Supplementary Information for Leake et al.

THE ELASTICITY OF SINGLE TITIN MOLECULES USING A TWO-BEAD OPTICAL TWEEZERS ASSAY

Expanded Materials and Methods

Polymer models and data handling

Most force-extension data obtained at stretch rates of 1 Hz exhibited no evidence of hysteresis. There was some evidence in a small number of instances (2 for each dataset of 150 mM and 300 mM ionic strength out of a total of 11 and 12 respectively) for short-term hysteresis for stretches on the whole titin molecule, (an example is given in Fig. 1, in which individual stretch and release half-cycles are marked), but which in all cases was absent from cycle number 10 and beyond. All the other curves had a similar degree of hysteresis to that in Fig1*D*, which was not measurable above the level of noise. Fits to the force-extension relationships, in general based on the time-average of 20-30 consecutive stretch-release cycles subsequent to the initial transient-hysteresis phase, were based on either a pure entropic model of elasticity or an entropic model with enthalpic component. One of the pure entropic models consisted of one or more independent worm-like chains (Marko and Siggia, 1995) acting in series, characterized by different contour and persistence lengths, so that the total end-to-end extension of the segment of titin stretched is given by:

$X = \sum_{i=1}^{n} x_i$ where the extension x_i at a force F of the i-th spring satisifies:

$F = \frac{k_B T}{L_{pi}}\left(\frac{1}{4(1-x_i/L_{ci})^2} + \frac{x_i}{L_{ci}} - \frac{1}{4}\right)$ where k_B is the Boltzmann constant, T the absolute temperature, with L_{pi} and L_{ci} the persistence and contour lengths respectively of the i-th spring. The number of independent worm-like chains in series, *n*, was at most three, as suggested from earlier single molecule stretch experiments (Li et al., 2002; Watanabe et al., 2002). Analogous multi-component freely-jointed chains in series (Florey, 1969) were also tried:

$$X = \sum_{i=1}^{n} x_i = \sum_{i=1}^{n} L_{ci}\left(\coth\left(\frac{FL_{ki}}{k_B T}\right) - \frac{k_B T}{FL_{ki}}\right)$$

Here, the L_{ki} values are the so-called Kühn lengths, corresponding to the lengths of stiff links (each of which are universally jointed) in the respective *i*-th chain in the series.

The presence of an enthalpic component to elasticity was investigated by replacing the fractional extension X/L_c in the single worm-like chain and freely-jointed chain models with (X/L_c-F/K) where K is the elastic modulus (Marko and Siggia, 1995; Smith et al., 1996) to give:

$$F = \left(\frac{k_B T}{L_p}\right)\left(\frac{1}{4(1-X/L_c+F/K)^2} - 1/4 + X/L_c - F/K\right) \qquad \text{for the worm-like chain.}$$

$$X = L_c\left(\coth\left(\frac{FL_k}{k_B T}\right) - \frac{k_B T}{FL_k}\right)(1+F/K) \qquad \text{for the freely-jointed chain.}$$

The worm-like chain interpolation formulation of Marko and Siggia (1995) is an approximation, which reduces to the exact solution as the fractional extension tends either to zero or to the contour length. However, higher-order improvements to the model were also tried, the coefficients α_i of the 7th-order polynomial being determined numerically by Bouchiat et al., (1999):

$$F = \left(\frac{k_B T}{L_p}\right)\left(\frac{1}{4(1-x/L_c)^2} - 1/4 + x/L_c + \sum_{i=2}^{7} \alpha_i (x/L_c)^i\right)$$

Solutions for all equations were found using a standard numerical interpolation technique. To perform the fits, code was written in the commercial mathematical analysis package Matlab which utilized a downhill modified Simplex Method (Nelder and Mead, 1965), modeling the *N* parameters to be varied as independent, orthogonal eigenvectors in *N*-dimensional space and varying the associated eigenvalues so as to minimize the goodness-of-fit parameter, σ_{fit}, taken as the normalized standard deviation between the theory curve and experimental data parallel to the force axis. With slight modification the routines could be made to find the minimum parallel to the extension axis, and the solutions obtained were consistent with those from the previous approach.

To test the fitting performance code was written to generate theoretical force-extension curves from known parameters, which were then run through the fitting routines (Fig. 2) under various levels of added noise. In most cases it was verified that the fitting algorithms would find the correct parameters to within <1% once a given iteration termination tolerance limit (usually set at the Matlab default of 0.0001, irrespective of units) had been reached. The situation worsened with increasing added noise and increasing number of parameters in the model; one-component pure entropic models (i.e. two variable parameters) performed the best since the function-error landscape (i.e. variation of σ_{fit} with the changes in the variable parameters) always had only one minimum and so more iterations would generally allow the true solution to be found. For models where the number of variable parameters was >2, a series of false minima were always present in the error landscape, and convergence would sometimes occur on these at values of noise >1 pN. However, for most actual

stretch-release data, 20-30 consecutive force-extension cycles were time-averaged prior to fitting, and so the actual level of noise on the final fitted data trace was more typically ~0.05 pN. Only the transient hysteresis PEVK force-extension curves were not time-averaged, but these were fitted with only a one-component worm-like chain model with low-pass filtering reducing the rms noise to a level of ~0.5 pN. For a one-component worm-like chain, the error ΔL_p on L_p, as a function of the errors on force, extension and fractional extension, ΔF, Δx and Δz respectively, for an individual fit could be estimated from:

$$F = g(z)/L_p$$
$$\therefore \Delta L_p = \sqrt{\left(L_p \Delta F / F^2\right)^2 + \left(g'(z)\Delta z / F\right)^2}$$
$$\equiv L_p k_m \Delta x \sqrt{2} / F$$
$$\therefore \Delta\left(1/L_p\right) = \left(1/L_p\right) k_m \Delta x \sqrt{2} / F$$

where k_m is the molecular stiffness and $g(z)$ is a function from the worm-like chain formulation.

