induced phosphorylation of polypeptides of M_r 20K in UCHT1 immunoprecipitates, but these were not necessarily specific for T3 as similar ^{32}P-labelled polypeptides were observed in HLA-A,B,C immunoprecipitates (Fig. 3*b*, tracks 5, 6).

In conclusion, these data show that stimulation of human T cells with PHA or antigen is associated with phosphorylation of the γ subunit of the T3 antigen. A possible candidate for the immune regulated kinase is protein kinase C since activators of protein kinase C such as phorbol esters also initiate phosphorylation of the T3 antigen[14,15]. The stimulation of protein kinase C moreover is anticipated as a normal consequence of T cell activation in that immune triggering of T cells is known to induce phosphatidylinositol breakdown and thus generate diacylglycerol, the proposed physiological activator of C kinase[25,26].

It has been shown previously that antigen and lectin stimulation of murine T cells are associated with phosphorylation of a M_r 20K-21K polypeptide that co-precipitates with the idiotypic antigen receptor[16,17]. A major question is whether, in spite of the apparent difference in M_r, this is analogous to the human T3 γ chain[16,17,27]. Interestingly, one difference between the murine and human systems is that in murine T cells, phorbol esters induce phosphorylation of a M_r 25K component which, because of its lack of *N*-glycosylation, could be the murine T3 ε subunit[16]. In the human there is no evidence for phorbol ester-induced *in vivo* ^{32}P-labelling of the T3 ε chain[15], although *in vitro* phosphorylation of the ε subunit can be readily demonstrated (unpublished data).

The functional relevance of intracellular signalling between protein kinase C and the T3 antigen is unclear. One possibility is that T3 γ chain phosphorylation is important in the primary activation pathways that initiate T cell growth and differentiation. It is noteworthy, however, that the immune activation of T cells by antigen, lectin, or phorbol ester is associated with a down-regulation of the surface levels of the Ti/T3 complex with a concomitant loss of antigen-regulated functions[28-31]. Thus, one function of protein kinase C mediated T3 γ chain phosphorylation could be as a negative feedback signal that controls the level of expression and/or functions of the Ti/T3 complex.

We thank Gail Essery-Rice and Carol Greenall for technical assistance, Dr G. Janossy for providing the monoclonal antibody RFT11 and Kim Richardson for typing the manuscript.

Received 28 August; accepted 25 November 1986.

