FEBS Letters 352 (1994) 27–31

FEBS 14524

Immunoglobulin-type domains of titin are stabilized by amino-terminal extension

A.S. Politou, M. Gautel*, C. Joseph, A. Pastore

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

Received 28 July; revised version received 9 August 1994

Abstract We have recently suggested that similarly folded titin modules located at different sarcomeric regions have distinct molecular properties and stability. Could our selection of module boundaries have potentially influenced our conclusions? To address this question we expressed amino-terminally extended versions of the same modules and determined, with the use of CD and Fluorescence techniques, key thermodynamic parameters characterizing their stability. We present here our results which confirm our previous observations and show that, while amino-terminal extension has a profound effect on the stability of individual modules, it does not affect at all their folding pattern or their relative stabilities. Moreover, our data suggest that the selection of module boundaries can be of critical importance for the structural analysis of modular proteins in general, especially when a well-defined intron–exon topography is absent and proteolytic methods are inconclusive.

Key words: Modular protein; Muscle protein; Boundary; Protein stability; Connectin

1. Introduction

The existence of proteins composed of various, repeated structural and functional units is widely recognized nowadays [1–3]. With the number of known protein sequences increasing and the sensitivity of sequence comparison methods being improved, the family of the so-called 'modular' proteins keeps expanding [4]; it now includes extracellular, as well as intracellular proteins, such as muscle proteins [5–7]. The evolutionary implications of this modular assembly have been extensively discussed in the past few years and are still at the center of an ongoing debate. A considerable number of protein modules, classified into several superfamilies, have been shown to be autonomously folded and their three-dimensional structures have been determined [8–12]. However, the selection of module boundaries poses a serious problem, especially in the case of intracellular proteins when it has to be based on criteria other than the topography of intron–exon junctions, such as limited proteolytic digestion or multiple sequence alignment methods.

We have undertaken a longterm structural analysis of the modular protein titin in solution using a variety of techniques and mainly NMR spectroscopy. Titin (also known as connectin) is a 3 MDa protein thought to form a fibrous intracellular system in vertebrate striated muscle and to play an important role in sarcomere alignment during muscle contraction [13–17]. It has also been implicated as a 'molecular ruler', regulating the assembly and the precise length of the thick filaments [18]. Sequencing of titin-encoding cDNAs showed that the protein is organized in a modular fashion, containing two classes of ~100-residue repeats [7]. These repeats, referred to as type I and type II modules, show sequence homology to the fibronectin III and immunoglobulin superfamilies, respectively. Their arrangement along the titin molecule shows considerable variation (Fig. 1a). Since Ig motifs represent the most diversified class of modules in animals [19] and along the titin molecule, we expressed several domains of this type in *E. coli*, spanning different regions of the sarcomere and studied their structure and stability. The results of this study have been described in detail elsewhere [20].

Briefly, we showed that all the fragments examined are independently folded in solution and share an overall folding pattern related to the Ig fold as modelled after the telokin structure [21]. However, despite the overall structural similarity, the stability of the modules seemed to differ considerably; the module corresponding to the M-line was found significantly more stable than the modules corresponding to the A-band. Different stability might parallel different mechanical roles of domains located in different regions of the sarcomere or might simply reflect the influence of function-specific sequence consensus corresponding to the specific sarcomeric origin.

Most of the A-band titin motifs are encoded by a compact continuous gene, while along the M-line there are a few intron insertions, which are in general not related to the boundaries of structural domains (Kolmerer and Labeit, personal communication). Thus, our selection of module boundaries could not be based on the topography of intron–exon junctions. Limited proteolytic digestion of expressed multi-module constructs was attempted but proved inconclusive. The criterion we chose was a thorough alignment of the available titin sequences [7,22] between themselves and against telokin (Fig. 1b).

To establish that the different stability is indeed an intrinsic feature of the studied modules and not a result of a different length and to facilitate the ongoing structure determination, we decided to investigate in a systematic way the effect of peptide length on stability. We extended N-terminally the M-line domain, M11 (the most stable one according to our previous data) and the A-band module, Ab1 (the least stable one) and determined, with the use of CD and Fluorescence techniques, key thermodynamic parameters characterizing their stability. Ab1 was extended by 11 residues and by 4 residues, while 6 residues were added to M11.

