Biochem. J. (1997) **328**, 717–720 (Printed in Great Britain)

BIOCHEMICAL JOURNAL LETTERS

A new class of cytochrome b_5 fusion proteins

Cytochrome b_5 is a small (M_r 14000) haem-binding protein usually associated with the endoplasmic reticulum of higher plants, fungi and animals. In animals, cytochrome b_5 is the ultimate electron donor in a microsomal electron-transport chain from NAD(P)H via cytochrome b_5 oxidoreductase; hence it is involved in oxidation–reduction reactions, such as the desaturation of acyl-CoA substrates, the reduction of cytochrome P-450 species and cholesterol biosynthesis [1]. In higher plants, cytochrome b_5 has been clearly shown to be involved in the desaturation of acyl-complex lipids in oilseeds [2], as well as the related process of fatty acid hydroxylation [3]; it may also have roles in sterol biosynthesis and cytochrome P-450 reactions. Recently, cytochrome b_5 domains or folds have been found in a number of unrelated proteins, such as nitrate reductase, sulphite oxidase and L-lactate dehydrogenase [4]. These domains include the diagnostic cytochrome b_5 motif, His-Pro-Gly-Gly-Xaa$_8$-Gly-Xaa$_6$-Phe-Xaa$_{3–6}$-His, where His-Pro-Gly-Gly is the haem-binding domain. This has led to the hypothesis that such proteins may have evolved by recombination of the enzymes with an ancestral haemoprotein.

Recently we isolated a cDNA clone encoding the microsomal fatty acid Δ^6-desaturase from developing seeds of borage (*Borago officinalis*; also known as starflower) [5]. This cDNA was functionally characterized by ectopic expression in transgenic tobacco (*Nicotiana tabacum*) plants, which resulted in the accumulation of two types of Δ^6-unsaturated fatty acids [γ-linolenic acid ($C_{18:3}\Delta^{6,9,12}$) and octadecatetraenoic acid ($C_{18:4}\Delta^{6,9,12,15}$)] not present in wild-type or control tobacco lines. However, the protein encoded by the Δ^6-desaturase cDNA was unusual in that it contained a sequence related to cytochrome b_5 at the N-terminus. The diagnostic cytochrome b_5 motif was present, with the His-Pro-Gly-Gly box at residues 41–44 of the polypeptide, which was 448 amino acids in length. The microsomal form of cytochrome b_5 cloned from borage was 32% identical with the N-terminal extension of the Δ^6-desaturase, with the haem-binding domain at residues 41–44 of the 132-amino-acid polypeptide. Microsomal fatty acid desaturases (Δ^{12} and Δ^{15}), which have been previously characterized from higher plants, do not contain this cytochrome b_5 domain (see [6] for example), even though the requirement for cytochrome b_5 has been demonstrated *in vitro* [2,3]. The borage Δ^6-desaturase, therefore, represents a new class of plant fatty acid desaturase enzymes, and also perhaps a new class of cytochrome b_5 fusion proteins.

The presence of a cytochrome b_5 extension in fatty acid desaturase proteins is, however, not without precedent. Deletion of the yeast (*Saccharomyces cerevisiae*) microsomal cytochrome b_5 gene failed to produce the expected fatty acid auxotroph predicted by the requirement of cytochrome b_5 for desaturation [7], even though this is a single-copy gene. It was also observed that rescue of deletion mutants of the yeast microsomal Δ^9-desaturase gene (*ole1*) by the rat Δ^9-desaturase cDNA required the presence of a functional microsomal cytochrome b_5 gene, whereas this was not required for complementation by *OLE1* itself [8]. Comparison of the sequence of the yeast microsomal Δ^9-desaturase with the corresponding rat microsomal Δ^9-desaturase showed that the yeast OLE1 sequence was $\approx$ 100 residues larger [8]. The additional sequence showed homology with cytochrome b_5 and contained invariant residues characteristic of the haem-binding domain [4,8]. Since OLE1 complements *ole1* mutants in a cytochrome b_5-deletion strain, unlike the rat Δ^9 desaturase, this strongly implies that the 100-residue extension present in OLE1 functions as the electron donor, alleviating or replacing the need for cytochrome b_5. However, in the yeast OLE1 *microsomal* Δ^9-desaturase this cytochrome b_5 domain was present as a C-terminal extension, with the His-Pro-Gly-Gly box at residues 444–447 of the 510-amino-acid *ole1* gene product. This contrasts with the N-terminal extension in the higher-plant microsomal Δ^6-desaturase. Subsequently, putative C-terminal cytochrome b_5 domains have been found in a number of homologous fungal microsomal Δ^9-desaturases (see, e.g. [9]). Although genetic evidence in *S. cerevisisiae* indicates that this domain can allow the desaturase to function in the absence of a microsomal cytochrome b_5, it not clear whether this domain is essential for enzymic function in cells containing endogenous cytochrome b_5, or whether the domain is able to act in *a trans* manner.

