Proteins with Spectrin Motifs Which Do Not Belong to the Spectrin-α-Actinin-Dystrophin Family

Maciej Kukuła[a], Beata Hanus-Lorenz[a,§], Ewa Bok[a], Jacek Leluk[b], and Aleksander F. Sikorski[a,c,*]

[a] Institute of Biochemistry and Molecular Biology, University of Wrocław, ul. Przybyszewskiego 63/77, 51-148 Wrocław, Poland. Fax: (+4871) 3756208. E-mail: afsbc@ibmb.uni.wroc.pl

[b] Interdisciplinary Centre for Mathematical and Computational Modelling, Warsaw University, ul. Pawińskiego 5a, 02-106 Warsaw, Poland

[c] Institute of Biotechnology and Environmental Sciences, University of Zielona Góra, ul. Monte Cassino 21b, 65-561 Zielona Góra, Poland

[§] Present address: MRC Muscle and Cell Motility Unit, Randall Centre for Molecular Mechanisms of Cell Function, King's College London, New Hunt's House Guy's Campus, London SE1 1UL, England

* Author for correspondence and reprint requests

Z. Naturforsch. **59 c**, 565–571 (2004); received March 22/April 20, 2004

Using several consensus sequences for the 106 amino acid residue α-spectrin repeat segment as probes we searched animal sequence databases using the BLAST program in order to find proteins revealing limited, but significant similarity to spectrin. Among many spectrins and proteins from the spectrin-α-actinin-dystrophin family as well as sequences showing a rather high degree of similarity in very short stretches, we found seven homologous animal sequences of low overall similarity to spectrin but showing the presence of one or more spectrin-repeat motifs. The homology relationship of these sequences to α-spectrin was further analysed using the SEMIHOM program. Depending on the probe, these segments showed the presence of 6 to 26 identical amino acid residues and a variable number of semihomologous residues. Moreover, we found six protein sequences, which contained a sequence fragment sharing the SH3 (sarc homology region 3) domain homology of 42–59% similarity. Our data indicate the occurrence of motifs of significant homology to α-spectrin repeat segments among animal proteins, which are not classical members of the spectrin-α-actinin-dystrophin family. This might indicate that these segments together with the SH3 domain motif are conserved in proteins which possibly at the early stage of evolution were close cognates of spectrin-α-actinin-dystrophin progenitors but then evolved separately.

Key words: Animal Spectrin-related Proteins, Genetic Semihomology, Spectrin Motif

Introduction

For the first time spectrin was discovered in human erythrocytes, in which it was restricted only to the membrane of the erythrocyte. Goodman *et al.* (1981) discovered nonerythroid spectrins. Subsequently, spectrins were demonstrated in cells of all vertebrates. Further studies revealed that spectrins are members of the subfamily that is a member of the family of spectrin-α-actinin-dystrophin, characterised by the presence of multiple α-helical tandem ~ 110 amino acid residue repeat segments as well as the presence of EF-hand (a Ca^{2+}-binding motif, consisting of a helix, a loop and a second helix, present in Ca^{2+}-binding proteins, such as calmodulin or parvalbumin) and actin-binding domains. Spectrins' α- and β-subunits are associated to form antiparallel heterodimers, which are assembled head-head to form heterotetramers. They are extended, flexible molecules, ~ 200–260 nm in length, and 3–6 nm in diameter, with actin-binding domains at each end of the molecule located at the amino terminus of the β-subunit (for review see *e.g.*: De Matteis and Morrow, 1998; Bennett and Baines, 2001). Known spectrins in humans include two α-subunits (α1, α2) (Sahr *et al.*, 1990; Wasenius *et al.*, 1989), four 'regular' β-subunits (β1, β2, β3, β4) (Berghs *et al.*, 2000; Ma *et al.*, 1993; Winkelmann *et al.*, 1990), and a β-V subunit (Stabach and Morrow, 2000) encoded by distinct genes. Alternative mRNA splicing provides additional diversity among α- and β-spectrins. A common feature of the primary structures of both major types of subunits (α and β) is the presence of sequential segments that use the ~ 106 amino acid residue motif characteristic of all subunits (Baines, 2003).

