FEBS 16170

FEBS Letters 374 (1995) 149–151

Hypothesis

Does Vav bind to F-actin through a CH domain?

Jose Castresana, Matti Saraste*

European Molecular Biology Laboratory, Postfach 102209, D-69012 Heidelberg, Germany

Received 12 September 1995

Abstract An actin-binding protein domain we call here 'calponin-homology' or CH is present in signalling proteins such as Vav which are involved in activation and inactivation of small G-proteins. Using profile methods, we have detected two repeats of this domain in the actin-binding region of α-actinin and related proteins. Based on this, we propose that CH domain in Vav and other signalling proteins is employed for association with filamentous actin, and that this function correlates with their control on the G-proteins Rac and Rho which are involved in the organization of cytoskeleton.

Key words: Calponin-homology; Cytoskeleton; Signal transduction; Vav; Actin binding

1. Introduction

Calponins are an emerging family of proteins, mainly involved in the regulation of smooth muscle contraction. Their interaction with actin filaments is an essential condition for the regulation of the actin-myosin machinery [1,2]. Calponins consist of a unique N-terminal domain followed by one to three calponin repeats. It has been proposed that the N-terminal domain is homologous to the actin-binding domain of α-actinin [1]. It has been further suggested that some members of the calponin family share sequence similarity with the N-terminal regions of Vav [3] and a human protein related to Ras-GAP [4]. However, these similarities have not been tested by methods that can give statistical confidence on sequence homology.

During the evolution of eukaryotes, domains carrying a distinct function have been shuffled between otherwise unrelated proteins [5]. The N-terminal calponin domain ('calponin-homology' or CH) appears to be such a protein module with a characteristic function. We define here the boundaries of the CH domain and demonstrate that it is present in a number of different proteins. The common feature of these cytoskeletal and signalling proteins may be binding to F-actin.

2. Database searches

The database searches using the N-terminal calponin domain with a variety of methods resulted in significant matches (Fig. 2) to sequences in other actin-binding proteins (see below), a Ras-GAP-related protein, a protein expressed in neurons, the proto-oncogene product Vav and two open reading frames which probably code for signalling proteins (human hsorfp_1 and nematode cec35b8_2). The latter conclusion is based on the presence of other domains peculiar to signal-transducing proteins such as SH2, SH3, PH, DH and DAG/PE-binding domains (Fig. 1). Some of the previous statements [3,4] about similarities between calponins and other proteins pass our statistical test after proper alignment. The N- and C-terminal boundaries of the CH domain are based on the examination of multiple alignment and on its terminal location in several proteins (Figs. 1 and 2).

3. CH domain: putative function and predicted structure

The actin-binding domain of α-actinin has approximately 250 amino acid residues and is found in three protein families, namely in the spectrin family (which includes β-spectrin, α-actinin, dystrophin and utrophin), the filamin family (including the actin gelation factor of *Dictyostelium discoideum*) and the fimbrin family [6]. These cytoskeletal proteins are known to cross-link actin filaments, binding to the same region of actin [7]. Sequence alignments and image reconstitution with electron microscopy have proposed that the actin-binding domain consists of two sub-domains with a similar size [8,9], and it has been shown that the N-terminal half alone is able to bind to F-actin [10]. Actin-binding residues (Fig. 2) have been mapped to both sub-domains [6,11]. Our sequence comparison further indicates that both share significant homology with the calponin domain. Thus, when the pieces of the puzzle are put together, CH emerges as the prototypic domain which can be assigned a potential actin-binding function.

