ELSEVIER

FEBS Letters 343 (1994) 1–5

FEBS 13893

FEBS LETTERS

Minireview

New era of calpain research

Discovery of tissue-specific calpains

Hiroyuki Sorimachi[a], Takaomi C. Saido[b], Koichi Suzuki[a,*]

[a]*Laboratory of Molecular Structure and Functions, Department of Molecular Biology, Institute of Molecular and Cellular Biosciences, University of Tokyo, 1-1-1 Yayoi, Bunkyo-ku, Tokyo 113, Japan*

[b]*Department of Molecular Biology, Tokyo Metropolitan Institute of Medical Science, Bunkyo-ku, Tokyo 113, Japan*

Received 28 February 1994

Abstract

The recent discovery of several new calpain species other than the two species thus far studied reveals that calpain, especially the calpain large subunit, constitutes a family comprising at least six members that can be classified into ubiquitous (μ-, m- and μ/m-types) and tissue-specific (p94 or nCL-1 specific for skeletal muscle, and nCL-2 and -2′ specific for stomach) calpains. The newly identified tissue-specific calpains have various characteristics distinct from conventional calpains in structure, manner of expression, and enzyme activity. Unique features of tissuespecific calpains are discussed together with the evolutionary view of the calpain large subunit.

Key words: Calpain; Novel species; Molecular diversity; Evolution; Tissue-specific expression

1. Introduction

Calpain (EC 3.4.22.17), the most typical cytosolic calcium-dependent cysteine proteinase, has been studied extensively, since its ubiquitous expression, at least in animal cells, suggests its indispensable physiological function as one of the cellular receptors of calcium ions [1–6]. Two molecular species, μ- and m-calpains, have been identified. As their names suggest, their [Ca^{2+}] requirements for protease activity differ, at least in vitro (ca. μM and mM, respectively). Both species consist of distinct large subunits (M_r: ca. 8×10^4) and a common small subunit (ca. 3×10^4). The large subunit is responsible for protease activity and Ca^{2+} dependency, and can be structurally divided into 4 domains. As shown in Fig. 1, the second domain (II) is a cysteine protease domain, and the fourth domain (IV) is a Ca^{2+}-binding domain. Although the functions of the first and third domains are not clear, the domain structure clearly indicates that the calpain large subunit was generated by fusion of ancestral genes for cysteine protease and calmodulin [7]. Recent studies on calpain have been directed toward clarifying its physiological functions, because structural studies are thought to be complete [3,8,9]. In 1989, however, a novel calpain species, p94 (nCL-1), was discovered. p94 has the same fundamental domain structure as shown in Fig. 1 [10], but is different from μ- and m-calpains in that it is expressed only in skeletal muscle. In other words, p94 is the first discovery of tissue-specific calpain. A second tissue-specific calpain expressed predominantly in stomach was found in 1993 [11]. Moreover, a third ubiquitously expressed calpain was identified quite recently. Thus, the simple concept once raised for calpain, that it comprises two ubiquitous μ- and m-types, must be totally changed. In this review, unique features of the novelly identified calpains are summarized mainly with respect to structure. For discussion of the physiological function and other details of calpain, see recent reviews and the references cited therein [12,13].

2. Evolution of the calpain large subunit family

The six family members including the two species (μ- and m-types) thus far identified are named following the nomenclature described in [11], and listed in Fig. 1 and Table 1. n-Calpains-1, -2 and -2′ are assumed to be enzymes containing nCL-1 (p94), -2 and -2′ as the catalytic large subunits, respectively, but their regulatory sub-

*Corresponding author. Fax: (81) (3) 3813 0654.