All the models were compared with each other within each dataset on a pairwise basis using statistical testing: the mean and standard deviation values of σ_{fit} within each dataset were used to generate a t-statistic and a Student t-test was performed at a confidence level of $P<0.01$ to test the null hypothesis "were the pair of mean values of σ_{fit} sampled from the same distribution". If the null hypothesis was answered affirmatively in the direction of an improvement in the mean value of σ_{fit} then this was taken to indicate a statistically-significant improvement in fit.

Our approach cannot exclude non-statistically-significant improvements to the models as in fact being real, since there is clearly some arbitrariness as to the level of confidence set. Also, at some level of improvement to any model involving the introduction of further components there is the danger of fitting the intrinsic noise as

opposed to the underlying signal, which our simulations suggest may lead to systematic deviations in fitting.

Evidence for tether specificity and the presence of single titin monomer tethers

Many of the tethers formed were found to be transient, lasting less than 10 stretch-release cycles, whereas others would have a significantly longer lifetime, generally lasting until the experiment had been terminated by either one of the beads being pulled out of a trap or by an untethered bead entering one of the traps. Clearly there is the possibility that transient tethers had been generated by weak, non-specific associations between the molecule and the beads, and if so would need to discounted from subsequent analysis. If these transient tethers were non-specific then there will be a difference in contour and persistence lengths, and consequently a difference in molecular stiffness, k_m, at a given extension. The freely-jointed chain approximation of elasticity in the low force, low extension regime (Florey, 1969) gives, replacing α for FL_k/k_BT, where L_k is the characteristic Kühn length and z for x/L_c:

$$x \underset{F\to 0}{=} L_c\left(\coth\alpha - 1/\alpha\right) = L_c\left(\frac{e^{\alpha}+e^{-\alpha}}{e^{\alpha}-e^{-\alpha}} - 1/\alpha\right)$$

$$= \frac{L_c\alpha^3\left(1/2! - 1/3! + \ldots\right)}{\alpha^2\left(1+\alpha^2/3! + \ldots\right)} \approx \frac{L_c\alpha}{3} \equiv \frac{L_c F L_k}{3k_BT}$$

Inserting this result into the worm-like chain approximation:

$$F \underset{x\to 0}{\approx} \frac{k_BT}{L_p}\left(\left(1+2z-\ldots\right)/4 - 1/4 + z\right) \equiv \frac{3k_BTx}{2L_pL_c} \cong \frac{3k_BT}{2L_pL_c}\left(\frac{L_cFL_k}{3k_BT}\right) = \frac{FL_k}{2L_p}$$

$$\Rightarrow L_k \underset{x\to 0, F\to 0}{=} 2L_p$$

Similarly $k_m = \dfrac{\partial F}{\partial x} \underset{x\to 0, F\to 0}{=} \dfrac{3k_BT}{2L_pL_c} \cong \dfrac{3k_BT}{L_kL_c}$

A population of single tethers bounded only by the same *specific* interactions (the same pair of antibodies in the case of native titin) will all have roughly the same stiffness in the low force regime, and thus such specific tether force-extension curves may be selected from those due to non-specific tether (for example due to one end of the protein in question binding weakly to the hydrophobic latex of the microsphere surface) on the basis of the product L_pL_c or L_kL_c. Hence, the reciprocal scatter plot of $1/L_p$ versus L_c should form a clump of points; this approach was used as an initial mechanism to identify putatively specific tethers.

This method was also relevant to identifying long-lived tethers putatively due to either n multiple parallel tethers between a bead pair (resulting in an apparent persistence length of L_p/n but with the same contour length, thereby forming population clumps on the reciprocal scatter plot all lying close the same value of L_c but spaced at roughly equal intervals on the $1/L_p$ axis) or long-lived tethers due multimers of titin molecules formed from n single molecules joined end-to-end (resulting in the same apparent persistence length but a contour length of n L_c, thus manifest on the reciprocal scatter plot as population clumps lying close to the same value of L_p but spaced at roughly equal intervals on the L_c axis).

The information contained within a reciprocal scatter plot can also be presented as a histogram of the distribution of persistence lengths alone, examples of which are shown for whole skeletal muscle titin in Fig. 4*A* (150 mM ionic strength) and 4*B* (300 mM). Both approaches showed that all long-lived tether datasets had a distribution of persistence lengths consistent with a single population, strongly suggesting that multiple parallel tethers were not present. Short-lived tethers, for example the green bars of Fig. 4*A*, showed some evidence for a bimodal distribution which a more detailed analysis suggests is due to the presence or absence of the

PEVK region in the segment between the beads. Fig. 4*C* and 4*D* are histograms for the entire datasets of all regions stretched of native titin (both skeletal muscle and cardiac), normalizing the persistence length in each dataset to the mean persistence length within that dataset. Thus if multiple tethers (either in the form of parallel n-mers or n separate parallel monomers) are present there should be peaks in the distribution at 1/n (i.e. 1/2 for dimers, 1/3 for trimers etc.), however this is clearly not the case. An alternative albeit unlikely explanation is that most of the long-lived tethers are multimers of the same order (e.g. most of them are in fact titin dimers). However, in such a population there would clearly be some monomers, whose presence is confirmed from EM data which, even if present at a small percentage of the whole, would be manifest with an apparent normalized persistence length twice that of the titin parallel dimer, and so would presumably show a peak at around 2 on the scale of Fig4*C* and Fig4*D* if the dimer were the most abundant species for long-lived tethers. This is quite clearly not the case, though for short-lived tethers it is not possible to distinguish between random non-specific binding of different segments of a titin monomers and transient binding of a dimer on the basis of persistence length alone.

Reciprocal scatter plots revealed some variation in contour length, with a total range as high as ~20% in some cases. This range was significantly smaller than that expected for end-to-end titin multimers, and is more consistent with changes in contour length due to variation in isoform differential expression, as suggested from high-resolution SDS gel electrophoresis data (Neagoe et al., 2003): the range of variation in molecular weight of titin prepared from rabbit longissimus dorsi muscle is as high as ~200kDa.

