The Elasticity of Individual Titin PEVK Exons Measured by Single Molecule Atomic Force Microscopy*

Received for publication, December 13, 2004, and in revised form, January 3, 2005
Published, JBC Papers in Press, January 4, 2005,
DOI 10.1074/jbc.C400573200

Atom Sarkar§¶†, Sofia Caamano‡, and Julio M. Fernandez‡**

From the ‡Department of Biological Sciences, Columbia University, New York, New York 10027 and the §Department of Neurological Surgery, Mayo Clinic College of Medicine, Rochester, Minnesota 55905

The I-band region of the giant muscle protein titin contains a large domain enriched for the amino acids proline, glutamate, valine, and lysine and is denoted the PEVK domain. The PEVK domain of titin encodes a random coil shown to be an important factor in the passive elasticity of titin. Muscle-specific splicing of 116 PEVK exons encodes this domain. It has been proposed that proline contents determine the elasticity of the PEVK polypeptide, where the individual exons code for “flexibility cassettes.” To test this hypothesis, we have measured the elasticity of three distinct polypeptides encoded by individual PEVK exons (161, 120 and 184) that varied greatly in their proline contents (7, 14, and 37%) and total PEVK contents (55, 70, and 87%).

We used single molecule atomic force microscopy techniques to measure the persistence length, \( p \), of the engineered PEVK proteins. Surprisingly, all three exons 161, 120, and 184 coded for proteins with similar values of persistence length, \( p = 0.92 \pm 0.38, 0.89 \pm 0.42, \) and 0.98 ± 0.4 nm, respectively. We conclude that the PEVK exons encode polypeptides of similar elastic properties, unrelated to their total PEVK contents. Hence, alternative splicing solely adjusts the length of the PEVK domain of titin.

The giant muscle protein titin is crucial to the structural organization in striated muscle sarcomeres (1–3). In addition to the vital role of titin in myofibril architecture by organizing the half-sarcomere, titin is also responsible for the elasticity of myofibrils (4, 5). The elastic region of skeletal muscle titin, corresponding ultrastructurally to the I-band of the sarcomere, is composed of two distinct elastic motifs, tandem repeats of immunoglobulin-like modules (Ig), and random-coiled domains (6–12). While 363 exons encode the entire titin gene, 116 exons encode PEVK domains (2, 13). The expression of PEVK alternative splicing variants has been shown to be a function of factors ranging from tissue specificity, developmental, physiology, or pathologic cues and leads to myofibrils with distinct passive elasticity (14–17). Prior AFM and electron microscopy single molecule studies showed that a cardiac PEVK isoform behaved as a random coil with structural variations that resulted in persistence length values \( (p) \) ranging widely between 0.2 and 2.5 nm (8, 9). Given that the persistence length of a polypeptide chain determines its elasticity, it was surprising that the values of \( p \) varied so much between molecules. In this earlier work, it was proposed that proline content, resulting into individualized mixtures of cis and trans conformations, might account for the broad range of persistence lengths observed.

Here we present a study designed to gain a better understanding for the role various exons play in the passive elastic properties of PEVK and whether this is influenced by either proline content alone or rather the overall proline, glutamate, valine, and lysine amino acid content of the PEVK domain. We have used molecular biology techniques to generate three distinct polypeptide chimeras containing PEVK domains encoded by a single exon, fused together with the titin I27 domain. We examined these proteins with single molecule atomic force spectroscopy. The PEVK polypeptide chimeras design was essential to unambiguously identify single molecules, a crucial step in correctly measuring the persistence length of the PEVK domains (8, 9, 18). Application of the worm-like chain (WLC) fit to our AFM force-extension curves describes the restoring force as a result of stretching the polymer and is given by,

\[
F(x) = \frac{k T}{p} \left( \frac{1}{4} \left( 1 - \frac{x}{L_c} \right)^2 - \frac{1}{4} + \frac{x}{L_c} \right) \]  