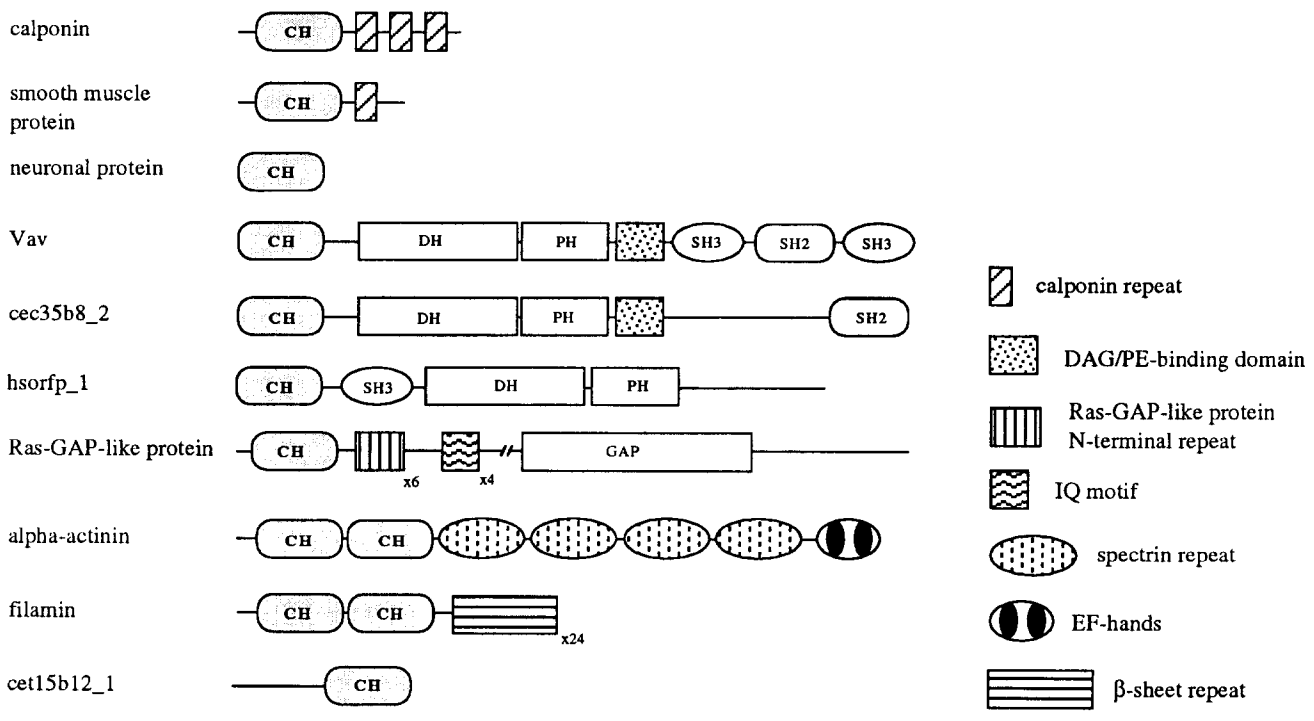
The structure of the CH domain is probably mainly α-helical as the structure predictions with the PHD program [12] propose four α-helices. There is a tendency that the CH domain(s) is (are) located at the N-terminus of the host protein. The independent nature of the CH domain is demonstrated by a neuronal protein which appears to be exclusively made of this domain and by a protein of unknown function (cet15b12_1 in Fig. 1) that contains a single CH domain which seems to be evolutionarily closely related to the second domain in α-actinin.

4. CH domain in signalling proteins

The CH domain is present in a human Ras-GAP (GTPase-activating protein of Ras), proteins encoded by open reading frames called hsorfp_1 and cec35b8_2, and in Vav. The three latter proteins contain the Dbl homology (DH) domain that is a guanine-nucleotide-exchange factor for small G-proteins

*Corresponding author. Fax: (49) (6221) 387 306.

