# Dissecting Titin into Its Structural Motifs: Identification of an α-Helix Motif near the Titin N-Terminus[†,‡]

Giovanna Musco,*[,§] Christos Tziatzios,[||] Peter Schuck,[||] and Annalisa Pastore[§]

*EMBL, Meyerhofstrasse 1, W-69012 Heidelberg, Germany, and Institut für Biophysik der Johann Wolfgang Goethe-Universität, Frankfurt am Main, Germany*

*Received June 13, 1994; Revised Manuscript Received September 20, 1994*[⊗]

ABSTRACT: Titin, also known as connectin, is a giant modular protein specifically found in vertebrate striated muscle. Since the huge size of titin does not allow a direct structure determination, we have started a long-term project to characterize the protein by cutting it into smaller domains or structural units. The major part of the titin sequence is assembled by modules ≈100 amino acids long that belong to two major protein superfamilies. Most of these modules are linked together by stretches of variable length with unique sequence. No direct structural characterization has been achieved so far for any of these linkers. We present here a study of a stretch located in the titin N-terminus and part of a linker between two modules. Our attention was drawn toward this region because it shows 100% probability to form a coiled coil when analyzed by a prediction program. A synthetic 38 amino acid peptide spanning such a sequence was studied in aqueous solution by circular dichroism, nuclear magnetic resonance, and analytical ultracentrifugation at various pH, salt, and peptide concentrations. Under all conditions, it shows a strong tendency to form α-helical structures. In the presence of salt, this conformation is associated with the formation of helical bundles below pH 5. Above pH 5, any aggregate breaks, and the titin peptide is a monomeric helix in equilibrium with its random coil conformation. We discuss the factors which stabilize the helical conformation and the possible role of this stretch *in vivo*.

Titin—the largest protein yet described (about 3 MDa in mass and 1 μm long)—is an abundant myofibril component of vertebrate striated muscles (Maruyama et al., 1984; Wang, 1985; Kurzban & Wang, 1988; Fuerst et al., 1988; Maruyama, 1994). Titin filaments span the region between the M-line and the Z-disk in each half sarcomere (Fuerst et al., 1988; Nave et al., 1989). Although its function in the muscle is still largely unknown, accumulating evidence suggests that titin contributes to the assembly and maintenance of the sarcomere and regulates the precise length of the thick filament during myofibrilogenesis (Trinick et al., 1984; Whiting et al., 1989; Fulton & Isaacs, 1991; Isaacs et al., 1992; Gautel et al., 1993). It is also thought to maintain the thick filament in register and to resist, with its elastic properties, any overstretch of the sarcomere (Wang et al., 1993).

Recent sequencing of titin has shown that it has a quite diversified and modular architecture (Labeit et al., 1990, 1992; Gautel et al., 1993). Most of the sequence is assembled by ≈100 amino acid repeats, named type I and type II, which have been suggested by sequence analysis to belong to the fibronectin III and immunoglobulin superfamilies, respectively (Labeit et al., 1990). The modularity of titin offers the possibility of dissecting it into several independent domains of size suitable for structure determination. The study of selected modules and their interfaces can then be related to the structure and the function of the overall protein. Such an approach has been successfully applied to the characterization by NMR[1] of titin type II modules (Politou et al., 1994).

In addition to the modular repeats, the remaining 15% of the titin sequence is composed of insertions of unique sequence which do not share any similarity with other proteins in the sequence database. Structural information about any of these regions is therefore extremely valuable for understanding the structural/functional role of these linkers as well as the way the repetitive modules are assembled. While analyzing the titin sequence, a short nonrepetitive region of about 40 amino acids was identified near the protein N-terminus (Labeit et al., 1994; the accession number for the EMBL data library is X83270). It is part of an approximately 150 residue insertion which links two type II modules (Figure 1a). This stretch not only has a helical propensity with strong amphiphilic character but also shows an unusually high score (90−100% probability) when the whole titin sequence was analyzed by a coiled coil prediction program (Figure 1b; Lupas et al., 1991). A coiled coil is a widespread structural motif which often acts as a highly stable polymerization domain. It is composed of amphiphilic α-helices that are highly unstable in hydrophilic media and therefore tend to dimerize or form higher aggregates in which the hydrophobic faces pack against each other (Cohen & Parry, 1986, 1990). Coiled coils are best known in several fibrous proteins such as tropomyosin, myosin, and keratin (Cohen & Parry, 1990; Phillips et al., 1986) as well as in a

[†] This work was supported by the Human Frontier Science Program Organization (A.P.), by the Deutsche Forschungsgemeinschaft (SFB169) (C.T. and P.S.), and by a DAAD fellowship (G.M.).

[‡] The titin peptide sequence in this paper has been submitted to the EMBL data library under the accession number X83270.

* To whom correspondence should be addressed. Telephone: +49 6221 387457. Fax: +49 6221 387306.

[§] EMBL.

[||] Institut für Biophysik der Johann Wolfgang Goethe-Universität.

[⊗] Abstract published in *Advance ACS Abstracts,* November 15, 1994.

[1] Abbreviations: CD, circular dichroism; NMR, nuclear magnetic resonance; TPPI, time-proportional phase incrementation; 1D and 2D, one and two-dimensional; TOCSY, total correlation spectroscopy; NOE, nuclear Overhauser enhancement; NOESY, 2D NOE spectroscopy; TFE, trifluoroethanol; $T_m$, melting temperature.

0006-2960/95/0434-553$09.00/0 

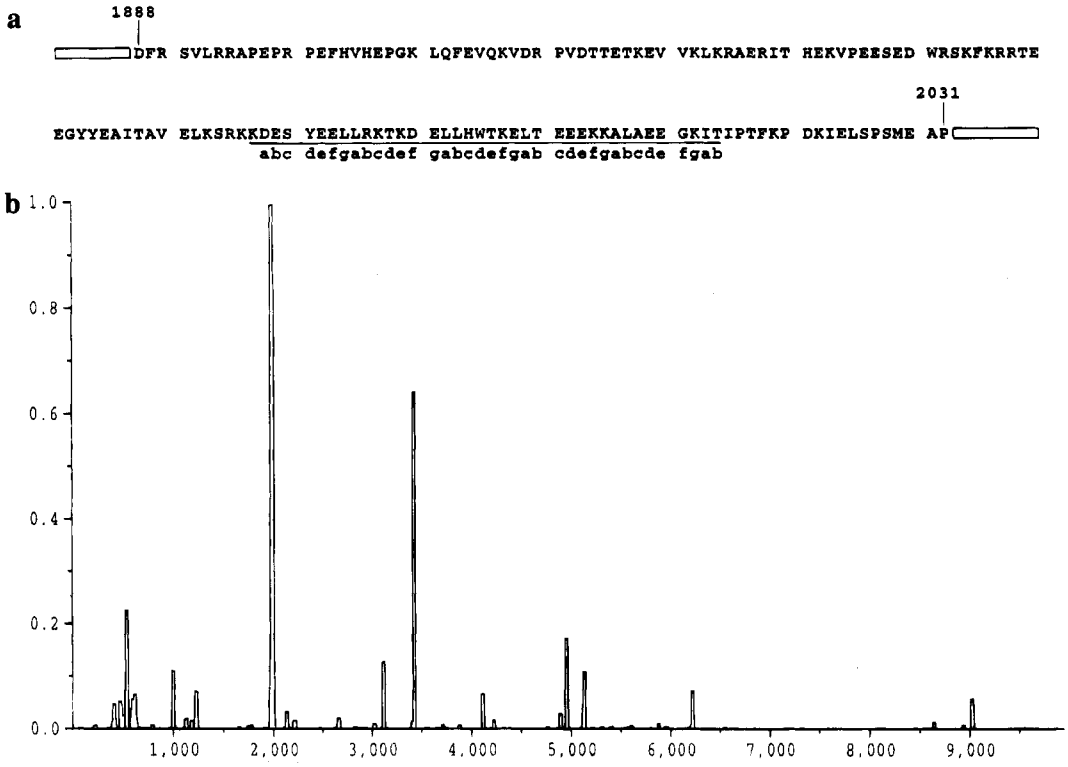


FIGURE 1: (a) Partial sequence of titin in the region around the stretch studied in this work. The linker connects two immunoglobulin-like motifs here represented by boxes flanking both ends. The sequence of the putative coiled coil is underlined. The coiled coil frame is indicated below the sequence with lower case alphabetic letters from a to g. (b) Plot representing the coiled coil score as calculated according to Lupas et al. (1991) using a 28-residue window versus the titin N-terminal sequence. The region studied in this work corresponds to the peak at residues 1977−2014 from the N-terminus. According to this prediction, the probability to have a coiled coil in this region is 1.0 (100% probability). If the program is run on the rest of the titin sequence, only maxima below 30−40% are detected. However, the same plot obtained using a window of 14 shows a much lower probability (10−20%) at L17-H18-W19-T20.

class of DNA-binding proteins (Landschulz et al., 1988; O'Shea et al., 1991; Saudek et al., 1991).