Using the borage Δ^6-desaturase sequence [5] to search databases, a related sunflower (*Helianthus annuus*) cDNA sequence was found (60% identity; haem-binding domain at residues 51–54 of a 458-amino-acid polypeptide), which had previously been characterized on the basis of an N-terminal domain with homology to cytochrome b_5 [10]. The sunflower cDNA showed homology across the entire sequence to the borage Δ^6-desaturase. It is unlikely, however, that the sunflower cDNA encodes a Δ^6-desaturase, since this species does not contain Δ^6-unsaturated fatty acids, and the function of this cDNA remains unknown. The presence of related sequences in other plant species is also indicated by a number of ESTs from *Arabidopsis thaliana* (thale cress), which also does not accumulate Δ^6-unsaturated fatty acids. It seems highly likely, therefore, that plants contain a family of enzymes, at least one of which is a fatty acid Δ^6-desaturase, characterized by having N-terminal extensions related to cytochrome b_5. That these domains are indeed functional haem-binding domains has been clearly demonstrated in the case of the sunflower protein, where the first 122 residues were expressed in *Escherichia coli* and the spectral properties examined [10]. When compared with the spectra produced by recombinant tobacco cytochrome b_5 (as determined by us in a previous study [11]), the recombinant sunflower protein exhibited a very similar redox absorbance spectra, with absorbance maxima at 426, 527 and 557 nm.

It is intriguing that none of the other higher-plant microsomal desaturases [6] (Δ^{12} and Δ^{15}) or the related oleate hydroxylase [12] (which introduces a hydroxy group at the Δ^{12} position to form ricinoleic acid) have a cytochrome b_5 domain, presumably using the 'free' microsomal form of the protein. Whether it is more efficient to have an endogenous cytochrome b_5 haem-binding domain is not known, though it is noteworthy that the Δ^{12}- and Δ^{15}- desaturases are much more prevalent than Δ^6-desaturases in the Plant Kingdom. One possible explanation for the differences between the plant microsomal desaturases may be due to the fact that the borage enzyme carries out a 'front-end'

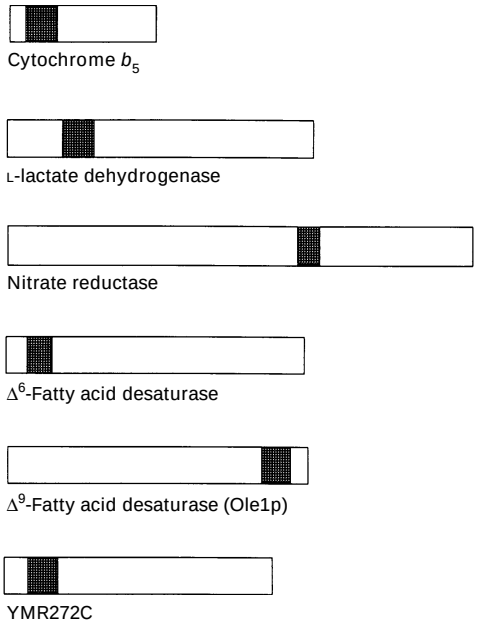

Figure 1 The positions of the diagnostic haem-binding domain in various members of the 'cytochrome b_5 fusion protein' family

YMR272C is an ORF predicated by the sequence of the yeast genome. The haem-binding domain is shown by ▦.

desaturation reaction compared with the Δ^{12} and Δ^{15} desaturases. Double bonds between C-3 and C-7 from the COOH group of a fatty acid are usually introduced as the final desaturation reaction: this is carried out by the so called 'front-end desaturases' [13]. The borage Δ^6-desaturase sequence [5] is the only current example of a plant microsomal 'front-end' desaturase, so it will be of interest to examine the sequences of similar microsomal desaturases, as they become available, for the presence of a haem-binding domain.