0939–5075/2004/0700–0565 $ 06.00

Spectrins play a crucial role in the structural integrity, morphology and organisation of cell membranes (for review see *e.g.*: De Matteis and Morrow, 1998; Bennett and Baines, 2001). They are responsible for maintaining the shape and flexibility of cells, stability of the membrane, regulation of the movement of the integral proteins. Several recent reports have implicated that spectrins are also engaged in regulatory and signal transduction pathways (De Matteis and Morrow, 2000; Xu *et al.*, 2000, 2001). Besides several domains responsible for interactions with membrane and membrane attachment proteins, spectrins possess two domains which are involved in regulatory and signalling pathways. the SH3 (Src homology 3) domain, first discovered in the Src protein tyrosine kinase and present in many proteins engaged in cell signalling, mediates interactions with proline-rich stretches in a number of target proteins (Pawson and Scott, 1997). This domain was found mostly in α-spectrins (Djinovic-Carugo *et al.*, 2002). The COOH-terminal regions of long isoforms of β-spectrins have a PH (pleckstrin homology) domain. The PH domain, responsible for phosphatidylinositol 4,5-bisphosphate-binding is approx. 100 residues in length, and it is present in a variety of proteins (Mayer *et al.*, 1993; Haslam *et al.*, 1993).

Spectrin regulates the function of the Golgi apparatus. It organises and stabilises this organelle. Moreover, spectrin is engaged in exocytosis (vesicular trafficking) and is also responsible for the dynamics of the cisternae skeleton (De Matteis and Morrow, 2000).

A motif of the spectrin repeat was also identified in proteins that do not belong to the abovementioned family (Djinovic-Carugo *et al.*, 2002) using the SMART program (Schultz *et al.*, 1998) within the several transcripts or putative proteins with the evidence of the spectrin repeat (*e.g.* DPN or kalirin protein). These proteins do not have the other hallmark domains of the spectrin superfamily (Djinovic-Carugo *et al.*, 2002), but some of them contain signalling domains in addition to spectrin repeat(s).

Our approach to answer the question whether proteins of limited similarity to the α-spectrin subunit, not belonging to spectrin-α-actinin-dystrophin, could be found among animal proteins, was to use several consensus sequences of the α-spectrin repetitive segment as probes to search databases. Position-specific analyses of homology of selected sequences to spectrin, in particular to its 106 amino acid residue repeat segments, using the SEMIHOM algorithm, revealed seven proteins of 20–70% similarity to α-spectrin, six of them also containing region(s) homologous to the SH3 domain.

Materials and Methods

GenBank database searches

The sequence searches were performed using the National Center for Biotechnology Information BLAST programs and the specific databases for: *Mammalia*, *Primates*, *Homo sapiens*, *Rodentia*, *Rattus norvegus*, *Mus musculus*. As probes, the consensus sequences for α-spectrin given by Sahr *et al.* (1990) and the consensus sequence for the 5th α-spectrin segment (Hartwig, 1995), as well as the previously described consensus sequences L106α for the 106 amino acid residue repeat segment and L-SH3 for the SH3 domain obtained using the genetic semihomology algorithm (Leluk *et al.*, 2001; Leluk, 1998) were used. All the consensus sequences are included in Fig. 1. The parameters were as follows: ungapped alignment, matrix: BLOSUM 62, no filters. As the number of the homologous sequences was rather high, we used a graphic analysis provided by BLAST.

Sequence comparisons

Sequence comparisons between α-spectrin and proteins showing homology in short fragments were carried out using the SEMIHOM program based on the genetic semihomology algorithm, published earlier by Leluk (1998).

Results

To find homologous sequences to α-spectrin, the following probes were submitted to sequence databases by means of the BLAST program: a) a consensus sequence of the 106 amino acid residue segment L106α (Leluk *et. al.*, 2001; Leluk, 1998); b) a consensus sequence of the α-spectrin repeat reported by Sahr *et al.* (1990) (*i.e.* S106α); and c) a consensus sequence of the 5th α-spectrin segment 5–98H (Hartwig, 1995), using the parameters described in Materials and Methods. For further analyses the sequences characterised by following features were selected: a) presence of tryptophan at positions 12 and 45; b) presence of tryptophan only at position 45; and c) presence of at least one segment homologous to α-spectrin lacking conser-

Table I. Animal sequences revealing similarity to consensus sequences of α-spectrin found in databases. Protein databases were searched with the BLAST® programme as described in Materials and Methods and sequences were chosen as described in the text.