If the region identified in titin were a coiled coil, it could play a significant role in titin self-association. Sequence analysis shows that this region has some but not all of the requirements for the formation of a coiled coil, such as the presence of heptad repeats with hydrophobic amino acids at every first ('a' position) and fourth ('d' position) residue (Figure 1a). Amino acids in these positions form the interface between the α-helical chains of the coil. The 'a' positions of the titin region are occupied by bulky hydrophobic residues (LLLAI). However, most of the residues at the 'd' position (Y, T, and W) are rarely found at such a position in parallel coiled coils (Cohen & Parry, 1990).

In order to characterize the structure of this fragment and to verify whether it represents an autonomous substructure, a 38-residue peptide containing the putative coiled coil region was synthesized, and its structural properties were studied by CD, NMR, and ultracentrifugation.

## MATERIALS AND METHODS

*Peptide Synthesis.* The peptide was synthesized and purified by the Peptide Synthesis Group at EMBL using standard solid-phase methods. The N- and C-termini were free and amidated groups, respectively. Peptide purity (≥98%) was determined by HPLC, and composition was confirmed both by amino acid analysis and electrospray mass spectrometry.

*Circular Dichroism.* CD measurements were performed on a Jasco J-710 spectropolarimeter equipped with a thermostatically controlled cell holder stabilized by circulating water from a Neslab RTE-110 water bath. Rectangular quartz cuvettes with 0.02−0.1 cm path length were used (Hellma). The instrument was calibrated with d-10-camphorsulfonic acid solution. All the spectra were baseline corrected by subtracting buffer spectra, and the observed ellipticity was converted to mean residue weight ellipticity $[\theta]/(\text{deg} \cdot \text{cm}^2 \cdot \text{dmol}^{-1})$. The helical populations were estimated from the measured mean residue ellipticity at 222 nm assuming a value of $-39500 \times (1-2.57/n)$ (where $n$ is the peptide length) for 100% helicity (Chen et al., 1974). The peptide concentration was determined from the tyrosine and tryptophan absorbance at 280 nm in 6 M guanidine hydrochloride. Extinction coefficients of 5690 and 1280 $M^{-1}$ $cm^{-1}$ reported for *N*-acetyl-L-tryptophanamide and for Gly-L-Tyr-Gly at 280 nm were used (Edelhoch, 1967). The pH was adjusted by diluting aqueous solutions of peptide either in 10 mM $NaH_2PO_4/Na_2HPO_4$ or in 10 mM $CH_3COOH/CH_3COONa$ buffer. The effect of different salts was screened by using NaF, NaCl, and $MgSO_4$. When possible, NaF was preferred to other salts because of its higher transparency in the far-UV region. CD spectra were also obtained in 50% TFE/$H_2O$ at pH 7.0. Five scans obtained by collecting data at 0.5-nm intervals from 250 to 185 nm were usually recorded and averaged for each spectrum. The temperature-dependent folding and unfolding were followed by monitoring the ellipticity at 222 nm from 4 to 80 °C at a heating rate of 20 °C/h. The spectra were processed with the program SNARF (Frans van Hoesel, University of Groningen) and plotted with GNUPLOT.

*Nuclear Magnetic Resonance.* The NMR measurements were carried out using 1−2 mM samples in 90% $H_2O$/10% $D_2O$. 1D NMR spectra to study protein aggregation were recorded both on a Bruker AM-300 and on a Bruker AMX-500. Salt titration experiments were repeated using different salts (KCl, NaCl, and $MgSO_4$). All 2D NMR spectra were acquired on a AMX-600 spectrometer in phase sensitive mode (TPPI) (Marion & Wüthrich, 1986) either with pre-irradiation of the water resonance or with selective excitation (Watergate pulse sequence, Piotto et al., 1992). Clean TOCSY spectra (Griesinger et al., 1988) were measured using the TOWNY composite pulse cycle (Kadkhodaei et al., 1993). Mixing times used were in the 50−100-ms range for the TOWNY and 80−200 ms for the NOESY experiments. 2D spectra were recorded at 17 and 27 °C, with 2048 data points in the acquisition domain and 512 data points in $t_1$. Prior to Fourier transformation, the data were zero-filled to 2048 points in the $t_1$ dimension and weighted with a Gaussian window in $t_2$ and a sine window in $t_1$. A baseline correction was performed in both dimensions using a polynomial. Data were processed on a Bruker X-32 station using the UXNMR program. The AURELIA program was used for displaying and plotting spectra. Spectra were referred to the 2,2-dimethyl-2-silapentane-5-sulfonate (DSS) resonance.

*Analytical Ultracentrifugation.* Sedimentation equilibrium experiments were performed in a Beckman Optima XL-A analytical ultracentrifuge, using an An-60 Ti rotor and Epon double sector cells of 12 mm path length. The sample volume was 200 μL. The measurements were recorded at a rotor temperature of 4 °C and a rotor speed of 35000 and 25000 rpm. The absorption profiles were recorded at 275 nm. The partial specific volume of the peptide of 0.743 mL/g was calculated according to the method suggested by Durchschlag (1986). The experimental sedimentation profiles were analyzed according to the following equation:

$$c(r) = \sum c_i(r_0) \exp M_i(1 - V\varrho_L)(r_0^2)(\omega^2/RT)$$

where $c$ is the peptide concentration; $r$ is the distance from the center of rotation; $r_0$ is an arbitrary reference radius; $M_i$ is the relative mass of component $i$; $V$ is the partial specific volume; $\varrho_L$ is the buffer density; $\omega$ is the rotor speed; $R$ is the gas constant; and $T$ is the absolute temperature. Least-squares fits and the analysis of the statistical accuracy of the parameters (including the baseline) were performed according to the method suggested by Schuck (1994). Concentration distributions which could not be fitted using the above equation were analyzed applying the following equation, which takes into account thermodynamic nonidealities:

$$c(r) = c(r_0) \exp M(1 - V\varrho_L)(r^2 - r_0^2)(\omega^2/RT) - B_1 M + V)(c(r) - c(r_0))$$

where $B_1$ is the second virial coefficient. This equation describes the concentration distribution of one peptide component.[2] The dependence of $B_1$ on the electrolyte concentration was used to calculate the number of net charges per molecule (Roak & Yphantis, 1971). The experiments were performed at pH 3.1 and 7.0, under different conditions of peptide concentration and ionic strength. At pH 7.0, loading concentrations were between 0.22 and 0.73 mg/mL in buffers containing 44 mM potassium phosphate with 0, 43, and 87 mM KCl. At pH 3.1 (17.5 mM acetic acid, 0−2.2 M KCl), loading concentrations were between 0.31 and 0.42 mg/mL.