Since the yeast microsomal Δ^9-desaturase contains a C-terminal cytochrome b_5 domain, a computer-assisted examination of the open reading frames (ORFs) encoded by the yeast genome was carried out, using the haem-binding domain [4] as a probe. This revealed the presence of a number of proteins containing cytochrome b_5-like domains, most notably ORF YMR272C, which shows weak similarity to 3-oxo-5-α-steroid 4-dehydrogenase. This ORF contains a putative N-terminal cytochrome b_5 domain (haem-binding domain at residues 55–58 of the 385-amino-acid polypeptide), indicating the possible presence in different yeast enzymes of both N-terminal and C-terminal cytochrome b_5 domains. It may be that these examples of cytochrome b_5 fusion proteins have arisen at a later date than some of the other cytochrome b_5-like proteins such as nitrate reductase and L-lactate dehydrogenase. The position of the haem-binding domains in these proteins are more central; residues 571–574 of the 911-amino-acid tomato (*Lycopersicon esculentum*) nitrate reductase polypeptide and residues 123–126 of the 591-amino-acid yeast lactate 2-mono-oxygenase polypeptide. This is shown diagrammatically in Figure 1.

Clearly the new class of cytochrome b_5 fusion proteins described above is of great interest, especially in relation to their role in fatty acid desaturation reactions. Further examples need to be identified and characterized in order to understand more precisely their specialized features, mode of action and possible evolutionary relationships, and also to verify if they are confined to 'front-end'-type reactions.

Note added in proof (received 17 October 1997)

Since this manuscript was accepted for publication it has come to our attention that the yeast ORF YMR272C has been identified as a ceramide hydroxylase involved in sphingolipid metabolism. This gene has been designated SCS7 (updated entry on ORF YMR272C on the Yeast Protein Database site: http://quest7.proteome.com/YPD/SCS7.html14). Therefore SCS7 is the second example of a fatty acid desaturase/hydroxylase having an N-terminal cytochrome b_5 domain.

IACR-Long Ashton Research Station receives grant-aided support from the Biotechnology and Biological Sciences Research Council (U.K.).

Johnathan A. NAPIER*[1], Olga SAYANOVA*†, A. Keith STOBART† and Peter R. SHEWRY*

*IACR-Long Ashton Research Station, Department of Agricultural Sciences, University of Bristol, Long Ashton, Bristol BS18 9AF, U.K., and †School of Biological Sciences, University of Bristol, Woodland Road, Bristol BS8 1UG, U.K.

[1] To whom correspondence should be addressed.

1 Ozols, J. (1989) Biochim. Biophys. Acta **997**, 121–130
2 Smith, M. A., Cross, A. R., Jones, O. T. G., Griffiths, W. T., Stymne, S. and Stobart, A. K. (1990) Biochem. J. **272**, 23–29
3 Smith, M. A., Jonsson, L., Stymne, S. and Stobart, A. K. (1992) Biochem. J. **287**, 141–144
4 Lederer, F. (1994) Biochimie **76**, 674–692
5 Sayanova, O., Smith, M. A., Lapinskas, P., Stobart, A. K., Dobson, G., Christie, W. W., Shewry, P. R. and Napier, J. A. (1997) Proc. Natl. Acad. Sci. U.S.A. **94**, 4211–4216
6 Okuley, J., Lightner, J., Feldmann, K., Yadav, N., Lark, E. and Browse, J. (1994) Plant Cell **6**, 147–158
7 Truan, G, Epinat, J.-C., Rougelle, C., Cullin, C. and Pompon, D. (1994) Gene **149**, 123–127
8 Mitchell, A. G. and Martin, C. E. (1995) J. Biol. Chem. **270**, 29766–29772
9 Meesters, P. A. E. P., Springer, J. and Eggink, G. (1997) Appl. Microbiol. Biotechnol. **47**, 663–667
10 Sperling, P., Schmidt, H. and Heinz, E. (1995) Eur. J. Biochem. **232**, 798–805
11 Smith, M. A., Napier, J. A., Stymne, S., Tatham, A. S., Shewry, P. R. and Stobart, A. K. (1994) Biochem J. **303**, 73–79
12 van de Loo, F. N., Broun, P., Turner, S. and Somerville, C. R. (1995) Proc. Natl. Acad. Sci. U.S.A. **92**, 6743–6747
13 Aitzetmuller, K. and Tseegsuren, N. (1994) J. Plant Physiol. **143**, 538–543