1. Meuer, S. C. *et al. J. exp. Med.* **157**, 705-719 (1983).
2. Acuto, O. *et al. Cell* **34**, 717 (1983).
3. Meuer, S. C. *et al. Science* **222**, 1239-1241 (1983).
4. Meuer, S. C. *et al. Nature* **303**, 808-810 (1983).
5. Brenner, M. B., Trowbridge, J. S. & Strominger, J. L. *Cell* **40**, 183-190 (1985).
6. Kanellopoulus, J. M., Wigglesworth, N. M., Owen, M. J. & Crumpton, M. J. *EMBO J.* **2**, 1807-1814 (1983).
7. Borst, J., Alexander, S., Elder, J. & Terhorst, C. *J. biol. Chem.* **258**, 5135-5141 (1983).
8. Van Wauve, J. P., De Mey, J. R. & Goosens, J. G. *J. Immun.* **125**, 2708-2712 (1980).
9. Chang, T. W., Kung, P. C., Gingrus, S. P. & Goldstein, G. *Proc. natn. Acad. Sci. U.S.A.* **78**, 1805-1808 (1981).
10. Burns, G. R., Boyd, A. W. & Beverley, P. C. L. *J. Immun.* **129**, 1451-1457 (1982).
11. Reinherz, E. L. *et al. Cell* **30**, 735-743 (1982).
12. Oettgen, H. C., Terhost, C., Cantley, L. C. & Roscoff, P. M. *Cell* **40**, 583-590 (1985).
13. Roscoff, P. M. & Cantley, L. W. *J. biol. Chem.* **260**, 14053-14059 (1985).
14. Cantrell, D. A., Davies, A. A. & Crumpton, M. J. *Proc. natn. Acad. Sci. U.S.A.* **82**, 8158-8162 (1985).
15. Davies, A. A., Cantrell, D. A. & Crumpton, M. J. *Biosciences* **5**, 867-876 (1985).
16. Samelson, L. E., Harford, J. B., Schwartz, N. H. & Klausner, R. D. *Proc. natn. Acad. Sci. U.S.A.* **82**, 1969-1973 (1985).
17. Samelson, L. E., Harford, J. B. & Klausner, R. D. *Cell* **43**, 223-231 (1985).
18. Beverley, P. C. L. & Callard, R. E. *Eur. J. Immun.* **11**, 329-334 (1981).
19. Walsh, F. S. & Crumpton, M. J. *Nature* **269**, 307-311 (1977).
20. Barnstable, C. J. *et al. Cell* **14**, 9-20 (1978).
21. Chilson, O. P., Boylston, A. W. & Crumpton M. J. *EMBO J.* 3239-3245 (1984).
22. O'Flynn, K., Krensky, A. M., Beverley, P. C. L., Burakoff, S. J. & Lynch, D. C. *Nature* **313**, 686-687 (1985).
23. Ohashi, P. S. *et al. Nature* **316**, 606-609 (1985).
24. Lamb, J. R., Eckels, D. D., Lack, P., Woody, J. U. & Green, N. *Nature* **300**, 66-69 (1982).
25. Farrar, W. L. & Ruscetti, F. W. *J. Immun.* **136**, 1266-1273 (1985).
26. Imboden, J. B. & Stobo, J. O. *J. exp. Med.* **161**, 446-456 (1985).
27. Oettgen, H. C., Pettey, C. L., Maloy, L. & Terhorst, C. *Nature* **320**, 272-275 (1986).
28. Reinherz, E. L. *et al. Cell* **30**, 735-743 (1982).
29. Zanders, E., Lamb, J., Feldman, M., Green, N. & Beverley, P. *Nature* **313**, 318-320 (1983).
30. Delia, D. *et al. Int. J. Cancer* **29**, 23-31 (1982).
31. Ando, I., Hariri, G., Wallace, D. & Beverley, P. *Euro. J. Immun.* **15**, 196-199 (1985).

Use of a cDNA clone to identify a supposed precursor protein containing valosin

Kerry J. Koller* & Michael J. Brownstein

Laboratory of Cell Biology, NIMH, Bethesda, Maryland 20892, USA

Valosin, a novel 25-amino-acid peptide isolated recently from pig intestine[1], has several effects on the digestive system of dogs[2]. We report here that the valosin-specific complementary DNA clone from pigs codes for a polypeptide unlike most precursors of biologically active peptides. The predicted protein lacks a characteristic amino-terminal hydrophobic signal sequence and contains no processing signals of the type acted upon by endopeptidases to generate other active peptides from precursors. Antibodies to synthetic valosin have been used to show that nearly all valosin immunoreactivity is in the cytoplasm and that the protein detected (valosin-containing protein, VCP), although smaller than the predicted product of the cDNA sequence, is much larger than valosin. Valosin-specific messenger RNA is found in extracts from many pig tissues, which contrasts with the restricted occurrence expected of a biologically active peptide. We conclude that valosin is an artefact of the purification procedure and does not occur *in vivo*.

In recent years, Mutt and coworkers have isolated a number of biologically active peptides from pig gut extracts. These peptides have been identified by their biological activity or the presence of an amidated carboxy terminus[3]. The success of this approach is attested to by the large number of novel peptides that have been found, including peptide YY[4], peptide HI[5], galanin[6], cholecystokinin-33[7], secretin[8], vasoactive intestinal polypeptide[9] and motilin[10]. Many of the peptides discovered in the gut are also present in the brain illustrating the fact the transmitters used in the periphery are also often used in the central nervous system.