SSDI 0014-5793(94)00278-4

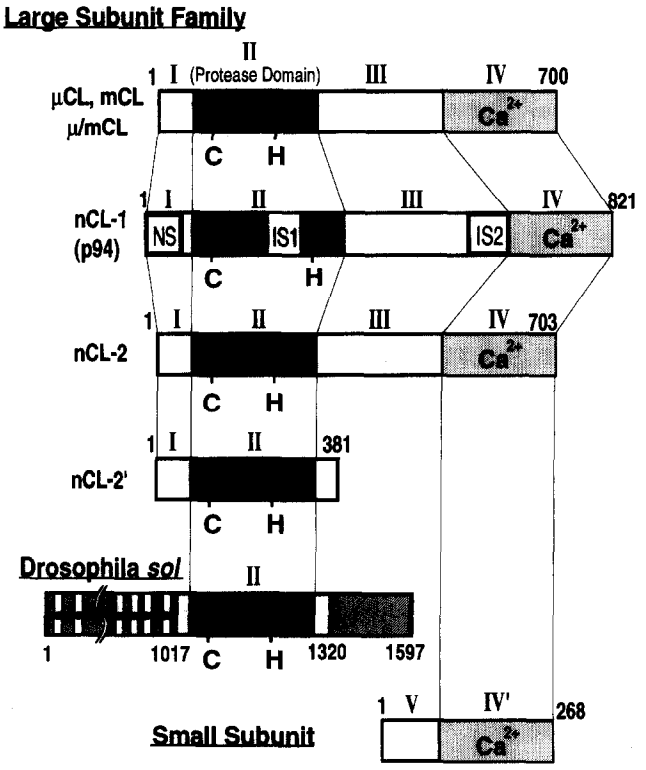

Fig. 1. Schematic structure of calpain subunits. Arabic and Roman numerals shown for each structure indicate amino acid residue No and domains, respectively. Ca^{2+} stands for the region containing Ca^{2+}-binding structures. C and H in the protease domain (Domain II) indicate the active site cysteine and histidine residues. Open boxes with ● in the structure of *Sol* stand for zinc-finger structures. Shadowed boxes show no similarity to calpain large subunits [18].

units, if any, have not yet been identified. Since the large subunits of μ- and m-calpains, μCL and mCL, can express protease activity, at least partly, by themselves in vitro, the possibility that the calpain large subunit functions in vivo without the small subunit cannot be excluded. Thus, the functional enzyme, x-calpain, and the large subunit, xCL (where x = μ, μ/m, m, or n) are distinct, but they might be identical in certain situations.

Before discussing the details of each family member, it is worthwhile to consider the evolutionary tree of the calpain family to provide a general view. Fig. 2 shows the evolutionary relationship between family members of various calpain large subunits constructed on the basis of amino acid sequence. The family members can be clearly divided into five sub-families excluding schistosome CL, which is somewhat far from animal CLs. Members of the same sub-family from different species are much closer than members of different sub-families from the same species. For example, the relative distance (RD) between chicken and human p94 is 21.2, while the RD between human mCL and μCL is 40.4. In other words, each xCL forms an evolutionary island, implying that each species of calpain possesses its own role in the cellular proteolytic system resulting in conservation of the structure. Since only limited information about calpains from lower animals and plants is available at present, further comparative studies are required to shed light on more detailed evolutionary aspects of the calpain family.

3. Skeletal muscle-specific p94 (nCL-1)

As shown in Fig. 1, p94 contains unique NS, IS1, and IS2 regions not found in μCL and mCL [10]. These three regions have no wide-range similarity to any sequences available at present. The IS2 region contains a sequence similar to the nuclear translocation signal, PxKKKKxKP, that is conserved in human, rat and chicken p94 [6,14]. Although the mRNA for p94 exists abundantly in skeletal muscle, its protein cannot be identified. Expression experiments in cultured cells have elucidated the reason [15]. The wild type p94 protein can be hardly detected when expressed in COS and L8 cells, but point mutants of the active site Cys or His residue are produced normally, revealing that p94 autolyzes very quickly after translation. In vitro transcription and translation experiments confirmed the fact showing a half-time for turnover of roughly 30 min in the system used. This very rapid autolysis is prevented by deletion of IS2, but not IS1, suggesting the following possibilities: (i) IS2 regulates protease activity directly, or (ii) indirectly by inducing conformational changes, or (iii) IS2 contains a signal for rapid degradation. When expressed in COS and L8 cells, p94 localizes in the nucleus possibly by the nuclear localization signal in IS2. p94 is also observed along the cytoskeleton and in cytosol. p94 might change its localization according to cell cycle and be involved in muscle differentiation by interacting with the MyoD family [16], as conventional calpain regulates transcription by controlling the level of c-Jun [8].

