FEBS Letters 431 (1998) 134–137 FEBS 20514

Minireview

# CH domains revisited

Theresia Stradal[a], Wolfgang Kranewitter[a], Steven J. Winder[b], Mario Gimona[a,*]

[a]*Institute of Molecular Biology, Austrian Academy of Sciences, Department of Cell Biology, Billrothstrasse 11, A-5020 Salzburg, Austria*
[b]*Institute of Cell and Molecular Biology, University of Edinburgh, Mayfield Road, Edinburgh EH9 3JR, UK*

Received 2 June 1998

**Abstract A sequence motif of about 100 amino acids, termed the 'calponin homology domain' has been suggested to confer actin binding to a variety of cytoskeletal and signalling molecules. Here we analyse and compare the sequences of all calponin homology domain-containing proteins identified to date. We propose that single calponin homology domains do not confer actin-binding per se and that the actin-binding motifs of cross-linking proteins, which comprise two disparate calponin homology domains, represent a unique protein module.**



*Key words:* Calponin; Calponin homology domain; Actin binding

## 1. Introduction

The organisation of the diversity of cellular actin structures is governed by a variety of actin modulatory proteins. Some have a preference for binding to the monomeric (G) form of actin, whilst others specifically target polymerised filamentous (F) actin structures (see [1] for review). The steady increase in sequence information available for actin modulatory proteins led to the proposal for a modular organisation of actin-binding proteins (reviewed in [2] and [3]). A 250 amino acid stretch called the 'actin binding domain' (ABD) consisting of two tandem repeats of 125 residues each was identified as a general motif in all actin cross-linking proteins. With the refinement of algorithms for the identification of sequence motifs, especially those using secondary structure predictions, new putative protein modules were proposed in actin binding proteins.

Significant similarities between the predicted secondary structure of the calponin family of actin-modulatory proteins and the amino terminal regions of the proto-oncogene Vav and other GTPase effector proteins were recently identified [4]. This 100 residue motif, hypothesised to comprise the actin-binding domain of calponin, was termed the 'calponin homology (CH) domain'. This domain was subsequently identified in other signalling molecules and was proposed as the region responsible for directly linking signal transduction molecules to the actin cytoskeleton via an association with F-actin [5–8]. This assumption was substantiated by the subdivision of ABDs of established actin cross linking proteins like α-actinin, filamin and fimbrin into two equivalent halves, each corresponding to one CH domain [9–11]. Together, this led to the commonly accepted misconception that CH domains function as an autonomous F-actin binding domain, and the CH domain was installed as the prototype for a novel actin-binding module [3,4]. Closer inspection of CH domain sequences, however, reveals that this region does not overlap the established actin-binding domain in calponin, delineated in several independent studies [12–14]. Moreover, a second actin-targeting site was recently identified in the C-terminal tandem repeats of calponin [15], and it was further demonstrated that the single CH domain is neither sufficient nor necessary for F-actin binding of calponin and SM22 [16].

## 2. Data base entries

Detailed analysis of all known CH domain-containing proteins defines three groups of CH domain containing molecules: (i) the fimbrin family of monomeric actin cross-linking molecules containing two ABDs; (ii) dimeric cross-linking proteins (α-actinin, β-spectrin, filamin etc.) and monomeric F-actin binding proteins (dystrophin, utrophin) each containing one ABD; and (iii) proteins containing only a single amino terminal CH domain. We have compared the corresponding sequences of 60 individual proteins including a total of 104 CH domains from human, chicken and all invertebrate entries. Using the search and alignment programs BLAST and CLUSTAL-W, five major branches with significant bootstrap values were identified as delineated in Fig. 1. The amino terminal (type 1) CH domains of proteins containing 1 ABD are clearly more similar to each other than they are to the carboxy terminal (type 2) CH domains of the motif. Likewise as indicated in Fig. 1, both N-terminal CH domains (CH 1.1 and CH 1.2) of Fimbrin's ABD 1 and ABD 2 group together, but are significantly different from the C-terminal CH domains (CH 2.1 and CH 2.2). All those proteins in which only a single CH domain was identified form a separate major subgroup, displaying a higher degree of sequence similarity to each other than to any of the ABD-constituting CH domains, although this group includes members of such functionally distinct protein families as the calponin family of actin-modulating proteins, the SM22 family, and a number of signalling molecules from a wide range of species.