Probe	Accession number	Protein	Organism	Highest similarity established for at least a 100 amino acid residue segment	Residues	SH3 homology domain(s)
L106α 5–98H	BAB27703	hypothetical protein	*Mus musculus*	68%	460–585	Yes
S106α 5–98H L106α	BAB13468	protein KIAA 1642	*Homo sapiens*	35%	980–1090	No
5–98H	AAH15620	unknown protein	*Homo sapiens*	28.3%	440–550	Yes
S106α	CAC15530	BA122O1.2	*Homo sapiens*	21.75%	340–450	Yes
S106α	XP018033	hypothetical protein	*Homo sapiens*	21.70%	310–430	Yes
5–98H	CAC12899	protein dJ361114.2	*Homo sapiens*	20.75%	320–380	Yes
S106α	BAB29485	FLJ12785 hypothetical protein	*Mus musculus*	20.75%	250–360	Yes

vative tryptophan residues. The sequences fulfilling the above-mentioned criteria, found in the databases using consensus sequences L106α, S106α and 5–98H are listed in Table I.

These sequences were analysed using the SEMIHOM program by their comparison to the entire α-spectrin sequence and to the consensus sequences L106α, S106α, 5–98H and L-SH3 (Fig. 1). The homology search was based on the genetic relationship between amino acid residues occupying the corresponding positions. The results are shown in Fig. 1. All the analysed sequences contain at least one ~100 amino acid residue segment homologous to at least one of the segments of the α-spectrin sequence. The exact number of homologous segments depends on setting the identity threshold and on the frame size. A further, detailed analysis of these sequences, *i.e.* identity and similarity search for all the three consensus sequences for the spectrin repeat unit as well as for the SH3 domain, revealed existence of a class of homologous proteins which could be subdivided into four groups:

1) Sequences that are characterised by the presence of at least one homologous segment containing tryptophan residues 12 and 45.

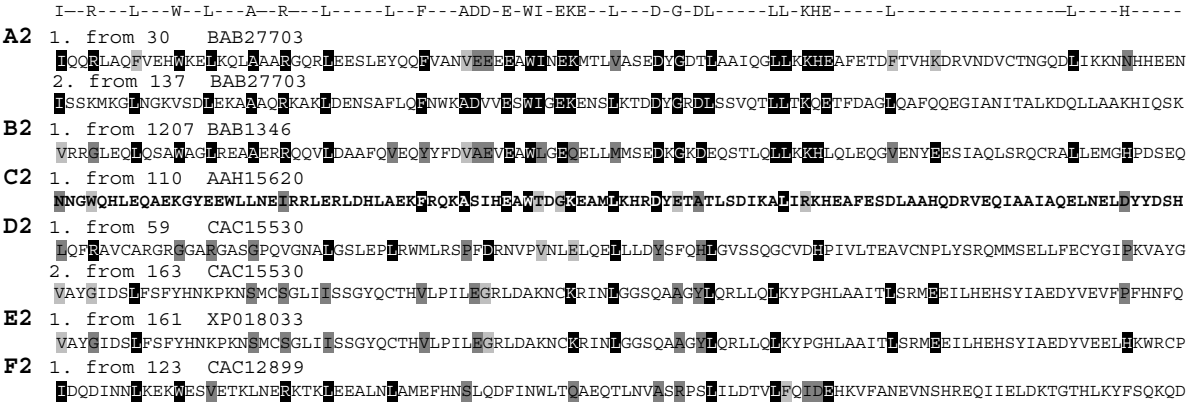
This group of protein sequences consists of only two members: sequence BAB27703.1, a hypothetical protein from *Mus musculus*, and sequence BAB13468, coding for the KIAA 1642 protein from *Homo sapiens* (Table I). Analyses of these sequences are shown in Fig. 1.

In the case of sequence BAB27703.1 (Fig. 1, lines A1–A4), among several regions, which are detected by consensus sequences for α-spectrin (depending on the probe), the most obviously similar is the region starting from residue 460 and ending at position 585, with the highest similarity degree to α-spectrin (68%; Table I). However, the region of the analysed sequence containing tryptophan at positions 12 and 45 is present in the fragment starting from residues 30 and 137, which is not detected when the entire α-spectrin sequence is compared to the analysed one (Fig. 1, lines A1.1 and A2.1). This sequence also contains a LLKKHE stretch, homologous to α-spectrin (Fig. 1, line A1.1). The segment starting from position 13 is relatively rich in residues identical or homologous to the consensus sequence of the 5th segment of α-spectrin: 22 identical and 36 homologous, altogether 58 per 98 residues (Fig. 1, line A3.1). The region (starting from residue 137) adjacent to the detected spectrin repeat segment (30–136) also shows a limited similarity to the SH3 domain (the domain is found to show 9 identical and 13 homologous residues to SH3; Fig. 1, line A4.1).

