Journal of Biomolecular NMR, 9 (1997) 2–10
ESCOM

J-Bio NMR 400

^{1}H and ^{15}N NMR resonance assignments and secondary structure of titin type I domains

Claudia Muhle-Goll, Michael Nilges and Annalisa Pastore*

European Molecular Biology Laboratory, Meyerhofstrasse 1, D-69117 Heidelberg, Germany

Received 12 August 1996
Accepted 8 October 1996

Keywords: Connectin; Modular proteins; Muscle; Fibronectin

Summary

Titin/connectin is a giant muscle protein with a highly modular architecture consisting of multiple repeats of two sequence motifs, named type I and type II. Type I modules have been suggested to be intracellular members of the fibronectin type III (Fn3) domain family. Along the titin sequence they are exclusively present in the region of the molecule located in the sarcomere A-band. This region has been shown to interact with myosin and C-protein. One of the most noticeable features of type I modules is that they are particularly rich in semiconserved prolines, since these residues account for about 8% of their sequence. We have determined the secondary structure of a representative type I domain (A71) by ^{15}N and ^{1}H NMR. We show that the type I domains of titin have the Fn3 fold as proposed, consisting of a three- and a four-stranded β-sheet. When the two sheets are placed on top of each other to form the β-sandwich characteristic of the Fn3 fold, 8 out of 10 prolines are found on the same side of the molecule and form an exposed hydrophobic patch. This suggests that the semiconserved prolines might be relevant for the function of type I modules, providing a surface for binding to other A-band proteins. The secondary structure of A71 was structurally aligned to other extracellular Fn3 modules of known 3D structure. The alignment shows that titin type I modules have closest similarity to the first Fn3 domain of *Drosophila* neuroglian.

Introduction

The fibronectin type III (Fn3) module is one of the most common motifs found in modular proteins (Campbell and Spitzfaden, 1994). Its main function seems to be to mediate protein–protein interactions and to act as a spacer to get the required biological function into the right place. The Fn3 motif was first characterized in fibronectin and, since then, has been identified in more than 50 unrelated proteins. Although these include mostly extracellular proteins, the Fn3 module has also been identified in an intracellular muscle protein family, which includes twitchin and titin (Benian et al., 1989; Labeit et al., 1990). Several high-resolution structures of extracellular protein domains with an Fn3-like fold are known from X-ray crystallography or NMR spectroscopy (De Vos et al., 1992; Leahy et al., 1992,1996; Main et al., 1992; Dickinson et al., 1994; Huber et al., 1994), but the structure of an intracellular Fn3-like domain is still missing.

Titin (also known as connectin) is the largest single-chain polypeptide identified to date (≈3000 kDa). The complete sequence of human cardiac titin has been recently published (Labeit and Kolmerer, 1995). It consists mainly of an array of two module types, identified either as the already mentioned Fn3 or as belonging to the immunoglobulin I (Ig) family (Labeit et al., 1990; Pfuhl and Pastore, 1995). The occurrence of type I (Fn3-like) or type II (Ig-like) modules varies with respect to the titin location in the sarcomere. Both the N- and the C-termini of the molecule (respectively located in the I-band and the

*To whom correspondence should be addressed.

Abbreviations: 2D, 3D, two-, three-dimensional; DSS, 3-(trimethylsilyl)-1-propanesulfonic acid sodium salt; FN, fibronectin; Fn3, fibronectin type III module; FNFn3(10), 10th type III module of fibronectin; HGHR2, human growth hormone receptor; HSQC, heteronuclear single quantum coherence spectroscopy; Ig, immunoglobulin; NGFn1, NGFn2, first and second Fn3 domain from *Drosophila* neuroglian; NOE, nuclear Overhauser enhancement; NOESY, 2D NOE spectroscopy; TNFn3(3), third Fn3 domain of tenascin; TOCSY, 2D total correlation spectroscopy.

M-line regions of the sarcomere) consist mostly of type II modules. More than 60% of the modules in the A-band (roughly corresponding to the titin central region) are arranged in super-repeats of the pattern II-I-I-II-I-I-I-II-I-I-I (Labeit and Kolmerer, 1995). Towards the A/I junction, the super-repeats break into a reduced version which contains six II-I-I-II-I-I-I motifs. Near the M-line, a unique II-II-I-I-II-II-I motif is present. These three sequence motifs and the distance between them correlate well with the ultrastructure organization of vertebrate thick filaments, which appears at the electron microscopy as divided into three distinct zones known as D-, C- and P-zones (Craig, 1977; Sjöström and Squire, 1977). It was therefore suggested that the region of titin which is located in the A-band has a key role in determining the whole muscle ultrastructure (Labeit and Kolmerer, 1995). In vitro binding studies indicate that myosin and the C-protein interact with titin in the A-band region (Labeit et al., 1992; Houmeida et al., 1995).

Titin type I domains contain no Arg-Gly-Asp (RGD) motif, a sequence known to form a binding site for integrins in FNFn3(10) and related domains (Campbell and Spitzfaden, 1994). Instead, one of the most noticeable features of titin type I modules is that they contain an unusual number of prolines in their sequences, that is between 8 and 10 per domain. The prolines are mostly found at the N-terminus of the domains, where they form a conserved stretch of Pro-Gly-Pro-Pro or Pro-Ile-Asp-Pro-Pro, and towards the C-terminus. The precise positions of the prolines are semiconserved within the type I module alignment.

The structures of several domains of titin belonging either to type II (Ig-like) or to unique stretches in the sequence have been recently solved (Politou et al., 1994a,b; Musco et al., 1995; Pfuhl and Pastore, 1995). A step further in the direction towards a molecular understanding of titin is to determine the structure of type I modules. A knowledge of this structure might be used as a guideline to rationalize the interactions between titin and other A-band proteins. In the present paper we give a full account of the secondary structure determination of a representative type I module and compare the results with the alignment of Fn3 sequences both from titin and from extracellular proteins.