Received 5 August 1997

A growing family of the Ca^{2+}-binding proteins with five EF-hand motifs

The EF-hand motif, the Ca^{2+}-binding helix–loop–helix structure, has been identified in numerous Ca^{2+}-binding proteins [1]. The number of repetitive EF-hand motifs in protein molecules, regardless of whether they are capable of Ca^{2+} binding, ranges from two to eight (see [1] and references cited therein). From X-ray-crystallographic analyses, we and others have recently independently revealed that the Ca^{2+}-binding domains of pig and rat calpain small subunits have five, instead of the previously believed four, EF-hand motifs, and that the calpain large subunits also possess the penta-EF-hand motifs based on the sequence alignment [2,3]. Whereas EF-1–EF-4 bind Ca^{2+} ions with high and low affinity, the C-terminal EF-5, has a two-residue insertion and does not bind Ca^{2+} ions. We have re-evaluated the reported primary structures of various proteins with multiple EF-hand motifs in comparison with the Ca^{2+}-binding domain of the calpain small subunit.

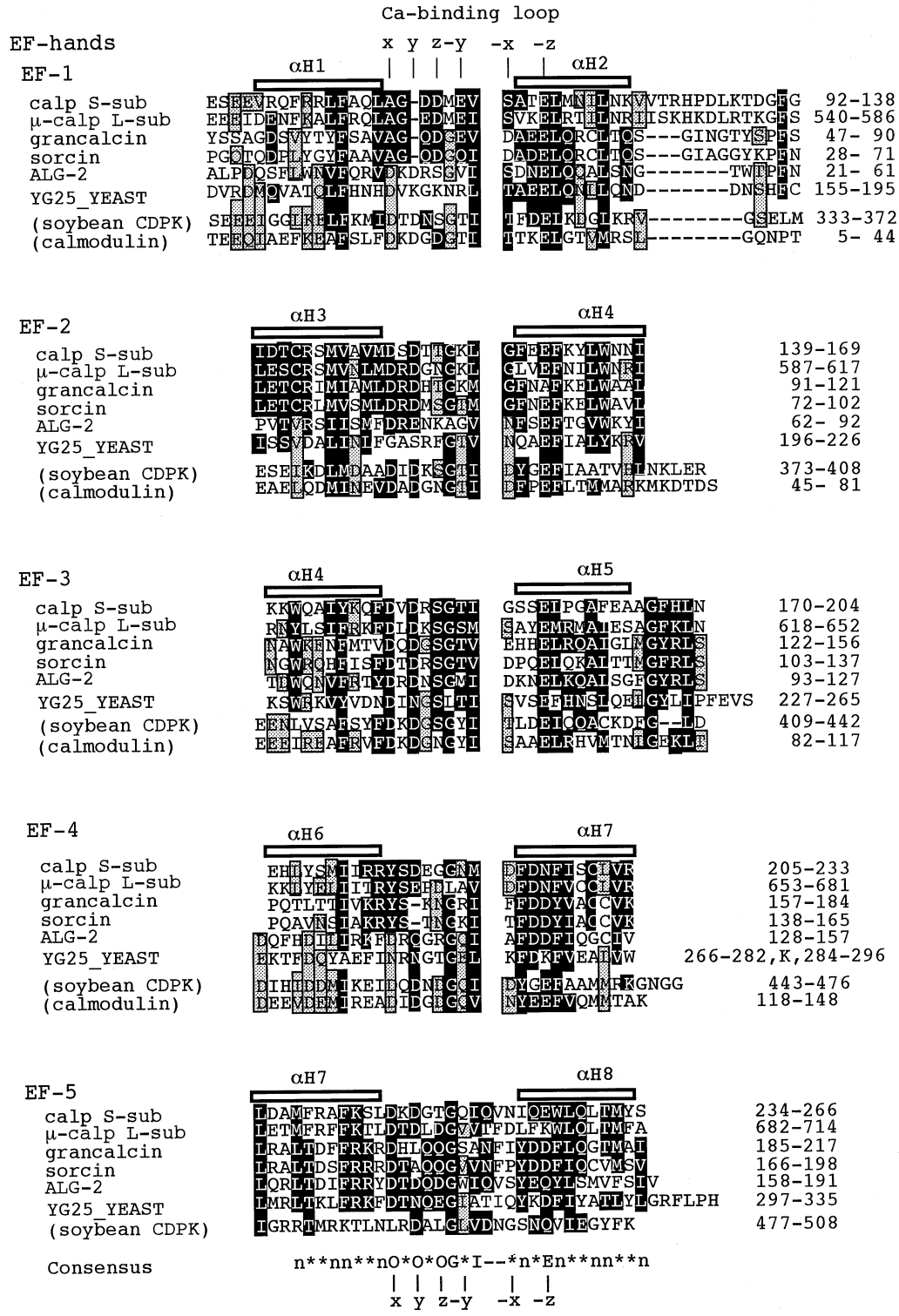

Figure 1 Sequence alignment of the penta-EF-hand motifs

The α-helices of the Ca^{2+}-bound pig calpain small subunit (calp S-sub) domain VI elucidated by X-ray crystallography [3] are indicated by open bars above the sequence (αH1–αH8). Ca^{2+}-co-ordinating positions are indicated by x, y, z, -y, -x, and -z. Identical or similar residues among the proteins with the penta-EF-hand motifs (PEF, top six sequences) and non-PEF proteins (CDPK and calmodulin) are indicated in two different patterns: highlighted (white letters on black), conserved in four or more PEF proteins; stippled, conserved in four or more proteins including non-PEF proteins. Amino acid similarity groupings are: aromatic (F, Y and W); hydrophobic large side chain (L, I, V and M); acidic and amide (E, D, Q and N); hydroxy small side chain (S and T). The amino acid sequence are from the Genbank/EMBL/DDBJ Nucleotide Sequence Databases (pig calpain small subunit, M1179; humani μ-calpain large subunit, X04366; mouse ALG-2, U49112; soybean CDPK, M64987) and the SwissProt Protein Sequence Database (human