[2] It is obtained by integrating eq 5.13 of Fujita et al. (1975) and is equivalent to eq A-6 of Roak & Yphantis (1971).

Table 1: Residue Molar Ellipticity [θ] at Different pH and, for Comparison, in TFE[a]

| pH | [θ] | helical % |
|---|---|---|
| 3.0 | −14800 | 40 |
| 4.5 | −13272 | 36 |
| 7.1 | −11067 | 30 |
| 8.0 | −10811 | 30 |
| 10.8 | −10868 | 30 |
| 12.0 | −8037 | 22 |
| TFE | −31200 | 85 |

[a] CD spectra were recorded at 8.8 μM peptide concentration and 4 °C.

*Predictions and Database Search.* The coiled coil prediction was achieved using the approach of Lupas et al. (1991) with 14, 21, and 28 residue windows. Data base homology searches were run using the BLITZ program (Rice et al., 1993). Helical propensity calculations were kindly performed by L. Serrano according to the method proposed by Muñoz and Serrano (1994; Jiménez, 1994). The conditions imposed in the calculations were pH 7.0 and 3 °C. The N- and C-termini were assumed to be free and amidated groups, respectively, according to the peptide protective group.

## RESULTS

*The Titin Peptide Has Significant Helical Content in Aqueous Solution.* The conformational properties of the titin peptide under several conditions were first assessed by CD spectroscopy. The high sensitivity of this technique allows the screening of a wide range of experimental conditions without consuming large amounts of sample. Far-UV spectra (185−250 nm) of the peptide recorded in the pH range between 3.0 and 12.0 are typical of α-helical structures with a maximum at 193 nm and a local double minimum at around 208 and 222 nm (Greenfield & Fasman, 1969). A significant α-helical content as assessed by the depth of the minimum at 222 nm (Chen et al., 1974) is present in the whole pH range explored and increases at lower pH (see Table 1). Addition of 50% TFE, a solvent which stabilizes helical conformations, increases the helical content up to 85%.

*Helicity Is Stabilized by Increasing Concentrations of Salt at Low pH but Is Independent of the Ionic Strength at Neutral pH.* Salt contribution to the secondary structure was screened by the addition of increasing amounts of different salts to peptide solutions at pH 3.1 and 7.0. Comparable results were observed when using NaF, NaCl, KCl, and $MgSO_4$. As an example, Figure 2a shows the effect of NaF on the CD spectrum at pH 3.1. The molar residue ellipticity at 222 nm is strongly enhanced by salt. At this pH, the helical percentage increases noticeably and has been estimated to vary from 40 up to 70% in the salt concentration range from 0 to 200 mM at 4 °C. At pH 7.0, the addition of salt does not affect the CD spectrum (data not shown), and the α-helical content at 4 °C remains around 30% both in the presence and in the absence of salt.

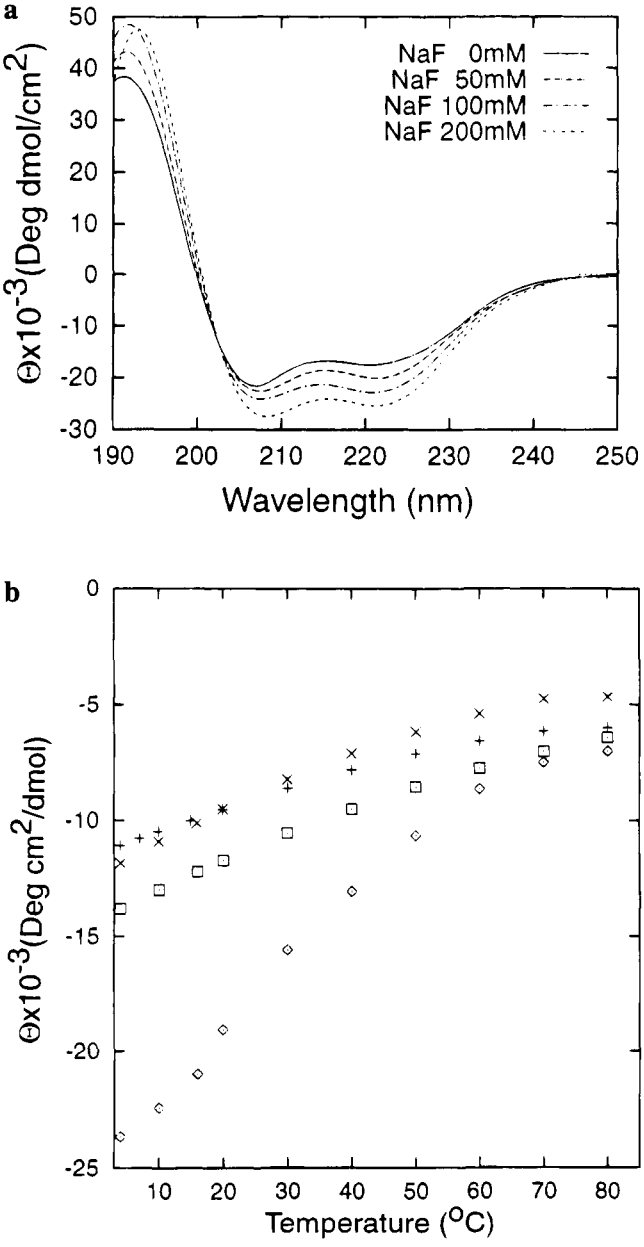


FIGURE 2: Influence of the ionic strength as recorded by far-UV CD experiments. (a) Influence of the ionic strength on the mean molar ellipticity in the 180−250-nm range at pH 3.1 as measured at 4 °C. The sample concentration was 88 $\mu$M. (b) Temperature dependence of the mean residue ellipticity ([$\theta$]) at 222 nm recorded for a 88 $\mu$M peptide sample at different pH and salt conditions. The following symbols were used: diamonds and squares indicate measurements performed at pH 3.1 (10 mM acetate buffer) in the presence and in the absence of 100 mM NaF, respectively; stars and crosses indicate measurements performed at pH 7.0 (10 mM phosphate buffer) in the presence and in the absence of 100 mM NaF, respectively.

Temperature denaturation curves were obtained by recording CD spectra at different pH's and ionic strengths as a function of temperature in the range between 4 and 80 °C. The spectra were all reversible by lowering the temperature back to the initial value. At increased temperatures, the CD spectra at neutral and acidic pH (data not shown), both in the presence and in the absence of salt, show an isodichroic point at 203 nm. This is consistent with a two-state transition from a folded state to a disordered one (Greenfield & Fasman, 1969). Figure 2b shows a comparison between denaturation curves, again detected at both pHs and in the presence and in the absence of salt. For all curves except the one collected at low pH with 100 mM NaF, the helical content decreases gradually with temperature without sharp transitions. This lack of cooperativity has been observed before for the denaturation of single helices where the unfolding is thought to proceed from the ends inward or else randomly and segmentally along the entire helix (Wang et al., 1991). Addition of salt does not have any influence on the thermal unfolding at neutral pH (Figure 2b). On the contrary, the thermal denaturation curve at pH 3.1 and 100 mM NaF shows a smooth cooperative helix to coil transition with an estimated $T_m$ (the temperature at which 50% of the peptide is in its unfolded form) around 30 °C.

*A Concentration Dependence on the Helix Content Is Observed Only at Acidic pH and High Ionic Strength.* Tendency to aggregation was checked by screening the concentration dependence of the CD signal both under native conditions and during thermal unfolding. As a first step, the ellipticity at 222 nm was monitored as an indication of the helical content in subsequent experiments at different concentrations. Figure 3a shows that the helix percentage is independent of concentration both at pH 3.1 in the absence of salt and at pH 7.0 in the whole range of salt concentration checked. At pH 3.1 and in the presence of 100 mM NaF, the helicity increases with the peptide concentration up to a plateau above 100 $\mu$M.

