The EMBO Journal vol.13 no.8 pp.1782 – 1789, 1994

Nebulin, a helical actin binding protein

Mark Pfuhl, Steven J.Winder[1] and Annalisa Pastore

EMBL Heidelberg, Meyerhofstrasse 1, 69012 Heidelberg, Germany and [1]MRC Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2HQ, UK

Communicated by B.Bullard

Nebulin, a giant protein (molecular mass 800 kDa) specific for the skeletal muscle of vertebrates, has been suggested to be involved in the length regulation of the thin filament as a 'molecular ruler'. Despite its size, nebulin appears to be composed mainly of small repeats of ~35 amino acids. We have characterized in this study the conformational and functional properties of single repeats. Complete repeats were found to bind to F-actin while a truncated one did not. One repeat is therefore the smallest unit for nebulin–actin interaction. Circular dichroism and nuclear magnetic resonance spectra measured for the peptides in water indicated a transient helical conformation. The folded region is located for them all around the conserved sequence SDxxYK. The helical conformation is strongly stabilized by anionic detergents and trifluoroethanol while uncharged or positively charged detergents have no effect. Since the surface of the actin filament is known to contain clusters of negative charges, anionic detergents may mimic the effect of an actin environment. 3D structures were calculated for three representative peptides in SDS. *In vivo*, the nebulin helices should form a complex with the actin filament. Based on the assumed importance of charge interactions between nebulin and actin, we propose a model for the structure of the F-actin–nebulin complex *in vivo*. According to that, two nebulin molecules occupy symmetrical positions along the central cleft of the actin filament bridging the two strands of the actin two-start helix. The consistency of this model with experimental data is discussed.

Key words: muscles/nmr/structure/thin filament

Introduction

Specialized filamentous proteins play an important role in the cytoskeleton and in muscles. Their biological function requires interaction with several other proteins and the knowledge of their interaction sites is the key to the understanding of their function. An important step towards a thorough investigation of these interactions is the elucidation of high resolution structures of the proteins involved. However, this is often nearly impossible for at least two reasons: because of their large molecular weight; the folding of filamentous proteins is usually more strongly influenced by their environment (e.g. interacting proteins, salt concentration and pH) than is the case for globular proteins. What has made an investigation possible is the recent assessment that several filamentous proteins are built from repeating units or domains. The size problem can then be mitigated by investigating small regions and their boundaries. The global structure will then possibly be reconstructed as an assembly of mostly independent building blocks.

Following this approach we have started the structural characterization of nebulin, one of the filamentous proteins involved in complex protein assemblies. It is a huge protein (molecular mass 800 kDa) abundant in vertebrate skeletal muscle (Wang and Williamson, 1980; Hu *et al.*, 1989; Kruger *et al.*, 1991). Biochemical and structural characterization of the isolated protein has been impossible because it could not be purified in the native state (Wang, 1982). As a consequence, native nebulin on its own has not yet been observed by electron microscopy. However, indirect data suggest that it is associated with the thin filament (Wang and Wright, 1988). A variation of nebulin molecular weight that is proportional to sarcomer length has been observed in different species. These data have suggested that nebulin functions as a 'molecular ruler' which defines the length of thin filaments by its own length (Kruger *et al.*, 1991; Labeit *et al.*, 1991). The amino acid composition (Wang and Wright, 1988) and a partial sequence corresponding to ~8% of the complete sequence has been determined (Labeit *et al.*, 1991) showing an unusually high amount of positively charged residues and no homology to any known sequence in the database. In spite of its large overall size, nebulin is made up of small repeats of ~35 amino acids each (Figure 1). Unpublished work suggests that the 35 residue repeat/super-repeat pattern extends over most of the nebulin sequence (Wang *et al.*, 1990; Wright *et al.*, 1993). Seven small repeats are assembled to form a super-repeat. The overall similarity between the super-repeats is stronger than between the single 35 amino acid-containing repeats. Multiple alignment of single repeats indicate an α-helical periodicity of conserved amino acid properties which breaks around the conserved proline at the repeat boundary. A short consensus segment consisting of a PD at the N-terminus and a SDxxYK in the middle of a single repeat was identified. The 7-fold periodicity in the super-repeat has been suggested to reflect an interaction with tropomyosin or troponin while the single repeat might reflect a 1:1 binding of the 35 amino acid repeat to actin (Wang *et al.*, 1990; Kruger *et al.*, 1991; Labeit *et al.*, 1991). The assumed single helix of nebulin along the thin filament might then be stabilized by interactions with other proteins in the thin filament.

Binding experiments performed on recombinant fragments containing between six and 15 repeats have shown high affinity binding to actin (Jin and Wang, 1991). More recently, Chen *et al.* (1993a,b) have demonstrated that a two repeat construct binds to F-actin. These authors have also shown that the two repeat construct accelerates the polymerization of G-actin and reduces the depolymerization rate

```
                    1   1   2   2   3   3
          1    5    0   5   0   5   0   5
   s1r1      IDAMSLEHTRNVNQIQSDNVYKDEYNSFLKGIGWI
   s2r1      PDALNVEHSRNAMQIQSDNLYKSDFTNWMKGIGWV
 * s3r1      PDAMSFELAKNMMQIQSDNQYKADYADFMKGIGWL
   s4r1      PDAMNIELTRNMNRIQSDNEYKQDYNEWYKGLGWS
   s1r2     PIGSLEVEKVKKAGDALNERKYRQHPDTVKFTSV
   s2r2     PIESLEVEKAKKAGEILSEKKYRQHPEKLKFTYA
   s3r2     PLGSLEAEKNKKAMEIISEKKYRQHPDTLKYSTL
   s4r2     PAGSLEVEKAKKATEYASDQKYRQHPSNFQFKKL
   s1r3      PDSMGMMLAQHNTKQLSDLNYKVEGEKLKHKYTID
   s2r3      MDTMEQALNKSNKLNMDKRLYTEKWNKDKTTIHV
   s3r3      MDSMNMVLARNNAKIMNEHLYKQAWEADKTKVHIM
   s4r3      TDSMDMVLAKQNAHTMNKHLYTIDWNKDKTKIHVM
   s1r4      PELPQFIQAKVNALNMSDAHYKADWKKTIRKGYDLR
   s2r4      PDTPDILLSRVNQITMSDKLYKAGWEEEKKKGYDLR
   s3r4      PDIPQIILAKANAINISDKLYKLSLEESKKKGYDLR
   s4r4      PDTPDILQAKQNQTLYSQKLYKLGWEEALKKGYDL
   s1r5      PDAIPIVAAKSSRNIASDCKYKEAYEKAKGKQVGFLS
 * s2r5      PDAIAIKAARASRDIASDYKYKKAYEQAKGKHIGFRS
   s3r5      PDAIPIKARKASRDIASDYKYKYNYEKGKGKMVGFRS
   s4r5     PVDAISVQLRKASRDIASDYKYKQAGI
 * s1r6     LQDDPKLVHYMNVAKIQSDREYKKGYEASKTKYHT
   s2r6     LEDDPKLVHFMQVAKMQSDREYKKGYEKSKTSFHT
   s3r6     LEDDPKLVHSMQVAKMQSDREYKKNYENTKTSLPP
 * s1r7     PLDMVSVTAAKKSQEVATNANYRQSYHHYTLL
   s2r7     PVDMLSVVAAKKSQEVATNANYRNVIHTYNML
   s3r7     PADMLSVTAAKDAQANITNTNYKHLIHKYILL

   CONS     .PD.@.@...+......SD..YK..@........
```

