Analysis of the three-α-helix motif in the spectrin superfamily of proteins

David A. D. Parry,* Tony W. Dixon* and Carolyn Cohen‡
*Department of Physics and Biophysics, Massey University, Palmerston North, New Zealand; and ‡Rosenstiel Basic Medical Sciences Research Center, Brandeis University, Waltham, Massachusetts 02254-9110 USA

ABSTRACT Members of the spectrin superfamily of proteins contain different numbers of homologous repeats arranged in tandem. Each of these consists of a three-α-helix motif, comprising two similarly and one oppositely directed α-helical segment joined by nonhelical linkers of characteristic length. The right-handed α-helices each display a heptad repeat in their amino acid sequences indicative of left-handed coiled-coil-like packing. We have calculated the potential number of inter-helix ionic interactions that specify the spatial arrangement of the helices in the motif in terms of both the handedness of helix connectivity (left or right) and the relative axial stagger between the three α-helices. All of the models examined were constrained to have optimal coiled-coil packing. For α-spectrin and α-actinin the results provide strong support for a left-handed connectivity of the three helices and axial repeat lengths of 5.05 and 6.24 nm, respectively. Furthermore, the axial staggers between homologous segments in the preferred models are identical. The insights provided into the topography of this widespread tertiary fold may prove of value to those concerned with the problem of de novo protein design.

INTRODUCTION

Spectrin, α-actinin and dystrophin comprise a class of structurally related single chain fibrous proteins, termed the "spectrin superfamily," with important mechanical roles in the cytoskeleton (Byers et al., 1989; Davison et al., 1989). On the basis of amino acid sequences, it has been shown that a major portion of each molecule contains multiple repeats of length 106 and ~122 and 109 residues, respectively, with a significant degree of interprotein homology. Using various predictive methods, a number of models have been advanced for the structure of these repeats. A bundle of three α-helices of similar length, connected by a flexible loop to the next bundle to form a rodlike domain, was first proposed for α-spectrin by Speicher and Marchesi (1984). Analysis of the most regular repeats of dystrophin by Koenig et al. (1988) led to a modification of this scheme, such that the motif consisted of one long helix (rather than two short helices) followed by a β-turn, followed by a short helix and another β-turn. A three-α-helix motif was then generated from this arrangement by each motif incorporating part of the long helix from the successive repeat. A similar scheme for spectrin and α-actinin was advanced by Dubreuil et al. (1989).

It has also been recognized that the α-helical regions in the spectrin superfamily contain a heptad substructure of the form $(a\text{-}b\text{-}c\text{-}d\text{-}e\text{-}f\text{-}g)_n$ where *a* and *d* are generally occupied by apolar residues (Cohen and Parry, 1986; Koenig et al., 1988; Cross et al., 1989). Such a pattern is characteristic of the α-fibrous class of proteins in which two or three right-handed α-helices wind around one another in a left-handed manner to generate a coiled-coil stabilized (in large part) by optimal knob-hole packing of the apolar residues in the *a* and *d* positions (Cohen and Parry, 1986, 1990). The heptad repeat is also a feature of globular proteins containing bundles of α-helices (Cohen and Parry, 1986, 1990). In a previous report (Parry and Cohen, 1991), we prepared consensus sequences for the most regular repeats of α-spectrin, α-actinin and dystrophin, and used these to delineate those pieces of structure predicted to be α-helical and which also displayed a heptad repeat (Fig. 1). Although no simple algorithm currently exists to predict helix ends we used the approach of Richardson and Richardson (1988) in our analysis to specify the most probable terminal residues. A consistent structural model for the members of the spectrin superfamily was then postulated where the α-helical segments pack together in a common way involving antiparallel coiled-coils, and differences in the chemical repeat length were localized to the linking regions (Fig. 2). The three stretches of α-helix found by this method (and designated *A*, *B* and *C*) are similar to those derived by Koenig et al. (1988) and Dubreuil et al. (1989), because the location of the α-helices and β-turns in the spectrin superfamily is consistently predicted by a number of methods. (We note, however, that a variety of other arrangements have also been put forward (e.g., Davison et al., 1989; Xu et al., 1990) but we believe that these are based on less plausible combinations of predictive methods.) Cross et al. (1989) have recently proposed a related

0006-3495/92/04/858/10 $2.00

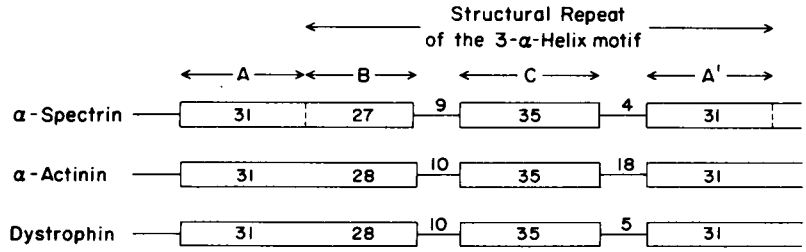


FIGURE 1 The predicted secondary structure of each repeat in the α-spectrin, α-actinin and dystrophin rod domains consists of three α-helices, designated *B*, *C* and *A'*. The number of residues in each of the helices is marked. On the basis of homology, the sequence of the *A* helix in α-spectrin and dystrophin is displaced by one heptad with respect to that in α-actinin (Parry and Cohen, 1991). The dotted line separating helices *A* and *B* (or *A'* and *B'*, et cetera) in α-spectrin indicates a discontinuity in the phasing of the heptad substructure. In the case of α-actinin and dystrophin helices *A* and *B* form a continuous piece of structure but the *A*/*B* nomenclature is maintained to allow comparisons between the various members of the spectrin superfamily. The number of residues in the link between helices *B* and *C* is almost constant (i.e., either 9 or 10 residues) whereas there is a considerable difference in the length of the link between helices *C* and *A'* (i.e., either 4 or 5 residues for α-spectrin and dystrophin, but 18 residues for α-actinin). Secondary structure prediction techniques suggest that each link contains a β-turn consistent with the 3-α-helix motif adopting an antiparallel conformation (Parry and Cohen, 1991).

model for dystrophin based on coiled-coil packing of both regular and irregular repeats. Our analysis of the regular repeats in terms of coiled-coil interactions has the advantage, however, of revealing additional features of packing in the bundle. Because our model was advanced a study of the expressed fragments of *Drosophila* α-spectrin has revealed a stable, highly α-helical 109 residue segment very closely similar in sequence to the motif we have predicted for the spectrin superfamily (Winograd et al., 1991). The NH_2 and COOH termini of this fragment correspond almost exactly to the motif between two A/B junctions (Fig. 2), thus providing the first experimental evidence in support of key aspects of our model.