Quantitative predictions as to the prevalence of multiple parallel tethers were made by modeling tether formation as an essentially random process, with therefore a Poisson probability distribution. Supposing that at each contact between the beads a different number of tethers is formed, 0, 1, 2, etc, and the average is n_{av}. Thus, for a Poisson distribution:

$$p(0) = \exp(-n_{av})$$
$$p(1) = n_{av} \exp(-n_{av})$$
$$p(2) = n_{av}^2 \exp(-n_{av})/2!$$
$$p(b) = n_{av}^b \exp(-n_{av})/b!$$

and so forth, where b is a positive integer or zero. But if binding is observed once in n_{trial} trials, then the probability of not binding is equal to (1-1/ n_{trial}).

$$\therefore p(0) = \exp(-n_{av}) = 1 - 1/n_{trial}$$
$$\therefore n_{av} = \log_e(n_{trial}/(n_{trial} - 1))$$

Fraction of binding events in which >1 tether is formed:

$$= p(>1)/(p(1) + p(>1))$$
$$= (1 - p(0) - p(1))/(1 - p(0))$$
$$= (1 - \exp(-n_{av}) - n_{av} \exp(-n_{av}))/(1 - \exp(-n_{av}))$$

For example, suppose there is one observed binding per 10 trials:

$$n_{trial} = 10$$
$$n_{av} = \log_e(10/9) = 0.1054$$
$$p(0) = 0.9$$
$$p(1) = 0.1054\, p(0) = 0.0948$$
$$p(>1) = 1 - 0.9 - 0.0948 = 0.0052$$
$$p(>1)/(p(1) + p(>1)) = 0.0052/0.1 = 0.052$$

So ~5% of bindings will be multiple tethers. The theoretical dependence on the probability of multiple tethers as a function of number of trials per binding event is shown in Fig. 5. Under the standard conditions of bead incubation it was observed that a tether would be generated once every 10-20 minutes using the conventional

mating protocol of tapping the bead pair against each other at a frequency of 1 Hz. This implies that the oscillated bead will enter the "capture-zone" on average 600-1200 times before a specific binding event is observed, i.e. n_{av} is between 600 and 1200. From the previous analysis this would suggest that the probability of multiple tether formation is less than 10^{-3}, in other words <0.1% of all tethers could be attributed to multiple interactions.

The detection of steps

To detect steps in noisy data, especially where the steps are of comparable size to the noise, specifically for stretches on the titin construct PEVK1, required a filter that reduced the noise amplitude whilst preserving the edge of the step. A median-filter satisfied both criteria. However, a window consisting of W data points in which noise is normally-distributed with variance σ^2 about a mean of μ, has a sample median equal to the sample mean, namely μ, but a sample median variance of $\pi\sigma^2/2W$, which compares with the sample mean variance of σ^2/W (Mood et al., 1974). The significance of this increased variance is a reduction in the t-statistic between a median-filter window immediately before and immediately after a step compared to that using mean-filter windows. The implication is that although the step edge has been preserved there is a consequent reduction in confidence that it is a true step, though a simple mean-filter on its own will not preserve an edge.

A better approach was to use the Chung-Kennedy filter (Chung and Kennedy, 1991). Here two adjacent windows of size W are run across the data point by point measuring sample means, μ_1 and μ_2, and sample variances, S_1 and S_2, in each. Weighting functions are simultaneously calculated such that:

$g_1 = \frac{S_2^r}{S_1^r + S_2^r}$ and $g_2 = \frac{S_1^r}{S_1^r + S_2^r}$ where the sensitivity factor r satisfies 1«r<100. The filtered output then consists of x_{filt}=g$_1$ μ_1+ g$_2$ μ_2.

As the windows run across the points only the second window will initially encounter a step. In this case the first window spans points before a step whereas the second will span the step itself as it moves forward. The result is that S_2 will be greater than S_1 due to a shift in some of the points within the second window away from the previous mean upon stepping, and therefore $g_1 \to 1$ and $g_2 \to 0$, thus $x_{filt} \to \mu_1$. As the windows move forward the first will eventually encounter the step whereas the second no longer spans the edge, thus S_1 is greater than S_2, implying $g_1 \to 0$, $g_2 \to 1$ and $x_{filt} \to \mu_2$. The net effect therefore is to switch between the windows so as to use the mean value corresponding to the window with the smallest variance, which will generally be the one uncontaminated by the step, which thus not only preserves the edge of the step but also results in a larger effective t-statistic compared to the median-filter, and hence greater confidence that the step is true. The size of the step (μ_1- μ_2), and the t-statistic (T=(μ_1- μ_2)/S) can be computed, where S is the uncontaminated variance such that: $S^2 = g_1 S_1^2 + g_2 S_2^2$.

Acceptance of the step as true is made by setting a positive threshold T_{thresh} and insisting that $|T|>T_{\text{thresh}}$. Setting T_{thresh} too low results in excessive false-positives, whereas setting it too high increases the number of false-negatives, and thus it becomes important to have some knowledge of the likely size of the step before starting (Smith, 1998). Greater true detection, namely an increase in true-positive and true-negative and a decrease in false-positive and false-negative, can be achieved by increasing the window size. The only limit to this is the mean

separation of neighboring steps, since increasing the window size beyond this will result in some true steps being undetected.

An algorithm was written in Matlab to incorporate the Chung-Kennedy filter and automate the process of step detection, and the filter was applied to data following square wave stretches. It was similarly applied to the raw force and extension time courses of triangle wave data. Here, it was important to ensure that the window size was much less than the ramp duration. In applying the technique to instances of very high ramp-rate, anomalously high values of T may be generated even in the absence of true steps, which are therefore falsely detected. If the ramp-rate in extension is R_x, with W measured in time, then the size of this false step will be of the order of $R_x W$, which should thus be less than the anticipated size of a true step, Δx. Therefore $R_x W < \Delta x$, or $R_x < \Delta x / W$, and so making the window size too high limits the range of ramp-rate that can be employed. Setting W= 20 ms, equivalent to 20 data points if the normal sample rate of 1 kHz was used, implied a maximum ramp-rate of ~0.5 μm s^{-1} for steps of size ~10 nm.