4. Stomach-specific nCL-2 family

nCL-2 is also a tissue-specific calpain, but unlike p94 has no insertion sequences similar to NS, IS1 or IS2 [11]. Further, as shown in Fig. 2, nCL-2 is more similar to mCL (RD = 39.9) than p94 (RD = 49.7).

A most significant feature is that by alternative splicing the same gene generates mRNAs for not only nCL-2 but also nCL-2′, which lacks the Ca^{2+}binding domain as shown in Fig. 1. Since the calpain small subunit binds to the large subunit through domain IV, nCL-2′ cannot associate with the small subunit [17]. nCL-2′ thus seems to be active without Ca^{2+}, and provides an example suitable for analyzing the functions of the protease domain (II), Ca^{2+}-binding domain (IV), and small subunit.

Interestingly in this respect, nCL-2′ is similar to the

calpain homologous region of the drosophila *sol* protein (res. 1017–1320) [18]. *Sol* (Small Optic Lobes), a DNA-binding regulatory protein with 6 zinc-finger structures in the nucleus, has an alternatively spliced form without the region similar to calpain, and is involved in the formation of drosophila optic lobes. The position where *Sol* loses its similarity to calpain is near the C-terminus of nCL-2′. This suggests that domain II with slight modification forms a module structure and functions as a protease by itself. A homologue to *Sol* has not yet been identified other than in drosophila, but it is worth searching for since drosophila has two other calpains in addition to *Sol* [19].

5. The third ubiquitously expressed species – Intermediate type

In 1984, the cDNA for the chicken calpain large subunit was first cloned, and the complete primary structure was determined [7]. This molecular species has long been regarded as 'm'-type, because the second species corresponding to mammalian μ-type was later identified at the protein level in chicken [20]. Quite recently, however, we cloned a novel chicken cDNA clone distinct from chicken 'm'CL, μCL and p94. The encoded protein comprising 699 amino acid residues shows 80% sequence identity with human mCL, whereas the 'm'CL first cloned shows 66% identity. This indicates that the newly identified molecule, rather than the old 'm'CL, must now be called the large subunit of chicken m-calpain on the basis of sequence identity. The old 'm'-calpain with intermediate sequence identity and calcium sensitivity between mammalian μ- and mcalpains must be named intermediate type or 'μ/m-calpain'. The newly identified mCL originally found in the chicken muscle cDNA library is expressed ubiquitously in all tissues like μCL and 'm'CL (or μ/mCL).

In summary, three ubiquitously expressed calpain large subunits μCL, mCL and μ/mCL have been identified in chicken. Among them μ- and μ/m-calpains have been extensively studied at the protein level, but m-cal-

Table 1. Members of the calpain family

Nomenclature is according to ref. 11 with slight modifications. N.D., not determined; Hm, Human; Rb, Rabbit; Rt, Rat; Ck, Chicken; Bv, Bovine; Pr, Porcine.

Class	Tissue Expressed	Functional Enzyme	[Ca^{2+}] Sensitivity	Subunit Structure	Gene	Species Identified
Ubiquitous or Conventional	ubiquitous	**μ-calpain**	μM	**μCL**	*cls1*	Hm[32], Rb[33], Rt*, Ck*
				+ C30	*css1*	Hm[34], Rb[35], Rt*, Bv[36], Pr[37]
		μ/m-calpain	μ/mM	**μ/mCL**	*cls5*	Ck[7]
				+ C30	*css1*	same as C30 shown for mCL
		m-calpain	mM	**mCL**	*cls2*	Hm[38], Rb[33], Rt[39], Ck*
				+ C30	*css1*	same as C30 shown for mCL
Tissue-Specific or Novel	Skeletal Muscle	**n-calpain-1**	(nM)	**p94 (nCL-1)**	*cls3*	Hm[10], Rt[10], Ck*
				+ N.D.	N.D.	N.D.
	Stomach	**n-calpain-2**	N.D.	**nCL-2**	*cls4*	Rt[11]
				+ N.D.	N.D.	N.D.
		n-calpain-2'	-	**nCL-2'**	*cls4*	Rt[11]

* Unpublished. Manuscript is in preparation.