Fig. 1. Alignment of the nebulin repeats together with the consensus pattern. S1-4 stands for super-repeat 1 to 4, R1-7 stands for repeat 1 to 7 in the corresponding super-repeat. The repeats that were used in this study are labelled by a star. The truncated (S2R5t) was constructed by synthesizing a peptide starting at Lys8 and ending at Gln28 of the S2R5 sequence. The overlapping repeat (R4R5) was made from the C-terminal part of S1R4 starting with Asp19 and adding the N-terminal part of S2R5 up to Ser18. The combination S1R4+S2R5 was chosen instead of the correct one S2R4+S2R5 to achieve a higher solubility of the peptide. The S3R1 peptide was insoluble in water.

of F-actin (Chen *et al.*, 1993b). Based on these results, they have suggested that nebulin could work as an 'actin zipper'.

As a basis for the understanding of the relationship of nebulin folding and binding we describe here a study of the conformation of nebulin using circular dichroism (CD) and nuclear magnetic resonance (NMR) spectroscopy on peptides containing single repeats as well as the repeat boundary. In addition the binding properties of these peptides are explored.

Results and discussion

Six peptides were synthesized: four complete single repeats (S3R1, S2R5, S1R6 and S1R7), a truncated repeat (S2R5t) and an overlapping repeat (R4R5) (Figure 1). All but S3R1 were soluble in water, the exception probably being related to the considerably smaller number of charged residues in this sequence. The peptide S3R1 will not be discussed further. The peptide S2R5t was studied to test the influence of the N- and C-terminal residues on folding and actin binding and to help assignment of the NMR spectra. The overlapping repeat was chosen to investigate the repeat boundary. It is constructed from repeats 4 and 5 from two different super-repeats to achieve higher solubility. No relevant influence on the results is expected since the first four residues are identical and the first 10 highly homologous (75%) in the two S1R5 and S2R5 sequences leading to an almost identical boundary region. The same applies to the C-terminus of repeats S1R4 and S2R4. The conformation

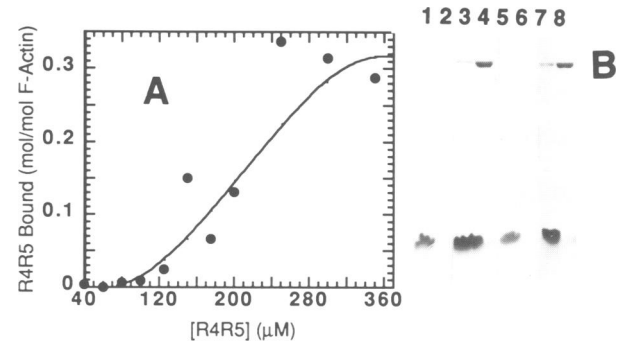

Fig. 2. (**A**) Actin binding curve for R4R5 binding to 10 μM F-actin as described under Materials and methods. The curve was fitted with the Hill equation giving a Hill coefficient of 3.6 ± 0.5. K_D was determined to be 220 ± 22 μM and stoichiometry 2.6 ± 0.5 mol actin per mol of peptide. (**B**) Tris/Tricine gels of the supernatant (lanes 1, 3, 5 and 7) and pellet fractions (lanes 2, 4, 6 and 8) of R4R5 (lanes 1–4) and S2R5 (lanes 5–8) sedimented alone (lanes 1 and 2 and 5 and 6) or in the presence of F-actin (lanes 3 and 4 and 7 and 8). The band of actin is on the top of the gel strips; the peptide bands are on the bottom.

and binding properties of the overlapping repeat should therefore be unaltered.

Binding assays

Densitometric analysis of Tris/Tricine gels of F-actin binding experiments revealed that peptides R4R5, S1R7 and S2R5 bound specifically to F-actin but with rather low affinity (200–500 μM), whilst S1R6 and S2R5t did not bind at all. A detailed description of the actin binding properties of these peptides will be given elsewhere. R4R5, S2R5 and S1R7 bound to F-actin in a concentration-dependent manner (examples shown for S2R5 and R4R5 in Figure 2B) and binding was independent of the presence of calcium (data not shown). These peptides were recovered solely in the supernatant fraction following centrifugation (Figure 2B). In the presence of actin , however, a small portion of peptide was recovered in the actin-containing pellet fraction (Figure 2B). After correction for trapping, the apparent dissociation constant, based on the peptide concentration required to achieve 50% maximal binding, was determined to be 220 ± 22 μM for R4R5 and 420 ± 25 μM for S2R5, with saturation of binding occurring at 1 mol of peptide per 2.5 ± 0.5 and 1.1 ± 0.1 mol of actin for R4R5 and S2R5 respectively (see Figure 2A). Hill constants of 3.6 ± 0.4 for R4R5 and 2.5 ± 0.3 for S2R5 indicate an extremely co-operative interaction. The repeat S1R7 gave similar results (data not shown). These data suggest that the single repeat is the smallest unit of nebulin–actin interaction.