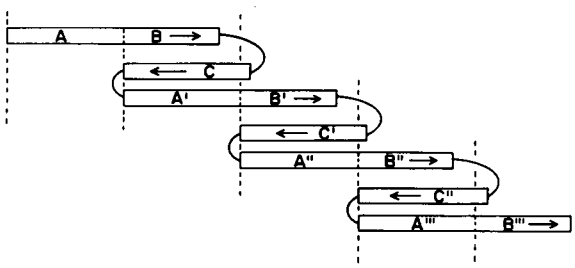


FIGURE 2 Model for the rod region of members of the spectrin superfamily of proteins. The simplest structural repeat contains helices *B*, *C*, and *A'*, where *A'* is the *A* helix of the next unit. This is a consequence of the prediction that helix *A*/*B* is ~9 nm in length, a value approximately twice that of the axial repeats in these proteins (~5 nm). It follows that helices *A* and *B* are incapable of interacting with one another whereas helices *B* and *A'* are ideally located spatially to do so. Note that the short linker has effectively zero length in a turn conformation. The dotted lines indicate the axial repeat length per motif. In the cases of α-actinin and dystrophin the repeat length will almost certainly vary from one motif to the next (due to variations in the length of the chemical repeat).

We have examined here the spatial relations of the three α-helices that appear to form the structural motif in the spectrin superfamily and have used the criterion of idealized coiled-coil packing to place limitations on the possible axial arrangements of the helices in this motif. In particular, we have deduced the handedness of the connectivity between consecutive α-helices, as well as the likely axial staggers between each of them (and, by inference, the axial repeat length of the motif). A preliminary study of this problem based on consensus sequences was described in our previous report (Parry and Cohen, 1991).

Another feature of potential importance with regard to helix packing was also recognized by Parry and Cohen (1991). They showed that while the length of the *A* helix in α-spectrin, α-actinin and dystrophin was the same, the homology between them was maximised only when the sequence of the *A* helix in α-spectrin and dystrophin was displaced by one heptad with respect to the corresponding sequence in α-actinin. Helix *B*, however, was directly homologous in all three proteins, as was helix *C*. These points are significant for the present study (see below).

The basis for our analysis is the computation of interhelix ionic interactions. We assume that the apolar core interactions will not be significantly different for left- and right-handed connectivities of the bundles, nor in the various axial alignments. For both parallel and antiparallel two-stranded coiled-coils, the residues likely to be involved in interhelix ionic interactions have been specified by McLachlan and Stewart (1975) to be those in the *e* and *g* positions in the heptad. Subsequent calculation of the numbers of such interactions for a

wide range of two-stranded α-fibrous proteins has, without exception, revealed that the chains prefer to be parallel and in axial register and that significant numbers of interhelix ionic interactions per heptad pair are made in all cases (see Conway and Parry, 1990 for a summary). These conclusions have been verified experimentally for myosin, tropomyosin and keratin intermediate filament proteins (Stewart, 1975; Phillips et al., 1986; Parry et al., 1985). More recently a so-called leucine zipper having the coiled-coil conformation has been solved to atomic resolution (O'Shea et al., 1991). In this structure interchain ionic interactions play a critical role in determining the relative orientation of the α-helices. Similarly, in an antiparallel coiled-coil-like helical "arm" described by Cusack et al. (1990) in the atomic structure of seryl tRNA synthetase, ion pairs play a key role in stabilizing the structure. Taken together these findings support the assumption made here that the specificity of helix alignment depends primarily on such ionic interactions; and that these interactions together with hydrophobic and hydrogen bonding stabilize the structures.

While it has been established that the specificity of the packing of the chains in two-stranded parallel coiled-coils lies with the inter-helix ionic interactions primarily involving residues in the *e* and *g* positions (Conway and Parry, 1990; Cohen and Parry, 1990), no comparable detailed studies have been reported for antiparallel coiled-coils. This analysis may therefore have implications for a wider class of α-helix-containing structures than the spectrin superfamily alone.

METHODS

The amino acid sequences used in these analyses were as follows: erythrocyte α-spectrin—human (Speicher and Marchesi, 1984), *Drosophila* (Dubreuil et al., 1989), chicken brain (Wasenius et al., 1989): α-actinin—chicken (Baron et al., 1987), *Drosophila* (Dubreuil et al., 1989), *Dictyostelium discoideum* (Noegel et al., 1987) : dystrophin—human (Koenig et al., 1985). In each case the *B*, *C* and *A'* helices were delineated by reference to the previously established consensus sequences, and the *a–g* positions were assigned accordingly (Parry and Cohen, 1991).

Those residues sterically able to participate in ionic interactions between the three α-helices (even with one of them arranged in an antiparallel manner with respect to the other two) are essentially identical to those previously formulated for two-stranded structures. Hence, the scheme adopted for specifying the residues involved in the ionic interactions at the three interfaces between the constituent α-helices was as follows (Fig. 3): (*a*) Between helices $A' \uparrow$ and $B \uparrow$ the interacting sites are of the type 2*e'*–1*g*, 2*a'*–1*g* and 1*e'*–1*d* for left-handed connectivity, and 1*g'*–2*e*, 1*g'*–2*a* and 1*d'*–1*e* for a right-handed arrangement. The terminology used here is that formulated by McLachlan and Stewart (1975). The designation of the terms "left-handed" and "right-handed" connectivity for the motif will be taken to mean that helices *B*, *C* and *A'* form a clockwise or anticlockwise set (respectively) when viewed from the NH_2-terminal end of helix *B* in a direction towards the COOH-terminal end of helix A' (Fig. 3). (*b*) Between helices $B \uparrow$ and $C \downarrow$ the interacting sites are of the type *g'*–*g*, *d'*–*g* and *g'*–*d*. For left-handed connectivity, these sites have the same spatial relationship as the 2*e'*–1*g*, 2*a'*–1*g* and 1*e'*–1*d* sites between helices $A \uparrow$ and B ↑. For right-handed connectivity the sites are of the type *e'*–*e*, *e'*–*a* and *a'*–e and, likewise, have the same spatial relationship as the 1*g'*–2*e*, 1*g'*–2*a* and 1*d'*–1*e* sites defined between helices $A \uparrow$ and $B \uparrow$. (*c*) Between helices $A \uparrow$ and $C \downarrow$ the interacting sites are of the type *e'*–*e*, *e'*–*a* and *a'*–*e*. For left-handed connectivity, these sites have the same spatial relationship as the 1*g'*–2*e*, 1*g'*–2*a* and 1*d'*–1*e* sites for helices $A \uparrow$ and $B \uparrow$. For right-handed connectivity the sites are of the type *g'*–*g*, *d'*–*g* and *g'*–*d* and, once more, have the same spatial relationship as 2*e'*–1*g*, 2*a'*–1*g* and 1*e'*–1*d* between helices $A \uparrow$ and $B \uparrow$.

The potential ionic interactions between oppositely charged residues in the sites listed above have thus been calculated computationally for each 3-α-helix motif in α-spectrin, α-actinin and dystrophin using the actual (not consensus) sequences. In all cases, the α-helices were packed together in a left-handed coiled-coil-like manner.