Temperature denaturation curves recorded at different concentrations show virtually superimposable unfolding curves at pH 7.0 (data not shown). The observed gradual decrease in ellipticity together with the concentration independence of the band intensity at 222 nm suggest the absence of cooperative interactions and indicate that, under these conditions, helix formation is a monomolecular process and is not stabilized by aggregation. Similar behavior is observed at low pH in the absence of salt. At pH 3.1 and in the presence of salt, a clear concentration dependence is observed (Figure 3b). This evidence, in agreement with all the other CD data, indicates that aggregation is only promoted by salt at low pH.

*Aggregation Which Depends on pH and Salt Concentration Was Monitored by 1D NMR Spectroscopy.* 1D NMR experiments were carried out both to compare the peptide behavior at NMR concentrations with the data collected by CD and to explore the pH range under which salt-dependent aggregation is observed. Under suitable conditions, NMR experiments may in fact give a faster answer, avoiding the much more time-consuming denaturation curves. We recorded two sets of experiments. The first (shown in Figure 4a−d) was carried out at a fixed pH (pH 3.0) adding increasing quantities of KCl. The addition of salt caused a noticeable broadening of the line widths over the whole spectrum (for clarity only the amide region is shown in Figure 4). The effect is particularly evident for the indolic proton resonance of W19 (resonance at 10.15 ppm) which completely disappears at 200 mM KCl. The line width observed in an NMR spectrum is inversely proportional to the relaxation time T2, which itself depends on the peptide size. The observation of a line broadening could therefore correspond to the formation of large aggregates. The alternative, i.e., the possibility that a line broadening may be related to exchange phenomena, can be excluded in our case because comparable line broadenings were observed at two different fields (300 and 500 MHz, data not shown).

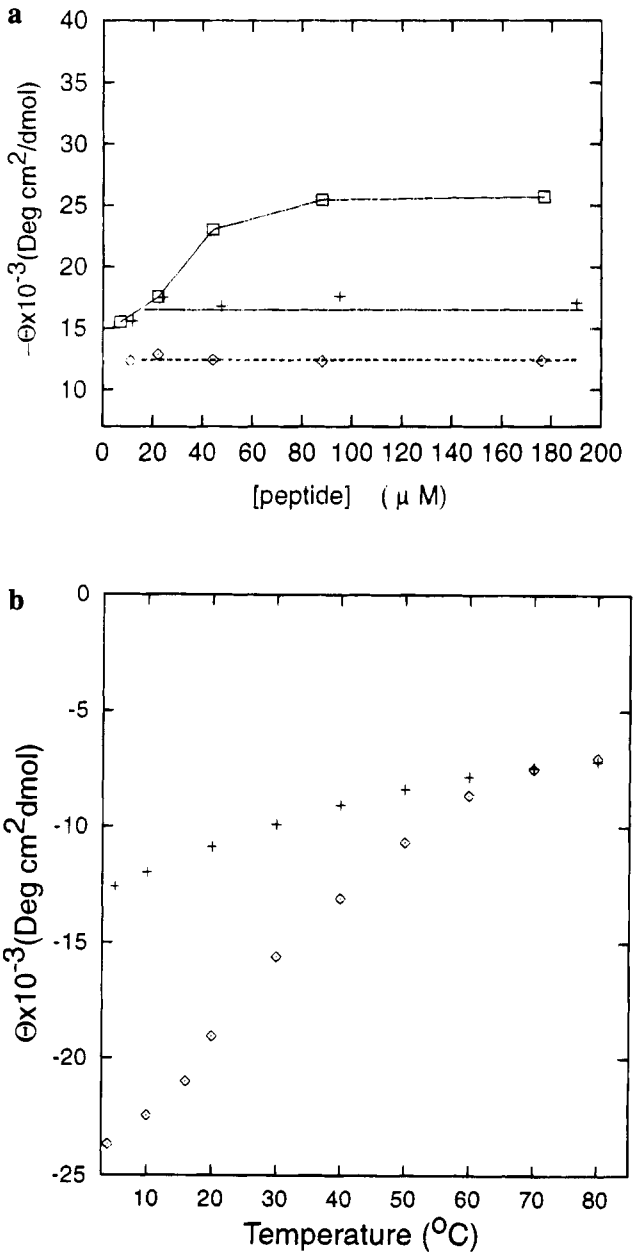


FIGURE 3: Concentration dependence recorded by far-UV CD measurements. (a) Mean residue ellipticity ([$\theta$]) reported as a function of the peptide concentration at 4 °C for different pH's and ionic strengths. Diamonds indicate the data recorded at pH 7.0. Crosses and squares are used for data recorded at pH 3.1 in the absence and in the presence of salt (100 mM NaF), respectively. (b) Temperature dependence of the mean residue ellipticity ([$\theta$]) at 222 nm recorded for different peptide concentrations at pH 3.1 in 10 mM acetate buffer and 100 mM NaF. Diamonds and crosses indicate 88 and 8 $\mu$M, respectively.

Exchange phenomena in the intermediate rate would lead instead to a field-dependent broadening, which would tend to disappear at high temperatures (Neuhaus & Williamson, 1989). Disappearance of the tryptophan indolic signal may be explained by the progressive formation of aggregates that trap this side chain in a rigid environment.

A second set of experiments (shown in Figure 4e−g) was performed with fixed salt concentration at different pHs. The spectrum at pH 5.0 shows evident line broadening roughly comparable with that obtained at low pH in the presence of salt. However, changing the pH from 5.0 to 6.0 is sufficient to obtain sharper lines with a line width comparable to that

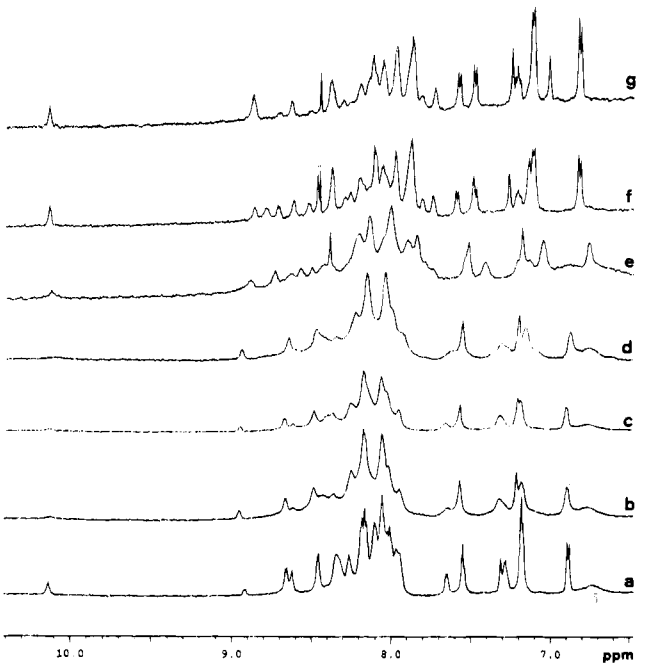


FIGURE 4: Effect of the ionic strength and of the pH on the amide region of the 1D NMR spectrum as detected at 500 MHz, 17 °C, and 1.5 mM peptide concentration. (a−d) Spectra recorded at pH 3.0 in the presence of 0, 50, 100, and 200 mM KCl, respectively. (e−g) Spectra recorded at pH 5.0, 6.0, and 6.7, respectively, in the presence of 100 mM NaCl and 20 mM MES buffer.

expected for peptides of similar size. Further increase of the pH to 6.7 results only in the slightly higher saturation of the amide signals expected at this pH.

From these results, we may conclude that the turning point for aggregation is between pH 5 and 6. At pH 5.0 and below, salt induces aggregation of helical species in agreement with the enhanced helical signal observed by CD under these conditions. Above pH 5.0, the presence of a helical monomeric species is not influenced by salt.

*Ultracentrifugation Studies Were Carried Out To Quantify the Nature of the Aggregate.* Sedimentation equilibrium experiments by analytical ultracentrifugation were carried out to determine the state of association of the peptide under various conditions.