Recently, Schmidt *et al.* described the isolation, from porcine intestinal extracts, of valosin (Fig. 2*b*, residues 493-517), a novel 25-amino-acid peptide[1]. It was purified from side fractions of an earlier preparation of peptide HI and secretin by its chromatographic and ultraviolet (UV)-absorbance characteristics and amino-terminal sequence. Its amino-acid sequence had no convincing sequence similarity with any known gut hormones (R. F. Doolittle, Databasem, San Diego), and it has several biological activities in dogs namely, release of gastrin; augmentation of pentagastrin-induced gastric secretion; stimulation of pancreatic protein secretion; and suppression of the

	1				5				10			
Valosin:	Val	Gln	Tyr	Pro	Val	Glu	His	Pro	Asp	Lys	Phe	Leu

Deduced mRNA: 5' GTN- CAPu-TAPy-CCN - GTN- GAPu-CAPy-CCN-GAPy-AAPu-TTPy-PyTN 3'

Oligonucleotide probe: 3' CAT/G-GTC- ATG- GGT/G-CAT/G-CTT-GTT-GGT/G-CTG-TTT -AAG-GA 5'

cDNA: 5' GTC - CAG-TAT - CCT - GTG - GAG-CAC-CCA - GAC-AAA -TTC-CTC 3'

Fig. 1 Oligonucleotide probe for valosin. The first 12 amino acids in valosin were used to design the 35-residue oligonucleotide probe. All possible mRNAs encoding this portion of valosin are represented in the second line of the figure. N is A, C, G and T; Pu is A and G; Py is T and C. The 16-fold degenerate oligonucleotide (line 3) was synthesized on an Applied Biosystems 380A DNA synthesizer according to the manufacturer's instructions and purified on a preparative denaturing acrylamide gel. The resulting mixture of probes was labelled with [γ-^{32}P]ATP using T4 DNA kinase[21]. Line four shows the segment of the sequence of the isolated valosin cDNA complementary to the probe. Nucleotides that differed from the probe, which is 85.7% homologous to the cDNA, are underlined.