SSDI 0014-5793(95)01098-X



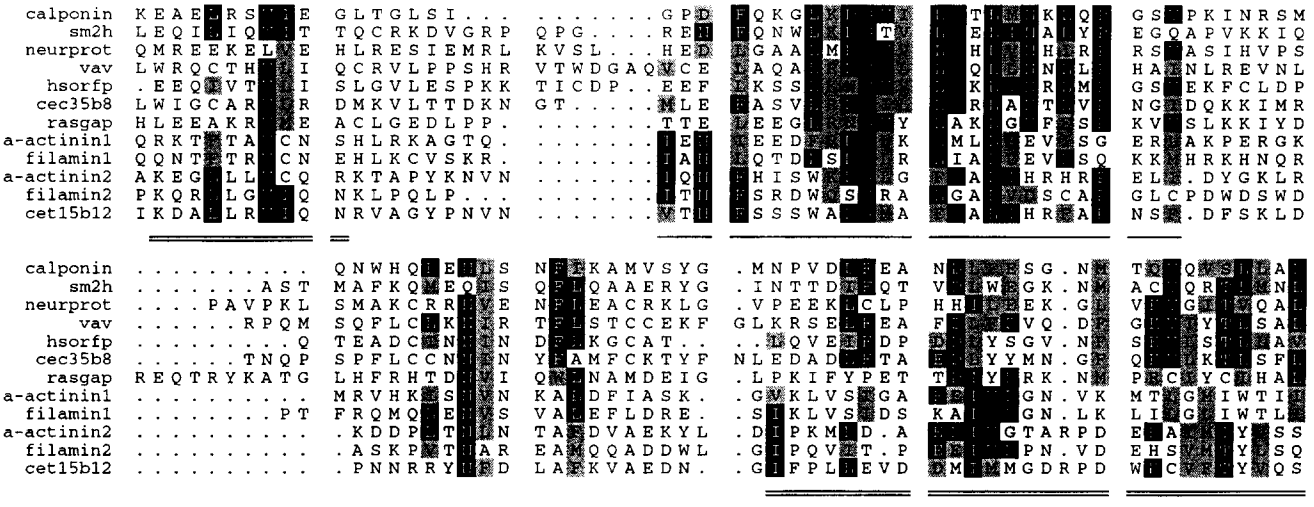


Fig. 2. Multiple alignment of CH domains. The actin-binding regions mapped by Winder et al. [11] are double-underlined. There is an obvious consistency in the alignments within calponin family and within the first and second sub-domains of the actin-binding domain in the spectrin/α-actinin family. In order to test whether the three sets of sequences (the calponin domain and the two sub-domains of spectrin/α-actinin) are homologous to each other, we submitted the single-underlined 26 amino acid motifs (with no gaps) of the three subfamilies to independent searches in protein databases using the MoST program [18]. Each separate block identified the members of the two other subfamilies with $r < 0.005$, where r is the ratio of the expected number of sequence segments with a given score to the observed number. With profile methods based on the full alignment (using the PROFILESEARCH program [19]), independent profile searches using the first half of the spectrin/α-actinin domain found calponins as the best matches after the proteins in the spectrin family. The same happened with profile searches using the second half of the spectrin/α-actinin domain, and with calponin profiles which identified the members of the spectrin/α-actinin proteins next after calponins. The domains in signalling proteins were found during these searches. They gave the highest scores with calponin profiles. The multiple alignment was constructed with the CLUSTALW program [20] and manually refined. Positions with conserved residues in 50% or more of the sequences are shadowed. Light shadowing for similar amino acids is applied according to the Gonnet matrix [21] of amino acid substitutions (only values greater than 1.5 are considered for shadowing). Identification names for the sequences containing the CH domain (with the position of the domain) are calponin, CLP2_MOUSE (29–127), sm2h, SM2H_HUMAN (26–132), neurprot, NEUR_FELCA (3–113), vav, VAV_HUMAN (3–115), α-actinin, AAC1_HUMAN (33–131 and 146–244) and filamin, ABP2_HUMAN (45–145 and 168–263) in the SWISSPROT database, and hsorfp, HSORFP (first orf, 1–103), cec35b8, CEC35B8 (second orf, 7–114), ras-GAP, HSIQGA (first orf, 46–155) and cet15b12, CET15B12 (first orf, 115–213) in the EMBL database. All seqeuces are from human exept CLP2_MOUSE from mouse, NEUR_FELCA from cat and CEC35B8 and CET15B12 from *Caenorhabditis elegans*. We have become aware that the CH domain has also been discovered by Philipp Bucher and Kay Hofmann (Swiss Institute of Experimental Cancer Research, Lausanne) in a similar search. They have submitted a profile to the PROSITE database under accession number PS50021.