grancalcin, P28676; human sorcin, P30626; YG-YEAST, P53238). YG25-YEAST has a one-residue insertion of lysine at position 283, but this is not displayed in the Figure. The consensus sequence, modified from Ca^{2+}-[1], has preferred residues at n (hydrophobic), O (oxygen containing), I (Ile, Val, Leu or Met), G (Gly), and E (acidic or amide). Asterisks indicate variable residues (often hydrophilic). Blanks in EF-1–EF-4 indicate a two-residue insertion in EF-5 (hyphens in the consensus sequence).

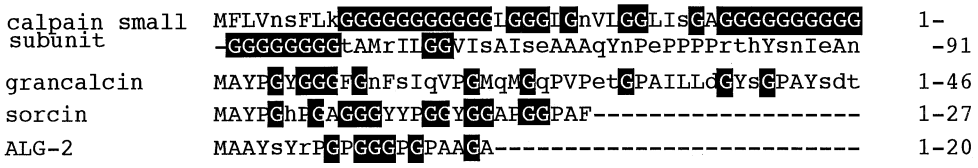

Figure 2 Sequence alignment of the N-terminal domains of the calpain small subunit, grancalcin, sorcin and ALG-2

Glycine residues are highlighted (white letters on black). Hydrophobic and hydrophilic residues are indicated by upper-case and lower-case letters respectively.

Sorcin, amplified together with P-glycoprotein gene in multi-drug-resistant cancer cells [4], and grancalcin, which is thought to be associated with granule–membrane fusion and degranulation of neutrophils [5], have been shown to be similar to the calpain small subunit. These two proteins also lack one residue in the Ca^{2+}-binding loop of the EF-1 region, as in the case of the calpain small subunit [2] (Figure 1). Here we report the existence of penta-EF-hand motifs in two additional proteins: the apoptosis-linked gene product ALG-2 [6] and an open reading frame of a yeast gene (a hypothetical protein of 38.4 kDa) (SwissProt accession no. P53238: YG25-YEAST). In contrast with calpains, sorcin and grancalcin, they do not have the one-residue deletion in the Ca^{2+}-binding loop of EF-1, but have two additional residues in the EF-5 region. The soybean (*Glyucine max*) Ca^{2+}-dependent protein kinase (CDPK) has a C-terminal extension following a calmodulin-like sequence [7]. Although high divergence in this region has been observed among the isoforms in thale cress (*Arabidopsis thaliana*) [8], a weak sequence similarity in the EF-5 region was also observed between the soybean CDPK and the calpain-small-subunit family. Like calmodulin, however, CDPK has a longer α-helix between EF-2 and EF-3, suggesting that it belongs to a family different from the one with the penta-EF-hand motifs.