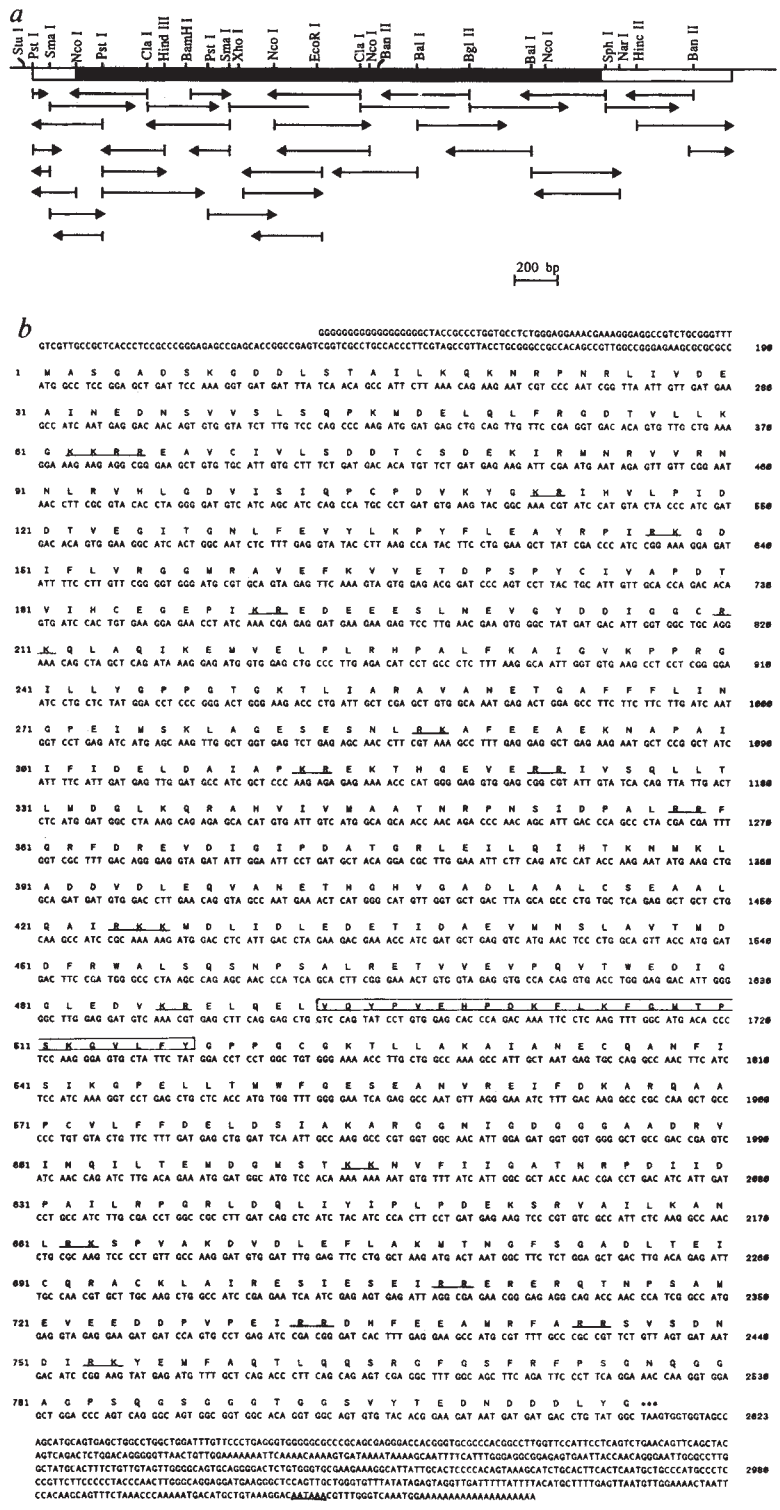

Fig. 2 Determination of the primary sequence of the valosin precursor mRNA. *a*, Restriction map and sequencing strategy used to characterize the largest valosin precursor clone, pcD-VQY 3-1. Horizontal arrows indicate the direction of sequencing and the location and length of sequenced regions. DNA fragments were subcloned into M13mp18 or 19, and the isolated single-stranded templates were used for sequencing by the dideoxy chain-termination method[22]. Open bar, total cDNA insert; dark bar, open reading frame. *b*, Nucleotide sequence of the cDNA insert from clone pcD-VQY 3-1 and the predicted amino-acid sequence of the valosin precursor. The sequence of the isolated valosin peptide is boxed. Possible paired basic amino-acid processing signals are underlined, as is the polyadenylation signal. Amino acids are numbered on the left and nucleotides on the right.

migrating myoelectric complexes of the small bowel in the fasting state[2].

Here we describe the characterization of the mRNA encoding valosin. An Okayama–Berg porcine adrenal medulla complementary DNA library[11] was screened with an oligonucleotide probe (Fig. 1) by filter hybridization[12]. Six positive clones were identified and purified by sequential low-density plating. The largest cDNA (pcD-VQY 3-1, 3.2-kilobase (kb) insert size) was studied in detail. The restriction map and sequencing strategy used to characterize pcD-VQY 3-1 are shown in Fig. 2*a*, and the nucleotide and corresponding deduced amino-acid sequences in Fig. 2*b*. The cDNA contains a 2,466 base pair (bp) open reading frame flanked by 143 and 550 bp of 5′ and 3′ untranslated sequences, respectively. The nucleotide and the predicted amino-acid sequences were not homologous to known sequences in either the Genebank or the NBRF Georgetown Protein databases.

Every tissue examined by Northern blot analysis had a single RNA message for valosin of ~3.4 kb (Fig. 3). Thus, clone pcD-VQY 3-1 is nearly full length. Message was found not only in porcine adrenal medulla, ileum and frontal cerebral cortex, but also in cerebellum, caudate, spinal cord, adrenal cortex, heart and liver. Valosin message was also present in many neuronal cells of porcine spinal cord, cerebellum and hippocam-

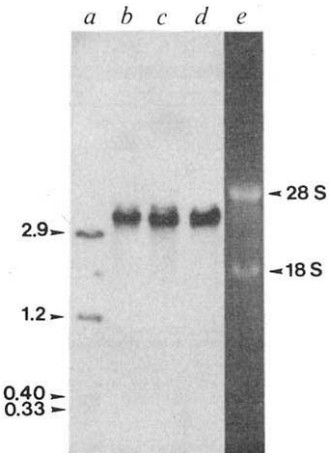

Fig. 3 Northern blot analysis of the tissue distribution of valosin mRNA. Lanes: *a*, θX 174 RF DNA digested with *Taq*I and end-labelled with ^{32}P-dCTP and large fragment DNA polymerase[21], denatured in formamide at 56 °C for 15 min before loading. Sizes in kb, *b*, ileum extract; *c*, adrenal medulla extract; *d*, frontal cerebral cortex extract; *e*, ethidium bromide staining of total ileum extract; RNA.