Materials and Methods

Expression and purification

A representative set of three consecutive type I domains out of the second super-repeat (C-zone in the A-band) of titin (A69–A71) (Labeit and Kolmerer, 1995) was chosen and their cDNAs were PCR-amplified. The PCR-amplified cDNAs encoding for A69, A70 and A71 were subcloned into a modified pET8c vector (Studier et al., 1990) and fused N-terminally with an oligonucleotide linker encoding a His_6-tag sequence, where two additional serines were introduced as linkers. Transformed BL21 [DE3] pLysS cells (Studier, 1991) were induced with 0.2 M IPTG and harvested after 4 h. Cells were lysed by sonication and the proteins were purified by metal chelate affinity chromatography on Ni^{2+}-NTA agarose (Quiagen), essentially as previously described (Politou et al., 1994a). Unbound fractions contained the pure type I domains as judged by SDS-PAGE and mass spectrometry. Out of the three domains, A71 showed the best expression behavior and was used for further NMR experiments. A71 eluted as a monomeric fraction on a 50 ml DEAE-Sephacel column in a buffer containing 20 mM KH_2PO_4, 50 mM NaCl, 1 mM EDTA and 1 mM DTT at pH 7.0. Uniformly ^{15}N-labeled samples were obtained from bacteria grown in M9 medium with $^{15}NH_4Cl$ as the sole nitrogen source.

NMR studies

NMR spectra were recorded on a Bruker AMX-600 using 1–1.5 mM samples in 20 mM deuterated acetate buffer containing 50 mM NaCl or 20 mM phosphate buffer at pH values between 4.0 and 7.0 and at temperatures between 290 and 325 K. Optimal conditions for NMR were found at pH 5.6, 50 mM NaCl, 20 mM acetate buffer and 315 K. To record spectra in D_2O, 0.5 ml of a 1.4 mM sample in deuterated acetate buffer was lyophilized and redissolved in the same amount of D_2O. ^{15}N-edited spectra were recorded on a 1.8 mM ^{15}N-labeled sample at pH 5.6 and 310 K. Water suppression was achieved using WATERGATE (Piotto et al., 1992) in both homo- and heteronuclear spectra (Sklenář et al., 1993) for measurements in H_2O and using low-power presaturation in the D_2O experiments. No significant changes occurred if no or low-power presaturation was used. The proton spectra are referenced to internal DSS. ^{15}N resonances are calculated from the ^1H frequency according to Wishart et al. (1995). 2D and 3D TOCSY spectra were acquired using the TOWNY composite pulse cycle (Kadkhodaei et al., 1993). The mixing times used were in the range of 20–80 ms for the TOCSY and 50–120 ms for the NOESY experiments. The spectral widths used in the 3D TOCSY- and NOESY-HSQC were 8196.72 Hz (13.658 ppm) in both the direct and indirect proton dimension, and 2000.00 Hz (32.885 ppm) for the nitrogen dimension. A matrix of $296 \times 128 \times 1K$ complex data points was acquired that was zero-filled and Fourier transformed to a matrix of $256 \times 512 \times 512$ points.

Homology model building

Homology model building was achieved by the WHAT-IF program. It was based on a multiple alignment of the titin A71 sequence with others for which the structures are known. The sequences other than the titin one were aligned according to their secondary structure elements as derived by the DSSP program (Kabsch and Sander, 1983).

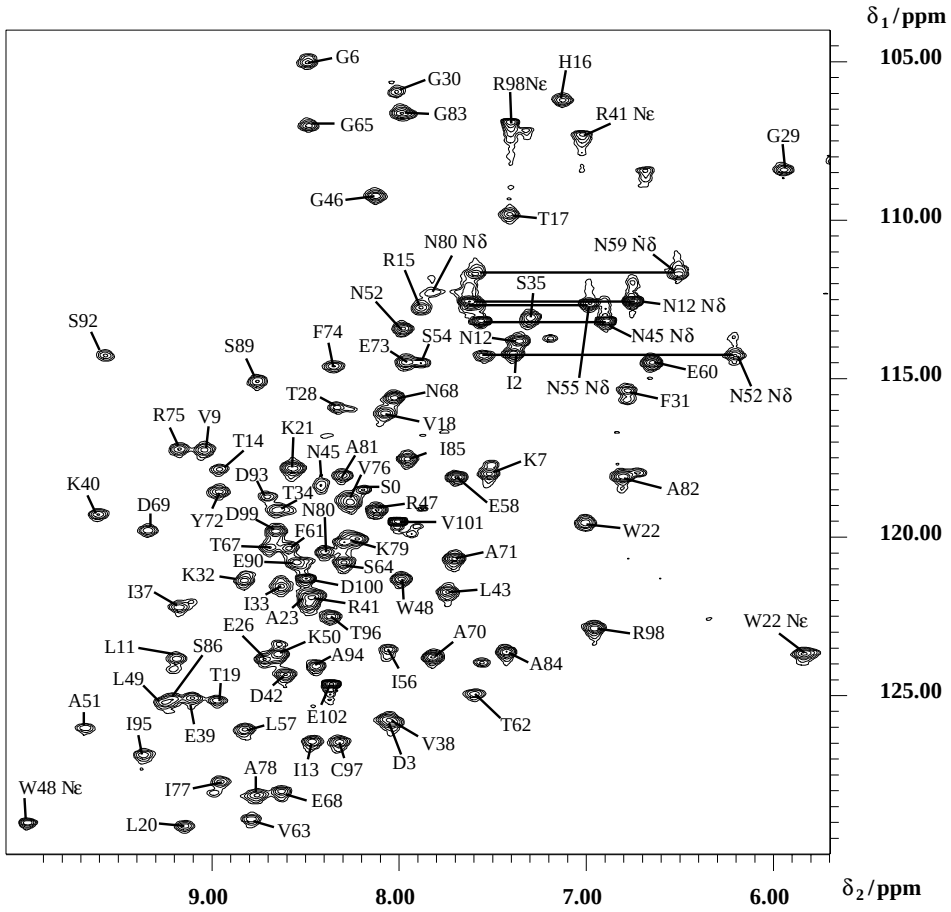

Fig. 1. ^{15}N HSQC spectrum of A71 in 20 mM d_3-sodium acetate buffer, 50 mM NaCl, pH 5.6 at 310 K. Cross peaks are labeled according to residue sequence number for main-chain nitrogens and additionally for atom type for side-chain nitrogens. One should note the different positions of the $H^{\varepsilon 1}$ resonances of Trp^{48} and Trp^{22}.