The other sequence in this group is BAB13468 (coding for the KIAA 1642 protein from *Homo sapiens*), which shows a single spectrin motif if compared to the entire α-spectrin sequence. The fragment showing the highest similarity to α-spectrin encompasses residues 980–1090 and its identity degree reaches 34.90% (Table I). Again, tryptophan residues 45 and 12 are present in the

Consensus sequence L106α

Consensus sequence S106α



segment starting from residue 1207 (Fig. 1, lines B1.1 – B2.1). The segment starting from position 785 is relatively rich in residues identical or homologous to the consensus sequence of the 5th segment of α-spectrin: 26 identical and 24 homologous, altogether 50 per 98 residues (Fig. 1, line B3.1).

2) Sequences characterised by at least one segment homologous to α-spectrin with tryptophan 45 residue and a region with homology to SH3.

There is only one sequence that complies with this rule: sequence AAH15620 coding for an unknown protein from *Homo sapiens*. As presented in Table I, one spectrin-like motif with the highest degree of similarity is detected in sequence AAH15620 if it is compared to the entire α-spectrin molecule. This motif starts from residue 440 and ends at residue 550, and its level of similarity to α-spectrin is 28.30%. Tryptophan 45 is present in the segment detected by S106α (starting from position 110; Fig. 1, line C2.1). The similarity degree of this region detected by consensus sequences 5–98H reaches 62% (51 identical and semihomologous residues; Fig. 1, line C3.1). A fragment of the sequence AAH15620 starting from residue 284 has 11 identical and 17 residues homologous to the L-SH3 consensus sequence (Fig. 1, line C4.1).

3) Sequences characterised by the presence of at least one segment homologous to α-spectrin lacking conservative tryptophan residues.

This class consists of four sequences: CAC15530, XP018033, CAC12899.1 coming from *Homo sa-*

Consensus sequence 5-98H

Consensus sequence L-SH3

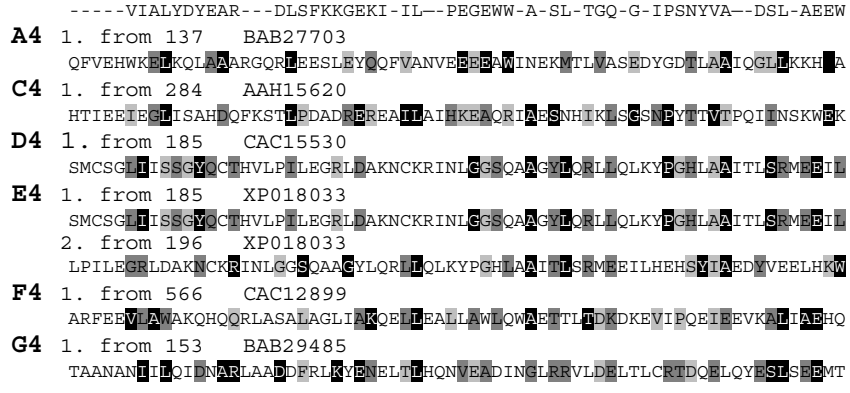


Fig. 1. Comparison of sequences: BAB27703 (A), BAB13468 (B), AAH15620 (C), CAC15530 (D), XP018033 (E), CAC12899 (F) and BAB29485 (G) from sequence databases (see Materials and Methods) to the consensus sequences L106α (A1–G1), S106α (A2–F2), 5–98H (A3–F3) and L-SH3 (A4–G4). 1, 2 … indicate the segments detected by the probe. Identical and semihomologous residues are indicated. Code: W, identical residues; W, semihomology of 1st order; W,semihomology of 2nd order (Leluk, 1998).