Methods. Porcine tissue (from the NIH animal facility, Poolesville, Maryland) was frozen on dry ice and stored at −70 °C until used for RNA extraction in guanidium thiocyanate[23]. Total RNA (5 μg) was suspended in denaturing buffer and electrophoresed on a 1% agarose formaldehyde denaturing gel and the RNA transferred to Gene Screen (NEN). The *Eco*RI-*Sph*I (1.3 kb) fragment (see Fig. 2*a*) was ^{32}P-labelled using the Feinberg random primer method[24] to a specific activity of 9.3×10^8 d.p.m. per μg. The filter was pre-hybridized for 8 h at 45 °C in 5 × SSPE, 1 × Denhardt's, 50% formamide, 0.1% SDS, 250 μg ml^{-1} tRNA, 124 μg ml^{-1} single-stranded, denatured salmon sperm DNA. After addition of denatured, labelled probe, hybridization was continued for 40 h in the same conditions. The filter was washed in 0.1 × SSPE, 0.1% SDS at 65 °C and placed against X-ray film in a cassette equipped with two intensifying screens for 36 h at −70 °C.

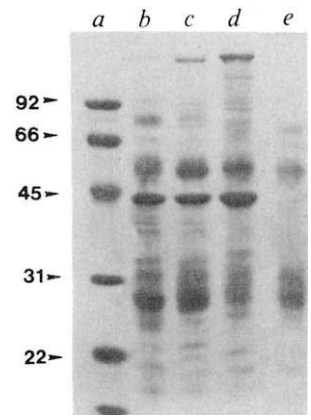

Fig. 4 Immunoprecipitation of valosin-containing proteins from porcine tissue. Lanes: *a*, Biorad protein size standard, M_r as shown; *b*, ileum; *c*, adrenal medulla; *d*, frontal cerebral cortex; *e*, control.

Methods. Frozen porcine ileum, adrenal medulla and frontal cerebral cortex were ground into powder under liquid N_2 and homogenized in RIPA buffer (150 mM NaCl, 1% Triton X-100, 1% sodium deoxycholate, 0.1% SDS, 10 mM Tris-HCl, *p*H 7.2) with 1% trasylol as protease inhibitor in a glass-glass homogenizer. For controls, RIPA buffer (1 μl) plus 1% trasylol without additional tissue was carried through the following procedures. Samples were spun at 100,000 *g* to clarify. Supernatants were removed and precleared with 200 μl of Staph A cell-suspension. The samples were incubated with 100 μl of valosin-specific protein A- and BSA-affinity purified serum (206 AP-686) on ice for 2.5 h. Staph A cells (200 μl) were added to the samples and incubated overnight at 4 °C. Samples were spun at 4 °C in a microcentrifuge for 10 min and pellets washed three times with RIPA buffer (1 ml) plus trasylol. The final pellet was suspended into 40 μl Laemelli sample buffer[25] with 5% β-mercaptoethanol and placed on ice for 1 h. Samples were boiled for 3 min, placed on ice for 5 min, and spun in microcentrifuge for 10 min at 4 °C. Supernatant was removed and samples (equivalent to 25 mg of soluble protein before immunoprecipitation) were run on a 10%, 1.5-mm polyacrylamide protein gel[25]. Gel was stained for 10 min in 0.25% Coomassie blue in 50% methanol, 10% acetic acid, and destained for 3 h in 20% methanol, 7% acetic acid.

pus, and in most cells of the adrenal gland, liver, and intestinal mucosa as determine by *in situ* hybridization histochemistry[13]. This almost ubiquitous distribution of valosin mRNA is unusual; a message that encodes a biological active peptide would be expected to be present in selected neurons and neuroendocrine cells.