The titin sequence was aligned to the others according to the experimental secondary structure definition. The structure of the first Fn3 domain of the *Drosophila* neuroglian was used as a template (Huber et al., 1994). No additional refinement was attempted on the model.

Results

Choice of domain

The sequence of cardiac titin shows 132 type I repeats, with sequence identity ranging between 65% (for domains in corresponding positions in the super-repeat) and 25% (Labeit and Kolmerer, 1995). Among them, the sequence of a representative type I domain from the super-repeat was selected according to the following criteria: (i) The position in the super-repeat had to be representative of the entire A-band, that is belonging to a motif present both in the D- and C-zones. (ii) By selecting more than one sequence we would have the possibility of choosing the best behaved construct. Previous experience with type II domains had shown that interdomain interactions may play an important role in stabilizing the structure of isolated modules (Politou et al., 1994b). A subset of three contiguous domains would therefore contain all the possible domain–domain interactions both between type I and type II modules, thus giving us the possibility of choosing the domain that is intrinsically the most stable. (iii) It is always wise to select sequences with as few cysteine residues as possible in order to avoid self-aggregation from disulfide bridge formation. The triplet of tandem modules from A69 to A71 was finally chosen. The domain boundaries were chosen according to the consensus obtained from a multiple alignment of the titin type I sequences. Out of the three domains, A71 showed the best expression behavior and was used for structure determination.

NMR assignment

Sequential assignment was achieved in the conventional manner following the standard protocol (Wüthrich, 1986). 2D TOCSY and DQF-COSY spectra at temperatures between 310 and 320 K were used to identify the spin systems. At all temperatures, the spectral dispersion is still comparable to that observed at room temperature or be-

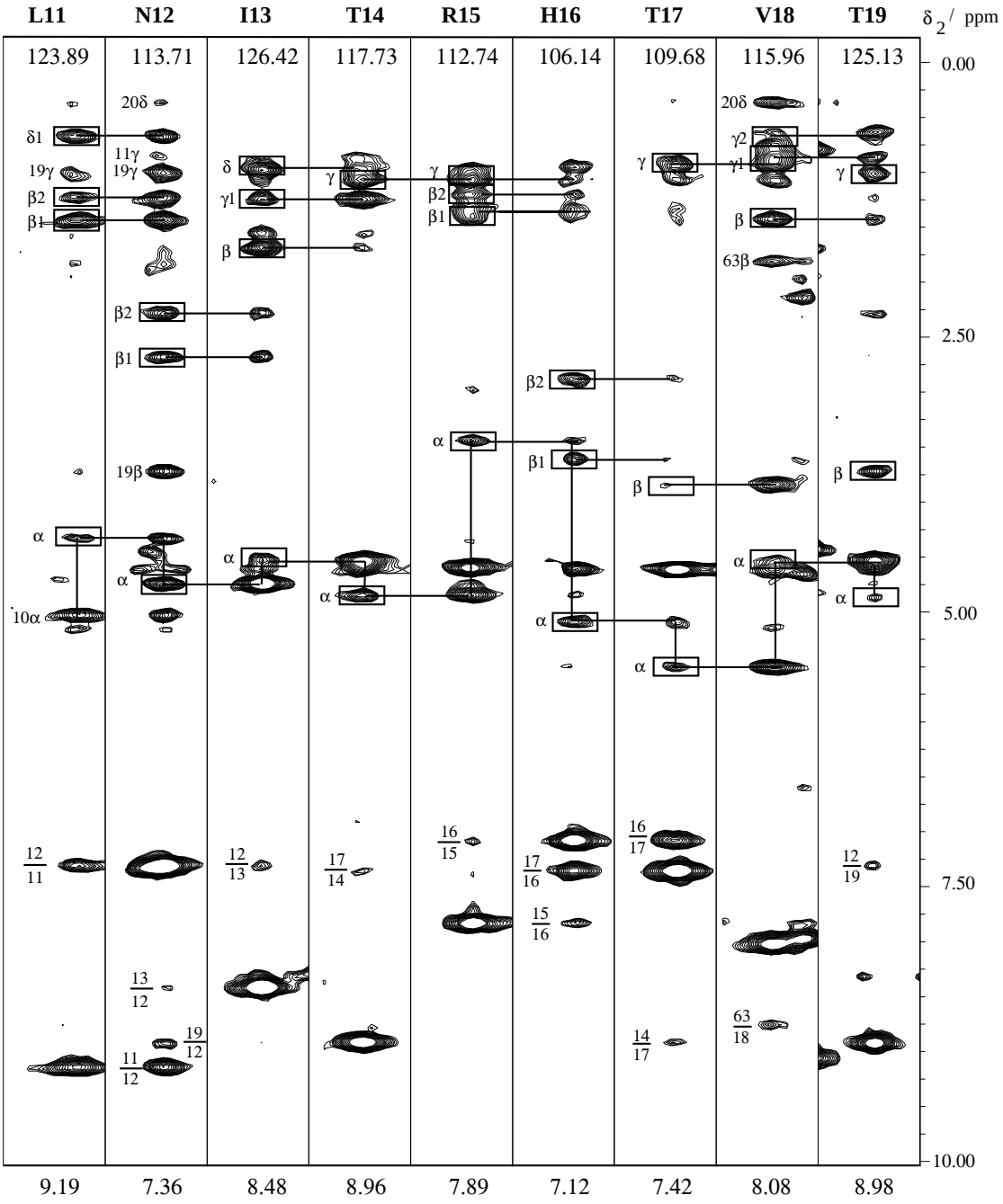

Fig. 2. Strips along F1, taken from a 3D ^{15}N NOESY-HSQC experiment, measured at 310 K. The F2 (^{15}N) coordinate of each strip is given at the top; the F3 (^{1}H) coordinate is given at the bottom. Intraresidual NOEs are marked with boxes and the corresponding labels; sequential NOEs are indicated by lines. Amide–amide NOEs are labeled explicitly. The strips correspond to the antiparallel β-sheet formed by strands A and B and connected by a β-turn (Thr^{14}–Thr^{17}).