piens, and BAB29485 from *Mus musculus*. The results of comparison of these sequences to the entire α-spectrin are shown in Table I. In the case of CAC15530 and XP018033, the highest similarity level reaches 21.75%, and is detected in the *N*-terminal region of α-spectrin. In the case of sequences CAC15530 and XP018033, the level of similarity of one of the three fragments detected with 5–98H is ~ 48% (47 of 98; Fig. 1, lines D3 and E3). The region which could be homologous to SH3 [starting from residues 185 and 196 for sequence XP01803 (Fig. 1, lines E4.1–4.2), and residue 185 for sequence CAC15530 (Fig. 1, line D4.1)] has relatively low similarity (41%). Fig. 1, line F1–F4 shows the results of comparison of sequence CAC12899.1 (coding for the dJ361114.2 protein from *Homo sapiens*) to the consensus sequences of a spectrin segment. Comparison of this sequence to the entire α-spectrin shows the identity of at most 20.75%. The segment starting from

position 123 is detected by three consensus sequences: L106α, S106α, and 5–98H (Fig. 1, lines F1.1–F3.1). A lower similarity level (37 identical and semihomologous residues) is observed when sequence BAB29485.1 is compared to consensus 5–98H. Both sequences CAC12899.1 and BAB29485.1 contain a region showing limited similarity to the SH3 domain (Fig. 1, line F4.1–G4.1).

Discussion

The aim of this study was to search animal gene and protein sequence databases using consensus sequences of the spectrin repeat segment as the probes in order to find polypeptide or putative polypeptide sequences which show limited similarity that proves their homology to α-spectrin but they do not belong to the well-known spectrin-α-actinin-dystrophin family of proteins (Hartwig, 1995; Matsudaira, 1991). In our first approach we used consensus sequences for α-spectrin 106 residue segments L106α (Leluk *et al.*, 2001; Leluk, 1998), S106α (Sahr *et al.*, 1990) and 5–98H (Hartwig, 1995). By means of these probes we searched the databases using the BLAST program to find sequences showing homology of at least their fragments to the α-spectrin 106 amino acid residue segment, which is a characteristic spectrin-like motif. Then analyses of the homology of selected sequences to α-spectrin were performed using the SEMIHOM program, which takes into account the genetic relationship. These sequences were also tested for the presence of the SH3 domain. For this purpose we used a consensus sequence, L-SH3 (Hanus-Lorenz *et al.*, 2004). The searches resulted in finding seven sequences of animal proteins (Table I). They contained at least one segment similar to the typical α-spectrin sequence motif. We classified these sequences into three categories: one which contained fragments with tryptophan at positions 45 and 12; then sequences that have tryptophan at position 45; and finally sequences that do not contain the above-mentioned residues. For control purposes, several proteins, including filamin (P21333), hemoglobin (XP-007935), albumin (NP–000468) and actin (ATHUC), were analysed using the SEMIHOM program and the entire spectrin and consensus sequences (results not shown). Apart from filamin, the similarity degree of these proteins did not exceed 18% (all the sequences discussed here show similarity higher than 20%). The similarity of filamin was 22%, but filamin is one of the members of the actin-binding superfamily of proteins (Matsudaira, 1991). When we compared nesprins (a novel family of spectrin-repeat-containing nuclear-membrane-associated proteins) (Zhang *et al.*, 2001) with α-spectrin using the SEMIHOM program, we observed a rather low level of similarity (19.81–20.75%) to human α-spectrin. Also, we tested four randomly chosen sequences in which spectrin motif(s) were discovered: KIAA0796 (from *Homo sapiens*), LP06350p (from *Drosophila melanogaster*), delta kalerinin-7 (from *Rattus norvegicus*) and a hypothetical 113.3 kDa protein (from *Homo sapiens*) found in the SMART Website. The program SEMIHOM confirmed the presence of spectrin motif(s) with a similarity ranging from 23.3–43.4% which is in principle similar to the level of relationship of sequences found using the SEMIHOM.

Baines A. J. (2003), Comprehensive analysis of all triple helical repeats in β-spectrins reveals patterns of selective evolutionary conservation. Cell. Mol. Biol. Lett. **8**, 195–214.

Bennett V. and Baines A. J. (2001), Spectrin and ankyrin-based pathways: metazoan inventions for integrating cells into tissues. Physiol. Rev. **81**, 1353–1392.

Berghs S., Aggujaro D., Dirkx R. Jr., Maksimova E., Stabach P., Hermel J. M., Zhang J. P., Philbrick W., Slepnev V., Ort T., and Solimena M. (2000), BetaIV spectrin, a new spectrin localized at axon initial segments and nodes of ranvier in the central and peripheral nervous system. J. Cell. Biol. **151**, 985–1002.