## 3. Multiple CH domain proteins

The CH domains of ABD proteins are clearly split into 4 branches with all type 1 and type 2 CH domains from single ABD proteins fall into their own distinct groups (Fig. 1). Fimbrin, from yeast through *Arabidopsis* and *Drosophila* to man, however, groups separately from these CH domain con-

*Corresponding author. Fax: (43) (662) 63 961 29.
E-mail: mhimona@imb.oeaw.ac.at

*Abbreviations:* CaP, calponin; CH domain, calponin homology domain; ABD, actin binding domain; CLR, calponin-like repeat


PII: S0014-5793(98)00751-0

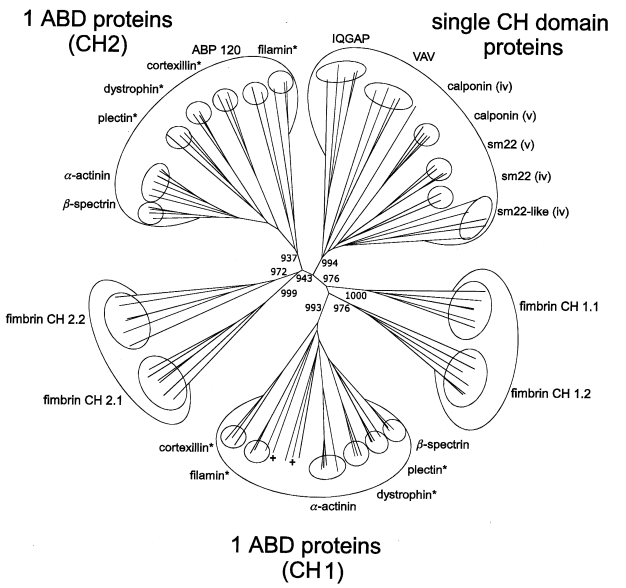


Fig. 1. Phylogenetic tree of CH domains. The analysis includes 104 individual sequences from 60 single and multiple CH domain-containing proteins. Note the specific clustering of the N- and C-terminal CH domains of ABDs from proteins containing 1 or 2 ABDs, respectively. Single CH domain containing molecules create an independent subpopulation. Bootstrap values are given for the major branches. ABD, actin binding domain; CH, calponin homology domain; v, vertebrates; iv, invertebrates. * indicates diversity in the current nomenclature of proteins falling into the respective subgroups. + marks the unusual grouping of the type 1 CH domains of two hypothetical α-actinin/spectrin-like proteins from *C. elegans* and *S. pombe* together with those of ABP 120 (open branches in 1 ABD proteins).

taining ABDs, with the amino-terminal type 1 CH domain in each pair forming one branch and the type 2 CH domains another. Structurally all of these domains appear equivalent, with the core elements of the second CH domains from β-spectrin, fimbrin and utrophin ([9,11]; Keep and Winder, unpublished observations) being conserved to a rmsd of 1.4 Å despite considerable primary sequence divergence. Coupled with the primary sequence divergence in the core elements of fimbrin CH domains is the insertion of longer intra-helical linkers, not seen in other CH domains, which appear to set them aside as a distinct group from other CH domain containing ABDs.

Despite the overall structural conservation of the unique CH domain fold, the first and second CH domains within one ABD group separately, and furthermore are functionally not equivalent. Each single ABD, comprising two CH domains, is able to bind one actin monomer in the filament, with affinities in the low μM range. When CH domains from ABDs have been analysed in isolation, it is apparent that their actin binding characteristics are quite different. All of the amino terminal CH domains so far analysed have the intrinsic ability to bind to actin albeit with lower affinity than the complete ABD ([17,18]; S.J. Winder, unpublished observations). Carboxy terminal CH domains on the other hand bind actin extremely weakly or not at all ([9,17,18]; S.J. Winder, unpublished observations). Nevertheless both N- and C-terminal CH domains are required for a fully functional ABD, with the second CH domain contributing to the overall stability of the complete ABD through inter-domain helix-helix interactions [11].