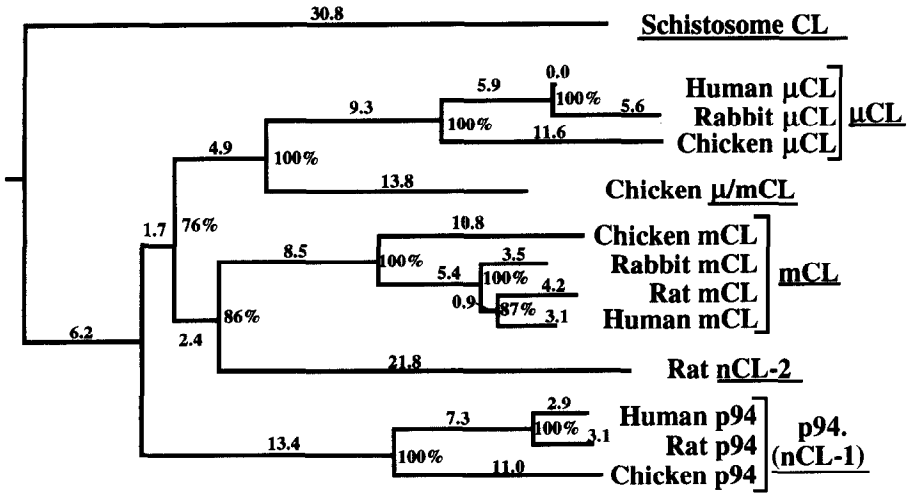

Fig. 2. Phylogenetic tree of calpain large subunits. Genetic distance was calculated by Clustal V Multiple Sequence Alignment Software using the neighbor-joining method [29,30]. Lengths of horizontal lines represent calculated relative distance (RD); vertical lengths have no meaning. The RD between two CLs can be calculated by adding the numbers on the lines connecting the two CLs. For example, the RD between chicken p94 and human p94 is 11.0 + 7.3 + 2.9 = 21.2, and the RD between chicken p94 and μ/mCL is 11.0 + 13.4 + 1.7 + 4.9 + 13.8 = 44.8. Percentages indicated at branches stand for occurrence (%) in bootstrap sampling [31]. Schistosome (*S. mansoni*) CL is from ref. [24] and all other sequences are from references listed in Table 1.

pain has not yet been identified in chicken tissues. A third Ca^{2+}-dependent protease activity was found in chicken muscle extract, but it remains to be known whether this activity corresponds to the newly found m-calpain [21]. Although μ/m-calpain is known only in chicken at present, the high conservation of other calpain species among mammals and birds strongly suggests its presence in mammals and other animals. Further precise studies are required to confirm μ/m-calpain in other species together with its characterization as an enzyme.

6. Molecular diversity of calpain and perspectives

Calpain constitutes a family comprising at least six homologous molecular species. In skeletal muscle, for example, four members, μ-, m-, μ/m-types and p94, are expressed. Why are so many similar calpain species required? Analysis of the total activity as a mixture of the four members, as well as differences in properties among molecular species, will reveal the reason for the family and will be important for clarifying the physiological function.

Calpain is an inactive proenzyme. Various lines of evidence indicate that pro-calpain is activated at the biological membrane in the presence of calcium and phospholipids such as phosphatidylinositol-bis-phosphate (PIP_2), especially when intracellular Ca^{2+} concentrations increase by stimulation of cell surface receptors. In the presence of PIP_2, μ-calpain is active at 100 nM Ca^{2+}, whereas m-calpain requires ca. 10 μM or higher Ca^{2+} [22]. Autolysis of p94 is observed even without the addition of Ca^{2+}, suggesting that p94 is active at very low Ca^{2+} concentrations. Further, it should be noted that the expression of mCL in cells is responsive to phorbol ester treatment, whereas that of μCL is not [23]. This suggests that m-calpain functions in response to outside signals, while μ-calpain plays house-keeping roles. Similar gene expression studies have not yet been done with other calpain species.