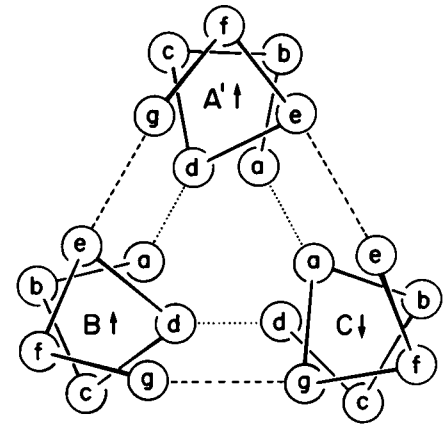


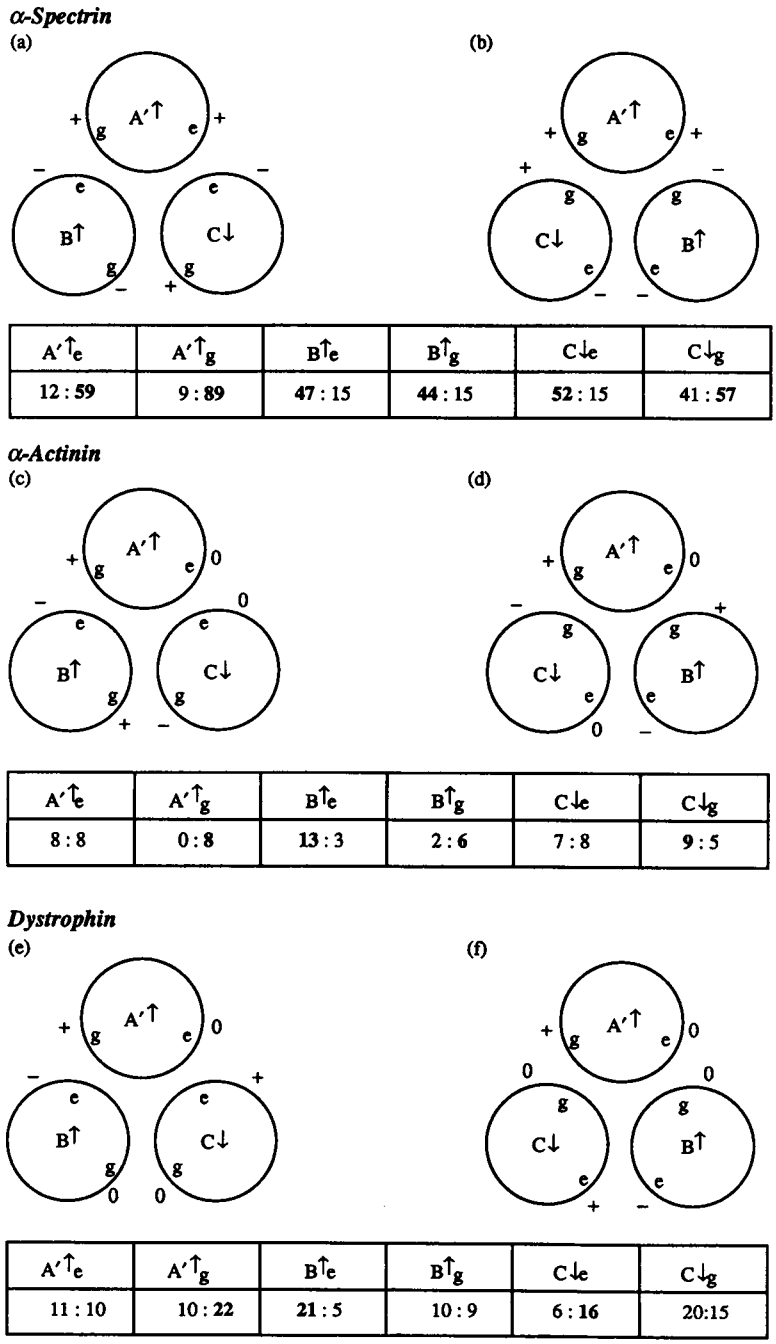
FIGURE 3 Plan view of the up-pointing α-helices ($A' \uparrow$ and $B \uparrow$) and the down-pointing α-helix ($C \downarrow$) that together constitute the 3-α-helix motif in the rod domains of the spectrin superfamily. This projection shows left-handed connectivity of *B*, *C*, and *A'* when viewed from the NH_2-terminal end of helix *B* (below the page) towards the COOH-terminal end of helix *A'* (above the page). Each of the sequences has a heptad substructure $(a\text{-}b\text{-}c\text{-}d\text{-}e\text{-}f\text{-}g)_n$ characteristic of α-fibrous coiled-coil structures, in which the *a* and *d* residues (usually apolar) pack closely along the axis of the motif and are shielded from the aqueous environment. Residues in the *e* and *g* positions are often charged, and are capable of forming stabilizing interhelix ionic interactions. These specify a particular spatial arrangement of the helices with regard to both the relative polarity of the helices and their relative axial stagger.

RESULTS AND DISCUSSION

Before carrying out detailed calculations of the numbers of interhelix ionic interactions that occur within the motif we first calculated the heptad distributions for the individual *B*, *C* and *A'* α-helices in α-spectrin, α-actinin, and dystrophin separately. Of particular interest were the number of acidic, basic and apolar residues in the *e* and *g* positions, because these sites are of special importance at the helix–helix interfaces. It immediately

became apparent that a left-handed connectivity was likely to be strongly favored for α-spectrin (Fig. 4 *a*). The *e* and *g* positions in *A'* were both highly basic, and in *B* both were highly acidic. In helix *C*, however, *e* was strongly acidic and *g* was basic. In principle, therefore, optimal charge pairing will occur for left-handed (Fig. 4 *a*) rather than right-handed connectivity (Fig. 4 *b*).

For α-actinin, the distributions showed some striking differences, albeit with a smaller database. For example the *e* position in helix *A'* was neutral (and not strongly



A'↑e	A'↑g	B↑e	B↑g	C↓e	C↓g
12 : 59	9 : 89	47 : 15	44 : 15	52 : 15	41 : 57

A'↑e	A'↑g	B↑e	B↑g	C↓e	C↓g
8 : 8	0 : 8	13 : 3	2 : 6	7 : 8	9 : 5

A'↑e	A'↑g	B↑e	B↑g	C↓e	C↓g
11 : 10	10 : 22	21 : 5	10 : 9	6 : 16	20:15

FIGURE 4 Plan projection of the 3-α-helix motif in α-spectrin (*a* and *b*), α-actinin (*c* and *d*) and dystrophin (*e* and *f*). Models with left-handed connectivity are shown in *a*, *c*, and *e* and those with right-handed connectivity in *b*, *d*, and *f*. The nature of the *e* and *g* position are indicated by +, −, and 0 respectively for basic, acidic and neutral residues. The data on which these designations are made are given below in the form *a*:*b*, where *a* and *b* are the numbers of acidic (*D*, *E*) and basic (*K*, *R*) residues respectively in the sequences used in this work.

basic), *g* in helix B was basic (and not strongly acidic) and e and g were neutral and acidic in helix *C* (as compared to strongly acidic and basic respectively). Reference to Fig. 4, *c* and *d*, however, shows that optimal interhelix ionic interactions were still most likely to occur for left-handed rather than right-handed connectivity. In other words, there have been complementary charge changes in some of the *e* and *g* positions of the different helices in α-spectrin and α-actinin, even though these are homologous proteins.