low, ensuring that the domain is still folded. However, the quality of TOCSY transfer is undoubtedly better at higher temperature. The line widths in the NOESY spectra improved noticeably. A 3D ^{15}N TOCSY-HSQC helped to resolve spectral overlap for the α/amide proton cross peaks. The 2D ^{15}N HSQC spectrum of A71 is shown in Fig. 1. The assignment of all residues was achieved through sequential H^{α}/HN and side chain/HN NOEs in a 3D ^{15}N NOESY-HSQC spectrum except for the prolines and Phe[53], for which the amide chemical shift could not be determined. Residues of the His_6-tag were only identified in the TOCSY spectra from their characteristic side-chain cross peaks, but remained invisible in the NOESY spectra. Such a behavior was also observed for titin type II domains (Pfuhl and Pastore, 1995; Improta et al., 1996) and suggests that the His-tag residues are flexible and do not interfere with the rest of the structure. Figure 2 shows selected strips from a ^{15}N NOESY-HSQC arranged se-

quentially in order to exemplify part of the sequential assignment. Phe[53] lies at the end of a β-strand, most probably in a flexible bulge or pseudo-bulge (Richardson et al., 1978), and the following residues Ser[54] and Asn[55] either show no long-range NOEs (Ser[54]) or exhibit weak TOCSY and NOESY transfer (Asn[55]). The aromatic ring systems (four phenylalanines, three tyrosines and two tryptophans) were identified, except for Trp[22], in a 2D TOCSY acquired at 315 K and connected to their amide protons through the β-protons in the corresponding 2D NOESY. The chemical shift of the δ-proton of the aromatic ring of Trp[22] resonates at 5.86 ppm (at 315 K). Because of this unusual position, this residue was identified only in the later stage of the assignment from the ^{15}N-edited NOESY-HSQC. Particularly strong TOCSY transfer from the amide to the side-chain protons is observed for residues Asp[99]–Ala[103]. These residues also show strong intraresidual as well as sequential NOEs. However, they do not exhibit long-range connectivities to the rest of the protein as judged from the fingerprint region alone.

Assignment of the prolines

The α-protons of prolines 5, 9, 10, 24, 88 and 91 were identified by their sequential cross peaks to the following residues and connected to their β-, γ- and δ-protons in the 2D TOCSY. Pro[44] was identified by its δ-protons, which give rise to a cross peak with Asn[45] in the NOESY-HSQC. The remaining three prolines show no NOEs with residues in the fingerprint region and were identified in the later stage of the assignment. Apart from Pro[44], all prolines are in the trans conformation as demonstrated by NOEs of the δ-protons with the α-resonance of the previ-

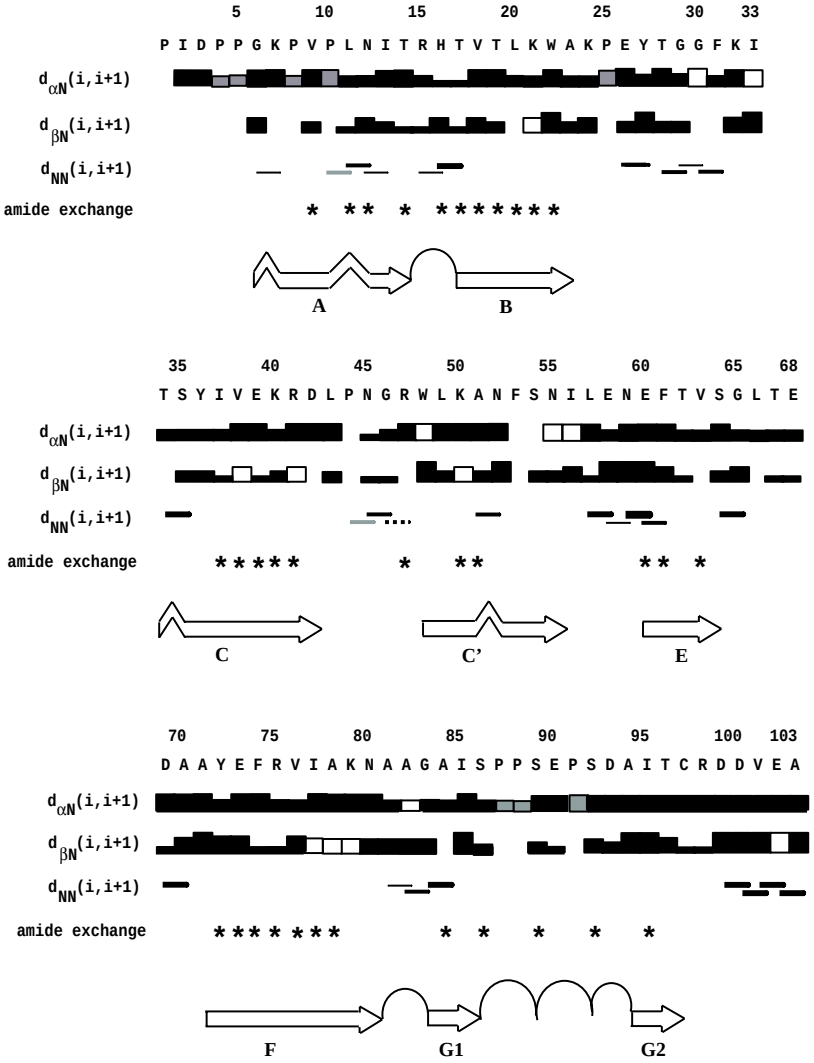

Fig. 3. Summary of sequential NOEs and slowly exchanging amides. The bar heights correspond to the relative strengths of the NOEs. Open boxes or dashed lines indicate ambiguous NOEs due to spectral overlap. Shaded boxes represent NOEs involving proline $C^{\delta}H$. Those amides that are still visible after 4 weeks in D_2O are marked with asterisks. The secondary structure of A71 is indicated underneath. Arrows represent regions of β-sheet structure and are marked according to the convention for Fn3 modules (Campbell and Spitzfaden, 1994). Turns are indicated with a ∩ and bulges with a ∧.