*Corresponding author. Fax: (49) (6221) 387 306;
E-mail: <politou@EMBL-Heidelberg.DE>.

Abbreviations: CD, circular dichroism; ΔG, Gibbs free energy change; NMR, nuclear magnetic resonance; NTA, nitrilotetra-acetic acid; PCR, polymerase chain reaction; T_m, melting temperature; UV, ultraviolet.

SSDI 0014-5793(94)00911-2

The N-terminus of the modules was extended for two reasons. First, the multiple alignment shows that the C-terminus is quite unambiguous and can be precisely defined by the type II module comprising the physical C-terminus of the entire titin molecule and by the very well conserved N-terminus (*PgPP*) of the adjacent type I modules in the A-band region (Fig. 1b). Second, our NMR data, in particular from the Ab1 domain, are inconclusive only in the N-terminal region which might represent the most flexible part of the protein.

We report here the results of this analysis which confirms our previous observations and shows that, while amino-terminal extension has a profound effect on the stability of individual modules, it does not affect at all their relative stabilities.

2. Materials and methods

2.1. Protein expression and purification

The original M11 and Ab1 modules were expressed and purified as previously described [20,23–24]. The extended versions (Ab1ex11, Ab1ex4 and M11ex6) were solubly expressed at 37°C and purified from cell lysates. Briefly, induced BL21 cells were lysed in Buffer A and the cytosolic fraction applied to a Ni^{2+}-NTA agarose column after centrifugation at 15000 rpm in a Sorvall SS34 rotor. The column was eluted with a gradient of 0–200 mM Imidazol-HCl pH 7, 20 mM β-mercaptoethanol and fractions containing pure proteins were pooled. After dialysis against several changes of Buffer B, the proteins were purified by conventional anion-exchange chromatography.

2.2. Fluorescence spectroscopy–chemical denaturation studies

Urea denaturation of the titin domains was monitored by measuring the intrinsic fluorescence intensity of solutions containing typically 50 μg/ml of protein and varying concentrations of urea. Urea stock solutions (10 M) were prepared by dissolving 'ultrapure' urea (Schwarz/Mann Biotech) in 10 mM acetate buffer, pH 4.2 and in 20 mM MOPS (or 10 mM phosphate) buffer, pH 7.2. Dilution with the appropriate buffer and addition of stock protein solutions resulted in a series of 20–28 solutions, each one of which represents a point in the obtained denaturation curves. Fluorescence measurements were made in 1.0 cm quartz cuvettes thermostatted at 25 ± 0.1°C with an SLM-Aminco Bowman Series 2 spectrofluorimeter, after the solutions were incubated at 25°C for at least 12 h. The slit widths were 4 nm for both excitation and emission. The intensity of fluorescence emission was monitored at 305, 309, 316, 352 and 357 nm with an excitation wavelength of 293 nm. The pH of five solutions near the midpoint of the transition was recorded and the average was considered as the pH of denaturation. The denaturation experiment was repeated with at least two different preparations of each protein module studied. The free energy of folding, $\Delta G(H_2O)$, and the concentration of urea at which 50% of the protein is denatured, $[U]_{1/2}$, were determined by postulating a two-state mechanism of unfolding, as described in detail elsewhere [20,25–29].

2.3. Circular dichroism–thermal unfolding

Circular dichroism (CD) spectra in the far UV were recorded on a Jasco J-710 spectropolarimeter, fitted with a thermostatted cell holder and interfaced with a Neslab RTE-110 water bath. The instrument was calibrated with a 0.10% aqueous solution of d-10-camphor-sulphonic acid. Quartz, thermostatted cuvettes with 1 mm or 0.2 mm path lengths were used (Hellma). Spectra were typically recorded in 10 mM acetate buffer, pH 4.2 and in 20 mM MOPS or 10 mM phosphate buffer, pH 7.2 and were baseline corrected by subtraction of the appropriate buffer spectra. The combined absorbance of cell, sample and solvent was kept less than 1 over the measured range. Thermal denaturation curves of at least two different preparations were obtained for each module by measuring the ellipticity at 201–204 nm with increasing temperature (20°C/h). Non-linear regression analysis of the curves yielded the value of T_m (midpoint of thermal transition).