Similar considerations for dystrophin using sequence data from the 10 regular repeats (and some of the incomplete repeats) are less conclusive and leave open the possibility of either left- or right-handed connectivity (Fig. 4 *e* and *f*). The charge characteristics of the *e* and *g* positions in the *A'* and *B* helices in dystrophin are similar to those in α-actinin (in particular, the *e* and *g* positions of helix *A'* and the *e* position of helix *B*) but differ significantly for helix *C*. The only features common to the *e* and *g* positions in all three proteins in the spectrin superfamily are the basic character of the *g* position in helix *A'* and the acidic character of the *e* position in helix *B*. As a consequence of this charge conservation we believe that left-handed connectivity is likely to be favored in all cases; the *A'–B* interaction may in fact represent the initial step in the formation of the three-α-helix motif (see Fig. 4, *a*, *c*, and *e*).

Also of structural relevance is the observation that the overall percentage of apolar residues in positions *e* and *g* (~30%) is much higher than that in either two- or three-stranded α-fibrous proteins (~12 and 20% respectively), thus indicating the probability of a proportionally greater role for the nonspecific but stabilizing apolar interactions in the 3-α-helix motif of the spectrin superfamily than in the two- and three-stranded rope structures. Note that in the case of the 4-α-helix motif the percentage of apolar residues in positions *e* and *g* is even higher (~35%); here dipole interactions may play a role in the antiparallel alignment of helices, so that the motif has a negligible net dipole moment (Cohen and Parry, 1990). In contrast, the 3-α-helix motif has a net dipole moment, so that dipolar interactions would probably be of less significance in the assembly of the motif. Irrespective of these points, however, the specificity of the packing of the α-helices in the 3-α-helix motif can only be determined by specific interactions, i.e., interhelix ionic interactions, as in the two- and three-stranded α-fibrous proteins.

It is interesting to observe that the numbers of structural repeats forming the rod domains in α-spectrin and α-actinin will be one less than the number of chemical repeats. This results from the partial structural repeat comprising the *A* helix at the NH_2-terminal end of the rod, and the *B* and *C* helices at the COOH-terminal end. However, when a dimer forms, the *A* helix of molecule 1 could pair with the *B* and *C* helices of molecule 2 (and vice versa). This arrangement would generate two more 3-α-helix motifs than in a single molecule (or one more than the number of chemical repeats). The precise geometry of these extra motifs would nonetheless be different from the others contributing to the rod.

It is worth commenting at this point on the range of possible axial repeat lengths in α-spectrin, α-actinin and dystrophin, as deduced from electron microscope observations. The α/β spectrin dimer has been measured as 98–101 nm (Shotton et al., 1979) and 95–125 nm in length (Matsudaira, 1991) including the NH_2-terminal domain of the α-subunit. Approximately 20 3-α-helix motifs constitute the rod domain, thus giving an estimate of the axial repeat length of the motif in the range 4–6 nm, with the most probable value close to 5 nm. Similarly the α-actinin dimer is reported to be ~30–40 nm long (Davison et al., 1989), including both the NH_2- and COOH-terminal domains. In this case five 3-α-helix motifs (corresponding to four chemical repeats) form the rod region (see earlier). Hence, the axial repeat length per motif will lie within the range 4–7 nm. The most probable axial repeat length would seem to be ~5 or 6 nm. The lengths of isolated dystrophin molecules have been measured as 175 ± 15 nm by Pons et al. (1990) and include a contribution from both the NH_2- and COOH-terminal domains. The extent of the irregular and hinge regions in the rod domain makes the determination of the axial repeat length corresponding to the regular 3-α-helix motifs difficult, but most likely it will be ~5 or 6 nm.

Nine topologically distinct models have been selected for detailed investigation in this study. All have at least 75% of the maximal overlap possible between the three constituent α-helices, *B*, *C*, and *A'* (Fig. 5). This constraint is consistent with the presence of continuous heptad substructure in an α-helical segment, which implies that knob-hole packing of apolar residues will occur over most if not all of their lengths. We note that the degree of overlap and interaction between consecutive motifs will differ significantly. However, all of our models satisfy the constraints of ideal coiled-coil packing within the motif. Consequently the axial repeat length of each of the models can be determined. The values for the models studied are close to 4, 5, or 6 nm (see Fig. 6 for more precise values), and cover the likely range determined from the molecular length measurements previously noted.

In the case of α-spectrin the maximum number of stabilizing ionic interactions occurs for the model with left-handed connectivity illustrated in Figs. 5 *d*, 6 *d*, and 7, and has (on average) 2.51 interactions per motif