CD spectroscopy

Peptides dissolved in water gave CD spectra with negative maxima in ellipticity around 190–200 nm which points to a dominant content of unfolded conformation as shown in Figure 3. Analysis of secondary structure based on the fitting of standard spectra of α-helix, β-sheet and random-coil (Brahms and Brahms, 1980) gave a maximal amount of α-helix of $<10\%$. Increase of the pH value from 4.4 to 8.0 using 25 mM phosphate buffer, slightly increased the amount of α-helix to 20% (spectra shown for S2R5 in Figure 3). Changes in ion concentration did not affect the amount of secondary structure in the peptides. Addition of SDS or de-

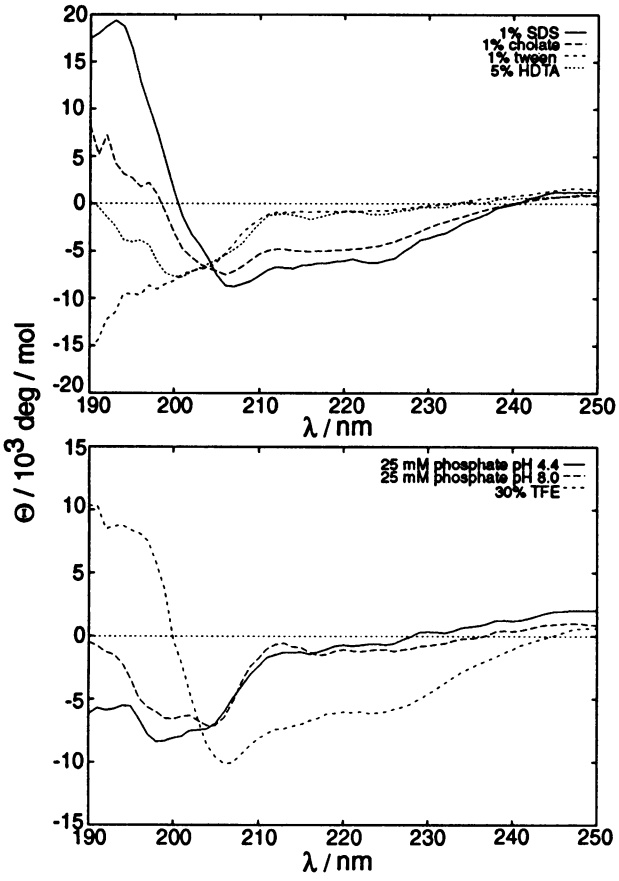

Fig. 3. CD spectra of S2R5 at room temperature under different conditions. SDS and deoxycholate produce a helix-like spectrum, while Tween and HDTA have only little effect (10 and 19% α-helix). A change in pH from 4.4 and 8.0 in 25 mM phosphate increased the helicity from 4 and 23% while TFE gave 45% of α-helix. All the peptides gave similar results.

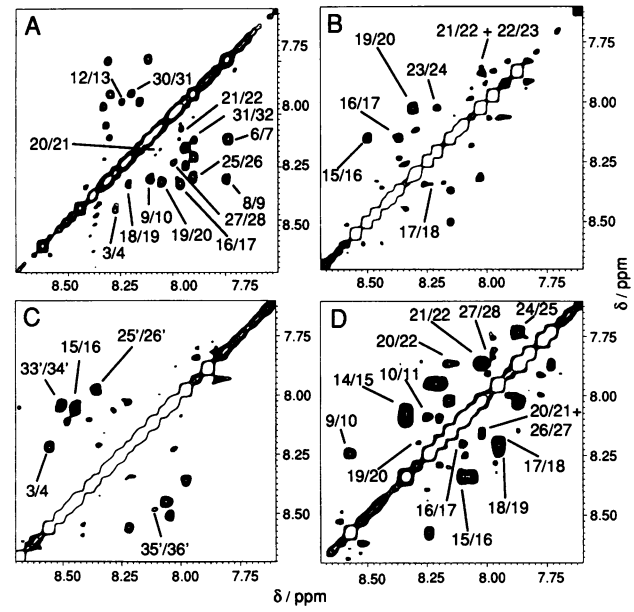

Fig. 4. Region of the NH−NH crosspeaks in NOESY spectra of peptides S1R6 (**A**), S2R5t (**B**) and R4R5 (**C**) at a concentration of 2 mM in water. The experiments were performed with a mixing time of 120 ms at a temperature of 290 K. Sequential helix-type peaks are labelled. For comparison the corresponding region from a spectrum of S2R5t at 4 mM with 600 mM d25-SDS measured at T = 305 K is shown in (**D**). Some expected peaks are not visible because they are too close to the diagonal. Chemical shifts are calibrated relative to 3-(trimethylsilyl)-2-propionic acid sodium salt (TSS).

oxycholate produced a CD spectrum having a strong positive peak around 195 nm and two negative peaks around 208 and 223 nm, indicative of a dominant contribution of helical conformation (see Figure 3). Secondary structure analysis gave ~80% α-helix with SDS and ~60% with deoxycholate. The effect of the positively charged detergent hexadecyl diamine tetraalkyl bromide (HDTA) was weak while the uncharged detergents Tween and Triton X-100 had almost no effect on the CD spectra (see Figure 3). The addition of 30−40% trifluoroethanol (TFE) (see Figure 3) gave a clear but smaller increase in the amount of α-helix than SDS or deoxycholate.

It is well known that SDS stabilizes helices, although the exact mechanism is still not completely understood (Motta *et al.*, 1991). On the other hand, the importance of electrostatic effects for the interaction of negatively charged lipids with positively charged polypeptides and the influence on secondary structure formation has been widely reported (Takahashi *et al.*, 1991; Reynaud *et al.*, 1993). The ionic interaction is thought to hold peptide and lipids/detergents together while the induced secondary structure of the peptide depends both on the intrinsic conformational propensities of the peptide and on the shape of the surface. In our case, a strong electrostatic contribution specific for negative charged micelles can be easily related to the large number of positively charged amino acids that are consistently more abundant than negative ones in the nebulin sequence (Wang and Wright, 1988; Labeit *et al.*, 1991). This correlates well with the large patches of negative charge on the surface of actin (Kabsch *et al.*, 1990). We may argue that the interaction with this molecule is the driving force for the folding of nebulin and therefore that SDS provides a suitable medium to mimic this environment *in vivo*.

NMR spectroscopy: results in water

Complete assignment of the NMR spectra in water was obtained for all peptides. They showed a number of nuclear Overhauser effects (NOE) typical of an α-helical conformation. Interestingly, the best resolved parts of the spectra, which allowed a straightforward assignment, contained mainly those residues that show α-helical type NOEs. An almost uninterrupted stretch of HN−HN connectivities was observed in homologous regions of S2R5t and S2R5 (from Asp15 to Lys24), S1R6 (from Lys6 to Gly31), and S1R7 (from Val16 and Lys23) (Figure 4). Only a small number of sparse HN−HN peaks was observable instead for R4R5. The latter is usually more indicative of a turn-like conformation than of a helix. In addition, β_i-NH$_{i+1}$ peaks were identified in the central region of all peptides with the exception of R4R5 (see Figure 5). Non-sequential helix-type peaks like α_i-β_{i+3} or long range side-chain−side-chain interactions were missing or were very weak in all peptides. The presence of some of the NOE effects typical of α-helices but the absence of others suggests that the peptides are folded as transient rather than stable helices, that is they have an intrinsic strong tendency to adopt a helical fold but this needs to be stabilized by the environment.