To test the performance of the step detection routines code was written in Matlab to simulate 'steppy' data. Pseudo-random noise was added to the trace and the step size could be varied as to a given factor multiplied by the amplitude of the noise. The position of the step could either be imposed or randomized on the basis of a given rate constant; it was recorded and an assessment of the proportion of true-positive, true-negative, false-positive and false-negative could be made. Fig. 3 shows an example of the output with randomized stepping at a rate of 10 s^{-1} with a step size five times the noise amplitude, using three different sizes of window.

Simulations on a time-series of generated steps of size 10 nm with T_{thresh}=8 and W=20 ms generated a false-negative probability <0.02 and a false-positive

probability <5 x 10^{-5} for an rms noise of 3 nm (equivalent to a molecular tension intermediate between zero and the highest amplitudes employed).

Changing the sensitivity factor r in the Chung-Kennedy filter caused a small bias in favor of either the first or second window, resulting in a marginal difference in location of the step's edge. Nominally a value of 50 was used throughout, at which the precision on edge location was found to be ~W/4 (~5 ms) from simulations.

Calculating changes in contour length due to stress-relaxation

The simplest approach to first deduce the effective change in contour length ΔL_c due to a change Δx in extension and ΔF in force was to calculate the anticipated contour length following these changes, $L_{c,\,post}$, from the single worm-like chain model and subtract this from that already determined prior to the stepping event, $L_{c,\,pre}$:

$$F_{post} = F_{pre} + \Delta F = \left(\frac{k_B T}{L_{p,post}}\right)\left(\frac{1}{4(1-z_{post})^2} - 1/4 + z_{post}\right)$$

This reduces to a cubic in the fractional extension following the event, z_{post}. Assuming a negligible change to the persistence length following a single small event (Rief et al., 1998) allows the cubic to be solved, which here generates two imaginary roots and one real, the latter of which is the one of physical relevance. Therefore $z_{post}=x_{post}/\,L_{c,\,post}=(x_{pre}+\Delta x)/\,L_{c,\,post}$. Thus $L_{c,\,post}$ and therefore ΔL_c can be determined.

Monte Carlo simulations of unfolding and refolding

A simulation algorithm was written in Matlab, modeling the Ig/Fn transitions as a simple two-state process similar to those described previously (Rief et al., 1998) polling at 1 ms time intervals (Δt) to calculate the probability of unfolding (P_u) and folding (P_f) based upon pre-set kinetic parameters, using the following equations:

$$P_u = k_u(0)\Delta t \exp\left(F\Delta x_u / k_B T\right)$$

$$P_f = k_f(0)\Delta t \exp\left(-F\Delta x_f / k_B T\right)$$

The widths of activation potential for the unfolding (Δx_u) and folding (Δx_f) transitions were set at 0.28 nm and 2.3 nm respectively with the spontaneous folding rate at zero force ($k_f(0)$) set at 1.2 s^{-1} (Carrion-Vazquez et al., 1998). The spontaneous unfolding rate at zero force ($k_u(0)$) was varied from 1 x 10^{-5} to 3 x 10^{-3} s^{-1} in separate simulations. The displacement of the optical trap whose relevant trapped bead was not position-clamped was assumed to be a triangle wave of frequency 1 Hz with amplitude varied between 0.5-5.0 μm for separate simulations. The equation of motion of the bead was solved at each polled displacement of the trap assuming a two-component worm-like chain model for titin elasticity, with ~300 Ig/Fn domains and ~1400 residues in the PEVK region (as suggested from the worm-like chain fits to force-extension data) to yield the force (F). The calculated probabilities were then compared with pseudo-random numbers between 0 and 1 to determine whether the transition had occurred, polling each relevant Ig/Fn domain in turn. A typical output in shown in Fig. 6*A*.

The same Monte Carlo algorithm was used to estimate the effect of PEVK hysteresis on whole skeletal muscle titin elasticity. The same parameters were used as above, however chain I (putatively the PEVK region) in the pre-stretched molecule was assumed to consist of (1400-28n) residues, with n being the number of putative PEVK structural motifs which, upon unfolding, result in an increase in contour length of ~10 nm. Lack of accurate kinetic data made estimations for the width of the activation barrier for the unfolding transition and the spontaneous rate of unfolding difficult, and so values of 0.3-0.6 nm and 1 x 10^{-5} to 3 x 10^{-3} s^{-1}, selecting only combinations which generated mean unfolding forces in the range 30-60 pN. The

number of putative domains n was varied in the range 0-40 (the maximum based on sequence alignment data for the 28-mer residue repeat), and the effect of refolding was assumed to be negligible. A typical output is shown in Fig. 6*B*. The impact of PEVK hysteresis to overall elasticity was generally found to be small (hysteresis loss <10% of total mechanical work done on the stretching molecule) due principally to a masking effect from the relatively large tandem Ig region dominating compliance at the low forces over which putative stress-relaxation events in the PEVK region occur.

Steady-state prediction of Ig domain unfolding in the I-band region of skeletal muscle titin

We used the same approach as Li et al. (2002), in assuming that the unfolding of a proximal Ig domain (I1-15 in the constitutively-expressed region N-terminal to the PEVK) can be modeled as a 2-state process (folded and unfolded) with a probability of $P_{prox}=A/(A+B)$, where A and B are force-dependent unfolding and folding rate constants, whereas unfolding of a distal Ig domain (constitutively-expressed tandem Ig region C-terminal to the PEVK) is modeled as a 3-state process (folded, intermediate and unfolded), with a probability of $P_{dist}=A_1A_2/(A_1A_2+B_1B_2+A_1B_2)$, where A_1 and B_1 are force-dependent unfolding and folding rate constants between the fully-folded state and an intermediate state, and A_2 and B_2 are the equivalent rate constants between the intermediate state and the fully-unfolded state. Mid Ig domains (those differentially-spliced in between the proximal and distal regions) were assumed to have properties similar to distal Ig domains. The values of parameters used are shown in Table 1.