## 4. Single CH domain proteins

Twenty-seven database entries were found to encode proteins containing a single amino-terminal CH domain (Fig. 2). The three variants of vertebrate calponin share a high degree of similarity in their CH domain sequences and form a separate subgroup, similar to the vertebrate SM22 family. NP 25 from rat brain aligns with this latter family, identifying itself as a vertebrate neuronal SM22 variant. The flight muscle protein mp 20 from *Drosophila* aligns with the invertebrate SM22s (also known as myophilins). Proteins considerably different from either calponin or SM22 but displaying a characteristic SM22 domain structure (one CH domain, one CLR; see Fig. 3), form a distinct subgroup, although its (exclusively invertebrate) members appear less closely related (see below). Particularly interesting is the identification of the first 'real' invertebrate calponin from *Schistosoma mansoni*. In addition to an amino terminal CH domain, it contains 5 C-terminal CLRs (calponin-like repeats; see also [19]). The CH domain appears significantly different from those of vertebrate calponin, but also from invertebrate SM22-like proteins and does not align with any of the other protein clusters. Notably, with the sole exception of a protein identified in the nematode *Meloidogyne javanica*, which contains only a single CH domain but lacks the C-terminal CLR, all of the invertebrate SM22-like members are hypothetical proteins derived from genomic sequences. Their CH domain sequences fail to place them into any of the calponin or SM22 subgroups and a definite classification may have to await their characterisation at the protein level.

Among the family of signalling molecules we found invertebrate homologs for both IQGAP (*S. pombe*) and Vav (*C. elegans*). Their CH domain sequences contain several conserved residues present only in these two protein classes. A few examples for group-specific amino acids are highlighted in Fig. 2. Vav, for example, contains two conserved Cys residues (marked in blue) which distinguishes them from IQGAPs. Similarly, the CH domains of IQGAP contain a unique Cys residue at the beginning of their CH domain (marked in yellow) whereas all CaPs contain a Gln at the same position. Interestingly, the calponin CH domains contain two conserved Gln residues unique to this family (marked in green). These specific residues define the signalling molecule family as a separate group among the single CH domain proteins and infers particular functions for individual CH domains.

## 5. CHD versus ABD

Confusion concerning the function of CH domains currently prevails the literature [4,6,8,10,20]. Single CH domains clearly function differently from ABDs containing two CH domains in tandem. The low sequence similarity between CH domains allows for functional diversity despite the predicted similarity in secondary structure. Moreover, data derived from the crystal structure of the ABD of fimbrin clearly demonstrate that the two dissimilar CH domains contribute differently to the hydrophobic core of an ABD [11] and CH domains within an ABD display significant differences in their respective affinities for actin [17,18]. Hence, the general assumption that single CH domain-containing proteins bind to actin via their CH domain and that a single CH domain is sufficient for F-actin binding, is inadequate. Although the 190-