Similar conclusions could be drawn using other NMR

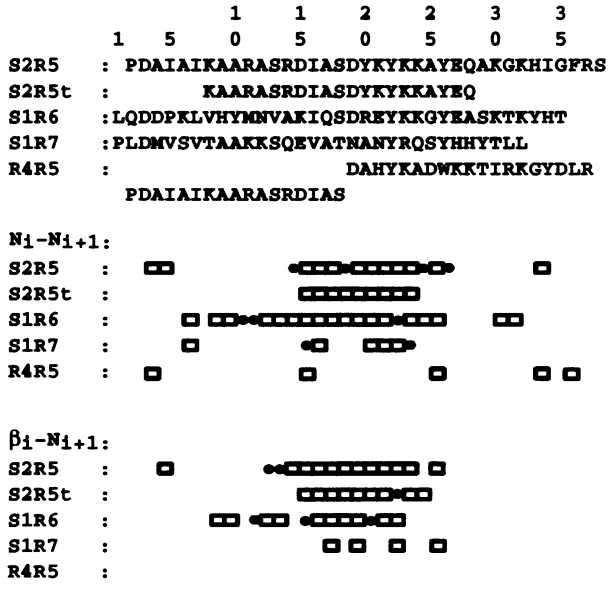

Fig. 5. Secondary structure specific NOEs identified in the full and truncated repeats in water. Open boxes indicate helix-specific NOEs between two sequentially connected amino acids. Filled circles are used when an NOE effect might be present but it is obscured by spectral overlap.

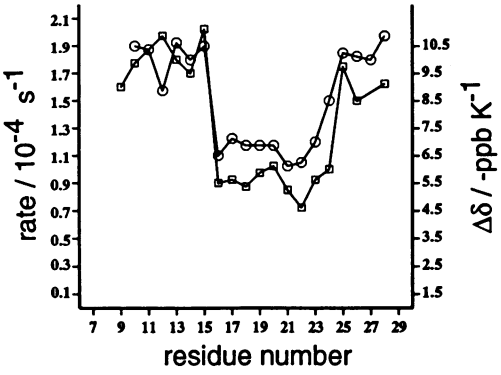

Fig. 6. D/H exchange rates (left axis and squares) at T = 7°C and temperature coefficients $\Delta\delta$ (right axis and circles) for the S2R5t peptide in water. Sample concentration was 6 mM.

parameters such as secondary chemical shifts, proton/deuterium exchange rates and temperature dependence coefficients of the amide protons. The secondary chemical shifts (that is the difference between the observed chemical shifts and the corresponding values for the same amino acid in a random coil conformation) of α-protons in amino acids in a protein are highly sensitive to the secondary structure elements (Pastore and Saudek, 1990). Positive values are indicative of α-helices while negative ones indicate the presence of β-sheets. The secondary chemical shifts of the α-protons gave consistently small positive values of $\sim 0.15-0.25$ p.p.m. for the central regions of all peptides. These values are comparable with those reported for isolated flexible helices of similar size (Dyson *et al.*, 1988, 1992). The presence of a well folded conformation in a peptide is manifested by protection of labile protons from exchange with protons from the solvent (Wüthrich, 1986). D/H exchange measurements for amide protons of the S2R5t peptide were consistent with the presence of local conformations. The exchange rates of the residues from Ile16 to Lys24

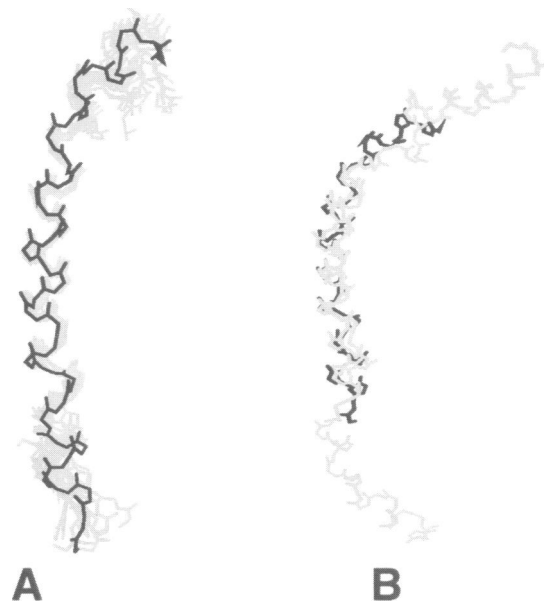

Fig. 7. (**A**) Superposition of the 10 best structures (in gray) for S2R5 after simulated annealing molecular dynamics. The r.m.s.-fit was done for all helical residues from 6 to 30. The final structure (in black) was obtained after minimization with XPLOR of the averaged conformation. (**B**) Superposition of the final average structures for the three peptides according to the overlapping sequence regions. Two copies of R4R5 are shown: the N-terminal half of the first R4R5 is aligned with the not identical but homologous C-terminal part of S2R5/S2R5t while the second is aligned with its C-terminal helix to the identical N-terminus of S2R5/S2R5t. S2R5 is in black, S2R5t and both copies of R4R5 are in gray.

differ significantly by a factor of almost two from the other amino acids (see Figure 6). Similarly, low temperature dependence of the amide proton chemical shifts also reflects protection from solvent (Wüthrich, 1986). The measurement of temperature coefficients for the S2R5t peptide in water showed only a short region with values around -5 ppb/K from Asp15 to Lys23. The rest of the peptide gave values around -9 ppb/K as shown in Figure 6. All peptides showed a very similar behaviour (data not shown).

From these data we may conclude that all secondary structure specific NMR parameters show a clear transition from the centre of the repeat towards both ends (see Figure 6). Neither pH increase nor addition of phosphate at constant pH has an effect on the pattern of helix-specific peaks (data not shown). The central part of nebulin repeats therefore folds in a transient helical conformation mostly independent from the environmental conditions such as solvent, salt content, pH and from the peptide length.

NMR spectroscopy: results in SDS

Since the behaviour of the investigated peptides is consistently very similar, we restricted a detailed structural characterization in SDS only to one representative complete repeat (S2R5), the truncated repeat (S2R5t) and the overlapping repeat (R4R5). Complete sequential assignments of these peptides in SDS were easily obtained for these three peptides. The sequence-specific assignment was facilitated by extensive comparison of the corresponding identical regions of the three peptides and by analogy with the spectra in water. Spectra at different temperatures were recorded

to reduce spectral overlap. A detailed description of the assignment will be published elsewhere.

In agreement with the results from CD spectroscopy, a larger number of α-helix-type peaks was present in the NOESY spectra measured in SDS than in water. In both S2R5t and S2R5 a nearly continuous stretch of helix-type NOEs was identified along their whole sequence with the exception of the termini, extending from residue Ala9 to Glu27 in S2R5t (see Figure 4) and from Ala4 to Ile34 in S2R5. In addition to the sequential connectivities observed in water (see Figure 5), medium-range $H\alpha_i$-$H\beta_{i+3}$ and a few HN_i-NH_{i+2} and $H\alpha_i$-HN_{i+3} peaks were also present. In R4R5, two distinct nearly continuous stretches of sequential amide−amide NOE peaks were identified from His21 to Gly33 and from Asp3 to Ile16 (data not shown). The Arg37−Pro2 peptide bond R4R5 assumes a *trans* conformation as shown by the presence of NOE peaks between the Hα and HN of Arg37 and the δ-protons of Pro2.