At pH 7.0, the experimental data can be fitted with high precision under the assumption that the peptide is present only as a monomer, independent of ionic strength and loading concentration conditions (Figure 5a). Significant formation of the dimer can be excluded to a lower limit of 15 mM for the monomer−dimer dissociation constant. At pH 3.1, strong nonideal behavior was observed at low salt concentration (below 20 mM KCl), which indicates repulsive electrostatic interactions. To correct for nonideality, the second virial coefficient (which represents a measure of the strength of the interactions between particles) had to be introduced into the fitting equation (Fujita et al., 1975). The dependence of the second virial coefficient on the electrolyte concentration corresponds to +10.8 ± 1.5 positive charges, in agreement with the number of charges expected at this pH (+10). The analysis of several concentration profiles shows that under these low salt conditions the peptide is monomeric. At a KCl concentration of 20 mM, the peptide sediments as an ideal monomeric component. At higher salt concentration, however, it aggregates increasingly into dimers and tetramers (see Figure 5b). The influence of the KCl concentration on

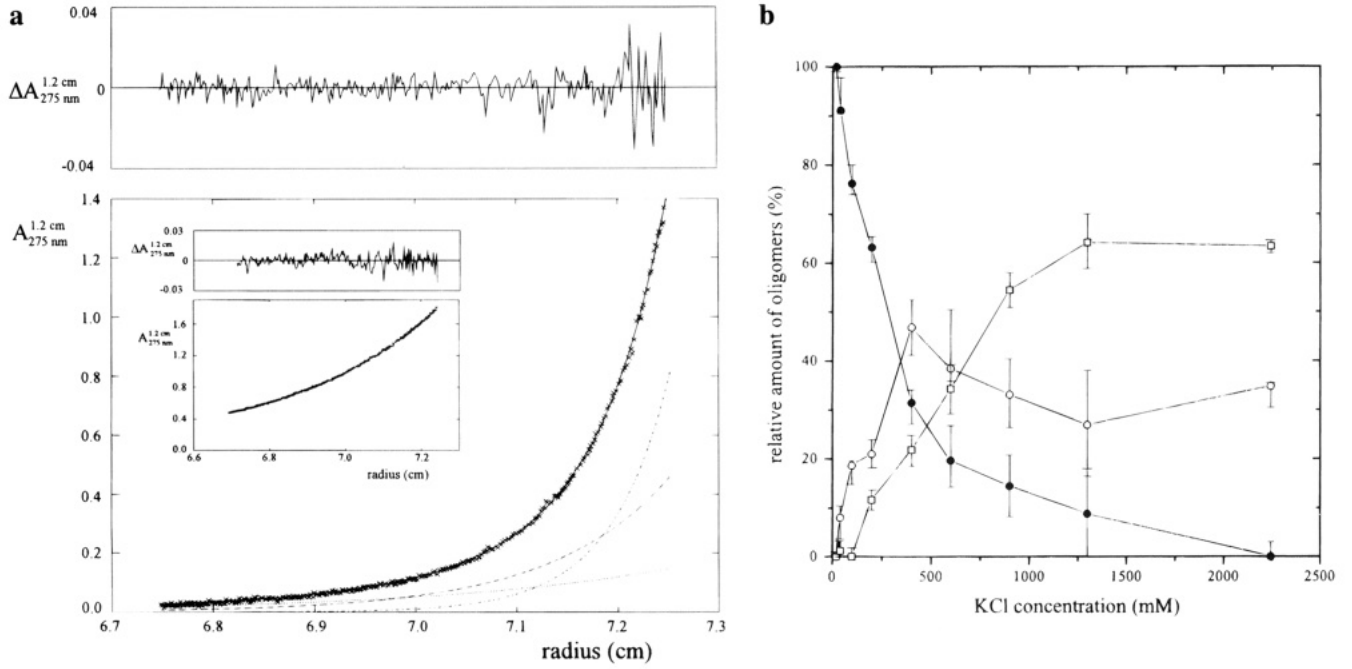


FIGURE 5: (a) Sedimentation equilibrium of the titin peptide in the analytical ultracentrifuge at 35000 rpm and 4 °C. The loading concentration was 75 $\mu$M ($A_{275nm}^{1.2cm} = 0.47$) in 17.5 mM acetic acid (pH 3.1) and 900 mM KCl. Lower plot: Experimental absorbance $A_{275nm}^{1.2cm}(r)$ as a function of the radial position (X), with the calculated distribution of titin oligomers. The plot also shows the calculated absorbance contributions of monomer (•••), dimer (- - -), and (— • —). (Upper plot) Local differences $\Delta A_{275nm}^{1.2cm}(r)$ between experimental and calculated values. (Inset, lower plot) Sedimentation equilibrium of the titin peptide in the analytical ultracentrifuge at 25000 rpm and 5 °C. The loading concentration was 0.16 mM ($A_{275nm}^{1.2cm} = 1.0$) in 43 mM potassium phosphate (pH 7.0) and 87 mM KCl. (Inset, upper plot) Local differences between experimental and calculated values of a fit assuming that the peptide sediments as a single ideal monomeric species (- - -). (b) Influence of the KCl concentration on the relative amount of monomer (closed circle), dimer (open circle), and tetramer (square) of the titin peptide in sedimentation equilibrium at 35000 rpm and 4 °C. Loading concentration was held constant at 80 $\mu$M with 17.5 mM acetic acid (pH 3.1). The analysis was based on a monomer/dimer/tetramer model of self-association. The relative amounts were calculated by averaging the components contributions to the fitted absorption profiles over the sample volume. The error bars include the correlation between the calculated component concentrations and the baseline.

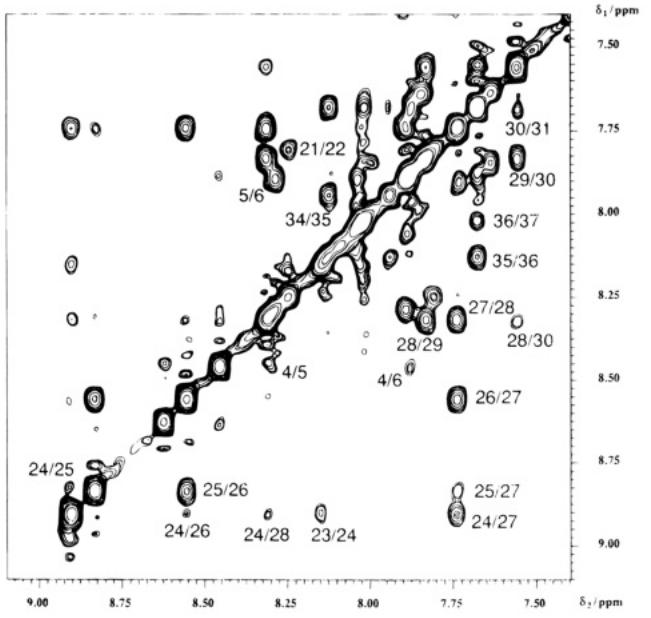


FIGURE 6: Amide region of a NOESY spectrum at 600 MHz with a 200-ms mixing time. The experimental conditions were 1.5 mM peptide concentration at 17 °C and pH 7.3. The assignment of only some of the amide—amide connectivities is shown in the figure.

the relative contribution of monomer, dimer, and tetramer to the fitted absorption profiles (averaged over the sample volume) is shown in Figure 5b. At approximately 2 M KCl, no monomeric peptide is left although, even under these conditions, no evidence of oligomers larger than tetramers is found. Although we cannot in principle exclude completely the presence of trimer formation, the introduction of a trimeric component in the data analysis does not improve the quality of the fitting at all salt concentrations. In contrast, dimer formation can be established with high significance in the range from 40 to 400 mM KCl, while tetramer formation becomes significant at salt concentration ≥900 mM KCl.