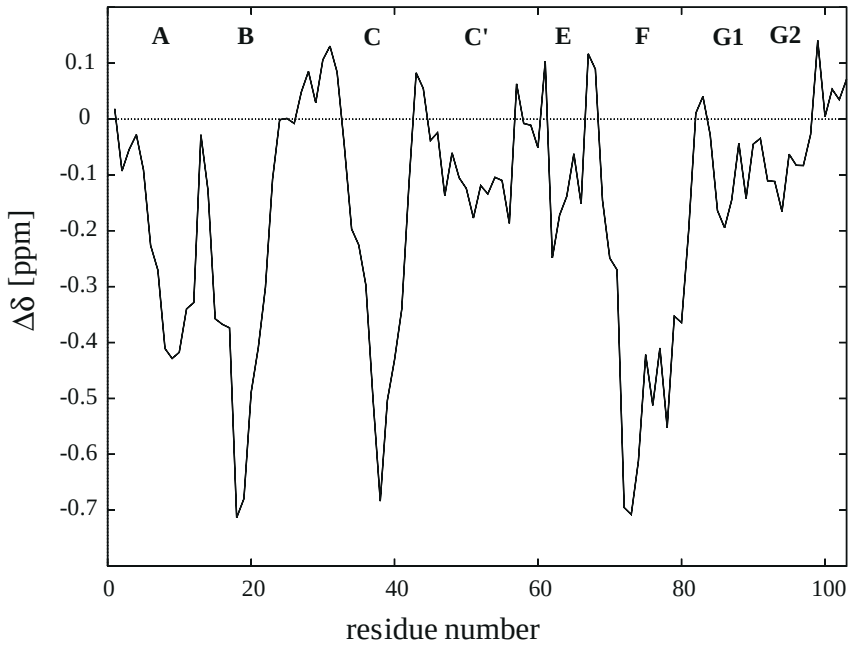

Fig. 4. Deviation of the chemical shifts of the $C^{\alpha}H$ resonances from random coil values (after smoothing using a window of ±2 residues (Pastore and Saudek, 1990; Wishart et al., 1991)). Low-field shifts hint at the existence of β-sheet structure. The regions of the seven β-sheets are marked with the corresponding letters (Campbell and Spitzfaden, 1994).

ous residue. The presence of an $H^{\alpha}(i-1)$-$H^{\alpha}(i)$ NOE of Pro^{44} with the previous residue Leu^{43} together with the absence of $H^{\alpha}(i-1)$-$H^{\delta}(i)$ NOEs are evidences that Pro^{44} is present in the cis conformation.

Serine γ-hydrogens

Two out of the eight serines, Ser^{86} and Ser^{92}, show partially protected γ-hydrogens that are visible in the 2D spectra measured in H_2O but are absent from those in D_2O.

Slowly exchanging NH protons

These were identified from D_2O exchange experiments and are shown in Figs. 3 and 5. The amide protons that are visible after 12 h at 310 K do not exchange much further in the next 48 h and remain stable if kept in D_2O for 4 weeks at 4 °C. The temperature coefficients of the HN chemical shifts calculated in the 300–320 K range were all linear.

The assignment is reported in the Supplementary Material (available on request from the authors).

Identification of secondary structure

The pattern of sequential and interstrand NOEs, slowly exchanging amide protons, and the downfield H^{α} chemical shift deviation from random coil values shown in Figs. 3 and 4 are consistent with a seven-stranded sandwich of two β-sheets (conventionally indicated with letters from A to G). Strands A (Gly^{6}–Thr^{14}) and B (Thr^{17}–Ala^{23}) and strands G (Ala^{84}–Ser^{86} and Ala^{94}–Cys^{97}) and F (Ala^{71}–Ala^{81}) are connected by a type I β-turn and a type II β-turn, respectively. Several β-bulges occur in the regular sheet structure as summarized in Fig. 5. Strand A contains a G1 type bulge formed by Gly^{6}, Lys^{7} and Ala^{23} and a wide bulge (Richardson et al., 1978; Chan et al., 1993) formed by Pro^{10} and Leu^{11} and Leu^{20} on the opposite strand. A classic β-bulge (Thr^{34}, Ser^{35} and Lys^{79}) is found at the beginning of strand C. One further β-bulge is located at the edge of strand D (Gly^{65}, Leu^{66} and His^{16}). Two out of the seven strands are rather short. Strand C' consists of only three residues (Leu^{49}–Ala^{51}) that show β-sheet connectivities followed by an adjacent narrow β-bulge (Ala^{51}, Asn^{52} and Val^{38}). Strand G, divided into G1 and G2, is interrupted by a sequence of seven residues (Pro^{87}–Asp^{93}) that are not in β-sheet conformation and show only two connectivities to the opposite strand F at residues Ser^{89} and Ser^{92}. The intervening sequence is rich in prolines (Pro^{87}, Pro^{88} and Pro^{91}) and is sandwiched between the two serines whose γ-hydrogens are partially protected. This arrangement is reminiscent of the polyproline helix found in the crystal structure of the first Fn3 domain of the neuroglian (Huber et al., 1994). The chemical shift deviation of the α-proton resonances of the C'- and G-strands, which hardly deviate from the random coil values, possibly reflects distortions of the β-sheet conformation in these regions.