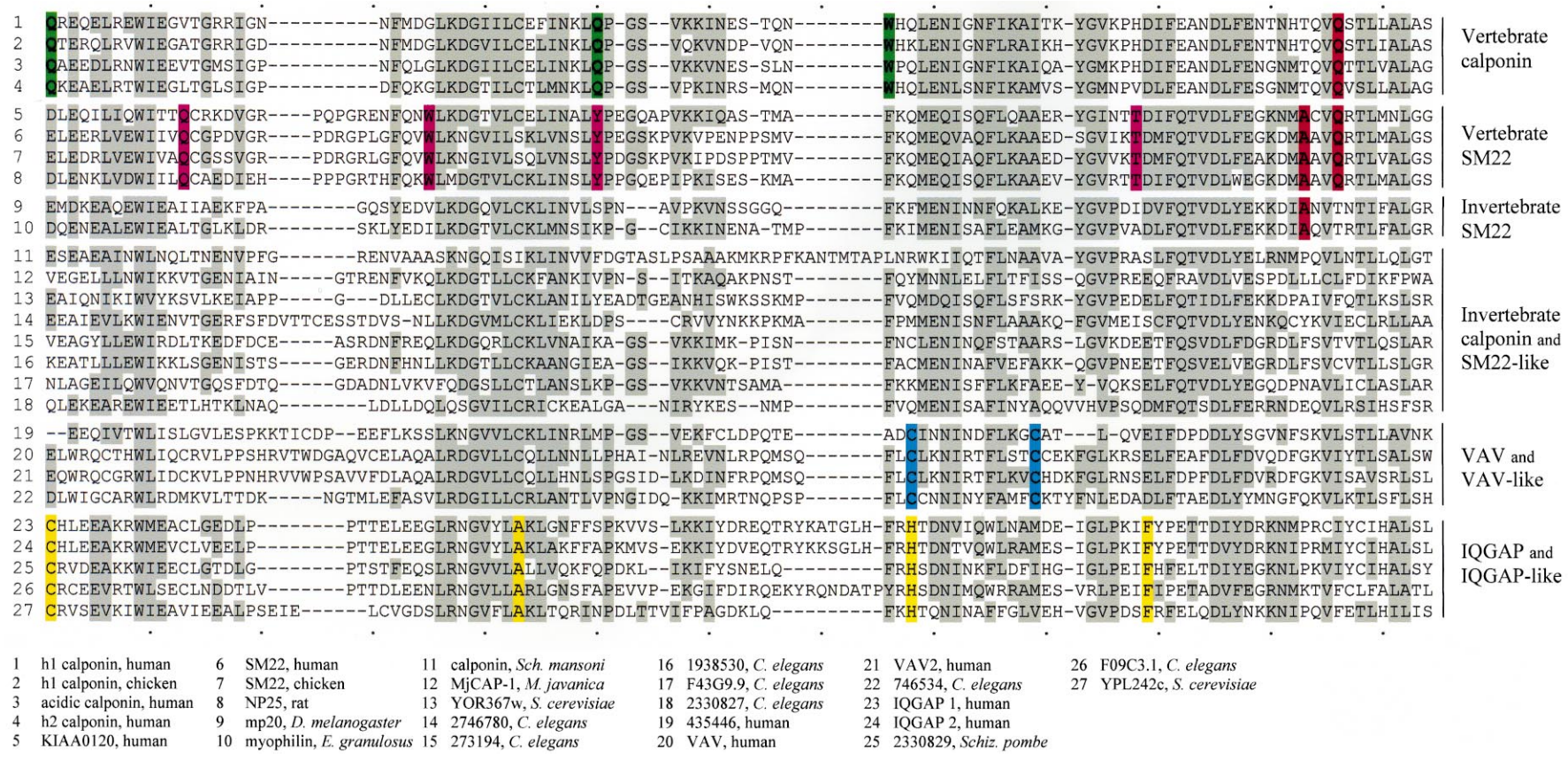


Fig. 2. Sequence alignment of CH domains from single CHD proteins. Residues on grey background indicate conserved amino acids in more than 50% of any given column, with the exception of hypothetical proteins. Residues marked in colour indicate amino acids unique for CH domains of different single CH domain subfamilies: yellow, IQGAP; blue, VAV; green, vertebrate CaP; pink, vertebrate SM22. Residues marked in red indicate residues conserved in vertebrate CaP and SM22, and vertebrate and invertebrate SM22, respectively.

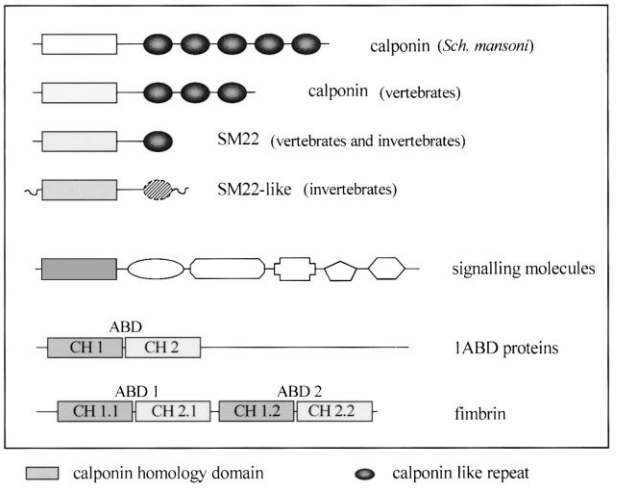


Fig. 3. Schematic representation of the molecular structure of CH domain-containing proteins. Hatched symbol indicates presence of a CLR motif of significantly divergent sequence or absence of a recogniseable CLR in some hypothetical proteins of nematodes and *S. cerevisiae*.