*At Neutral pH, the Titin Peptide Shows Two Distinct Helical Regions.* We decided to focus on the behavior of the titin peptide at neutral pH and further characterize it under these conditions of greater biological relevance.

Proton resonance assignment of the peptide at pH 7.3 was achieved using clean TOCSY and NOESY experiments following standard procedures (Wüthrich, 1986). Sequential assignment was accomplished by concatenation of HN—HN sequential connectivities (Figure 6). Peak overlaps caused several interruptions in the $H^{\alpha}$—HN connectivities. Residues unique in the sequence, such as S4, Y5, H18, W19, and G35 were used as starting points for the sequential assignment. A list of the assignments is shown in Table 2.

The observation of virtually all the amide connectivities at this pH is already an indication of the presence of some secondary structure. The exchange rate, unfavorable at neutral conditions, would otherwise result in broadening up to disappearance of these signals (Wüthrich, 1986). A noticeable upfield shift from the random coil values of most of the resonances and in particular the α protons as indicated in Figure 7B is consistent with a helical conformation

Table 2: Chemical Shifts in ppm of Assigned $^1$H NMR Resonances of Titin Peptide at 600 MHz and 17 °C

| | NH | $C_\alpha H$ | $C_\beta H$ | $C_\gamma H$ | $C_\delta H$ | $C_\epsilon H$ | others |
|---|---|---|---|---|---|---|---|
| 1 Lys | | | | | | | |
| 2 Asp | | 4.44 | 2.45 2.62 | | | | |
| 3 Glu | 8.63 | 4.12 | 1.72 | 1.83 2.08 | | | |
| 4 Ser | 8.45 | 4.23 | 3.75 3.89 | | | | |
| 5 Tyr | 8.29 | 4.17 | 2.82 2.95 | | | | 6.93(CH2,6) 6.62(CH3,5) |
| 6 Glu | 7.88 | 3.89 | 1.83 | 1.92 2.10 | | | |
| 7 Glu | 8.27 | 3.73 | 1.81 | 2.13 | | | |
| 8 Leu | 7.89 | 4.05 | 1.39 1.44 | 1.54 | 0.64 0.69 | | |
| 9 Leu | 7.83 | 3.87 | 1.40 | 1.29 | 0.60 0.68 | | |
| 10 Arg | 7.63 | 3.90 | 1.53 1.68 | 1.41 | 3.00 | | 7.29(NH) |
| 11 Lys | 7.85 | 4.10 | 1.69 | 1.30 | 1.40 | 2.74 | |
| 12 Thr | 7.98 | 4.10 | 4.04 | 1.06 | | | |
| 13 Lys | 8.02 | 3.96 | 1.66 | 1.29 | 1.46 | 2.75 | |
| 14 Asp | 8.08 | 4.32 | 2.46 2.51 | | | | |
| 15 Glu | 8.04 | 3.96 | 1.87 | 2.07 2.17 | | | |
| 16 Leu | 7.88 | 4.07 | 1.39 1.44 | 1.54 | 0.59 | | |
| 17 Leu | 7.73 | 4.04 | 1.38 1.44 | 1.22 | 0.63 0.71 | | |
| 18 His | 7.90 | 4.41 | 2.90 2.98 | | | | 7.90 (CH2) 6.89 (CH4) |
| 19 Trp | 7.88 | 4.47 | 3.09 3.12 | | | | 7.07 (CH2) 7.39 (CH4) 6.92 (CH5) 7.02 (CH6) 7.29 (CH7) 10.05 (NH) |
| 20 Thr | 7.73 | 3.98 | 3.90 | 0.89 | | | |
| 21 Lys | 7.81 | 3.97 | 1.55 | 1.19 | 1.54 | 2.80 | |
| 22 Glu | 8.24 | 4.07 | 1.71 1.82 | 2.09 | | | |
| 23 Leu | 8.14 | 4.35 | 1.38 | 1.17 | 0.60 | | |
| 24 Thr | 8.91 | 4.23 | 4.55 | 1.17 | | | |
| 25 Glu | 8.83 | 3.80 | 1.87 | 2.16 | | | |
| 26 Glu | 8.55 | 3.82 | 1.78 1.86 | 2.13 2.20 | | | |
| 27 Glu | 7.73 | 3.83 | 1.76 1.87 | 2.06 2.16 | | | |
| 28 Lys | 8.31 | 3.70 | 1.66 | 1.14 | 1.39 | 2.68 | |
| 29 Lys | 7.83 | 3.89 | 1.68 | 1.40 | 1.30 | 2.69 | |
| 30 Ala | 7.55 | 4.02 | | 1.28 | | | |
| 31 Leu | 7.67 | 3.96 | 1.40 | 1.17 | 0.68 | | |
| 32 Ala | 7.67 | 4.03 | | 1.30 | | | |
| 33 Glu | 7.88 | 3.97 | 1.92 | 2.10 2.21 | | | |
| 34 Glu | 7.93 | 4.01 | 1.90 | 2.10 2.20 | | | |
| 35 Gly | 8.12 | 3.76 | | | | | |
| 36 Lys | 7.67 | 4.14 | 1.68 | 1.40 | 1.55 | 2.78 | |
| 37 Ile | 8.01 | 4.13 | 1.73 | 1.30 | 0.73 | 0.66 | |
| 38 Thr | 8.01 | 4.13 | 4.04 | 1.02 | | | |

(Pastore & Saudek, 1990; Wishart et al., 1991). The NOESY spectrum of the peptide is characteristic of an α-helical conformation. From a survey of sequential and medium-range NOE connectivities (Figure 7A), it is again possible to identify two helical regions extending approximately from residues 5 to 17 and from residues 25 to 32. The N-terminal region from 2 to 5 shows strong $H^\alpha-HN(i,i+1)$ and weak $HN-HN(i,i+1)$ connectivities, which are indicative of a short extended region in the N-terminus. A succession of several intense sequential NOESY cross peaks between amide protons is indicative of regions with helical structure (Wüthrich, 1986). Strong NOEs corresponding to short distances between adjacent amide protons starting from residues 3 to 12 and from 25 to 31 are present (the interruption at residue 7, 8, and 11 is due to cross peaks overlapping or to connectivities hidden under the diagonal). In addition, the presence of both medium-range NOEs of $H^\alpha-HN(i,i+3)$ and $H^\alpha-H^\beta(i,i+3)$ is another indication of helix formation (Wüthrich, 1986). No connectivities involving more than four positions apart in the sequence are observed, the only exception being an unusual NOE between HN of K28 and the $H^\delta$ of L23. This NOE has been typically found in capping boxes and involves the residue before the S (or T) at position $n$ of the capping box and residue $n+4$ (Serrano, personal communication). In the region around residue T24, a number of nonhelical contacts are present between T24 and E27 (between HN of T24 and the HN, $H^\beta$, and $H^\gamma$ of E27). Additional contacts are also present between the $H^\beta$ of T24 and backbone amides of E25 and E26 (although the latter contact could also involve the amide of E27 instead of E26). These NOEs are consistent with what is expected for capping boxes (Lyu et al., 1993; Zhou et al., 1994). The presence of a capping box involving residues S4 and E7 is suggested by the sequence but could not be supported experimentally because of resonance overlap.

Accordingly, the presence of a break both in the NOE pattern and in the secondary chemical shift can be compared with helix forming tendency coefficients as calculated for instance by Muñoz and Serrano's (1994; Jiménez et al., 1994) method (Figure 7C). These coefficients, calculated for each amino acid in a given sequence and corrected for their local environment have been found to correlate well with experimental results. They show two helical regions interrupted around residues 22−24 that have a lower tendency to form an α-helix. Overall, the C-terminal region shows lower helical propensity than the N-terminus.