Discussion

The data described in the present work provide conclusive evidence that A71, a representative of titin type I modules, has a typical Fn3 fold with seven β-strands

arranged in a β-sandwich. A71, the first type I module of titin to be structurally investigated, is also the first intracellular fibronectin type III module for which the secondary structure has been reported.

Our data allow us to obtain a reliable alignment, based on the secondary structure information, of the A71 sequence with that of other Fn3 modules, the structures of which are available (Fig. 6). Such an alignment is not entirely straightforward on the basis of the sequence information only, because of the different lengths of the sequences with a number of expected insertions/deletions and because the β-strands are often distorted into bulges. A structure-based alignment is therefore valuable to model build by homology intracellular Fn3 domains from other structures.

The pattern of the most conserved residues in the Fn3 family remains mostly unchanged in A71 (Fig. 6), with the exception of Lys[40]. This residue replaces a conserved aromatic residue at the end of strand C in other modules. This position is occupied in more than 60% of the titin type I modules by a conserved basic residue (compare with the consensus sequence in Fig. 6 and Labeit and Kolmerer (1995)). A similar behavior is also observed in other muscle proteins, such as twitchin, protein C and myosin light chain kinase, and seems therefore a unique feature of this Fn3 subfamily.

Loosely conserved β-bulges within the A-strand and at the beginning of the C-strand, observed in all of the other Fn3 structures, are also present in A71. The BC loop in A71 is elongated by four residues and is comparable in length to both the Fn3 domains of the human growth hormone receptor and the *Drosophila* neuroglian. In the Fn3 modules of *Drosophila* neuroglian and in the interface between the seventh and the eighth Fn3 domain of fibronectin, the BC loop from the second domain together with the EF loop of the first domain form the majority of the domain interface, yet none of the amino acids involved in interdomain contacts is highly conserved. The G-strand includes two short stretches, denoted as G1 and G2, separated by a sequence of non-β-sheet structure. G2 corresponds to the original G-strand that, in some titin type I domains, is much longer in other Fn3 modules like FNFn3(10). G1 is also present in the other Fn3 modules, except for FNFn3(10), according to the DSSP definition (Kabsch and Sander, 1983), but in all cases it is only two residues long (one residue shorter than in A71) and thus was never explicitly mentioned before.

Overall, a comparison of A71 with the other known

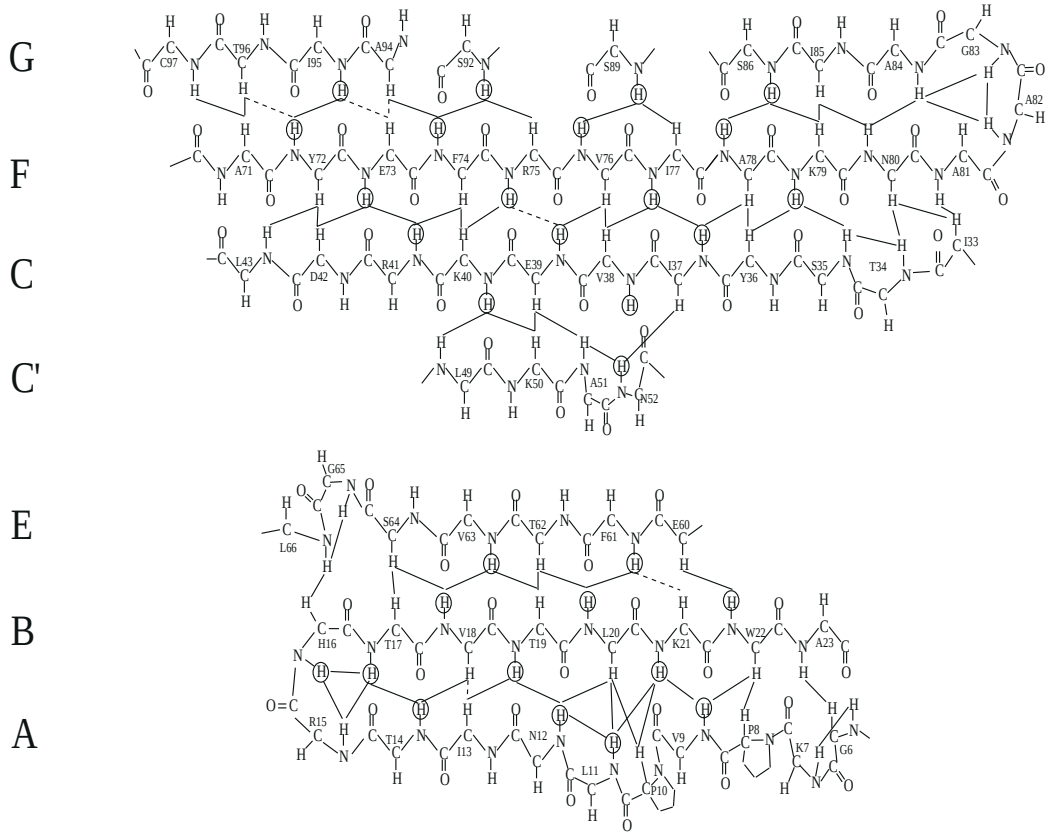

Fig. 5. Secondary structure of A71. NOEs that are indicative of secondary structure are depicted by thin lines. Amide protons that do not exchange in D_2O after 4 weeks are labeled with circles. Dashed lines represent NOEs that could not be observed because of spectral overlap. Strands A and B and strands F and G are connected via a type I β-turn and a type II β-turn, respectively. β-Bulges can be recognized for the residues 6–7, 10–12, 34–35, 51–52 and 65–66. Strand G is interrupted by three bulges that are reminiscent of a polyproline helix.

