is brought in close proximity to the adsorbed protein. It is this model which is believed to account for the high frequency (>50%) of observing all eight domain extensions after each retraction as well as the ability to perform extension and retraction experiments on a single molecule.

*See Back Side for References*



**Figure 3.** A) Unfolding and refolding cycles of an I27<sup>RS8</sup> protein probed with a double-pulse protocol (inset). The protein is first stretched to count the number of domains that unfold,  $N_{\text{total}}$  (a-c, top traces), and then it is relaxed to its initial length. A second extension after a delay,  $\Delta t$ , measures the number of refolded domains,  $N_{\text{refolded}}$  (a-c, bottom traces). B) Plot of the refolded fraction,  $N_{\text{refolded}}/N_{\text{total}}$  versus  $\Delta t$ . The solid line is a fit of the data (squares) to the function,  $P_f(t) = 1 - e^{-k_f^{\circ} t}$ , where  $k_f^{\circ} = 1.2 \text{ s}^{-1}$ . The filled circles correspond to a Monte Carlo simulation of a two state folding/unfolding kinetic model using a folding rate constant,  $k_f^{\circ} = 1.2 \text{ s}^{-1}$ .

## References

1. Czajkowsky, D. M. and Shao, Z. 1998. Submolecular resolution of single macromolecules with atomic force microscopy. *FEBS Lett.* 430:51-54.
2. Carrion-Vazquez, M., Oberhauser, A. F., Fisher, T.K., Marszalek, P. E., Li, H., Fernandez, J. M. 2000. Mechanical design of proteins studied by single-molecule force spectroscopy and protein engineering. *Prog. Biophys Mol. Biol.* 74(1-2):63-91.
3. Fisher, T. E., Oberhauser, A. F., Carrion-Vazques, M., Marszalek, P. E., and Fernandez, J. M. 1999. The study of protein mechanics with the atomic force microscope. *TIBS* 24:379-384.
4. Carrion-Vazquez, M., Oberhauser, A. F., Fowler, S. B., Marszalek, P. E., Broedel, Jr, S. E., Clark, J., and Fernandez, J. M. 1999. Mechanical and chemical unfolding of a single protein: A comparison. *Proc. Natl. Acad. Sci. (USA)*. 96:3694-3699.
5. Marko, J. F. and E. D. Siggia. 1995. Stretching DNA. *Macromolecules* 28:8759-8770.
6. Oberhauser, A.F., Marszalek, P.E., Carrion-Vazques, M. and Fernandez, J.M. (1999). Single protein misfolding events captured by atomic force microscopy. *Nature Struc. Biol.* 6(11):1025-1028.
7. Carrion.Vazquez, M., Marszalek, P.E., Oberhauser, A.F., and Fernandez, J.M. (1999). Atomic force microscopy captures length phenotypes in single proteins. *Proc. Natl. Acad. Sci. USA* 96:11288-11292.