Simulations of fast-stretch force-extension relations for cardiac muscle titin

The elastic properties of whole cardiac muscle titin were simulated by using a three-component worm-like chain model (Marko and Siggia, 1995) whose independent spring components are ascribed to the tandem Ig, PEVK and N2-B regions respectively (Li et al., 2002; Watanabe et al, 2002). Simulations assumed ~250 Ig/Fn domains, 186 residues in the PEVK region and 572 residues in the N2-B region. A persistence length of 0.91 nm was assumed for the PEVK region (measured from our present study and Li et al. (2002)), 0.66 nm for the N2-B region (Li et al., 2002) and either 2.9 nm (our present study) or 10.0 nm for tandem-Ig's (Li et al., 2002). Stretches were assumed to be fast enough to make the likelihood for Ig/Fn unfolding negligible. The predictions are shown in Fig. 7 along with the results of fitting the prediction to a one-component worm-like chain model.

Identification of repeating sequences in the PEVK region

Sequence homology was identified using the Basic Local Alignment Search Tool (BLAST) algorithm (Altschul et al. 1990), running the peptide sequence of human soleus muscle titin (EMBL database accession number: X90569) from between residues 5000-8000 (incorporating the full-length PEVK sequence) against itself, using a window-length and stringency of 30 (Fig. 8*A*). A blast search was also performed running the PEVK1 construct sequence (residues 5594-6143, EMBL database accession number: X90569) against itself (Fig. 8*B*)**.**

REFERENCES:

Altschul,S.F., Gish,W., Miller,W., Myers,E.W. and Lipman,D.J. 1990. Basic local alignment search tool. *J. Mol. Bio*l., 215, 403–410.

Bouchiat, C., Wang, M.D., Allemand, J., Strick, T., Block, S.M., and Croquette, V. 1999. Estimating the persistence length of a worm-like chain molecule from force-extension measurements. *Biophys. J.* 76:409-413.

Carrion-Vazquez,M., Oberhauser, A.F., Fowler, S.B., Marszalek, P.E., Broedel, S.E., Clarke,J. 1999. Mechanical and chemical unfolding of a single protein: A comparison. *Proc. Natl. Acad. Sci. U.S.A.* 96: 3694-3699.

Chung,S.H. and Kennedy,R.A. 1991. Forward-backward non-linear filtering technique for extracting small biological signals from noise. *J. Neurosci. Meth*., 40, 71-86.

Florey, P.J. 1969. FJC. In: *Statistical Mechanics of Chain Molecules*, New York:Interscience Publishers.

Freiburg, A., Trombitas, K., Hell, W., Cazorla, O., Fougerousse, F., Centner, T., Kolmerer, B., Witt, C., Beckmann, J.S., Gregorio, C.C., Granzier, H., and Labeit, S. 2000. Series of exon-skipping events in the elastic spring region of titin as the structural basis for myofibrillar elastic diversity. *Circ. Res.* 86:1114-1121.

Li, H., Linke, W.A., Oberhauser, A.F., Carrion-Vazquez, M., Kerkvliet, J.G., Lu, H., Marszalek, P.E. and Fernandez, J.M. 2002. Reverse engineering of the giant muscle protein titin. *Nature*. 418:998-1002.

Marko, J.F. and Siggia, E. 1995. Stretching DNA. *Macromolecules* 28:209-212.

Mood, A.M., Graybill,F.A. and Boes, D.C. 1974. *Introduction to the Theory of Statistics*, 3rd Ed., McGraw-Hill International Editions, p.257.

Neagoe C., Opitz C.A., Makarenko I. and Linke W.A. 2003. Gigantic variety: expression patterns of titin isoforms in striated muscles and consequences for myofibrillar passive stiffness. *J. Muscle. Res. Cell Motil.* 24:175-89.

Nelder, J.A. and Mead, R. 1965. A simple method for function minimization. *Computer J.* 7:308-313.

Rief, M., Gautel, M., Schemmel, A., and Gaub, H.E. 1998. The mechanical stability of immunoglobulin and fibronectin III domains in the muscle protein titin measured by atomic force microscopy. *Biophys. J.* 75:3008-3014.

Smith, D.A. 1998. A quantitative method for the detection of edges in noisy time-series. *Phil. Trans. Soc. Lond. B* 353:1969-1981.

Smith, S.B., Cui, Y., and Bustamante, C. 1996. Overstretching B-DNA: the elastic response of individual double-stranded and single-stranded DNA molecules. *Science* 271:795-799.

Watanabe, K, Nair, P., Labeit, D., Kellermayer, M.S.Z., Greaser, M., Labeit, S. and Granzier, H. (2002) Molecular mechanics of cardiac titin's PEVK and N2B spring elements. *J. Biol. Chem.* 277:11549-11558.

TABLES

TABLE 1 Parameters used in the prediction of force-extension relation for the I-band region delimited by antibodies T12 and anti-I105

	$A/\ s^{-1}$	$B/\ s^{-1}$	$A_1/\ s^{-1}$	$B_1/\ s^{-1}$	$A_2/\ s^{-1}$	$B_2/\ s^{-1}$	n	L_p/ nm	L_c/ nm
proximal Ig's	-	-	1.0×10^{-2}	100	3.3×10^{-3}	0.33	15	2.9	66
distal Ig's	8.0×10^{-5}	1.2	-	-	-	-	22	2.9	99
mid Ig's	8.0×10^{-5}	1.2	-	-	-	-	48	2.9	211
PEVK	-	-	-	-	-	-	1408	0.76	507

Total contour lengths of 374 and 507 nm have been assumed for tandem Ig and PEVK respectively from the 2-wlc fit to the single molecule stretches on the I-band region, but the values of persistence lengths are the average from all data sets for the chain I and chain II components. The number (n) of Ig's has been deduced by assuming 4.4 nm per domain and assuming 15 proximal Ig's, and the number of residues in the PEVK region by assuming 0.36 nm per residue. The properties of the non-constitutively expressed mid Ig's have been assumed to be the same as those of the constitutively expressed distal Ig's. All kinetic parameters taken from data published in Li et al. (2002).

FIGURE LEGENDS:

FIGURE 1 Stretch (*blue*) and release (*red*) force-extension relationships taken from single stretch-release cycles on the same long-lived tether of whole titin delimited by the anti-titin antibodies T12 and AB5 for cycle number (*A*) 1, (*B*) 2, (*C*) 5 and (*D*) 10.