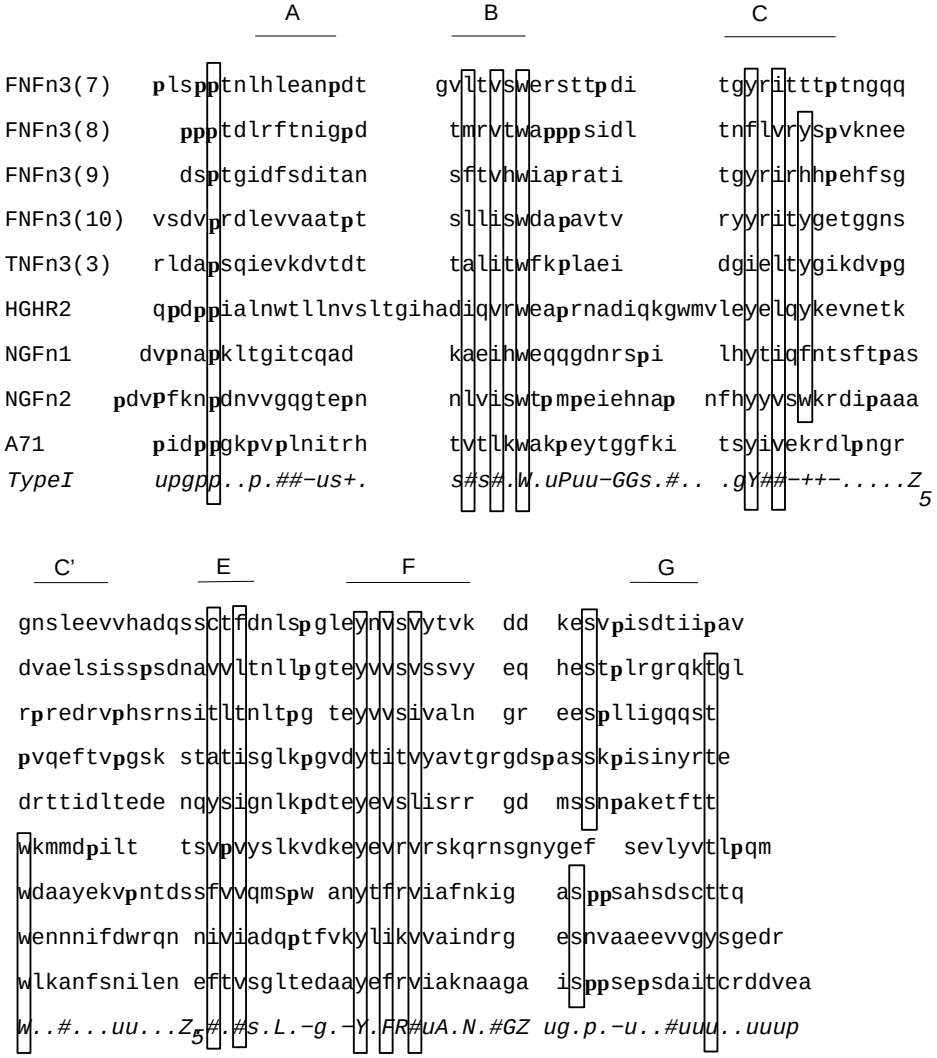

Fig. 6. Sequence alignment of A71 with Fn3 modules, the structures of which have been solved by X-ray crystallography or NMR. A consensus independently derived for all type I domains of titin is also shown in the bottom line (*Type I*). The abbreviations used for the sequences correspond to the following Brookhaven entries: FNFn3(7)–(9): 1fnf.brk (Leahy et al., 1996); FNFn3(10): 1fna.brk (Dickinson et al., 1994) and 1fnf.brk (Leahy et al., 1996) for the X-ray and 1ttg.brk for the NMR structure, respectively (Main et al., 1992); TNFn3(3): 1ten.brk (Leahy et al., 1992); HGHR2: 3hhr.brk (De Vos et al., 1992); NGFn1 and NGFn2: 1cfb.brk (Huber et al., 1994). The domains were aligned according to their secondary structure elements as derived by the DSSP program (Kabsch and Sander, 1983) and optimized by visual inspection of the structures. In two cases (FNFn3(7) and FNFn3(9)) the structural alignment was doubtful and the sequence alignment was used to align strand C to the other sequences. Residues that are highly conserved are enclosed in boxes. The secondary structure of FNFn3(10) is taken as reference for the position of the β-strands (marked above the alignment). Symbols used for the type I consensus are: g: >50% Gly; p: >50% Pro; s: >60% Thr + Ser + Asp; u: >45% hydrophobic residues; #: >80% hydrophobic residues; +: >60% Lys + Arg; −: >50% Asp + Glu. Capitals indicate residues with >90% conservation. Z_5 denotes insertions of five residues in some titin type I sequences.

Fn3 structures suggests that the structure of the titin domain most closely resembles the first Fn3 domain of the *Drosophila* neuroglian (Huber et al., 1994). These two sequences share 25% sequence identity as compared to 12–16% with any of the FNFn3 domains, 17% with TNFn3(3), 17% with HGHR(2) and 15% with NGFn2. Both modules contain irregular secondary structure elements, with additional β-bulges at the end of strand C' and a potential polyproline helix that separates G1 and G2 (Huber et al., 1994).