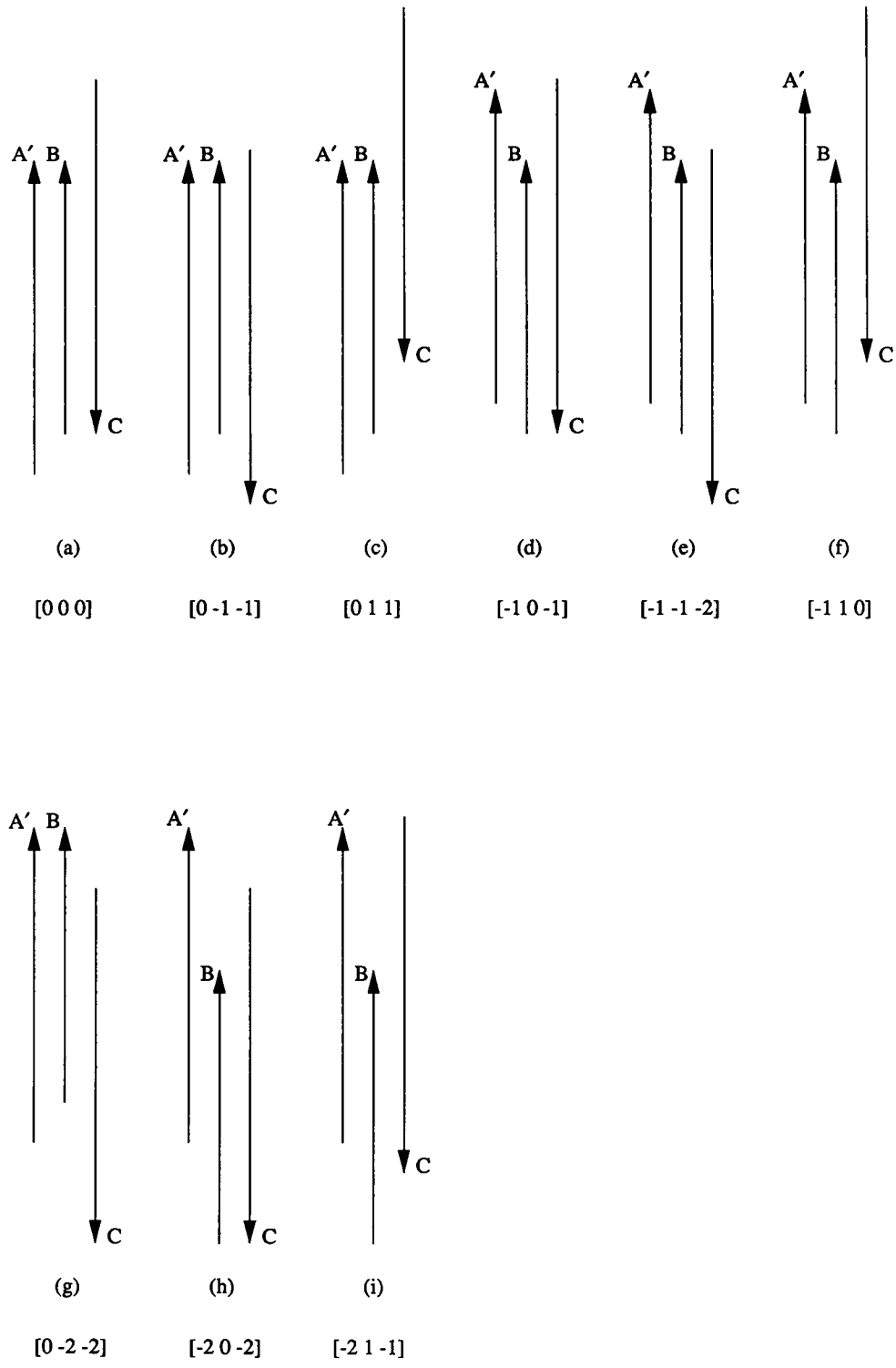


FIGURE 5 Nine topological arrangements of helices B, C, and A' were studied and are illustrated in a to i. Note that the arrow heads correspond to the COOH-termini of the helices. In all cases, the degree of overlap between the α-helices is at least 75% of the maximum possible. The terminology [lmn] refers to the relative axial staggers between the helices measured in multiples of heptads i.e., the stagger of B relative to A' is l heptads upwards, that of C relative to B is m heptads upwards and that of C relative to A' is n heptads upwards. All staggers are consistent with optimal coiled-coil packing and are measured relative to that of the model illustrated in a. As the sequence of the A helix in α-spectrin and dystrophin is displaced by one heptad with respect to that in α-actinin it follows that a structure for spectrin or dystrophin specified by [$l\ m\ n$] and a structure for α-actinin defined by [$l - 1\ m\ n - 1$] will result in the interaction of homologous segments in helices A', B, and C.

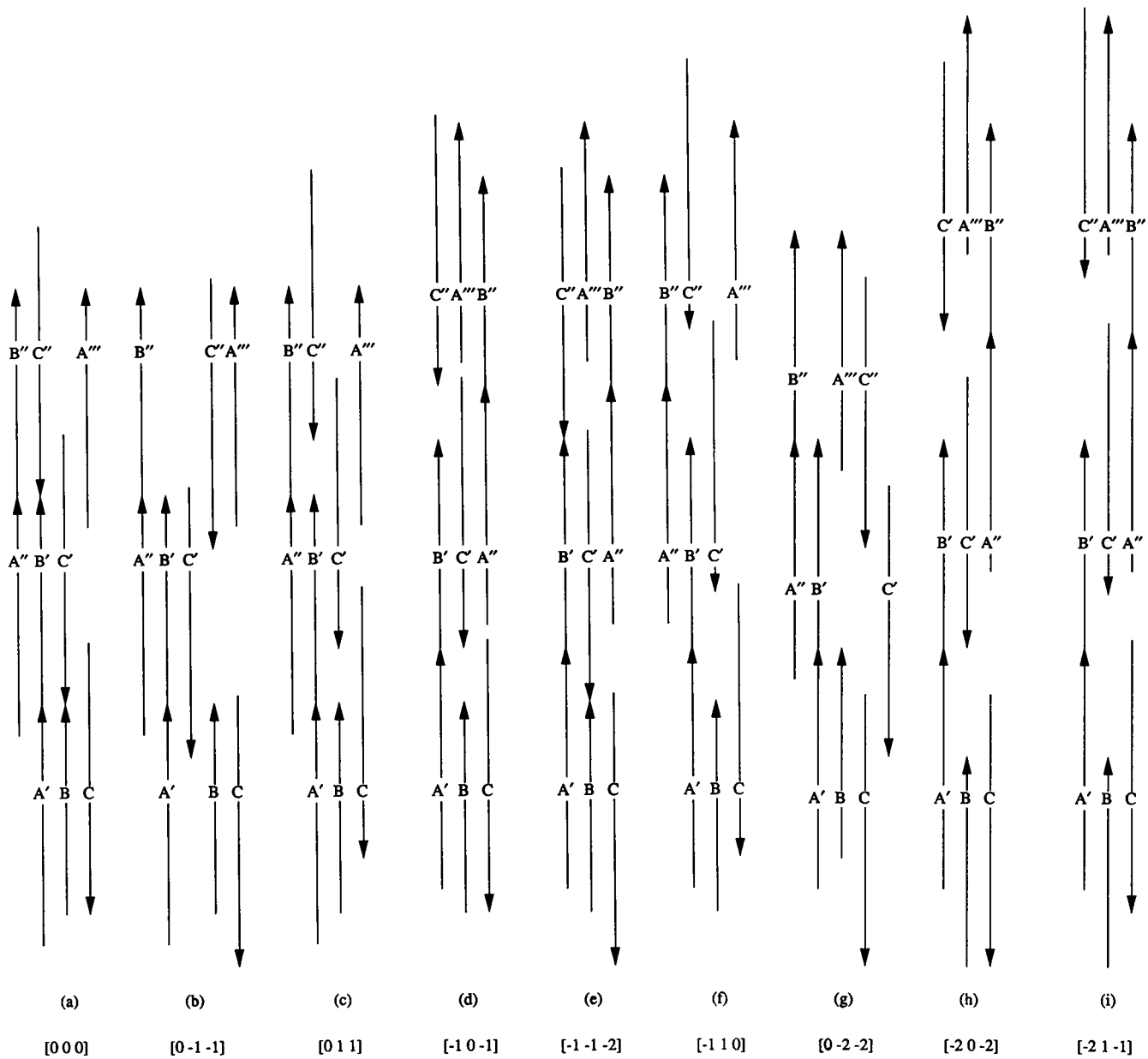


FIGURE 6 Topological arrangements of three consecutive 3-α-helix motifs for each of the models shown in Fig. 5. The number of strands at different axial positions within the repeat varies between one and five, depending on the model. The axial repeat length for each model is readily determined: for α-spectrin the axial repeat length for *a*, *b*, *c*, and *g* is 4.01 nm, that for *d*, *e*, and *f* is 5.05 nm and that for *h* and i is 6.09 nm. For α-actinin and dystrophin each of these values is increased by 0.15 nm.