(Eq. 1)

where \( F \) is the force, \( x \) is the extension, \( p \) is the persistence length, \( k \) is the Boltzmann constant, \( T \) is the temperature in degrees Kelvin, and \( L_c \) is the contour length of the polymer. The parameters of \( p \) and \( L_c \) are measured by fits to the measured force-extension curves (Fig. 1, A–C, and Fig. 2D, thin lines). In the region where the I27 fingerprint unfolds, consecutive fits measure the increase in contour length associated with each unfolding event (\( \Delta L_c \sim 28 \text{ nm} \)), which together with the peak unfolding forces, positively identify a single PEVK containing polypeptide.

EXPERIMENTAL PROCEDURES

Single Molecule Atomic Force Microscopy—Our AFM is based on a modified Digital Instruments detector head (AFM-889, Veeco Instruments, Santa Barbara, CA), the specifics of which have been described elsewhere (19). The force-extension mode of AFM operation has also been recently described (20). Data collection and control of the AFM head were done by means of data acquisition cards (6052E and 6703, National Instruments, Austin, TX) controlled by software written in IGOR PRO (WaveMetrics, Lake Oswego, OR). Cantilevers used for AFM experiments were Si3N4 levers (Digital Instruments, Santa Barbara, CA) having spring constants typically in the range of 40–50 pN nm \(^{-1} \) and were calibrated in solution with the equipartition theorem. A pulling speed for all measurements was set to 400 nm/s. Data analysis was done with IGOR PRO.

* This work was supported in part by funding from the National Institutes of Health (to J. M. F.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

† Supported in part by a fellowship from the Neurosurgery Research and Education Foundation.

‡ These two authors contributed equally to this work.

¶ The abbreviations used are: PEVK, proline, glutamate, valine, and lysine; AFM, atomic force microscope; I27, 27th immunoglobulin module from human titin; WLC, worm-like chain.

This paper is available on line at http://www.jbc.org
Construction of the [PEVK]I27 Polyproteins—The method for protein engineering, utilizing successive cycles of BamHI/BglII sticky end ligations, has been described previously (20, 21). Polyprotein constructs used in this study were based on the human titin I27 and PEVK sequence (4, 14). While the human titin gene contains 116 exons coding for PEVK sequences, we have confined our investigation to exons 120, 161, and 184. All three exons are approximately equal in size (encoding 27, 28, or 30 amino acids, respectively) but contrast in proline as well as overall PEVK content. Specifically exon 120 contains 4 prolines with an overall PEVK content equal to 70%, exon 161 has 2 prolines with an overall PEVK of 55%, and exon 184 has 11 prolines with 87% of its content made up by PEVK amino acids. Thus these three selected exons provide a range of proline and PEVK composition, features that make them ideal to use in an examination of the exon-specific contribution to PEVK flexibility. The core sequence for our polyprotein chimeras was always three identical tandem repeats of a PEVK exon (which we refer to as a single PEVK domain), followed by an I27 module, with the core sequence repeated four times. The finished constructs had the following structure: [exon120PEVK]I27, [exon161PEVK]I27, or [exon184PEVK]I27. Additionally all constructs contained two terminal cysteine-cysteine residues to orient a covalent protein attachment to the gold substrate during AFM experiments (20). Polyprotein expression in bacteria, as well as protein purification by Ni2+ chelation and size exclusion chromatography, have also been described previously (20).