These differences in Ca^{2+} sensitivity and gene expression among calpain species indicate their functional divergence. The expression of tissue-specific calpain in skeletal muscle or stomach suggests that the presence of ubiquitous μ- and m-calpains is not enough, at least in certain tissues and cells. It is interesting to imagine that each tissue such as brain, the immune system, lung, heart, etc. has its own tissue-specific calpain in addition to ubiquitous calpains to complement their function.

One difficulty in analyzing the physiological function of calpains resides in their ubiquitous distribution. Diseases caused by abnormal levels of calpain are not known, presumably due to the indispensable function of calpain. Significant changes in calpain levels are lethal and not permissible. Discovery of tissue-specific calpains such as p94 and nCL-2 in other tissues will facilitate analysis of the physiological function. Gene disruption, which is not suitable for ubiquitous calpain, will be applicable to tissue-specific calpains.

Calpain has been identified in lower animals, although not many, such as schistosoma [24] and *Caenorhabditis elegans* [25], but is not found in yeast, bacteria, or plants [26]. The recent discovery in yeast of protein kinase C, one of the typical endogenous calpain substrates in ani-

mals, may be an indication of the existence of yeast calpain [27,28]. Comprehensive comparative studies to identify calpain in other organisms, such as yeast, bacteria, and plants, will provide a powerful tool for functional analysis, although the genome projects that are now proceeding energetically will eventually identify all family members in each species.