De Matteis M. A. and Morrow J. S. (1998), The role of ankyrin and spectrin in membrane transport and domain formation. Curr. Opin. Cell Biol. **10**, 542–549.

De Matteis M. A. and Morrow J. S. (2000), Spectrin tethers and mesh in the biosynthetic pathway. J. Cell. Sci. **113**, 2331–2343.

Djinovic-Carugo K., Gauntel M., Ylanne J., and Young P. (2002), The spectrin repeat: a structural platform for cytoskeletal protein assemblies. FEBS Lett. **513**, 119–123.

Goodman S. R., Zagon I. S., and Kulikowski R. R. (1981), Identification of a spectrin-like protein in nonerythroid cells. Proc. Natl. Acad. Sci. USA **78**, 7570–7574.

Hanus-Lorenz B., Lorenz M. Hryniewicz-Jankowska A., Skała J., Leluk J., and Sikorski A. F. (2004), Spectrin motifs are present in plant and yeast cells (in press).

Hartwig J. H. (1995), Actin-binding proteins. 1: Spectrin super family. Protein Profile **2**, 703–799.

Haslam R. J., Koide H. B., and Hemmings B. A. (1993), Pleckstrin domain homology. Nature **363**, 309–310.

Leluk J. (1998), A new algorithm for analysis of the homology in protein primary structure. Computers Chem. **22**, 123–131.

Leluk J., Hanus-Lorenz B., and Sikorski A. F. (2001), Application of genetic semihomology algorithm to theoretical studies on various protein families. Acta Biochim. Polon. **48**, 21–33.

Ma Y., Zimmer W. E., Riederer B. M., Bloom M. L., Barker J. E., and Goodman S. R. (1993), The complete amino acid sequence for brain beta spectrin (beta fodrin): relationship to globin sequences. Mol. Brain Res. **18**, 87–99.

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

Mayer B. J., Ren R., Clark K. L., and Baltimore D. (1993), A putative modular domain present in diverse signaling proteins. Cell **73**, 629–630.

Pawson T. and Scott J. D. (1997), Signaling through scaffold, anchoring, and adaptor proteins. Science **278**, 2075–2080.

Sahr K. E., Laurila P., Kotula L., Scarpa A. L., Coupal E., Leto T. L., Linnenbach A. J., Winkelmann J. C., Speicher D. W., and Marchesi V. T. (1990), The complete cDNA and polypeptide sequences of human erythroid alpha-spectrin. J. Biol. Chem. **265**, 4434–4443.

Schultz J., Milpetz F., Bork P., and Ponting C. P. (1998), SMART, a simple modular architecture research tool: identification of signaling domains. Proc. Natl. Acad. Sci. USA **95**, 5857–5864.

Stabach P. R. and Morrow J. S. (2000), Identification and characterization of beta V spectrin, a mammalian ortholog of *Drosophila* beta H spectrin. J. Biol. Chem. **275**, 21385–21395.

Wasenius V. M., Saraste M., Salven P., Eramaa M., Holm., and Letho V. P. (1989), Primary structure of the brain alpha-spectrin. J. Cell Biol. **108**, 79–93.

Winkelmann J. C., Costa F. F., Linzie B. L., and Forget B. G. (1990), Beta spectrin in human skeletal muscle. Tissue-specific differential processing of 3′ beta spectrin pre-mRNA generates a beta spectrin isoform with a unique carboxyl terminus. J. Biol. Chem. **265**, 11827–11832.

Xu J., Ziemnicka D., Merz G. S., and Kotula L. (2000), Human spectrin Src homology 3 domain binding protein 1 regulates macropinocytosis in NIH 3T3 cells. J. Cell Sci. **21**, 3805–3814.

Xu J., Ziemnicka D., Scalia J., and Kotula L. (2001), Monoclonal antibodies to alphaI spectrin Src homology 3 domain associate with macropinocytic vesicles in nonerythroid cells. Brain Res. **898**, 171–177.

Zhang Q., Skepper J. N., Yang F., Davies J. D., Hegyi L., Roberts R. G., Weissberg P. L., Ellis J. A., and Shanahan C. M. (2001), Nesprins: a novel family of spectrin-repeat-containing proteins that localize to the nuclear membrane in multiple tissues. J. Cell Sci. **114**, 4485–4498.