RESULTS

The polyproteins used in this study were constructed from single PEVK exons and an I27 domain. Because the number of amino acids in the selected PEVK exons is approximately one-third the number of amino acids in the I27 domain (27–30 versus 89 amino acids, respectively), all the chimeric polyproteins were designed to include three repeats of a single PEVK exon linked to I27 as the core sequence, which was repeated 4-fold such that any of our recombinant constructs always contained 12 identical copies of a PEVK exon and four I27 domains. This can be appreciated in the cartoons of Fig. 1, A–C, where three identical PEVK exons are placed in tandem and represented by the green, lavender, and dark blue coiled structures, followed by the β-stranded I27 domain shown in gold, and repeated four times. By engineering such a polyprotein, we were able to confine our analysis to only those AFM force-extension traces in which the well described signature of I27 unfolding was present, namely a contour length increase of ~28 nm, as well as a peak unfolding force of ~200 pN (21, 22). Our method of analysis is highlighted in Fig. 1. Panel A in Fig. 1 illustrates a typical force-extension sawtooth pattern trace. In this experiment three I27 unfolding peaks are present, while the last peak represents the detachment of the protein from either the cantilever or the surface. The initial featureless region of the force-extension curve corresponds to the extension of the PEVK domain. We fit this region with the WLC model up to 90 pN, just before the first unfolding peak of the I27 fingerprint (Equation 1; Fig. 1, A–C, thin lavender trace). Following the extension of the PEVK domain, we fit the WLC to the unfolding peaks to identify the I27 fingerprint. We explicitly measure Lc per I27 unfolding event by measuring the difference in the contour length between the first and the last unfolding peaks and dividing by the total number of events (20). In Fig. 1A we measured Lc = 86.4 nm, which divided by three events, gives Lc = 28.2 nm per unfolding. This value together with unfolding forces of ~200 pN defines the mechanical unfolding of three I27 modules (8, 9, 20, 21). Given that we pick up molecules from random locations, the number of I27 peaks that we observe varies from molecule to molecule. In this study, we have only included recordings showing at least two I27 unfolding events. The three I27 unfolding peaks shown in Fig. 1A indicate that there must be at least two and at most four PEVK domains being stretched in the initial force-extension trace for this molecule (8, 21).

A single PEVK exon, repeated three times, represents a PEVK domain within our polyprotein of either 81 (exon 120), 84 (exon 161), or 90 (exon 184) amino acids. Since each amino acid contributes 0.38 nm/residue (23), this will result in the extension of each PEVK domain to a contour length of Lc = 30.8, 31.9, or 34.2 nm for exons 120, 161, or 184, respectively. Due to the characteristic lack of mechanical resistance of a random coil, the extension of a PEVK domain will always occur before any I27 unfolding event. A WLC fit to the last peak (gold trace) measures the contour length increment after unfolding and extending all its modules. A WLC fit up to 90 pN of the first unfolding peak (lavender trace) measures the contour length and persistence length of the PEVK modules, before any I27 unfolding occurs. A, the trace shows an initial extension to a contour length of Lc = 72.5 nm followed by the unfolding of three I27 modules. In the initial extension, the three folded I27 modules contribute to the initial contour length by ~4.5 nm/folded I27 domain. Subtracting this contribution we calculate a PEVK contour length of ~59 nm (see text). This extension is consistent with the elongation of two PEVK domains. B, the trace shows an initial extension corresponding to three PEVK domains (93 nm), followed by the unfolding of four I27 domains. C, while only three I27s unfold in this trace (similar to A), the initial extension corresponds to four PEVK domains (119 nm). D, histogram of I27 peak unfolding forces from all polyprotein recordings used in this study. The mean I27 unfolding force was 196 ± 30 pN, bin width = 10 pN, n = 823, collected from 239 force-extension traces.
traces obtained from the \([(\text{exon120PEVK})_3I_{27}]_4\) polyprotein are shown in gram of persistence length values (average persistence length values, tides from different PEVK exons.

The persistence length, elasticity to the PEVK domain force-extension relationships, measured in the histogram of Fig. 2, were 196 recordings used in our study. The mean value of the unfolding protein (sixty recordings) measured the persistence length shown this exon is \([(\text{exon161PEVK})_3I_{27}]_4\) polyprotein (94 recordings) gave an average persistence length of \(p = 0.98 \pm 0.4\) nm (Fig. 2C). Fig. 2D shows two recordings obtained from \([(\text{exon120PEVK})_3I_{27}]_4\) polyproteins, which are similar in length and have the same number of I27 unfolding events (four), yet their PEVK domains show very different values of persistence length \((p = 0.38\ \text{nm versus } p = 1.93\ \text{nm})\). These recordings highlight the variations in the persistence length of the PEVK domains.