The penta-EF-hand proteins, including ALG-2, have been shown to bind Ca^{2+} ions [6,9–11]. Whereas grancalcin and sorcin exist as homodiomers [10,12], calpain small subunits from various animals form heterodimers with large catalytic subunits [13]. The recombinant Ca^{2+}-binding domain of the calpain small subunit, however, forms a homodimer in the absence of the catalytic subunit [14]. One EF-5 pairs up with another EF-5 of the dimer counterpart, as revealed by X-ray crystallography [2,3]. Thus all the family members, probably including ALG-2, may form homo- or hetero-dimers through EF-5s in a Ca^{2+}-independent manner. Recently, sorcin has been shown to associate with cardiac ryanodine receptor, but the dimerization site remains to be established [15]. Ca^{2+}-induced translocation from cytoplasm to membranes is another common feature of this family [10,11,16]. The N-terminal regions of the calpain small subunit, grancalcin, sorcin and ARG-2 are rich in glycine and hydrophobic residues (Figure 2). As suggested for calpains [17], these hydrophobic N-terminal domains may also play an important role in their interactions with biological membranes. Thus the family members described in this Letter may function as hetero- or homo-dimers by interacting with target proteins on or near membranes in various aspects of cell regulation in response to Ca^{2+} ions.

Masatoshi MAKI*[1], Sthanam V. L. NARAYANA† and Kiyotaka HITOMI*

*Laboratory of Molecular and Cellular Regulation, Department of Applied Biological Sciences, School of Agricultural Sciences, Nagoya University, Furo-cho, Chikusa-ku, Nagoya 464-01, Japan and †Center for Macromolecular Crystallography, University of Alabama at Birmingham, BHS-286, 1918 University Blvd., Birmingham, AL 35294, U.S.A.

[1] To who correspondence should be addressed.

1 Kretsinger, R. H. (1996) Nat. Struct. Biol. **3**, 12–15
2 Blanchard, H., Grochulski, P., Li, Y., Arthur, S. C., Davies, P. L., Elce, J. S. and Cygler, M. (1997) Nat. Struct. Biol. **4**, 532–538
3 Lin, G., Chattopadhyay, D., Maki, M., Wang, K. K. W., Carson, M., Jin, L., Yuen, P., Takano, E., Hatanaka, M., DeLucas, L. J. and Narayana, S. V. L. (1997) Nat. Struct. Biol. **4**, 539–547
4 Van der Bliek, A. M., Meyers, M. B., Biedler, J. L., Hes, E. and Borst, P. (1986) EMBO J. **5**, 3201–3208
5 Boyhan, A., Casimir, C. M., French, J. K., Teahan, C. G. and Segal, A. W. (1992) J. Biol. Chem. **267**, 2928–2933
6 Vito, P., Lacanà, E. and D'Adamio, L. (1996) Science **271**, 521–525
7 Harper, J. F., Sussman, M. R., Schaller, G. E., Putnam-Evans, C., Charbonneau, H. and Harmon, A. C. (1991) Science **252**, 951–954
8 Hrabak, E. M., Dickmann, L. J., Satterlee, J. S. and Sussman, M. R. (1996) Plant Mol. Biol. **31**, 405–412
9 Minami, Y., Emori, Y., Kawasaki, H. and Suzuki, K. (1987) J. Biochem. (Tokyo) **101**, 889–895
10 Teahan, C. G., Totty, N. F. and Segal, A. W. (1992) Biochem. J. **286**, 549–554
11 Meyers, M. B., Zamparelli, C., Verzili, D., Dicker, A. P., Blanck, T. J. J. and Chiancone, E. (1995) FEBS Lett. **357**, 230–234
12 Hamada, H., Okochi, E., Oh-hara, T. and Tsuruo, T. (1988) Cancer Res. **48**, 3173–3178
13 Goll, D. E., Kleese, W. C., Okitani, A., Kumamoto, T., Cong, J. and Kapprell, H. (1990) in Intracellular Ca^{2+}-Dependent Proteolysis (Mellgren, R. L. and Murachi, T., eds.), pp. 3–24, CRC Press, Boca Raton, FL
14 Blanchard, H., Li, Y., Cygler, M., Kay, C. M., Arthur, J. S. C., Davies, P. L. and Elce, J. S. (1996) Protein Sci. **5**, 535–537
15 Meyers, M. B., Pickel, V. M., Sheu, S., Sharma, V. K., Scotto, K. W. and Fishman, G. I. (1995) J. Biol. Chem. **270**, 26411–26418
16 Mellgren, R. L. (1987) FASEB J. **1**, 110–115
17 Imajoh, S., Kawasaki, H. and Suzuki, K. (1986) J. Biochem. (Tokyo) **99**, 1281–1284

Received 18 August 1997