FIGURE 2 Function-error landscape for a single worm-like chain fit on a theoretical curve in the case of (*A*) no noise and (*B*) pseudo-random noise of amplitude 1 pN.

FIGURE 3 (*A*) Simulated steps of known size relative to pseudo-random noise used to test Chung-Kennedy edge detection routine. (*B*) t-statistic output at three different window sizes. A sample rate of 1 kHz was assumed throughout.

FIGURE 4 Distribution of persistence lengths for stretches of whole native rabbit longissimus dorsi muscle titin delimited by the anti-titin antibodies AB5 and T12 for short-lived (*green*) and long-lived (*red*) tethers. (*A*) single worm-like chain model, 150 mM ionic strength. (*B*) single worm-like chain model, 300 mM ionic strength. Histograms for the normalized distribution of persistence length for all native titin datasets, normalized by the mean persistence length within each individual data set, for (*C*) 150 mM ionic strength, and (*D*) 300 mM ionic strength.

FIGURE 5 Prediction of multiple tether generation probability based on a Poisson distribution model.

FIGURE 6 (*A*)Typical output from a Monte Carlo simulation of a single stretch (*blue*) and release (*red*) cycle of a single molecule of whole skeletal muscle titin, A trap

stiffness of 0.1 pN/nm was assumed with an amplitude of trap displacement of 4 μm at a frequency of 1 Hz. A spontaneous rate of unfolding ($k_u(0)$ of 5×10^{-4} s^{-1} was used in this instance. (*B*) Monte Carlo estimate of hysteresis due to unfolding of PEVK region using same color coding. An amplitude of 2 μm at a frequency of 1 Hz was been used to generate a maximum force of ~80 pN, consistent with the upper limit of our optical trap. For this example a spontaneous rate of unfolding of 5×10^{-3} s^{-1} with a width of activation potential for unfolding of 0.6 nm was used, assuming 10 putative unfolding motifs in the PEVK region.

FIGURE 7 Simulated force-extension relations (*blue circles*) for whole cardiac muscle titin using a three-component worm-like chain model (*A* and *B*). Fits to the relations using a single-component worm-like chain model are shown (*red*), using persistence lengths of 2.1 nm (*A*) and 4.1 nm (*B*), contour lengths for both of 1340 nm. Corresponding force-residuals (*C* and *D*) and variation of fractional extension (*E* and *F*) for the individual spring components of N2-B (*blue*), PEVK (*red*) and tandem Ig's (*green*) are shown. Figures in the left-hand column (*A*, *C* and *E*) use an average persistence length for the tandem Ig region in the three-component worm-like chain fit of 2.9 nm (our present study), those in the right-hand column (*B*, *D* and *F*) use 10.0 nm (Li et al., 2002).

FIGURE 8 BLAST of X90569 (5000-8000) run against (*A*) itself, repeat regions of ~28 residues length are indicated (*red*) as is the position of PEVK1 (*green*) and the full-length PEVK region (*magenta*), and (*B*) the peptide sequence of PEVK1, X90569 (5594-6143). The borders of the ~28-mer repeats are indicated (*red*).

FIGURES:

FIGURE 1 (SUPPLEMENTARY INFORMATION)

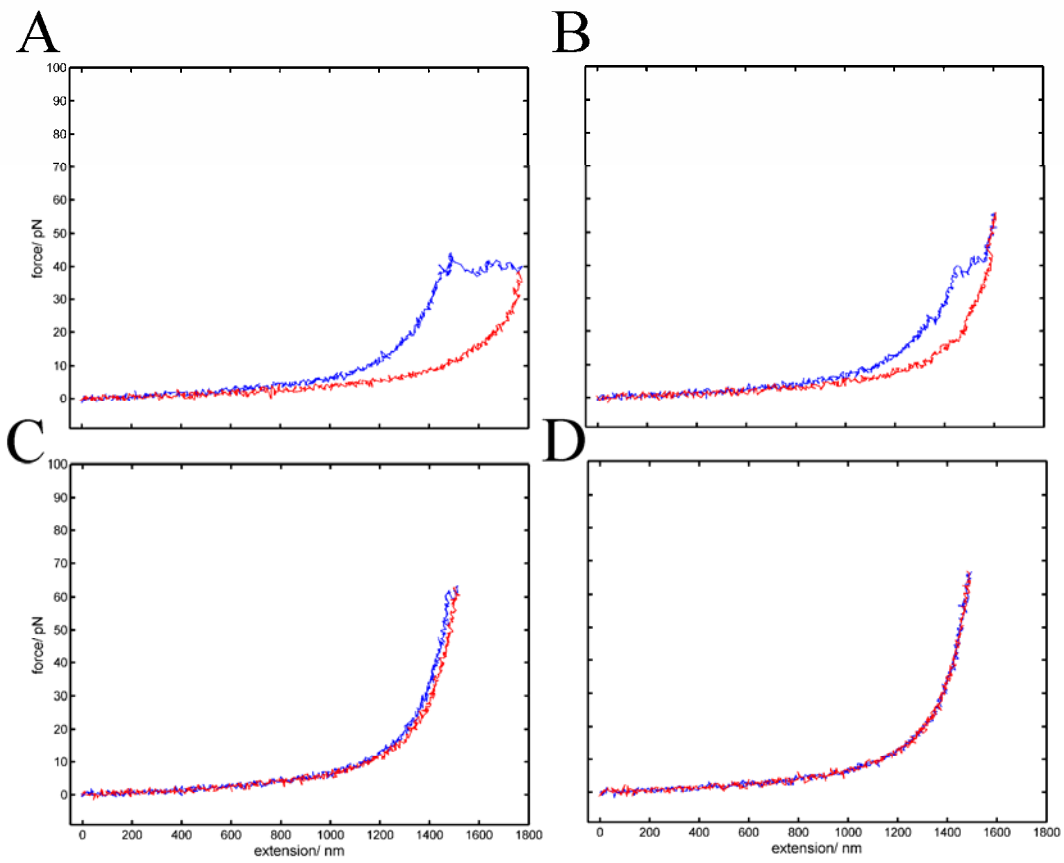

FIGURE 2 (SUPPLEMENTARY INFORMATION)