**DISCUSSION**

Of the 363 exons encoding titin, the elastic I-band region is encoded by a total of 251 exons, highlighting the importance of alternative splicing in controlling titin elasticity (2). The I-band comprises of 135 exons coding for immunoglobulin domains and unique sequences and there are 116 exons coding for PEVK sequences (2). It is believed that each of the approximately ~600 striated muscles of the human body will have a unique alternatively spliced form of titin (4). For example, cardiac titin possesses two extensible regions of unknown structure, the N2B unique region and the PEVK region. The unique cardiac-specific N2B and PEVK regions are 572 and 186 amino acids long, respectively. The mechanical properties of these regions were recently examined by single molecule AFM techniques and were found to behave like extensible random coils (7–9). In contrast to mechanically stable folded proteins like the immunoglobulin domains of titin (20), an unstructured protein domain generates a recoiling force that changes monotonically with its length (8). This recoiling force is well described by simple models of polymer elasticity, where the flexibility of the chain is defined by the persistence length, \(p\). For unstructured proteins of equal length a higher value of \(p\) predicts a lower force at any given extension. Hence we were surprised to find that the persistence length of the cardiac specific PEVK varied widely from molecule to molecule between values of \(p = 0.3–2.3\) nm (8), in sharp contrast with the persistence length of the N2B unique region, which was narrowly distributed around a mean value of \(p = 0.66\) nm (8, 9). The wide variations in the persistence length of the cardiac PEVK were attributed to result from the different conformations of its proline residues. Such conformations were thought to be the result of polyproline type I (cis) and type II (trans) helical structures (8), and the two allowed conformations of the trans prolyl residues (25). Extrapolating from these results, we proposed that the proline content encoded into each of the 116 PEVK exons would result into modular flexibility cassettes that through alternative splicing, tune the elasticity of a given myofibril.

In a more recent study, Granzier and colleagues used protein engineering and single molecule AFM techniques to measure the persistence length of multi-exon PEVK constructs (18). They showed a histogram of the measured persistence length for a PEVK protein encompassing exons 119 to 136. These combined PEVK exons showed a wide persistence length distribution, \(p = 0.3–2.7\) nm, which was very similar to the cardiac specific PEVK. Hence, these experiments did not resolve the question of whether individual exons code for protein sequences with well defined and narrowly distributed persistence length values.

The experiments that we report here were designed to test the “flexibility cassette” hypothesis. We engineered polyproteins that contained multiple repeats of a single PEVK exon, combined with an unfolding fingerprint used to identify single molecules (8, 9, 18). For our study, we chose the PEVK exons 161, 120, and 184, which contained 7, 14, and 37% proline, respectively. This range covered the proline content found in the cardiac titin N2B unique (7%) and cardiac PEVK (23%) extension of all four PEVK domains. Fig. 1D shows a histogram of the peak unfolding force for the I27 fingerprint from all 239 recordings used in our study. The mean value of the unfolding force was 196 ± 30 pN (n = 823), which was similar to our prior results (20, 21).

All single molecule traces used in our analysis had to meet the strict criteria outlined above, namely that a regular saw-tooth pattern of I27 unfolding events had to be preceded by a PEVK extension corresponding to integer multiples of a PEVK module. WLC fits to the PEVK domain (thin lavender lines in Fig. 1, A–C) are done before any I27 module unfolds. The persistence length of a folded I27 polyprotein is large and was measured to be \(p = 9.8\) nm (see supplementary information in Ref. 9), marking a rigid protein. Furthermore, an individual I27 module is likely to be much stiffer than the I27 polyprotein. Hence, it is reasonable to assume that in the case of the PEVK-I27 polyproteins studied here, I27 modules, prior to their unfolding, do not contribute to the flexibility of the PEVK domains being measured.