References

[1] Suzuki, K. and Ohno, S. (1990) Cell Structure Function 15, 1–6.
[2] Johnson, P. (1990) Int. J. Biochem. 22, 811–822.
[3] Mellgren, R.L. and Murachi, T. (1990) in: Intracellular Calcium-dependent Proteolysis, CRC Press, Boston, USA.
[4] Wang, K.K.W., Villalobo, A. and Roufogalis, B.D. (1989) Biochem. J. 262, 693–706.
[5] Mellgren, R.L. (1987) FASEB J. 1, 110–115.
[6] Sorimachi, H. and Suzuki, K. (1992) Biochim. Biophys. Acta 1160, 55–62.
[7] Ohno, S., Emori, Y., Imajoh, S., Kawasaki, H., Kisaragi, M. and Suzuki, K. (1984) Nature 312, 566–570.
[8] Hirai, S., Kawasaki, H., Yaniv, M. and Suzuki, K. (1991) FEBS Lett. 287, 57–61.
[9] Pontremoli, S., Melloni, E., Sparatore, B., Michetti, M., Salamino, F. and Horecker, B.L. (1990) J. Biol. Chem. 265, 706–712.
[10] Sorimachi, H., Imajoh-Ohmi, S., Emori, Y., Kawasaki, H., Ohno, S., Minami, Y. and Suzuki, K. (1989) J. Biol. Chem. 264, 20106–20111.
[11] Sorimachi, H., Ishiura, S. and Suzuki, K. (1993) J. Biol. Chem. 268, 19476–19482.
[12] Suzuki, K., Imajoh, S., Emori, Y., Kawasaki, H., Minami, Y. and Ohno, S. (1987) FEBS Lett. 220, 271–277.
[13] Saido, C.T., Sorimachi, H. and Suzuki, K. (1994) FASEB. J., in press.
[14] Sorimachi, H., Ohmi, S., Emori, Y., Kawasaki, H., Saido, T.C., Ohno, S., Minami, Y. and Suzuki, K. (1990) Biol. Chem. Hoppe-Seyler 371, 171–176.
[15] Sorimachi, H., Toyama-Sorimachi, N., Saido, T.C., Kawasaki, H., Sugita, H., Miyasaka, M., Arahata, K., Ishiura, S. and Suzuki, K. (1993) J. Biol. Chem. 268, 10593–10605.
[16] Weintraub, H. (1993) Cell 75, 1241–1244.
[17] Minami, Y., Emori, Y., Imajoh-Ohmi, S., Kawasaki, H. and Suzuki, K. (1988) J. Biochem. 104, 927–933
[18] Delaney, S.J., Hayward, D.C., Barleben, F., Fischbach, K.-F. and Miklos, G.L.G. (1991) Proc. Natl. Acad. Sci. USA 88, 7214–7218.
[19] Pinter, M., Stierandova, A. and Friedrich, P. (1992) Biochemistry, 31, 8201–8206.
[20] Murakami, T., Ueda, M., Hamakubo, T. and Murachi, T. (1988) J. Biochem. 103, 168–171.
[21] Wolfe, F.H., Sathe, S.K., Goll, D.E., Kleese, W.C., Edmunds, T. and Duperret, S.M. (1989) Biochim. Biophys. Acta 998, 236–250.
[22] Saido, T.C., Shibata, M., Takenawa, T., Murofushi, H. and Suzuki, K. (1992) J. Biol. Chem. 267, 24585–24590.
[23] Hata, A., Ohno, S. and Suzuki, K. (1992) FEBS Lett. 304, 241–244.
[24] Andresen, K., Tom, T.D. and Strand, M. (1991) J. Biol. Chem. 266, 15085–15090.
[25] Waterston, R., Martin, C., Craxton, M., Huynh, C., Coulson, A., Hillier, L., Durbin, R., Green, P., Shownkeen, R., Halloran, N., Metzstein, M., Hawkins, T., Wilson, R., Berks, M., Du, Z., Tomas, K., Thierry-Mieg, J. and Sulston, J. (1992) Nature Genetics 1, 114–123.
[26] Wolfe, F.H., Szpacenko, A., McGee, K. and Goll, D.E. (1989) Life Sci. 45, 2093–2101.
[27] Levin, D.E., Fields, F.O., Kunisawa, R., Bishop, J. M. and Thorner, J. (1990) Cell 62, 213–224.
[28] Toda, T., Shimanuki, M. and Yanagida, M. (1993) EMBO J. 12, 1987–1995.
[29] Higgins, D.G. and Sharp, P.M. (1988) Gene 73, 237–244.
[30] Saitou, N. and Nei, M. (1987) Mol. Biol. Evol. 4, 406–425.
[31] Felsenstein, J. (1985) Evolution 39, 783–791.
[32] Aoki, K., Imajoh, S., Ohno, S., Emori, Y., Koike, M., Kosaki, G. and Suzuki, K. (1986) FEBS Lett. 205, 313–317.
[33] Emori, Y., Kawasaki, H., Sugihara, H., Imajoh, S., Kawashima, S. and Suzuki, K. (1986) J. Biol. Chem. 261, 9465–9471.
[34] Ohno, S., Emori, Y. and Suzuki, K. (1986) Nucleic Acid Res. 14, 5559.
[35] Emori, Y., Kawasaki, H., Imajoh, S., Kawashima, S. and Suzuki, K. (1986) J. Biol. Chem. 261, 9472–9476.
[36] McClelland, P., Lash, J.A. and Hathaway, D.R. (1989) J. Biol. Chem. 264, 17428–17431.
[37] Sakihama, T., Kakidani, H., Zenita, K., Yumoto, N., Kikuchi, T., Sasaki, T., Kannagi, R., Nakanishi, S., Ohmori, M., Takio, K., Titani, K. and Murachi, T. (1985) Proc. Natl. Acad. Sci. USA 82, 6075–6079.
[38] Imajoh, S., Aoki, K., Ohno S., Emori, Y., Kawasaki, H., Sugihara, H. and Suzuki, K. (1988) Biochemistry 27, 8122–8128.
[39] DeLuca, C.I., Davies, P.L., Samis, J.A. and Elce, J.S. (1993) Biochim. Biophys. Acta 1216, 81–93.