Analysis of the secondary structure contribution of the chemical shift of the α-protons for S2R5 and S2R5t in SDS showed positive values (0.2−0.3 p.p.m.) indicative of an α-helix for the uninterrupted regions from residues 3 to 32 and 11 to 27, respectively (data not shown). In contrast, R4R5 gave two clearly distinct positive regions, the first from residue 2 to 12 and the second from 21 to 31.

The 40 best structures for each peptide calculated with DIANA gave target function values in the range of 9−13 Å^2. The 12 best structures after simulated annealing molecular dynamics with XPLOR gave no NOE distance violation greater than 0.6 Å. The bond lengths were all in an interval of 0.05 Å around the standard values as were the bond angles within an interval of 5°. The root mean square (r.m.s.) deviations of the average conformation for each peptide are 0.9 Å in S2R5, 0.6 Å for S2R5t and 1.1 Å for R4R5 (the r.m.s.-deviation was calculated for the α carbons only). The quality of the superposition can be assessed in Figure 7 for the S2R5 peptide. The helix in this peptide is slightly bent in the region of residue Gly31. A corresponding but slightly stronger bend is present in the first helix of peptide R4R5 which corresponds to the C-terminal part of S2R5. In the peptides S2R5t and S2R5 a single uninterrupted helix with 4.0 and 3.8 residues per turn was identified while the peptide R4R5 contained two distinct helices with 3.9 residues per turn each. The region in between these two helices around Pro2 assumed a more extended conformation. In contrast to the well defined helical regions in all three peptides, this part shows a broader variation.

The location of the helices in all three peptides was confirmed by the identification of hydrogen bonds from (9−13) to (24−28) in S2R5, (14−18) to (23−27) in S2R5t and (22′−26′) to (31′−35′) for the first half and (4−8) to (11−15) in the second half of R4R5.

Therefore SDS has the effect of stabilizing and extending the helical core present in the centre of the repeats. The effect breaks at the boundary region.

Conclusions

This report gives the first detailed experimental evidence of a helical conformation in nebulin in support of previous structural hypotheses (Labeit *et al.*, 1991). The measurements of peptides comprising single nebulin repeats, in pure water or phosphate buffer, using CD and NMR spectroscopy clearly exhibit the standard properties of a transient helix (nascent helix) (Dyson *et al.*, 1988, 1992). The central part of each repeat forms a more stable helical core. Interestingly, this contains among its eight residues the conserved motif SDxxYK. When this motif is replaced by TNxxYR as in S1R7 the helix core is, according to the NMR parameters, less stable than in the other peptides. Conversely, S1R6 has a stable core but does not bind to F-actin. This evidence, together with the observation that the overlapping repeat having the central motif disrupted at both ends binds as well as the full repeats, suggests that this region is important for folding rather than for binding. In addition, it could be involved in interactions with proteins other than actin.

Upon binding to F-actin the part of the repeat that is unfolded in solution might then fold to extend the helix towards the repeat boundary. Interactions must involve charges as well as hydrophobic effects as indicated by the results at different pH values as well as from adding SDS. This would explain some of the problems encountered when trying to isolate native nebulin (Wang, 1982).

From our results the extremely high cooperativity of binding to F-actin of even single repeats, strongly suggests that the whole nebulin molecule would exert a strong influence on the polymerization kinetics of actin and support its role as an 'actin zipper' (Chen *et al.*, 1993a,b). Similar cooperativity of binding to F-actin has been noted for tropomyosin (Wegner, 1979), also associated with the thin filament. Constructs containing two repeats, with binding constants in the range 20−40 μM and a stoichiometry of 1 or 2 mol of construct per mol actin have recently been characterized (Chen *et al.*, 1993a,b). The addition of extra units therefore increases the affinity of nebulin for actin so that high affinity binding is observed for constructs containing up to 15 repeats (Jin and Wang, 1991). Fluctuations of the binding affinity along the repeats may be expected because of the heterogeneity of their sequence. Ultimately full length nebulin should have a very high affinity for actin, coupled with highly cooperative binding. Actin−nebulin self-assembly would be a very rapid process leading to an extremely stable structure.

The overall conformation of the nebulin molecule in the sarcomere, according to the assembly shown in Figure 7, would then be contiguous rods of helices within the repeat boundaries as defined by the conserved prolines. Around the proline, from residue 36′ to 6, the conformation would form a 3_{10} helix or a more extended conformation. Previous work (Kruger *et al.*, 1991) has suggested that the 245 residue super-repeat of nebulin is likely to span the 38.5 nm structural repeat of the muscle thin filament as measured by electron microscopy (Ebashi *et al.*, 1969). This allows an average of 55 Å for each repeat. Assuming an all helical conformation for the whole super-repeat length, the total residual displacement would result in an axial translation of 1.6 Å per residue compared with 1.5 Å for an ideal α-helix (Labeit *et al.*, 1991). For comparison, 3_{10} helices and extended conformations have displacements of 2.0 and 3.5 Å respectively (Stryer, 1988). We can now reconsider these calculations in the light of our results.

The helix in S2R5 has a length of 24 residues but the helical limits are shifted slightly towards the repeat boundary in R4R5. The average α-helix in the repeat would then cover roughly three quarters of the whole repeat length. If we assume the remaining 10 residues at the boundary to be in a 3_{10} helix, the full length of a super-repeat would be 56 Å