We were able to collect 60, 94, and 85 single molecule recordings for the PEVK exons 120, 161, and 184, respectively. WLC fits to the initial extension of the \([(\text{exon120PEVK})_3I_{27}]_4\) polyprotein (sixty recordings) measured the persistence length shown in the histogram of Fig. 2A. The bin widths are set to 0.15 nm (8), and the mean value of the measured persistence length for this exon is \(p = 0.89 \pm 0.42\) nm. The same wide distribution of persistence length values was observed for the other two PEVK exons. WLC fits to the initial extension of the \([(\text{exon161PEVK})_3I_{27}]_4\) polyprotein (94 recordings) gave an average persistence length of \(p = 0.92 \pm 0.38\) nm (Fig. 2B), whereas the fits to the \([(\text{exon184PEVK})_3I_{27}]_4\) polyprotein (85 recordings) gave an average persistence length of \(p = 0.98 \pm 0.4\) nm (Fig. 2C). Fig. 2D shows two recordings obtained from \([(\text{exon120PEVK})_3I_{27}]_4\) polyproteins, which are similar in length and have the same number of I27 unfolding events (four), yet their PEVK domains show very different values of persistence length \((p = 0.38\ \text{nm versus } p = 1.93\ \text{nm})\). These recordings highlight the variations in the persistence length of the PEVK domains.
sequences. In contrast to our expectations, we found that in all three cases the persistence length values were distributed broadly, \( p = 0.26–2 \text{ nm} \), similar to all prior multi-exon single molecule measurements of PEVK protein flexibility (8, 9, 18). Given that there are still 113 exons to be examined at the single exon level with single molecule techniques, we cannot rule out the existence of persistence length distributions that are exon-specific. However, our current data does not support the flexibility cassette hypothesis for PEVK exons. Alternatively, the diversity of PEVK exons may simply represent a strategy for the fine length adjustment of a random coil-like protein, without ever repeating the same sequence. In this scenario, all PEVK proteins would show the same range of persistence length values, in agreement with the current literature and our latest observations (Fig. 3).

It is still possible that the single molecule AFM approach blurs the true values of the persistence length of PEVK. A possible explanation for the broad distribution in persistence lengths is that, in contrast to other proteins, PEVK proteins can accumulate torsional strain. In addition to their persistence length, molecules such as DNA show torsional stiffness (twist persistence length), which affects their flexibility (26). When single molecules are picked up by an AFM tip, they will extend by uncoiling from their resting conformation. While relaxed on the surface, different proteins will have varying degrees of coiling, which upon extending will result into different amounts of torsion for each molecule. If PEVK proteins can accumulate this torsional strain, their persistence length would vary over wide ranges. However, the ability of proteins to store torsional strain has not yet been measured.

Acknowledgments—We thank Drs. H. Li and A. S. R. Koti and R. B. Robertson for their help.