(Table 1), i.e., ~0.21 ionic interactions per heptad pair (cf 0.23–0.77 ionic interactions per heptad pair for two-stranded α-fibrous proteins). Furthermore, the difference between the models with left-handed and right-handed connectivity is the greatest for all the α-spectrins for this particular topological arrangement. This point may be significant in that it provides a measure of the net number of interactions that favor one hand of connectivity over the other. This result is consistent with our simple analysis (described above) on the distribution of charged residues in the *e* and *g* positions. It is noteworthy that left-handed connectivity is favored for all models. The predicted axial repeat length for the proposed model for α-spectrin is 5.05 nm and is in close agreement with the experimental estimate. Furthermore, the model illustrated in Fig. 5 *d* corresponds

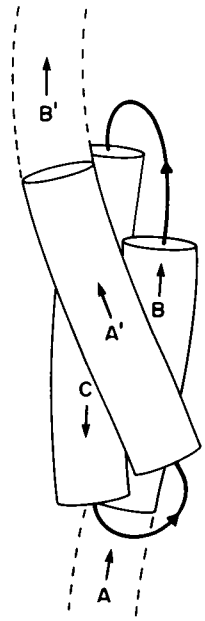


FIGURE 7 Representation of the spatial arrangement of *B*, *C*, and *A'* helices that form the 3-α-helix motif in the preferred model for α-spectrin. The pitch length of the left-handed coiled-coil-like structure is unknown. However the handedness of the connectivity of the motif is left-handed as based on the number of potential interhelix ionic interactions (see text).

exactly to the expressed fragment of *Drosophila* α-spectrin that is predicted by Winograd et al. (1991) to correspond to the structural unit.

For α-actinin, the model with the highest number of ionic interactions has left-handed connectivity and is illustrated in Figs. 5 *h* and 6 *h*. Once again, the handedness is consistent with our previous analysis of the *e* and *g* characters. The axial repeat length of the model is 6.24 nm, a value lying within the range of observed values. The number of ionic interactions per heptad pair is about 0.20, or 2.13 per motif (Table 1) and differs little from that calculated for α-spectrin. Although the models proposed for α-spectrin and α-actinin differ topologically (Fig. 5 *d* and 5 *h* respectively) they nonetheless incorporate identical packing environments for homologous segments from the three interacting α-helices. In other words, the models for α-spectrin and α-actinin, designated [−1 0 −1] and [−2 0 −2], respectively, are very closely related because the sequence of the *A* helix in α-spectrin is displaced by one heptad with respect to that in α-actinin (Parry and Cohen, 1991). It is interesting that the packing of homologous sequences in the three α-helical segments should be maintained for both proteins even though the repeat length per motif must necessarily differ as a consequence. Closely related conformations for two proteins with so high a sequence homology would certainly be expected on the basis of studies of other protein superfamilies. The models for α-spectrin and α-actinin result in the two conserved tryptophans in the motif (i.e. one in helix *A'* and one in helix *B*) occurring in near axial register. Whether or not this has structural or functional significance has yet to be established. We should also point out that the number of residues in the links between helices *C* and *A'* and helices *B* and *C* are compatible with the models proposed for both α-spectrin and a-actinin (Table 2).

The results for dystrophin are rather problematic. Although dystrophin contains about 25 3-α-helix motifs, 15 of these lack some part(s) of the three constituent α-helices as defined in Fig. 1 (see also Parry and Cohen, 1991). These partial motifs do not allow us to undertake an analysis directly comparable to that reported here for α-spectrin and α-actinin. Ten motifs do, however, contain full length *A'*, *B*, and *C* helices and only these have been used for the dystrophin calculations. The results show that the maximum number of ionic interactions (0.9–1.0 per motif; Table 1) occur for two left-handed and two right-handed models. The number of ionic interactions per heptad pair is thus less than half of that found for α-spectrin and α-actinin. While our methods do not produce a clearcut result for dystrophin, we note that one of the highest scoring models (Fig. 5 *d*) is the same as that found for α-spectrin, and has an axial repeat length (5.20 nm) compatible with that estimated from the electron microscope measurements. Furthermore the connectivity is also left-handed (i.e., the same as that determined for both α-spectrin and α-actinin). We note also that it is highly probable that the three homologous proteins that comprise the spectrin superfamily will have similar though not necessarily identical structures (as already shown for α-spectrin and α-actinin). Certainly proteins with 25% or less homology have been shown to be conformationally indistinguishable (Bowie et al. 1991). Clearly more sequence data will be required before the structure of the regular 3-α-helix motifs in dystrophin can be established with any degree of confidence. A modified analysis of the less regular motifs remains possible, however, provided that less stringent ground rules are permitted.

In summary, we have applied a simplified analysis of amino acid sequence data to analyze the conformation of the homologous motifs in the rod domain of the spectrin superfamily. Using optimisation of coiled-coil packing and maximization of interhelix ionic interactions, we have predicted that the motif has left-handed connectivity for α-spectrin and α-actinin, that the axial repeat lengths of the motif in the rod domain are about 5.05 and 6.24 nm, respectively, and that structural interactions are maintained between the favored helix arrangements in these two proteins. These data may

TABLE 1 Inter-helix ionic interactions stabilizing the 3-α-helix motif

α-spectrin

Model[1]	Human[2] (5)			Drosophila[2] (20)			Chicken Brain[2] (16)			Average[2] (41)		
	LH	RH	Diff	LH	RH	Diff	LH	RH	Diff	LH	RH	Diff
(*a*) [0 0 0]	2.60	0.60	2.00	2.05	0.00	2.05	2.13	0.25	1.88	2.15	0.17	1.98
(*b*) [0−1−1]	1.00	−0.80	1.80	0.35	−2.10	2.45	0.50	−2.00	2.50	0.49	−1.90	2.39
(*c*) [0 1 1]	0.80	−0.20	1.00	0.50	−0.50	1.00	1.06	0.25	0.81	0.76	−0.17	0.93
(*d*) [−1 0−1]	2.40	−2.00	4.40	2.30	−1.90	4.20	2.81	−2.00	4.81	2.51	−1.95	4.46
(*e*) [−1−1−2]	0.60	−0.20	0.80	−0.15	−0.35	0.20	0.31	−0.63	0.94	0.12	−0.44	0.56
(*f*) [−1 1 0]	0.20	0.20	0.00	0.45	0.15	0.30	1.31	−0.13	1.44	0.76	0.05	0.71
(*g*) [0−2−2]	3.00	0.80	2.20	1.90	−0.15	2.05	2.38	0.31	2.06	2.22	0.15	2.07
(*h*) [−2 0−2]	2.80	−0.20	3.00	1.60	0.00	1.60	1.81	0.63	1.19	1.83	0.22	1.61
(*i*) [−2 1−1]	0.80	−1.20	2.00	0.50	−1.60	2.10	1.19	−1.13	2.31	0.80	−1.37	2.17