3. Results

Far-UV CD spectra of the original and the extended titin domains at ambient temperature (25°C) are very similar, indicating that the secondary structure remains unchanged upon N-terminal extension of the modules. CD spectra recorded in the pH range of 4.2–7.2 show that there is no pH-dependent conformational change, at least in this pH range.

Fluorescence spectroscopy proved ideally suited for monitoring the unfolding, because in all cases, the emission spectra of the denatured state differed significantly from the ones of the native protein both in intensity and in the maximum emission wavelength (at least 21 nm red shift upon unfolding of the modules). The positions of the maxima did not change when the modules were amino-terminally extended and with pH.

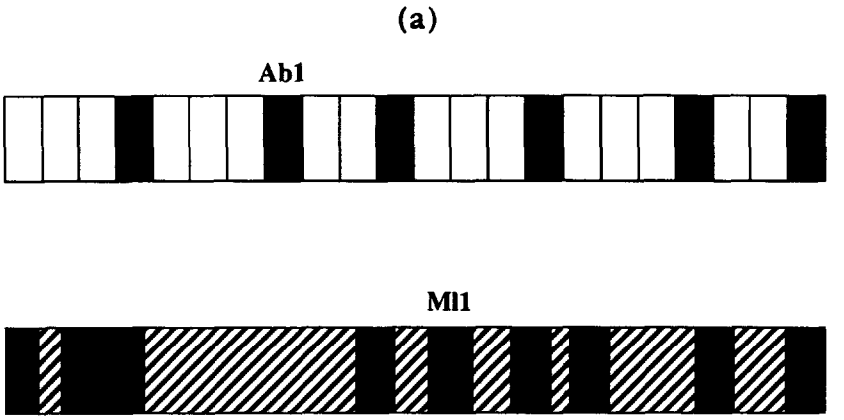
All denaturation curves obtained are characterized by the same sigmoidal shape (see Figs. 2 and 3 for representative curves).

The results of the chemical and thermal denaturation study are summarized in Table 1 and clearly show that all extended modules are significantly more stable than the corresponding original ones. The values of $\Delta G(H_2O)$ and $[U]_{1/2}$ for the extended modules are in general 1.2–1.6 times higher than those corresponding to the unextended modules. This is true for both pH values despite the fact that all modules are more stable at neutral pH. The difference in stability between the A-band and the M-line domains has not been affected by their amino-

Table 1
Thermodynamic parameters characterizing the stability of titin modules

Module	pH	T_m [a,c] (°C)	$[U]_{1/2}$ [b,c] (M)	m [d] (cal·mole^{-1}·M^{-1})	$\Delta G(H_2O)$ [e] (kcal·mole^{-1})
Ab1	4.2	43.9 (0.1)	1.65 (0.08)	1511 (135)	2.50 (0.32)
	7.1	46.1 (0.2)	2.51 (0.09)	1044 (162)	2.62 (0.44)
Ablex4	4.2	51.4 (0.2)	2.33 (0.02)	1615 (66)	3.77 (0.17)
	7.2	53.6 (0.3)	2.97 (0.03)	1282 (77)	3.81 (0.24)
Ablex11	4.2	52.5 (0.1)	2.79 (0.02)	1406 (60)	3.92 (0.17)
	7.2	54.3 (0.2)	3.33 (0.02)	1141 (45)	3.81 (0.15)
M11	4.2	50.5 (0.3)	2.12 (0.02)	1700 (72)	3.61 (0.16)
	7.2	52.9 (1.5)	4.38 (0.03)	1340 (99)	5.88 (0.45)
M11ex6	4.2	59.3 (0.9)	3.79 (0.01)	1443 (39)	5.47 (0.15)
	7.3	68.8 (0.5)	7.18 (0.05)	1418 (166)	10.19 (1.19)

[a] Calculated as the midpoint of the thermal denaturation curve.
[b] Determined as the midpoint of the urea unfolding curve.
[c] Parameters associated with the least error and the highest reproducibility.
[d] The slope of the urea unfolding curve.
[e] Conformational stability at 25°C.