Fig. 1. Molecular architecture of proteins containing the CH domain. The symbols for different domains are given on the right. A figure below a domain indicates the number of tandem repeats. The domains in open reading frames hsorfp_1, cec35b8_2 and cet15b12_1 were identified with BLAST [17]. See legend to Fig. 2 for entry names. ←

(Ras, Rho and Rac; see ref. [13]). The proto-oncogene product Vav is a complex modular protein that is specifically expressed in haematopoietic cells. It is interesting that cec35b8_2, a protein with a similar domain composition to Vav (Fig. 1), is present in nematodes which have no circulatory system. The N-terminus of Vav has been proposed to contain a helix-loop-helix DNA-binding domain, but our data based on profile methods clearly show that it has a CH domain. This is probably the last domain of Vav that has remained to be assigned (Fig. 1). Vav is known to become oncogenic when its first 65 N- terminal amino acids are deleted [14,15]. The actin-binding CH domain of Vav is therefore essential for normal function.

It is well established that cytoskeletal proteins such as spectrin share domains with signalling proteins [16]. The novel picture that now emerges shows the converse: signalling proteins may contain functional domains that have until now been found only in cytoskeletal proteins. The four examples discovered by our search have a domain that may be involved in binding to actin filaments. In particular, this could be required for the correct localization of factors activating or inactivating small G-proteins such as Rho and Rac that are involved in the organization of cytoskeleton and in cellular processes such as membrane ruffling [13].

References

[1] Vancompernolle, K., Gimona, M., Herzog, M., Van Damme, J., Vandekerckhove, J. and Small, V. (1990) FEBS Lett. 274, 146–150.
[2] Kolakowski, J., Makuch, R., Stepkowski, D. and Dabrowska, R. (1995) Biochem. J. 306, 199–204.
[3] Adams, J.M., Houston, H., Allen, J., Lints, T. and Harvey, R. (1992) Oncogene 7, 611–618.
[4] Weissbach, L., Settleman, J., Kalady, M.F., Snijders, A.J., Murthy, A.E., Yan, Y.-X. and Bernards, A. (1994) J. Biol. Chem. 269, 20517–20521.
[5] Doolittle, R.F. (1995) Annu. Rev. Biochem. 64, 287–314.
[6] Hartwig, J.H. (1994) Protein Profile. Actin Binding Proteins 1: Spectrin superfamily, Academic Press, London.
[7] Matsudaira, P. (1994) J. Cell Biol. 126, 285–287.
[8] de Arruda, M.V., Watson, S., Lin, C.-S., Leavitt, J. and Matsudaira, P. (1990) J. Cell Biol. 111, 1069–1079.
[9] Taylor, K.A. and Taylor, D.W. (1993) J. Mol. Biol. 230, 196–205.
[10] Way, M., Pope, B. and Weeds, A.G. (1992) J. Cell Biol. 119, 835–842.
[11] 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.
[12] Rost, B. and Sander, C. (1994) Proteins 19, 55–72.
[13] Boguski, M.S. and McGormick, F. (1993) Nature 366, 643–654.
[14] Coppola, J., Bryant, S., Koda, T., Conway, D. and Barbacid, M. (1991) Cell Growth Diff. 2, 95–105.
[15] Katzav, S., Cleveland, J.L., Heslop, H.E. and Pulido, D. (1991) Mol. Cell Biol. 11, 1912–1920.
[16] Cohen, G.B., Ren, R. Baltimore, D. (1995) Cell 80, 237–248.
[17] Altschul, S.F., Gish, W., Miller, W., Myers, E.W. and Lipman, D.J. (1990) J. Mol. Biol. 215, 403–410.
[18] Tatusov, R.L., Altschul, S.F. and Koonin, E.V. (1994) Proc. Natl. Acad. Sci. USA 91, 12091–12095.
[19] Gribskov, M., McLachlan, A.D. and Eisenberg, D. (1987) Proc. Natl Acad. Sci. USA 84, 4355–4358.
[20] Thompson, J.D., Higgins, D.G. and Gibson, T.J. (1994) Nucleic Acids Res. 22, 4673–4680.
[21] Gonnet, G.H., Cohen, M.A. and Benner, S.A. (1992) Science 256, 1443–1445.