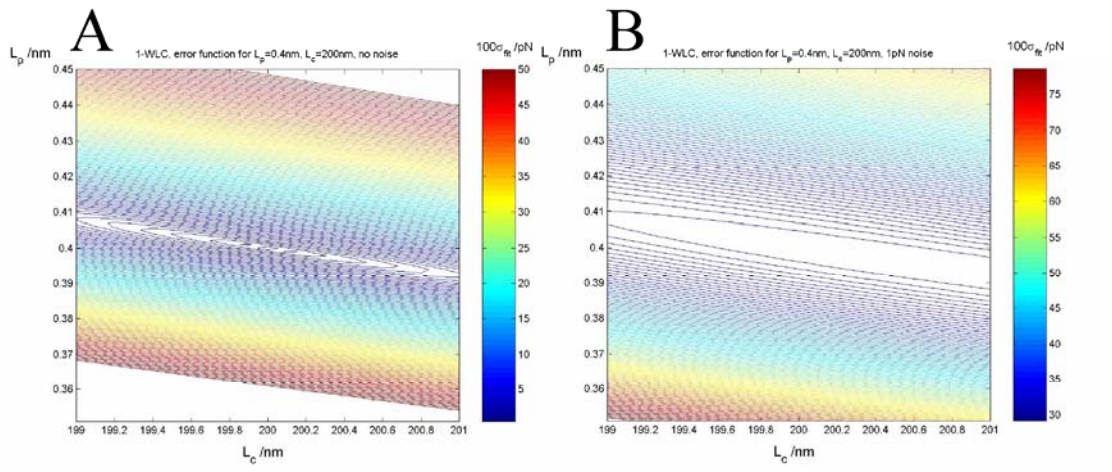

FIGURE 3 (SUPPLEMENTARY INFORMATION)

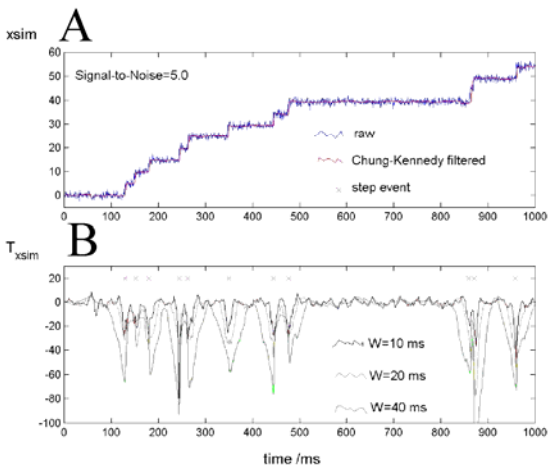

FIGURE 4 (SUPPLEMENTARY INFORMATION)

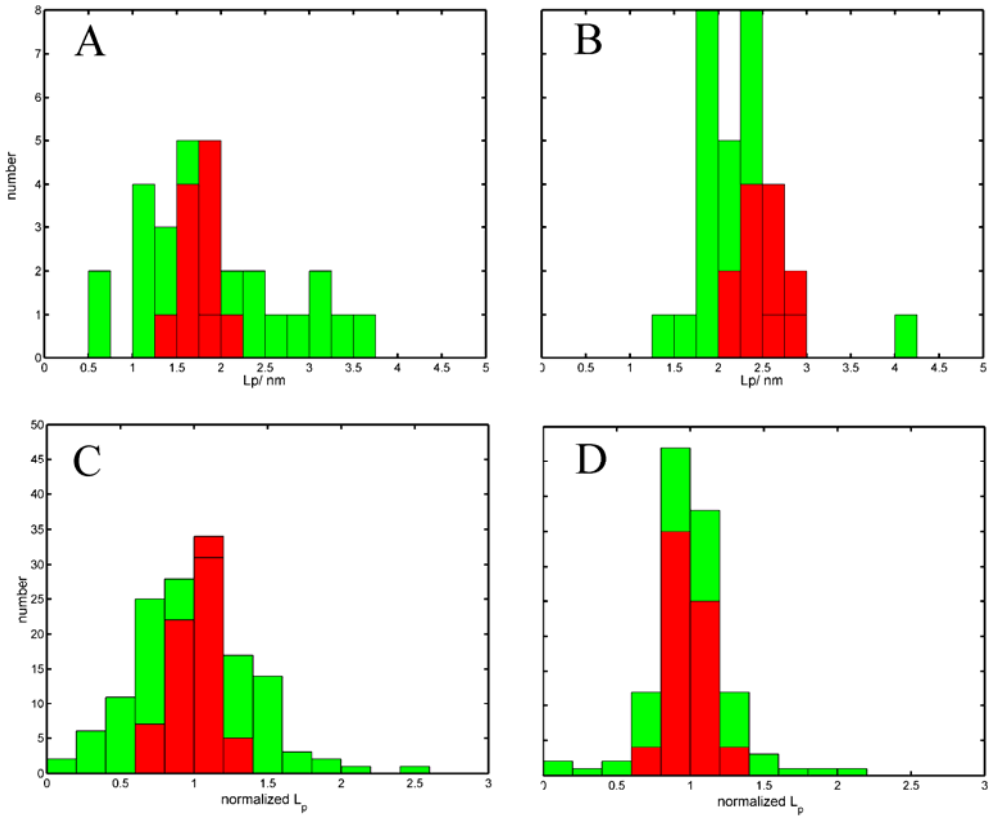

FIGURE 5 (SUPPLEMENTARY INFORMATION)