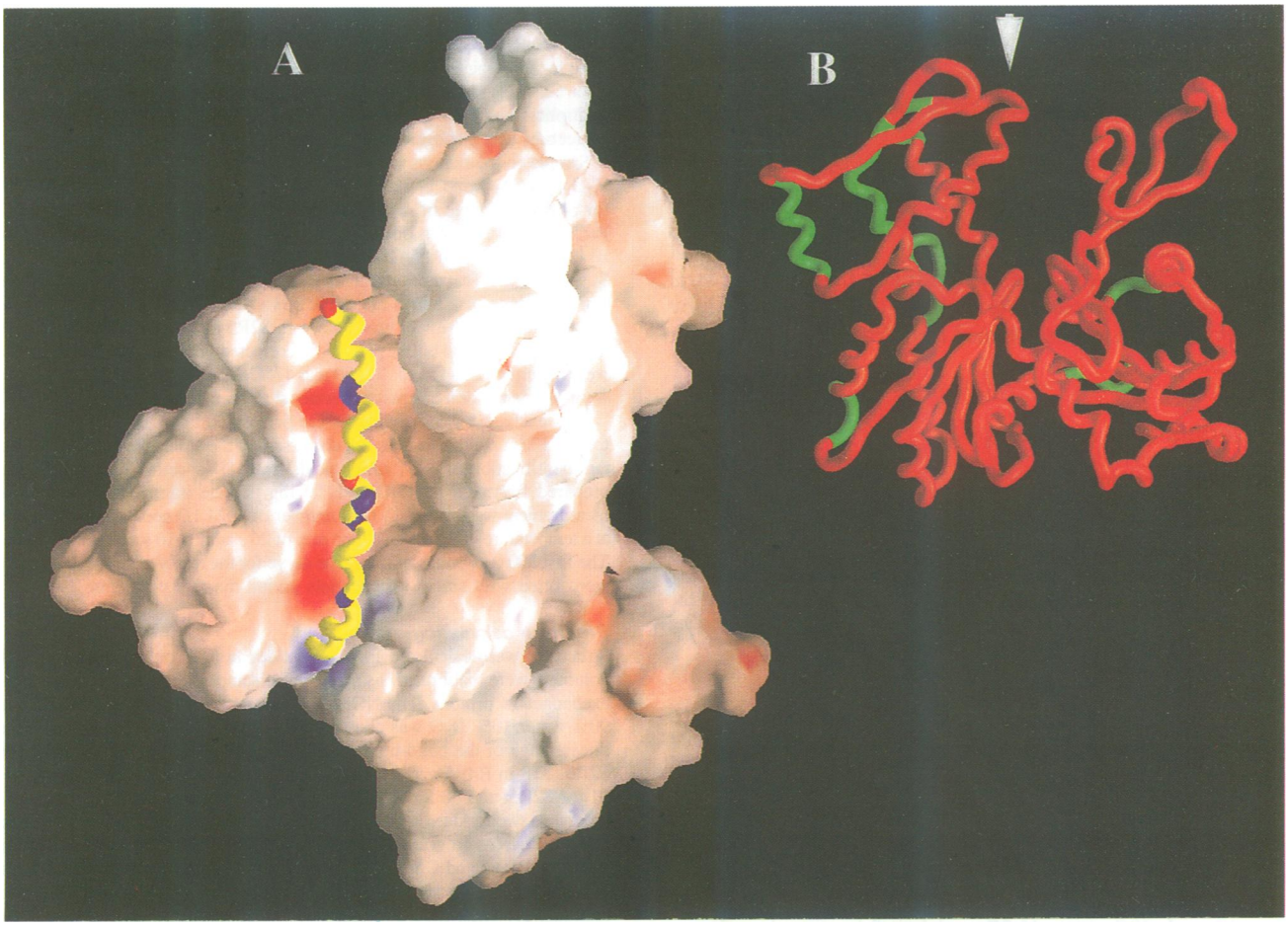

Fig. 8. Model of the interaction of nebulin with F-actin. (**A**) Three actin monomers are shown in the filament in the same orientation as published previously (Holmes *et al.*, 1990) with the barbed end at the bottom (toward the Z-line). The solvent accessible surface of actin is displayed coloured by the electrostatic potential (blue, +; red, −). One nebulin repeat is shown in yellow in the central cleft as an α-carbon trace. Conserved charges in the repeat are coloured in red (the negative ones at residues 3 and 19) and blue (the positive ones at residues 11+12, 20/21, 22/23 and 26−28). The nebulin repeat points down with its C-terminus. (**B**) Residues of actin in proximity to nebulin are shown in green on the actin monomer: residues 77−82 at the start of helix 2 and 111−115 in helix 3 on the N-terminal domain, 195−197 in the loop between helix 5 and helix 6, 223−231 in helix 7, 250−255 at the start of helix 8, 264−271 in the loop between helix 8 and helix 9 and 320−321 in the loop between helix 11 and β-sheet 16 in the C-terminal domain. This region has not yet been identified as the binding site of any other actin binding protein. The binding site for phalloidin is labelled with an arrow (Lorenz *et al.*, 1993).

with an average displacement of exactly the previously reported 1.6 Å. If the binding of nebulin to the thin filament occurs by means of a one to one interaction of a single repeat to an actin monomer, additional variability in the nebulin conformation is required to compensate for variations in the length of the repeats to the amount of approximately four residues.

Modelling of the nebulin binding site on actin

The nebulin binding site on actin should fullfil the following conditions. (i) The stoichiometry of the model should be consistent with the value of one nebulin repeat per actin monomer (Chen *et al.*, 1993a,b, as well as our results). This can only be achieved by arrangements having a 2-fold symmetry. (ii) The orientation of a single repeat should be almost parallel to the helix axis of the F-actin to match the length criterion of ~55 Å. (iii) It should have a higher number of negative charges as compared with positive ones.

A preliminary inspection of a plausible site on the actin filament as reconstructed from X-ray fibre diffraction measurements shows a cleft in the centre of the filament (Holmes *et al.*, 1991). This is the site where phalloidins, introduced to stabilize the filament, have been built in with a distance of ~27 Å from each other along the axis of the refined structure of F-actin (Lorenz *et al.*, 1993). A single nebulin repeat could be accommodated in this groove and would span the length of one actin monomer interacting with both domains of the actin monomer as indicated in Figure 8. Since both nebulin and phalloidin have been shown to alter actin polymerization properties (Wieland, 1977; Chen *et al.*, 1993b), they could achieve it by having a bridging function between the two strands of the actin filament. In this region, the number of negative charges is larger than positive ones and could match positive groups in nebulin. Since two grooves are present symmetrically on each of the two sides of the actin filament, this arrangement would be in agreement with the estimated stoichiometry. This model would be entirely consistent with both the proposed 'actin zipper' function of nebulin *in vivo* (Chen *et al.*, 1993b) and the recently described phalloidin binding experiments carried out on both skeletal and rabbit cardiac fibrils (Ao and Lehrer, 1994). According to these data, skeletal fibrils show both a fast and a slow uniform binding to phallotoxins which leads to complete uniform binding only after several hours. In

cardiac fibrils the slow process is absent and immediate uniform binding could be obtained. These results have been tentatively explained by the hypothesis that nebulin, present in skeletal fibrils but absent in rabbit cardiac fibrils, competes with phalloidins for actin sites.

In conclusion, our results confirm the original proposal that it is possible to obtain some insight into the folding, conformation and interaction properties of repeating modules of large filamentous proteins even when they are not globular. Finally we have demonstrated that some general properties of the complete protein can be obtained by extrapolating the results obtained from fragments of the intact molecule.