## DISCUSSION

We have presented here data which show that a 38 amino acid peptide spanning the sequence of a stretch of the titin N-terminus has a significant tendency to form α-helical conformations in aqueous solution. This conformation is associated with aggregation below pH 5.0 in the presence

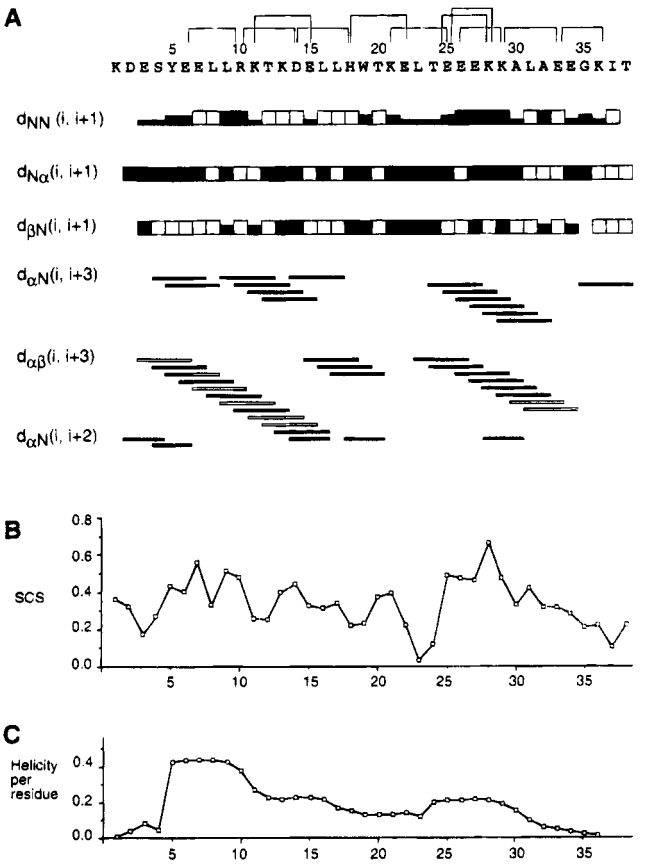


FIGURE 7: (A) Survey of NOE effects recorded for the titin peptide at pH 7.3. Different classes of NOE effects are indicated. Thicker bars correspond to stronger cross peak intensity. Open bars are used when the peak could not be observed by spectral overlap. (B) SCS indicates the secondary chemical shift profile for the $H^{\alpha}$ protons as calculated by subtracting the observed chemical shift from the corresponding random coil value. (C) The values of helix propensity are reported versus the sequence as calculated according to Muñoz and Serrano (1994; Jiménez et al., 1994). Residue pairs which are thought to stabilize the helix by salt bridge formation are indicated above the sequence. This effect should be stronger according to the pattern: $i + 4AB \geq i + 4BA \approx i + 3AB \geq i + 3BA$, where B and A represent a basic and an acidic residue, respectively (Huyghues-Despointes et al., 1993).

of salt. As shown by the enhancement of the helical content under these conditions (Table 1), the aggregate is not unfolded but forms helical bundles. Whether or not the helices in the bundle twist around each other to form a coiled coil is difficult to decide conclusively from our data. A ratio of the intensities of the minima at 208 and 222 nm around or greater than unity has been suggested as evidence for the presence of coiling (Zhou et al., 1992; Greenfield et al., 1993). Although the spectra in Figure 2a show this trend, the ratio remains below 1 for the salt concentrations explored by CD (it is 0.92 at 200 mM NaF). Above pH 5.0 and up to basic pH, the titin peptide is a monomeric helix in equilibrium with its random coil (30% helix present).

It is commonly agreed that helices are seldom stable in aqueous solutions unless stabilized by interactions either with the environment (e.g., self-association, packing to other structural elements, structuring solvents) or by specific intramolecular contributions (e.g., salt bridges, helix−dipole interactions) (Scholtz & Baldwin, 1992). This titin fragment displays both mechanisms in a pH-modulated fashion. At all pH's, the helix is stabilized by the large number of both positively and negatively charged residues, especially glutamates and lysines, with up to 11 possible intrahelical salt bridges (Figure 7a). Indeed, it is established that salt bridge formation between Glu and Lys or Arg three or, more importantly, four residues apart plays an important role in stabilizing α-helices even at extreme pH's (Sundaralingam et al., 1985; Marqusee & Baldwin, 1987, 1990; Merutka & Stellwagen, 1991; Gans et al., 1991; Lyu et al., 1992; Huyghues-Despointes et al., 1993). In comparison, the repulsive contribution of equally charged residues has a smaller destabilizing effect on helix formation (Lyu et al., 1992; Muñoz & Serrano, 1994a). Up to seven such pairs are present in the titin peptide, depending on pH. Negatively charged groups near the N-terminus (D2 and E3) and the positive K36 at the C-terminus could also strengthen the helix macrodipole (Hol, 1978; Shoemaker et al., 1987). Since leucines have an intrinsically high helical propensity (O'Neail & DeGrado, 1990; Padmanabhan & Baldwin, 1991; Muñoz & Serrano, 1994; Jiménez et al., 1994), their relatively high number present in the peptide may contribute positively to its stabilization even in the monomer where these residues are exposed in solution. Up to two capping boxes involving S4-x-x-E7 and T24-x-x-E27 could promote the starting of both distinct helical regions (Harper & Rose, 1993; Dargupta & Bell, 1993). A break in the central region of the sequence observed at neutral pH according to both the secondary chemical shifts and the NOE pattern can also be explained by the low propensity to helical formation of residues around this region (H18, W19). On the other hand, although electrostatic contributions stabilize the helicity of the monomer, they must at the same time prevent self-association, which is thus possible only at acidic pH and in the presence of salt. Under these conditions, the carboxyl groups are protonated and the remaining charges on the peptide surface are sufficiently shielded by the salt to allow hydrophobic interactions to become predominant leading to self-association.

We may now speculate about the role of this fragment in the whole titin molecule. The physiological pH in muscle is around 7.0 except in conditions of acute acidosis when the pH drops to 6.1 (Kemp et al., 1994; Seow & Ford, 1993; Lamb et al., 1992; Thompson et al., 1992). The isoelectric point of most of the proteins in muscle is around pH 4.5 (Rome, 1968; Collins & Edwards, 1971). It is therefore very unlikely that a pH below 6 is found in muscle even only locally. We may thus exclude the possibility that this fragment is the self-associating element of the titin filament. On the other hand, these 38 amino acids, together with the preceding ≈100 residues, could be part of a globular domain in which this helical motif could be a folding nucleation site. However, the whole linker between type II modules contains several relatively long stretches of charged residues interrupted only occasionally by more hydrophobic patches (Figure 1a). It would be difficult to accommodate such a charged surface into a globular core. On the basis of this consideration, we suggest that *in vivo* the sequence studied here forms a filamentous helical stretch possibly recognized and stabilized by other proteins present in the surroundings. Work to distinguish conclusively between these two possibilities and to identify possible partners is currently in progress.

## ACKNOWLEDGMENT

This work would have not been possible without the early communication of the titin sequence by S. Labeit and the

suggestion of the presence of a coiled coil by M. Gautel. We would also like to thank D. Schubert for his continual support and fruitful discussions, A. Lupas for critical evaluation of the coiled coil prediction, T. Gibson for helpful comments, L. Serrano for calculating the helix propensity of our sequence, and those who have given their suggestions during the manuscript preparation.