kDa protein IQGAP localises in the membrane ruffles of cultured cells, it remains to be shown whether this reflects an association with actin mediated by the amino terminal CH domain. Alternatively other domains like the PH domain may be responsible for this phenomenon [5]. In conclusion, the identification of a single CH domain should be treated with caution and special care taken in attributing actin-binding properties to proteins simply due to the presence of this motif.

*Acknowledgements:* Supported by Wellcome Trust Grant 042180 (S.J.W.). M.G. acknowledges the support of The Austrian Science Foundation (P-11845) and the Austrian National Bank (Grant #6518).

## References

[1] Vandekerckhove, J. (1990) Curr. Opin. Cell Biol. 2, 41–50.
[2] Matsudaira, P. (1990) Trends Biochem. Sci. 16, 87–92.
[3] Puius, Y., Mahoney, N.M. and Almo, S.C. (1998) Curr. Opin. Cell Biol. 10, 23–34.
[4] Castresana, J. and Saraste, M. (1995) FEBS Lett. 374, 149–151.
[5] Bashour, A.-M., Fullerton, A.T., Hart, M.J. and Bloom, G.S. (1997) J. Cell Biol. 137, 1555–1566.
[6] Brill, S., Li, S., Lyman, C.W., Church, D.M., Wasmuth, J.J., Weissbach, L., Bernards, A. and Snijders, A.J. (1996) Mol. Cell Biol. 16, 4869–4878.
[7] Hart, M.J., Callow, M.G., Souza, B. and Polakis, P. (1996) EMBO J. 15, 2997–3005.
[8] Fukata, M., Kuroda, S., Fujii, K., Nakamura, T., Shoji, I., Mastuura, Y., Okawa, K., Iwamatsu, A., Kikuchi, A. and Kaibuchi, K. (1997) J. Biol. Chem. 272, 29579–29583.
[9] Carugo, K.D., Banuelos, S. and Saraste, M. (1997) Nat. Struct. Biol. 4, 175–179.
[10] Hanein, D., Matsudaira, P. and DeRosier, D.J. (1997) J. Cell Biol. 139, 387–396.
[11] Goldsmith, S.C., Pokala, N., Shen, W., Fedorov, A.A., Matsudaira, P. and Almo, S.C. (1997) Nat. Struct. Biol. 4, 708–7212.
[12] Mezgueldi, M., Fattoum, A., Derancourt, J. and Kassab, R. (1992) J. Biol. Chem. 267, 15943–15951.
[13] Mezgueldi, M., Mendre, C., Calas, B., Kassab, R. and Fattoum, A. (1995) J. Biol. Chem. 270, 8867–8876.
[14] EL-Mezgueldi, M., Strasser, P., Fattoum, A. and Gimona, M. (1996) Biochemistry 35, 3654–3661.
[15] Mino, T., Yuasa, U., Nakamura, F., Naka, M. and Tanaka, T. (1998) Eur. J. Biochem. 251, 262–268.
[16] Gimona, M. and Mital, R. (1998) J. Cell Sci., in press.
[17] Way, M., Pope, B. and Weeds, A.G. (1992) J. Cell Biol. 119, 835–842.
[18] Winder, S.J., Hemmings, L., Maciver, S.K., Bolton, S.J., Tinsley, J.M., Davies, K.E., Critchley, D.R. and Kendrick-Jones, J. (1995) J. Cell Sci. 108, 63–71.
[19] Martin, R.M., Chilton, N.B., Lightowlers, M.W. and Gasser, R.B. (1997) Int. J. Parasitol. 27, 1561–1567.
[20] Epp, J.A. and Chant, J. (1997) Curr. Biol. 7, 921–929.