(b)

```
Telokin          KPHVKPYFTK--TILD-MDVVEGSAARFDCKVEGYPDPEVMWFKDDNPVKE--SRHFQIDYDEEGNCSLTISEVCGDDDAKYTCKAVNSLGEATCTAELLVETM
Ab1ex11   pgpviveekleaPDIDLDLELRKIINIRAGGSLRLFVPIKGRPTPEVKWGKVDGEIRD----AAIIDSTS-SFTSLVLDNVNRYDSGKYTLTLENSSGTKSAFVTVRVLDTpspp
Ml1ex6           kttlaaRILT--KPRS-MTVYEGESARFSCDTDGEPVPTVTWLRKGQVLST--SARHQVTTTK-YKSTFEISSVQASDEGNYSVVVENSEGKQEAEFTLTIQKArvte
C-terminus            PKIEA--LPSD-ISIDEGKVLTVACAFTGEPTPEVTWSCGGRKIHSQEQGRFHIENTD-DLTTLIIMDVQKQDGGLYTLSLGNEFGSDSATVNIHIRSI*
CONSENSUS             P-I-L---------V--G----------G-P-P-V-W-K-G--L-----------T------L-I----R-D-G-Y--T--N--G--S----V-VLD-PgPP
```

Fig. 1. (a) The arrangement of modules in the A-band (top) and in the M-line (bottom) part of the titin molecule. Black boxes indicate the type II and white ones the type I modules. Diagonally striped boxes represent non-repetitive 'linker' sequences separating the type II modules in the M-line. Also indicated are the positions of the modules expressed and studied in the present work. (b) Alignment of the original versions of the two titin modules (capital letters) against telokin (adapted from the one suggested by Y. Harpaz and C. Chothia). Bold characters indicate the residues critical for the selection of the original module boundaries, i.e. the conserved *Pro* residue around which the N-terminus of the Ab1 module was selected and the *Hydrophobic–X–Hydrophobic* stretch that defined the C-terminus. Four residues adjacent to the C-terminus of each module are included (small letters). Also shown are the C-terminus of the entire titin molecule and a consensus sequence (CONSENSUS) which includes the residues conserved in more than 55% of all type II modules and the vey well conserved N-terminus PgPP (underlined) of the adjacent type I modules in the A-band region. The residues later added amino-terminally to the original modules are shown in small letters. Two versions of extended Ab1 have been studied, one 4-residues and one 11-residues longer than the original module.

terminal extension. The titin domains differed appreciably in their thermal stability as well. In all cases, the extended ones were appreciably more stable than the original modules. More specifically: (i) at acidic pH, the melting temperature of the A-band domain increased by 7°C, when it was extended by 4 residues, while a marginal change was brought about by its further extension by 7 more amino acids. At neutral pH, Ab1 was more stable than at acidic pH; upon amino-terminal extension its thermal stability increased in a comparable manner as at low pH; (ii) the Ml1 module which is intrinsically more stable than Ab1, seemed to be affected most dramatically, especially at high pH (increase in the melting point of 9°C at acidic pH and 15°C at neutral pH).

The method of analysis used is known to yield the lowest estimates of $\Delta G(H_2O)$ [26] and could in part account for the low values of conformational stabilities obtained in general (Table 1). However, the proteins studied are all modules and it is reasonable to expect that single, isolated modules cannot be as stable as an intact protein.

4. Discussion

Modular proteins are a real blessing for the structural biologist, since the size, the stability and the efficient expression of their component modules facilitate their structure determination by a variety of spectroscopic techniques. Knowledge of the structure of individual modules is meaningful, because it can provide the molecular basis for important properties of the whole protein and could eventually lead to an understanding of the structure and function of the entire protein.

In dissecting a large protein into its modules, however, one should be fully aware of the potential influence of the boundaries on the stability of the isolated modules. Even in the favourable case of extracellular proteins with well-defined exon borders, domains can be unstable when expressed singly. In the case of intracellular proteins, like titin, the selection of boundaries poses a more serious problem and can be of critical importance.