## REFERENCES

Chen, Y. H., Yang, J. T., & Chau, K. H. (1974) *Biochemistry 13*, 3350−3359.
Cohen, C., & Parry, D. A. D. (1986) *Trends Biochem. Sci. 11*, 245−248.
Cohen, C., & Parry, D. A. D. (1990) *Proteins: Struct., Funct., Genet. 7*, 1−15.
Collins, E. W., & Edwards, C. (1971) *Am. J. Physiol. 221*, 1130−1133.
Dargupta, S., & Bell, J. (1993) *Int. J. Pept. Res. 41*, 499−551.
Durchshlag, H. (1986) *Thermodynamic Data for Biochemistry and Biotechnology* (Hinz, H. J., Ed.) pp 45−128, Springer-Verlag, Berlin.
Edelhoch, H. (1967) *Biochemistry 6*, 1948−1954.
Fuerst, D. O., Osborn, M., Nave, R., & Weber, K. (1988) *J. Cell Biol. 106*, 1563−1572.
Fujita, H. (1975) *Foundations of Ultracentrifugation Analysis*, pp 283−294 and 426−427, Wiley & Sons, New York.
Fulton, A. B., & Isaacs, W. B. (1991) *Bioessays 13*, 157−161.
Gans, P. J., Lyu, P. C., Manning, M. C., Woody, R. W., & Kallenbach, N. R. (1991) *Biopolymers 31*, 1605−1614.
Gautel, M., Leonard, K., & Labeit, S. (1993) *EMBO J. 10*, 3827−3834.
Greenfield, N. J., & Fasman, G. D. (1969) *Biochemistry 8*, 4108−4116.
Greenfield, N. J., & Hitchcock-DeGregori, S. E. (1993) *Protein Sci. 2*, 1263−1273.
Griesinger, C., Otting, G., Wüthrich, K., & Ernst, R. R. (1988) *J. Am. Chem. Soc. 110*, 7870−7872.
Harper, E. T., & Rose, G. D. (1993) *Biochemistry 32*, 7605−7609.
Hol, W. G. J., Van Duijnen, P. Th., & Berendsen, H. J. C. (1979) *Nature 273*, 443−446.
Huyghues-Despointes, B. M. P., Scholtz, J. M., & Baldwin, R. B. (1993) *Protein Sci. 2*, 80−85.
Isaacs, W. B., Kim, I. S., & Fulton, A. B. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 7496−7500.
Jiménez, M. A., Muñoz, V., Rico, M., & Serrano, L. (1994) *J. Mol. Biol. 242*, 487−496.
Kadkhodaei, M., Hwang, T., & Shaka, A. J. (1993) *J. Magn. Reson. 105*, 104−107.
Kemp, G. J., Thompson, C. H., Sanderson, A. L., & Radda, G. K. (1994) *Magn. Reson. Chem. 31*, 103−109.
Kurzban, G. P., & Wang, K. (1988) *Biochem. Biophys. Res. Commun. 155*, 1155−1161.
Labeit, S., Barlow, D. P., Gautel, M., Gibson, T., Holt, J., Hsieh, C. L., Francke, U., Leonard, K., Wardale, J., Whiting, A., & Trinick, J. (1990) *Nature 345*, 273−276.
Labeit, S., Gautel, M., Lackey, A., & Trinick, J. (1992) *EMBO J. 11*, 1711−1716.
Labeit, S., Kolmerer, B., & Gautel, M. (1994) (manuscript in preparation).
Lamb, G. D., Recupero, E., & Stephenson, D. G. (1992) *J. Physiol. 448*, 211−224.
Landschulz, W. H., Johnson, P. F., & McKnight, S. L. (1988) *Science 240*, 1759−1764.
Lupas, A., Van Dyke, M., & Stock, J. (1991) *Science 252*, 1162−1164.
Lyu, P. C., Gans, J. P., & Kallenbach, N. R. (1992) *J. Mol. Biol. 223*, 343−350.
Lyu, P. C., Wemmer, D. W., Zhou, H. X., Pinker, R. J., & Kallenbach, N. R. (1993) *Biochemistry 32*, 421−425.
Marion, D., & Wüthrich, K. (1983) *Biochem. Biophys. Res. Commun. 113*, 967−974.
Marqusee, S., & Baldwin, R. L. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 8898−8902.
Maruyama, K. (1994) *Biophys. Chem. 50*, 73−85.
Maruyama, K., Kimura, S., Yoshidomi, H., Sawada, H., & Kikuchi, K. (1984) *J. Biochem. (Tokyo) 89*, 701−709.
Merutka, G., Lipton, W., Shalongo, W., Park, S., & Stellwagen, E. (1990) *Biochemistry 29*, 7511−7515.
Muñoz, V., & Serrano, L. (1994a) *Nature Struct. 1*, 399−409.
Nave, R., Fuerst, D. O., & Weber, K. (1989) *J. Cell Biol. 109*, 2177−2187.
Neuhaus, D., & Williamson, M. (1989) *The Nuclear Overhauser Effect in Structural and Conformational Analysis*, VCH Publishers, Inc., New York.
O'Shea, E. K., Klemm, J. D., Kim, P. S., & Alber, T. (1991) *Science 254*, 539−544.
O'Neil, K. T., & DeGrado, W. F. (1990) *Science 250*, 646−651.
Padmanabhan, S., & Baldwin, R. L. (1991) *J. Mol. Biol. 219*, 135−137.
Pastore, A., & Saudek, V. (1990) *J. Magn. Reson. 89*, 165−177.
Phillips, G. N., Jr., Fillers, J. P., & Cohen, C. (1986) *J. Mol. Biol. 192*, 111−131.
Piotto, M., Saudek, V., & Sklenar, V. (1992) *J. Biomol. NMR 2*, 661−664.
Politou, A. S., Gautel, M., Pfuhl, M., Labeit, S., & Pastore, A. (1994) *Biochemistry 33*, 4730−4737.
Rice, C. M., Fuchs, R., Higgins, D. G., Stoehr, P. J., & Cameron, G. N. (1993) *Nucleic Acids Res. 21*, 2967−2971.
Roark, D. E., & Yphantis, D. A. (1971) *Biochemistry 10*, 3241−3249.
Rome, E. (1968) *J. Mol. Biol. 37*, 331−344.
Saudek, V., Pastore, A., Castiglione Morelli, M. A., Frank, R., Gausepohl, H., & Gibson, T. (1991) *Protein Eng. 4*, 512−529.
Scholtz, J. M., & Baldwin, R. L. (1992) *Annu. Rev. Biophys. Biomol. Struct. 21*, 95−118.
Schuck, P. (1994) *Prog. Colloid Polym. Sci. 94*, 1−13.
Seow, C. Y., & Ford, L. E. (1993) *J. Gen. Physiol. 101*, 487−511.
Shoemaker, K. R., Kim, P. S., York, E. J., Stewart, Y. M., & Baldwin, R. L. (1987) *Nature 326*, 563−567.
Sundaralingam, M., Drendel, W., & Greaser, M. (1985) *Proc. Natl. Acad. Sci. U.S.A. 82*, 7944−7947.
Thompson, L. V., Balog, E. M., & Fitts, R. H. (1992) *Am. J. Physiol. 262*, C1507−15012.
Trinick, J., Knight, P., & Whiting, A. J. (1984) *J. Mol. Biol. 180*, 331−356.
Wang, C. L. A., Chalovich, J. M., Graceffa, P., Lu, R. C., Mabuchi, K., & Stafford, W. F. (1991) *J. Biol. Chem. 266*, 13958−13963.
Wang, K. (1985) *Cell Muscle Motil. 6*, 315−369.
Wang, K., McCarter, R., Wright, J., Beverly, J., & Ramirez-Mitchell, R. (1993) *Biophys. J. 64*, 1161.
Whiting, A. J., Wardale, J., & Trinick, J. (1989) *J. Mol. Biol. 205*, 163−169.
Wishart, D. S., Sykes, B. D., & Richards, F. M. (1991) *J. Mol. Biol. 222*, 311−333.
Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, John Wiley and Sons, New York.
Zhou, H. X., Lyu, P., Wemmer, D. E., & Kallenbach, N. R. (1994) *Proteins 18*, 1−7.
Zhou, N. E., Kay, C. M., & Hodges, R. S. (1992) *J. Biol. Chem. 267*, 2664−2670.

BI941296M