α-actinin

Model[1]	Chicken[2] (3)			Drosophila[2] (2)			Dictyostelium[2] (3)			Average[2] (8)		
	LH	RH	Diff	LH	RH	Diff	LH	RH	Diff	LH	RH	Diff
(*a*) [0 0 0]	0.67	1.00	−0.33	2.50	1.00	1.50	0.67	−0.33	1.00	1.13	0.50	0.63
(*b*) [0−1−1]	−0.33	0.00	−0.33	−0.50	0.00	−0.50	0.33	0.67	−0.33	−0.13	0.25	−0.38
(*c*) [0 1 1]	0.33	0.33	0.00	0.00	0.00	0.00	−1.00	0.00	−1.00	−0.25	0.13	−0.38
(*d*) [−1 0−1]	−0.33	−0.33	0.00	−1.00	1.00	−2.00	1.67	0.67	1.00	0.25	0.38	−0.13
(*e*) [−1−1−2]	2.00	−1.33	3.33	2.00	−2.50	4.50	1.33	0.00	1.33	1.75	−1.13	2.88
(*f*) [−1 1 0]	−0.33	0.00	−0.33	0.50	−0.50	1.00	0.33	−0.33	0.67	0.13	−0.25	0.38
(*g*) [0−2−2]	2.00	0.00	2.00	2.50	0.00	2.50	0.00	−1.00	1.00	1.38	−0.38	1.75
(*h*) [−2 0−2]	2.33	−0.33	2.67	2.00	−0.50	2.50	2.00	0.33	1.67	2.13	−0.13	2.25
(*i*) [−2 1−1]	−1.00	0.00	−1.00	−2.50	0.50	−3.00	0.67	1.00	−0.33	−0.75	0.50	−1.25

dystrophin

Model[1]	Human[2] (10)		
	LH	RH	Diff
(*a*) [0 0 0]	0.90	0.60	0.30
(*b*) [0−1−1]	0.40	1.00	−0.60
(*c*) [0 1 1]	0.70	0.00	0.70
(*d*) [−1 0−1]	0.90	0.70	0.20
(*e*) [−1−1−2]	−0.40	0.70	−1.10
(*f*) [−1 1 0]	0.20	0.90	−0.70
(*g*) [0−2−2]	0.60	0.20	0.40
(*h*) [−2 0−2]	0.40	−0.20	0.60
(*i*) [−2 1−1]	0.50	0.40	0.10

[1]The models correspond to those illustrated in Fig. 4. The symbol [*l m n*] refers to the relative axial staggers (measured in terms of the number of heptads) between $A' \uparrow$ and $B \uparrow$, $B \uparrow$ and $C \downarrow$, and $A' \uparrow$ and $C \uparrow$, respectively. [2]The number of *B*, *C*, and *A'* motifs used in the calculations is given in brackets. LH refers to a left-handed arrangement of *A'*, *B*, and *C* helices when viewed from the NH_2-terminal end of helix *B* towards the COOH-terminal end of helix *A'*. RH represents the right-handed arrangement and Diff is the difference between the number of interactions for left- and right-handed arrangements. This value is a measure of the distinction between the two possible hands of helix connectivity. Positive and negative values indicate a preference for left-handed and right-handed connectivities respectively. The segments used in these calculations for the *B*, *C*, and *A'* helices are as follows: human α-spectrin, 46–72, 81–115 and 120–150; 151–177, 187–221 and 226–255; 257–283, 293–327 and 332–362; 363–389, 399–433 and 438–468; 469–495, 505–539 and 544–574; *Drosophila* α-spectrin, 46–72, 81–115 and 120–150; 151–177, 187–221 and 226–256; 257–283, 293–327 and 332–362; 363–389, 399–433 and 438–468; 469–495, 505–539 and 544–574; 575–601, 610–644 and 649–679; 680–706, 716–750 and 755–785; 786–812, 822–856 and 861–891; 892–918, 928–962 and 967–997; 1180–1206, 1216–1250 and 1255–1285; 1286–1312, 1322–1356 and 1361–1391; 1392–1418, 1428–1462 and 1467–1497; 1498–1524, 1533–1567 and 1574–1604; 1605–1631, 1641–1675 and 1680–1710; 1711–1737, 1747–1781 and 1786–1816; 1817–1843, 1853–1887 and 1892–1922, 1923–1949, 1959–1993 and 1999–2029; 2038–2064, 2074–2108 and 2113–2143; 2152–2178, 2188–2222 and 2227–2257; 2258–2284, 2294–2328 and 2333–2363; chicken brain α-spectrin, 43–69, 78–112 and 117–147; 148–174, 184–218 and 223–253; 254–280, 290–324 and 329–359; 360–386, 396–430 and 435–465; 466–492, 502–536 and 541–571; 572–598, 607–641 and 646–676; 677–703, 713–747 and 752–782; 783–809, 819–853 and 858–888; 1232–1258, 1268–1302 and 1307–1337; 1338–1364, 1374–1408 and 1413–1443; 1444–1470, 1480–1514 and 1519–1549; 1550–1576, 1587–1621 and 1631–1661; 1662–1688, 1698–1732 and 1737–1767; 1768–1794, 1804–1838 and 1843–1873; 1874–1900, 1910–1944 and 1949–1979; 1980–2006, 2016–2050 and 2056–2086; chicken α-actinin, 272–299, 309–343 and 361–391; 392–419, 430–464 and 476–506; 507–534, 544–578 and 597–627: *Drosophila* α-actinin, 36–63, 74–108 and 120–150; 151–178, 188–222 and 241–271: *Dictyostelium* α-actinin, 265–292, 302–336 and 354–384; 385–412, 422–456 and 470–500; 501–528, 538–572 and 584–614: Human dystrophin, 337–364, 375–409 and 417–447; 557–584, 596–630 and 637–667; 828–855, 866–900 and 904–934; 1046–1073, 1084–1118 and 1124–1154; 1155–1182, 1193–1227 and 1233–1263; 1569–1596, 1607–1641 and 1646–1676; 2209–2236, 2246–2280 and 2288–2318; 2578–2605, 2616–2650 and 2656–2686; 2687–2714, 2731–2765 and 2772–2802; 2803–2830, 2841–2875 and 2901–2931.