In a work reported recently [20] three Ig-type titin modules were expressed and studied, the similarity in their folding pattern was established and their relative stabilities compared and found to vary considerably. Limited proteolysis with trypsin, thermolysin, subtilysin and elastase of expressed multi-domain subunits proved not helpful regarding the selection of module boundaries which had, therefore, to be based on multiple alignment criteria. Amino-terminal extension of these modules did

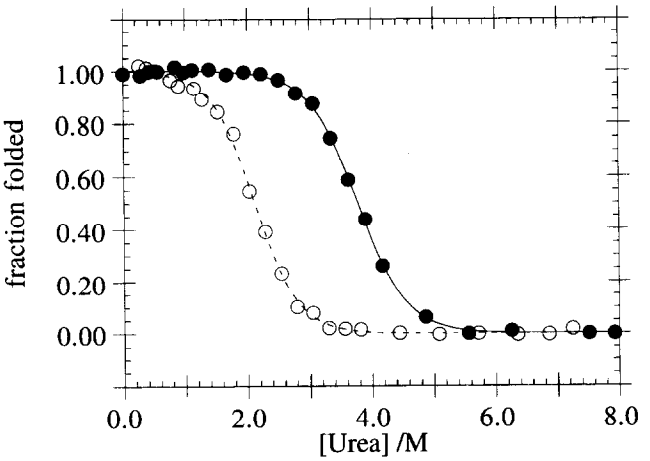


Fig. 2. Urea-induced equilibrium unfolding of the original and the extended MI1 module at pH 4.2, 10 mM acetate buffer, 25°C. MI1 (open circles); MI1 extended (filled circles). Unfolding was monitored by the intrinsic fluorescence intensity at 316 nm after excitation at 293 nm. These plots show normalized data and include the curve of best fit obtained from non-linear regression analysis of the data.

not seem to affect in any way the secondary structure, as indicated by the far-UV CD, fluorescence and NMR spectra recorded at two different pH values (data not shown). On the other hand, the effect on stability was striking (Table 1). Extension of the Ab1 module by 4 residues raised the melting point by >7°C at each pH, while futher extension by 7 additional residues had only a marginal influence (if at all) on the thermal stability. A similar effect was observed, when resistance to chemical (urea) denaturation was considered. The M-line domain showed an even more impressive stability enhancement as a result of its amino-terminal extension by 6 aminoacids (T_m increase of 15°C and $[U]_{1/2}$ 1.6 times higher). One should note that these changes were not accompanied by a change in the relative stabilities of the two modules, i.e. the A-band

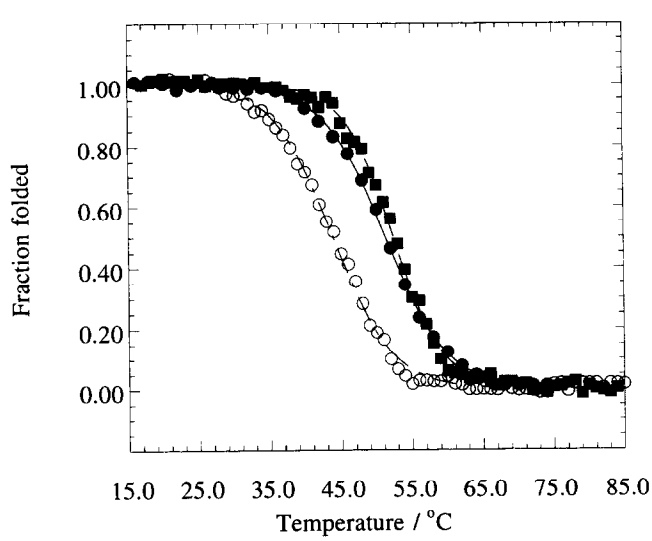


Fig. 3. Thermal denaturation curve of the original and the extended Ab1 modules in 10 mM acetate buffer, pH 4.2, monitored by circular dichroism at 203 nm. (Open circles) Ab1; (filled circles) Ab1ex4; (filled squares) Ab1ex11. Data shown are normalized and the curve of best fit is included.

domain remained considerably less stable than the M-line domain. Moreover, these changes are faithfully reproduced both at low and high pH values.

This paper is intended as a word of caution to those interested in studying the behaviour of large, modular proteins by dissecting them into their component modules. A careful limited proteolysis analysis, a rigorous multiple sequence alignment, even a clear-cut intron–exon topography might prove inadequate, when it comes to choosing the optimal module boundaries. Wrong choice could lead to a completely unfolded module or, in the best case, to autonomously folded, but very unstable ones. After all, same fold does in no way guarantee equally favourable to structural analysis behaviour. The time and the effort spent initially to explore several possible cutting positions, to prepare a few versions of the same module and to optimize their stabilities, will be repaid many times over in the process of their detailed structural analysis.