REFERENCES

1. Wang, K., McClure, J., and Tu, A. (1979) Proc. Natl. Acad. Sci. U. S. A. 76, 3698–3702
2. Granzier, H. L., and Labeit, S. (2004) Circ. Res. 94, 284–295
3. Furst, D. O., and Gautel, M. (1995) J. Mol. Cell. Cardiol. 27, 951–959
4. Labeit, S., and Kolmerer, B. (1995) Science 270, 293–296
5. Linke, W. A. (2000) Histochem. Histochem. 15, 799–811
6. Granzier, H., and Labeit, S. (2002) J. Physiol. 541, 335–342
7. Watanabe, K., Nair, P., Labeit, D., Kellermayer, M. S., Greaser, M., Labeit, S., and Granier, H. (2002) J. Biol. Chem. 277, 11549–11558
8. Li, H., Oberhauser, A. F., Redick, S. D., Carrion-Vazquez, M., Erickson, H. P., and Fernandez, J. M. (2001) Proc. Natl. Acad. Sci. U. S. A. 98, 10682–10686
9. Li, H., Linke, W. A., Oberhauser, A. F., Carrion-Vazquez, M., Kerviket, J. G., Lu, H., Marszalek, P. E., and Fernandez, J. M. (2002) Nature 418, 998–1002
10. Wang, K., McCarther, R., Wright, J., Beverly, J., and Ramirez-Mitchell, R. (1991) Proc. Natl. Acad. Sci. U. S. A. 88, 7101–7105
11. Rief, M., Gautel, M., Oesterhelt, F., Fernandez, J. M., and Gaub, H. E. (1997) Science 276, 1109–1112
12. Linke, W. A., Iemeyer, M., Mundel, P., Stockmeier, M. R., and Kolmerer, B. (1996) Proc. Natl. Acad. Sci. U. S. A. 93, 8052–8057
13. Bang, M. L., Centner, T., Fornoff, F., Geach, A. J., Gotthardt, M., McNabb, M., Witt, C. C., Labeit, D., Gregorio, C. C., Granier, H., and Labeit, S. (2001) Circ. Res. 89, 1065–1072
14. Freiburg, A., Trombitas, K., Hell, W., Cazorla, O., Fougerousse, F., Centner, T., Kolmerer, B., Witt, C. C., Beckmann, J. S., Gregorio, C. C., Granier, H., and Labeit, S. (2000) Circ. Res. 86, 1114–1121
15. Lahmers, S., Wu, Y., Call, D. R., Labeit, S., and Granier, H. (2004) Circ. Res. 94, 505–513
16. Trombitas, K., Redkar, A., Centner, T., Wu, Y., Labeit, S., and Granier, H. (2000) Biophys. J. 79, 3232–3234
17. Wu, Y., Bell, S. P., Trombitas, K., Witt, C. C., Labeit, S., LeWinter, M. M., and Granier, H. (2002) Circulation 106, 1384–1389
18. Labeit, D., Watanabe, K., Witt, C., Fujita, H., Wu, Y., Lahmers, S., Funck, T., Labeit, S., and Granier, H. (2003) Proc. Natl. Acad. Sci. U. S. A. 100, 13716–13721
19. Schlierf, M., Li, H., and Fernandez, J. M. (2004) Proc. Natl. Acad. Sci. U. S. A. 101, 7299–7304
20. Carrion-Vazquez, M., Oberhauser, A. F., Fowler, S. B., Marszalek, P. E., Broedel, S. E., Clarke, J., and Fernandez, J. M. (1999) Proc. Natl. Acad. Sci. U. S. A. 96, 3694–3699
21. Li, H., Oberhauser, A. F., Fowler, S. B., Clarke, J., and Fernandez, J. M. (2000) Proc. Natl. Acad. Sci. U. S. A. 97, 6527–6531
22. Oberhauser, A. F., Marszalek, P. E., Carrion-Vazquez, M., and Fernandez, J. M. (1999) Nat. Struct. Biol. 6, 1025–1028
23. Carrion-Vazquez, M., Marszalek, P. E., Oberhauser, A. F., and Fernandez, J. M. (1999) Proc. Natl. Acad. Sci. U. S. A. 96, 11288–11292
24. Improta, S., Krueger, J. K., Gautel, M., Atkinsk, R. A., Lefevre, J. F., Moulton, S., Trewhella, J., and Pastore, A. (1998) J. Mol. Biol. 284, 761–777
25. Schimmel, P. R., and Flory, P. J. (1967) Proc. Natl. Acad. Sci. U. S. A. 58, 52–59
26. Nelson, P. (1999) Proc. Natl. Acad. Sci. U. S. A. 96, 14342–14347