TABLE 2 **Length of links between helices**

	$C \rightarrow A'$		$B \rightarrow C$	
	Model[1]	Observed[2]	Model[1]	Observed[2]
α-spectrin [−1 0−1][3]	3	4	8	9
α-actinin [−2−0−2][3]	10	18	8	10

[1]Measured in terms of the number of residues in an α-helical conformation. [2]Measured as the number of residues in the consensus sequence. The conformation of these links is predicted to contain extended structure and a β-turn. [3]The models for α-spectrin [−1 0−1] and α-actinin [−2 0−2] both have left-handed connectivity but are topologically distinct with axial repeat lengths of 5.05 and 6.24 nm respectively. Nonetheless, the interactions between the homologous α-helices that constitute the 3-α-helix motif are precisely the same. This is a consequence of the *A* helix (and only the *A* helix) in spectrin having a relative sequence displacement of one heptad with respect to that in α-actinin (Parry and Cohen, 1991). See the text for details.

provide new and useful insights to those involved in the problem of protein design.

This work was supported by grants (to Carolyn Cohen) from the National Institutes of Health (AR17346), National Science Foundation (DMB90-04746) and the Muscular Dystrophy Association.

Note added in proof: O'Shea, Rutkowski and Kim (1992, *Cell,* in press) have recently undertaken a detailed study of Fos/Jun heterodimer formation. They provided convincing evidence that specificity of dimer assembly was mediated by the charged residues in the *e* and *g* positions of the leucine zipper structure. These results support the assumptions underlying the analysis presented here.

Received for publication 22 July 1991 and in final form 4 December 1991.

REFERENCES

Baron, M. D., M. D. Davison, P. Jones, and D. R. Critchley. 1987. The sequence of chick alpha-actinin reveals homologies to spectrin and calmodulin. *J. Biol. Chem.* 262:17623–17629.

Bowie, J. U., R. Lüthy, and D. Eisenberg. 1991. A method to identify protein sequences that fold into a known three-dimensional structure. *Science (Wash. DC)* 253:164–170.

Byers, T. J., A. Husain-Chisti, R. R. Dubreuil, D. Branton, and L. S. B. Goldstein. 1989. Sequence similarity of the amino-terminal domain of *Drosophila* beta-spectrin to alpha-actinin and dystrophin. *J. Cell Biol.* 109:1633–1641.

Cohen, C., and D. A. D. Parry. 1986. α-helical coiled coils; a widespread motif in proteins. *Trends Biochem. Sci.* 11:245–248.

Cohen, C., and D. A. D. Parry. 1990. α-helical coiled coils and bundles: how to design an α-helical protein. *Proteins Struct. Funct. Genet.* 7:1–15.

Conway, J. F., and D. A. D. Parry. 1990. Structural features in the heptad substructure and longer range repeats of two-stranded α-fibrous proteins. *Int. J. Biol. Macromol.* 12:328–334.

Cross, R. A., M. Stewart, and J. Kendrick-Jones. 1989. Structural predictions for the central domain of dystrophin. *FEBS (Fed. Eur. Biochem. Soc.) Lett.* 262:87–92.

Cusack, S., C. Berthet-Colominas, M. Hartlein, N. Nassar, and R. Leberman. 1990. A second class of synthetase structure revealed by x-ray analysis of *Escherichia coli* seryl-tRNA synthetase at 2.5 Å. *Nature (Lond.).* 347:249–255.

Davison, M. D., M. D. Baron, D. R. Critchley, and J. C. Wootton. 1989. Structural analysis of homologous repeated domains in α-actinin and spectrin. *Int. J. Biol. Macromol.* 11:81–90.

Dubreuil, R. R., T. J. Byers, A. L. Sillman, D. Bar-Zvi, L. S. B. Goldstein, and D. Branton. 1989. The complete sequence of *Drosophila* alpha spectrin: conservation of structural domains between alpha spectrins and alpha actinin. *J. Cell Biol.* 109:2197–2205.

Koenig, M., A. P. Monaco, and L. M. Kunkel. 1988. The complete sequence of dystrophin predicts a rod-shaped cytoskeletal protein. *Cell.* 54:219–228.

Matsudaira, P. 1991. Modular organization of actin crosslinking proteins. *Trends Biochem. Sci.* 16:87–92.

McLachlan, A. D., and M. Stewart. 1975. Tropomyosin coiled-coil interactions: evidence for an unstaggered structure. *J. Mol. Biol.* 98:293–304.

Noegel, A., W. Witke, and M. Schleicher. 1987. Calcium sensitive non-muscle alpha actinin contains EF hand structures and highly conserved regions. *FEBS (Fed. Eur. Biochem. Soc.) Lett.* 221:391–396.

O'Shea, E. K., J. D. Klemm, P. S. Kim, and T. Alber. 1991. X-ray structure of the GCN4 leucine zipper, a two-stranded, parallel coiled-coil. *Science (Wash. DC).* 254:539–544.

Parry, D. A. D., and C. Cohen. 1991. Structure of the spectrin superfamily: a three-α-helix motif. *In* The Living Cell in Four Dimensions. G. Paillotin, editor. AIP Conference Proceedings 226 American Institute of Physics, New York. 367–377.

Parry, D. A. D., A. C. Steven, and P. M. Steinert. 1985. The coiled-coil molecules of intermediate filaments consist of two parallel chains in exact axial register. *Biochem. Biophys. Res. Commun.* 127:1012–1018.

Phillips, G. N., J. P. Fillers, and C. Cohen. 1986. Tropomyosin crystal structure and muscle regulation. *J. Mol. Biol.* 192:111–131.

Pons, F., N. Augier, R. Heilig, J. Léger, D. Mornet, and J. J. Léger. 1990. Isolated dystrophin molecules as seen by electron microscopy. *Proc. Natl. Acad. Sci. USA.* 87:7851–7855.

Richardson, J. S., and D. C. Richardson. 1988. Amino acid preferences for specific locations at the ends of α helices. *Science (Wash. DC).* 240:1648–1652.

Shotton, D. M., B. E. Burke, and D. Branton. 1979. The molecular structure of human erythrocyte spectrin: biophysical and electron microscopic studies. *J. Mol. Biol.* 131:303–329.

Speicher, D. W., and V. T. Marchesi. 1984. Erythrocyte spectrin is comprised of many homologous triple helical segments. *Nature (Lond.).* 311:177–180.

Stewart, M. 1975. Tropomyosin: evidence for no stagger between chains. *FEBS (Fed. Eur. Biochem. Soc.) Lett.* 53:5–7.

Wasenius, V.-M., M. Saraste, P. Salvén, M. Erämaa, L. Holm, and V.-P. Lehto. 1989. Primary structure of the brain alpha-spectrin. *J. Cell Biol.* 108:79–93.

Winograd, E., D. Hume, and D. Branton. 1991. Phasing the conformational unit of spectrin. *Proc. Natl. Acad. Sci. USA.* 88:10788–10791.

Xu, Y., M. Prabhakaran, M. E. Johnson, and L. W. M. Fung. 1990. Secondary structure prediction for the spectrin 106-amino acid segment, and a proposed model for tertiary structure. *J. Biomolec. Struct. Dynamics.* 8:55–62.