It has been noticed before that sequences containing Fn3 motifs show an unusually large number of proline residues (Benian et al., 1989). While the average number of prolines in globular proteins is 5.1% according to Creighton (1993), in fibronectin Fn3 domains the occurrence increases to 9.3% and in titin type I modules it is 8.5%. In most of the known Fn3 structures, the prolines are more or less evenly distributed along the sequence with an obvious predominance in loop regions. In titin type I modules, semiconserved prolines seem to have a clear preference for specific positions: 8 of the 10 prolines of A71 cluster at the N-terminus of strand A and in the first half of strand G. The alignment of the A-strand with other members of the Fn3 family for which the structures

are known places the proline-rich stretch located at the N-terminus of the A71 domain to precede the beginning of the first strand. However, only Pro^{8} and Pro^{10} are part of strand A on the basis of the secondary structure pattern (Fig. 5). Whether Pro^{1}, Pro^{4} and Pro^{5} contribute to the hydrophobic core of the domain or whether they form a linker to the previous domain remains to be established. Preliminary model building by homology based on the structure of the first neuroglian Fn3 domain shows that these N-terminal prolines could form strong interactions with the other proline-rich stretch on the G-strand, which should pack closely when the sandwich is formed (not shown). This would explain why preliminary attempts to clone type I domains without the N-terminal proline-rich stretch led to inclusion bodies and, when subsequently redissolved, the domains remained unfolded (M. Gautel, personal communication). A similar contribution of the module N-terminus in the stabilization of the overall fold is also observed in titin type II (Politou et al., 1994b; Pfuhl et al., 1996). The presence of conserved patches of polyprolines on spatially close patches of the domain surface would strongly suggest that the exposed hydrophobic surface so formed might be of functional importance for binding to either myosin or other A-band proteins.

Conclusions

Full structure calculations of A71 are currently in progress. From a number of differences in the loop structure and in the consensus sequence, we may expect that, together with other known Fn3 modules from cell adhesion modules, titin type I modules form a distinct subset of the Fn3 domain superfamily.

Acknowledgements

The authors wish to thank S. Labeit and B. Kolmerer for the cDNA, M. Gautel for help in purification, M. Pfuhl for help in 3D processing, G. Musco for help in figure preparation, and C. Chothia for helpful discussions.

References

Benian, G.M., Kiff, J.E., Neckelmann, N., Moerman, D.G. and Waterston, R.H. (1989) *Nature*, **342**, 45–50.

Campbell, I.D. and Spitzfaden, C. (1994) *Structure*, **2**, 333–337.

Chan, A.W., Hutchinson, E.G., Harris, D. and Thornton, J.M. (1993) *Protein Sci.*, **2**, 1574–1590.

Craig, R. (1977) *J. Mol. Biol.*, **109**, 69–81.

Creighton, T.E. (1993) *Proteins: Structures and Molecular Properties*, 2nd ed., Freeman, New York, NY, U.S.A.

De Vos, A.M., Ultsch, M. and Kossiakoff, A.A. (1992) *Science*, **255**, 306–312.

Dickinson, C.D., Veerapandian, B., Dai, X.-P., Hamlin, R.C., Xuong, N., Ruoslahti, E. and Ely, K.R. (1994) *J. Mol. Biol.*, **236**, 1079–1092.

Houmeida, A., Holt, J., Tskhovrebova, L. and Trinick, J. (1995) *J. Cell Biol.*, **131**, 1471–1481.

Huber, A.H., Wang, Y.E., Bieber, A.J. and Bjorkman, P.J. (1994) *Neuron*, **12**, 717–731.

Improta, S., Politou, A.S. and Pastore, A. (1996) *Structure*, **4**, 323–337.

Kabsch, W. and Sander, C. (1983) *Biopolymers*, **22**, 2577–2637.

Kadkhodaei, M., Hwang, T.-L., Tang, J. and Shaka, A.J. (1993) *J. Magn. Reson.*, **A105**, 104–107.

Labeit, S., Barlow, D.P., Gautel, M., Gibson, T., Holt, J., Hsieh, C.-L., Francke, U., Leonard, K., Wardale, J., Whiting, A. and Trinick, J. (1990) *Nature*, **345**, 273–276.

Labeit, S., Gautel, M., Lakey, A. and Trinick, J. (1992) *EMBO J.*, **11**, 1711–1716.

Labeit, S. and Kolmerer, B. (1995) *Science*, **270**, 293–296.

Leahy, D.J., Hendrickson, W., Aukhil, I. and Erickson, H.P. (1992) *Science*, **258**, 987–991.

Leahy, D.J., Aukhil, I. and Erickson, H.P. (1996) *Cell*, **84**, 155–164.

Main, A.L., Harvey, T.S., Baron, M., Boyd, J. and Campbell, I.D. (1992) *Cell*, **71**, 671–678.

Musco, G., Tziatzios, C., Schuck, P. and Pastore, A. (1995) *Biochemistry*, **34**, 553–561.

Pastore, A. and Saudek, V. (1990) *J. Magn. Reson.*, **90**, 165–176.

Pfuhl, M. and Pastore, A. (1995) *Structure*, **3**, 391–401.

Pfuhl, M., Politou, A.S., Improta, S. and Pastore, A. (1997) *J. Mol. Biol.*, in press.

Piotto, M., Saudek, V. and Sklenář, V. (1992) *J. Biomol. NMR*, **2**, 661–665.

Politou, A.S., Gautel, M., Labeit, S. and Pastore, A. (1994a) *Biochemistry*, **33**, 4730–4737.

Politou, A.S., Gautel, M., Joseph, C. and Pastore, A. (1994b) *FEBS Lett.*, **352**, 27–31.

Richardson, J.S., Getzoff, E.D. and Richardson, D.C. (1978) *Proc. Natl. Acad. Sci. USA*, **75**, 2574–2578.

Sjöström, M. and Squire, J. (1977) *J. Mol. Biol.*, **109**, 49–68.

Sklenář, V., Piotto, M., Leppik, R. and Saudek, V. (1993) *J. Magn. Reson.*, **A102**, 241–245.

Studier, F.W., Rosenberg, A.H. and Dubendorf, J.W. (1990) *Methods Enzymol.*, **185**, 60–89.

Studier, F.W. (1991) *J. Mol. Biol.*, **219**, 37–44.

Wishart, D.S., Sykes, B.D. and Richards, F.M. (1991) *J. Mol. Biol.*, **222**, 311–333.

Wishart, D.S., Bigam, C.G., Yao, J., Abildgaard, F., Dyson, H.J., Oldfield, E., Markley, J. and Sykes, B.D. (1995) *J. Biomol. NMR*, **6**, 135–140.

Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, Wiley, New York, NY, U.S.A.