Materials and methods

Sample preparation

Five peptides were studied. The sequences of these peptides are given in Figure 1. Synthesis of the peptides was carried out by the EMBL peptide group using standard solid phase chemistry. The peptides were purified by HPLC on a preparative Vydac C18 column. The purity of the peptides was >99% each as checked by analytical HPLC. The completeness of the synthesis was checked by mass spectrometry. F-actin was prepared according to the standard method (Pardee and Spudich, 1982).

Binding assays

F-actin binding was assessed by cosedimentation in a Beckman TL100 benchtop ultracentrifuge using ~10 μM F-actin in a buffer of 20 mM Tris−HCl (pH 8.0), 100 mM NaCl, 2 mM $MgCl_2$, 1 mM ATP, 1 mM DTT and 1 mM EGTA (Winder and Walsh, 1990). Peptides were added at increasing concentrations from 2 to 350 μM. All peptides were clarified by centrifugation (386 000 *g*) before assay. All assay components were mixed together before addition of the actin. After a final mix, tubes were incubated for 10 min at room temperature before centrifugation. Equivalent volumes of supernatant and pellet fractions from each tube were run on Tris/Tricine gels (4% stack, 10% spacer, 16.5% resolving) (Schagger and von Jagow, 1987) with the following changes: Mini-Tris/Tricine gels (Matsudaira and Burgess, 1978) were run at 30 V constant voltage until the samples had completely entered the stacking gel (30 min). The voltage was then increased to 400 V for a further 45 min giving a total run time of 1 h and 15 min compared with 16 h for a full size gel (Schagger and von Jagow, 1987). Gels were fixed in 0.5% (v/v) glutaraldehyde solution before staining in Coomassie brilliant blue R250. Gels were diffusion destained in 10% (v/v) acetic acid 10% (v/v) methanol. Destained gels were scanned on a Molecular Dynamics 300A computing densitometer. Actin binding data were analysed by non-linear regression and fitted with the Hill equation (Enzfitter; Biosoft, Cambridge, UK) as described previously (Way *et al.*, 1992).

CD spectroscopy

CD spectra were measured on an ISA Jobin-Yvon CD6 spectrometer equipped with a Highscreen 386 IBM compatible computer. The temperature was kept constant with a Haake GH temperature unit. The temperature was checked inside the cuvette with a flexible temperature sensor. A round QS cuvette from HELMA having a 0.2 mm pathlength was used. Peptide concentrations were 50−150 μM. Detergents were added in the range of 0.01 to 5% w/v. Spectra were recorded at temperatures from 5 to 75°C with wavelengths between 190 and 250 nm. Twenty scans were added for each spectrum. Smoothing was performed by three point averaging after subtraction of the baseline spectrum. Data analysis was performed by the SNARF program (van Hoesel, 1992).

NMR spectroscopy

Spectra were recorded on Bruker AMX-500 and AMX-600 spectrometers and processed with UXNMR/P (Bruker AG, Karlsruhe). Spectra were analysed, plotted and integrated using AURELIA (Bruker AG, Karlsruhe). 2D spectra were recorded in phase sensitive mode (Marion and Wüthrich, 1983). Suppression of the water resonance was achieved by presaturation for 1.3 s. Total correlated (TOCSY) spectra (Bax and Davis, 1985) were measured using the MLEV-17 composite pulse cycle. The low-power 90° pulse had a length of 20 μs. Clean-TOCSY spectra (Griesinger *et al.*, 1988) were achieved by introducing a 'cleaning' delay of 18 μs. Mixing times of 60, 80 and 120 ms were used in the Clean-TOCSY experiments. NOESY spectra were recorded with mixing times of 40, 80, 100, 160 and 250 ms. Clean-TOCSY and NOESY experiments were recorded with 2048 data points in the acquisition domain and 512 data points in t_1. The t_1 domain was zero filled to 2048 points prior to Fourier transformation. The data were weighted by a Gaussian window with a line broadening of 20.0 Hz and Gaussian multiplication of 0.1 in t_2 and with a $\pi/8$ shifted sinebell in t_1. A baseline correction was performed in both dimensions using a polynomial.

Sample concentrations were 2 mM peptide and 300 mM d25-SDS in 90% H_2O/10% D_2O and in 100% D_2O solution for all peptides. In addition, for S2R5t a sample of 4 mM peptide with 600 mM d25-SDS was used. Perdeuterated SDS was obtained from Cambridge Isotope Laboratories (lot BL-120) at 98% purity. The pH value of all peptide samples containing SDS was 4.6. The measurements in water were performed with peptide concentrations of 2 and 6 mM. The pH value of these samples was 5.5. An additional experiment was done with S2R5 at a concentration of 6 mM in 50 mM phosphate buffer pH 7.0 (uncorrected pH meter reading). Spectra were acquired at 7, 17, 25, 32, 39 and 47°C. Temperature shift coefficients were calculated for the peptides in water at temperatures between 7 and 32°C.

D/H exchange measurements were performed in an extended way only for the S2R5t peptide as the good dispersion of the amide resonances allowed the identification of most amino acids from 1D spectra. The sample was prepared from 0.5 ml D_2O and 7.5 mg peptide immediately before the measurement was started. A total number of 1024 1D spectra was measured at a temperature of 7°C with a time step of 26 s adding 16 scans per spectrum. The first 300 data points of each time curve for the resolved amide resonances were fitted with single exponentials giving the rate constants.

Structure calculation

Structures were calculated for the peptides in SDS only. The cross peak volumes were calibrated using fixed distances between selected non-exchangeable protons. The NOESY spectrum with 100 ms mixing time in 90% water/10% D_2O was chosen for distance calculation as it was in the linear range of the build-up curve. From a total number of 362 NOEs in S2R5, 318 in R4R5 and 210 in S2R5t, roughly one third (112 for S2R5, 61 for S2R5t and 88 for R4R5) was caused by interactions of amino acids more than two residues apart. Distance geometry calculations were performed by the DIANA program (Güntert *et al.*, 1991) using the variable target function (Braun and Go, 1985). Pseudo-atoms with relative corrections on the observed distances were used in the case of methyl, methylene and aromatic protons where no stereo-specific assignment was possible. A set of 80 structures was calculated from which the 40 best ones were chosen for refinement using molecular dynamics. Simulated annealing calculations were performed with XPLOR (Brunger, 1993) using a standard protocol. The starting temperature was 3000 K. At this temperature 3000 steps (15 ps) of MD were calculated. The system was cooled down during 5000 steps (25 ps) to 300 K. At this temperature MD sampling was performed for 1000 steps (5 ps). The final structure for each run was obtained by averaging the last 1000 steps. The 12 best structures after this calculation were then averaged after r.m.s.-fit using a program (Diamond, 1992) kindly provided by Dr Robert Diamond based on a multiple structural alignment. For the peptide R4R5 separate r.m.s.-fits for the first and the second half had to be performed. The average structure was energy minimized for 1000 steps with XPLOR to remove geometric problems caused by the averaging to give the final structure for each peptide.