Acknowledgements: We would like to thank Toby Gibson for the alignment of titin, Yehouda Harpaz and Cyrus Chothia for the modelling of titin modules after telokin and Bernard Kolmerer and Siegfried Labeit for communicating their results on the genomic sequence of titin prior to publication. This work was upported by grants from the Deutsche Forschungsgemeinschaft (Ga 405/2–1, La 668/2–1), the European Community and the Human Frontiers Science Project. A.S.P. was supported by a BRIDGE Fellowship from EC.

References

[1] Doolittle, R.F. (1989) Trends Biochem. Sci. 14, 244–245.
[2] Baron, M., Norman, D.G. and Campbell, I.D. (1991) Trends Biochem. Sci. 16, 13–17.
[3] Patthy, L. (1991) Curr. Opin. Struct. Biol. 1, 351–361.
[4] Bork, P. (1992) Curr. Opin. Struct. Biol. 2, 413–421.
[5] Benian, G.M., Kiff, J.E., Neckelmann, N., Moerman, D.G. and Waterston, R.H. (1989) Nature 342, 45–50.
[6] Einheber, S. and Fischman, D.A. (1990) Proc. Natl. Acad. Sci. USA 87, 2157–2161.
[7] 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.
[8] Baron, M., Norman, D.G., Willis, A. and Campbell, I.D. (1990) Nature 345, 642–646.
[9] Driscoll, P.C., Cyster, J.G., Campbell, I.D. and Williams, A.F. (1991) Nature 353, 762–765.
[10] Main, A.L., Harvey, T.S., Baron, M., Boyd, J. and Campbell, I.D. (1992) Cell 71, 671–8.
[11] Brady, R.L., Dodson, E.J., Dodson, G.G., Lange, G., Davis, S.J., Williams, A.F. and Barclay, A.N. (1993) Science 260, 979–983.
[12] Williams, M.J., Phan, I., Harvey, T.S., Rostagno, A., Gold, L.I. and Campbell, I.D. (1994) J. Mol. Biol. 235, 1302–11.
[13] Maruyama, K., Kimura, S., Yoshidomi, H., Sawada, H. and Kikuchi, K. (1984) J. Biochem. (Tokyo) 89, 701–709.
[14] Trinick, J., Knight, P. and Whiting, A.J. (1984) J. Mol. Biol. 180, 331–356.
[15] Wang, K. (1985) Cell and Muscle Motility 6, 315–369.
[16] Fürst, D.O., Osborn, M., Nave, R. and Weber, K. (1988) J. Cell Biol. 106, 1563–1572.
[17] Kurzban, G.P. and Wang, K. (1988) Biochem. Biophys. Res. Commun. 155, 1155–1161.
[18] Whiting, A.J., Wardale, J. and Trinick, J. (1989) J. Mol. Biol. 205, 163–169.
[19] Doolittle, R.F. and Bork, P. (1993) Sci. Am. 269, 50–56.
[20] Politou, A.S., Gautel, M., Pfuhl, M., Labeit, S. and Pastore, A. (1994) Biochemistry 33, 4730–7.
[21] Holden, H.M., Ito, M., Hartshorne, D.J. and Rayment, I. (1992) J. Mol. Biol. 227, 840–851.
[22] Higgins, D.G., Labeit, S. and Gibson, T. (1994) J. Mol. Evol. 38, 395–404.

[23] Labeit, S., Gautel, M., Lakey, A. and Trinick, J. (1992) EMBO J. 11, 1711–6.
[24] Gautel, M., Leonard, K. and Labeit, S. (1993) EMBO J. 12, 3827–34.
[25] Pace, C.N. (1986) Methods Enzymol. 131, 266–279.
[26] Pace, C.N., Shirley, B.A. and Thomson, J.A. (1989) in: Protein Structure: A Practical Approach (Creighton, T.E., Ed.) pp. 311–330, IRL Press, Oxford.
[27] Shortle, D., Stites, W.E. and Meeker, A.K. (1990) Biochemistry 29, 8033–8041.
[28] Santoro, M.M. and Bolen, D.W. (1992) Biochemistry 31, 4901–4907.
[29] Jackson, S.E., Moracci, M., elMasry, N., Johnson, C.M. and Fersht, A.R. (1993) Biochemistry 32, 11259–11269.