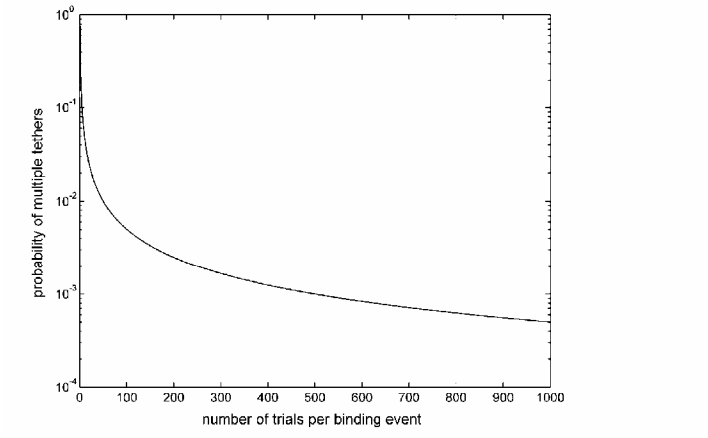

FIGURE 6 (SUPPLEMENTARY INFORMATION)

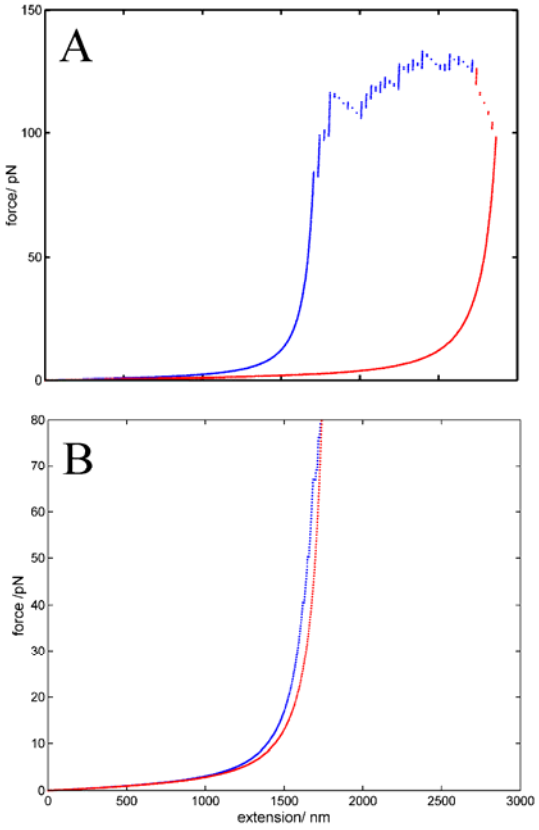

FIGURE 7 (SUPPLEMENTARY INFORMATION)

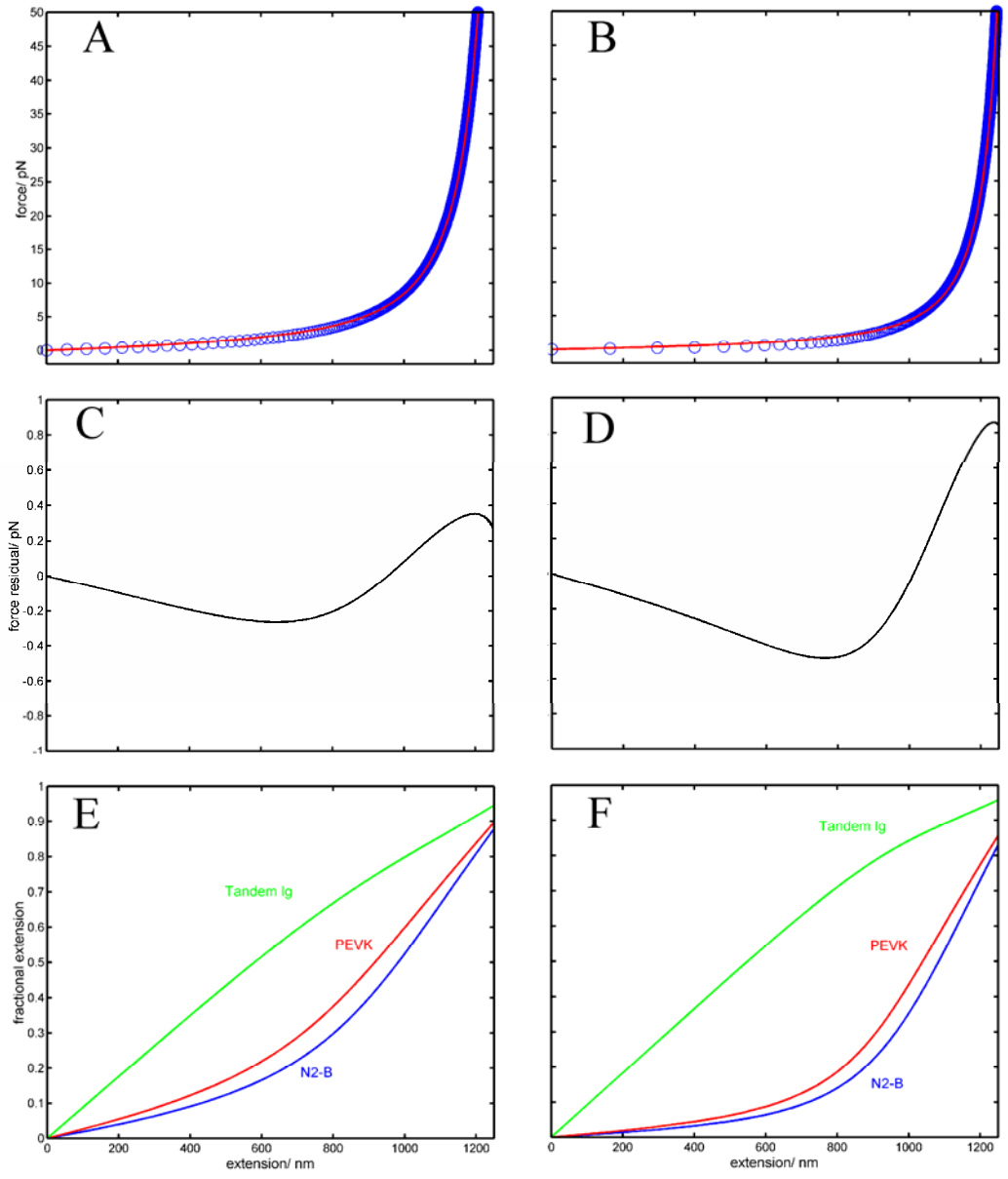

FIGURE 8 (SUPPLEMENTARY INFORMATION)