Model building

Modeling was done with the graphics program GRASP (Nicholls *et al.*, 1991). The vertical translation of the nebulin repeat was tentatively fixed by searching for the best match of complementary charged surfaces. For this purpose electrostatic potentials were calculated based on Poisson−Boltzmann statistics. Electrostatic potentials were mapped on a solvent accessible surface calculated with a test sphere of 1.6 Å. Both possible orientations of the nebulin repeat along the central cleft gave a similar quality of matching charges and surface contacts, so that we chose the recently determined orientation of nebulin with its C-terminus towards the Z-line (Wright *et al.*, 1993). The model in Figure 8 shows the best compromise obtained by maximizing the number of contacts between residues with opposite sign and minimizing the number of contacts between residues with the same sign.

Acknowledgements

The authors wish to thank D.Nalis and R.Jacob for synthesis of the peptides, M.Gautel, S.Labeit, T.Gibson, G.Stier, S.Georgatos and D.Suck for helpful discussions, P.Vibert for critical reading of the manuscript, B.Bullard for help with binding assays, M.Nilges for help with XPLOR and K.C.Holmes

for making the F-actin structure available to us and suggesting the central cleft in actin as a possible binding site for nebulin. S.J.W. is a recipient of a Molecular Dystrophy Group of Great Britain fellowship.

References

Ao,X. and Lehrer,S.S. (1994) In *Abstracts of the American Biophysical Meeting New Orleans.* in press.
Bax,A. and Davis,D.G. (1985) *J. Magn. Reson.*, **65**, 333–337.
Brahms,S. and Brahms,J. (1980) *J. Mol. Biol.*, **138**, 149–177.
Braun,W. and Go,N. (1985) *J. Mol. Biol.*, **186**, 611–626.
Brunger,A.T. (1993) *XPLOR, A System for X-ray Crystallography and NMR.* 3.1 edn. Yale University Press, New Haven.
Chen,M.J., Shih,C.L. and Wang,K. (1993a) *Biophys. J.*, **64**, A147.
Chen,M.-J.G., Shih,C.-L. and Wang,K. (1993b) *J. Biol. Chem.*, **268**, 20327–20334.
Diamond,R. (1992) *Protein Sci.*, **1**, 1279–1287.
Dyson,H.J., Rance,M., Houghten,R.A., Wright,P.E. and Lerner,R.A. (1988) *J. Mol. Biol.*, **201**, 161–200.
Dyson,H.J., Merutka,G., Waltho,J.P., Lerner,R.A. and Wright,P.E. (1992) *J. Mol. Biol.*, **226**, 795–817.
Ebashi,S., Endo,M. and Ohtsuki,I. (1969) *Q. Rev. Biophys.*, **2**, 351–357.
Griesinger,C., Otting,G., Wüthrich,K. and Ernst,R.R. (1988) *J. Am. Chem. Soc.*, **110**, 7870–7872.
Güntert,P., Braun,W. and Wüthrich,K. (1991) *J. Mol. Biol.*, **217**, 517–530.
Holmes,K.C., Popp,D., Gebhard,W. and Kabsch,W. (1991) *Nature*, **347**, 44–49.
Hu,D.H., Kimura,S. and Maruyama,K. (1989) *Biomed. Res.*, **10**, 165–171.
Jin,J.-P. and Wang,K. (1991) *J. Biol. Chem.*, **266**, 93–96.
Kabsch,W., Mannherz,H.G., Suck,D., Pai,E.F. and Holmes,K.C. (1990) *Nature*, **347**, 37–44.
Kruger,M., Wright,J. and Wang,K. (1991) *J. Cell Biol.*, **115**, 97–107.
Labeit,S., Gibson,T., Lakey,A., Leonhard,K., Zeviani,M., Knight,P., Wardale,J. and Trinick,J. (1991) *FEBS Lett.*, **282**, 313–316.
Lorenz,M., Popp,D. and Holmes,K.C. (1993) *J. Mol. Biol.*, **234**, 826–836.
Marion,D. and Wüthrich,K. (1983) *Biochem. Biophys. Res. Commun.*, **113**, 967–974.
Matsudaira,P.T. and Burgess,D.R. (1978) *Anal. Biochem.*, **87**, 386–391.
Motta,A., Pastore,A., Gould,N.A. and Castigliou Morelli,M.A. (1991) *Biochemistry*, **30**, 10444–10450.
Nicholls,A., Sharp,K.A. and Honig,B. (1991) *Proteins*, **11**, 281–296.
Pardee,J.D. and Spudich,J.A. (1982) *Methods Enzymol.*, **85**, 164–170.
Pastore,A. and Saudek,V. (1990) *J. Magn. Reson.*, **90**, 165–176.
Reynaud,J.A., Grivet,J.P. and Trudelle,Y. (1993) *Biochemistry*, **32**, 4997–5008.
Schagger,M. and von Jagow,G. (1987) *Anal. Biochem.*, **166**, 368–379.
Stryer,L. (1988) *Biochemistry.* W.H.Freeman and Company, New York.
Takahashi,H., Matuoka,S., Kato,S., Ohki,K. and Hatta,I. (1991) *Biochem. Biophys. Acta*, **1069**, 229–234.
van Hoesel,F.H.J. (1992) *SNARF processing software.*
Wang,K. (1982) *Methods Enzymol.*, **85**, 264–268.
Wang,K. and Williamson,C. (1980) *Proc. Natl Acad. Sci. USA*, **77**, 3254–3258.
Wang,K. and Wright,J. (1988) *J. Cell Biol.*, **107**, 2199–2212.
Wang,K., Knipfer,M., Huang,Q.Q., Hsu,L., van Heerden,A., Browning,K., Quian,X. and Stedman,H. (1990) *J. Cell Biol.*, **111**, 428A.
Way,M., Pope,B., Cross,R.A., Kendrick-Jones,J. and Weeds,A. (1992) *FEBS Lett.*, **301**, 243–246.
Wegner,A. (1979) *J. Mol. Biol.*, **131**, 839–847.
Wieland,T. (1977) *Naturwissenschaften*, **64**, 303–309.
Winder,S.J. and Walsh,M.P. (1990) *J. Biol. Chem.*, **265**, 10148–10155.
Wright,J., Huang,Q.-Q. and Wang,K. (1993) *J. Muscle Res. Cell. Motil.*, **14**, 476–483.
Wüthrich,K. (1986) *NMR of Proteins and Nucleic Acids.* Wiley-Interscience, New York.

Received on July 1, 1993; revised on January 31, 1994