The structure of the predicted valosin-containing protein (VCP) (Fig. 2*b*) is not typical for a peptide precursor. These proteins characteristically have a hydrophobic amino-terminal signal sequence required for their entry into the cisternal space of the rough endoplasmic reticulum from where the propeptide travels to the Golgi apparatus. In addition, within their precursors, the biologically active peptide or peptides are flanked by processing signals such as paired basic amino acids, at which a trypsin-like endopeptidase acts to liberate them. The cDNA sequence of VCP encodes a molecule of 806 amino acids (relative molecular mass (M_r) ~88,660 (Fig. 2*b*). The complete amino-acid sequence for valosin itself is encoded by nucleotides 1,667 to 1,741 (amino acids 493 to 517 of VCP). Valosin is not bracketed by canonical processing signals. Furthermore, although the amino terminus of VCP contains a few possible signal peptide cleavage sites[14], it is not very hydrophobic. Analysis of the complete amino-acid sequence of VCP using the algorithm of Hopp and Wood[15] revealed no putative internal signal sequences either. The above analyses suggest that VCP or its products are not secreted and that VCP is likely to be a cytoplasmic protein. Valosin-specific antiserum produced by injecting rabbits with synthetic valosin conjugated to bovine serum albumin (BSA) and purified on protein A- and BSA-affinity columns was used to immunoprecipitate subcellular fractions of porcine adrenal gland, and showed that >95% of the immunoreactivity was present in the cell cytoplasm (data not shown).

Although VCP does not appear to be a peptide precursor, it does seem to undergo post-translational modification. The mRNA predicts a protein product of M_r ~88,660. The *Stu*I-*Sph*I fragment of the cDNA encoding the complete open reading frame was subcloned into the GEM-3 riboprobe vector (Promega Biotec) and after *in vitro* transcription[16], the synthesized RNA was translated *in vitro* with nuclease-treated rabbit reticulocyte lysate (NEN)[17]. A number of proteins of different sizes were immunoprecipitated, presumably resulting from initiation of translation at various methionine residues. The largest protein band was of M_r ~90,000, the probable size of the initial protein translated *in vivo*, as the sequence flanking the proposed AUG initiating codon (GCCAUGG) is an optimal sequence for initiation of translation by eukaryotic ribosomes[18]. However, the protein species immunoprecipitated from tissue extracts of porcine adrenal medulla, ileum and frontal cerebral cortex was only of M_r 43,000 (Fig. 4). These results suggest that the VCP of M_r 90,000 is processed *in vivo* into a species of M_r 43,000 that is recognized by the valosin-specific antibody. We believe that valosin is a product of chymotryptic degradation of VCP and may have arisen in the course of extraction or purification.

Thus, our findings suggest that the recently described porcine gut peptide, valosin, is an artefactual product of a hitherto unknown cytoplasmic protein generated from a larger protein (VCP). Although preliminary experiments indicated that valosin may be biologically active, other protein degradation products

have also been shown to have biological activity (for example, fragments of the α chain of haemoglobin)[19,20], and care must be exercised before predicting that a novel peptide has a physiological role. Studies of cDNA of putative peptide 'precursors' may help identify strong candidates and eliminate weak ones.

We thank Anna Iacangelo, Alice Young and Drs Tom Bonner, Urs Affolter and Åke Rökaeus for technical insights; Dr Åke Rökaeus for providing the cDNA library; Dr Scott Young for his help with *in situ* hybridization histochemistry; Dr Wolfgang Schmidt for providing the complete structure of valosin before publication; Dr Dan Marshak and Joanne Gutierrez for providing the synthetic peptide; and Drs Dianne Rausch and Maribeth Eiden for critical review of the manuscript. K.J.K was supported by a National Institute of Neurological and Communicative Diseases and Stroke (NINCDS) NRSA postdoctoral grant no. NS07599.

Received 5 September; accepted 1 December 1986.

1. Schmidt, W. E. *et al. FEBS Lett.* **191,** 264–268 (1985).
2. Schmidt, W. E. *et al. Dig. Dis. Sci.* **29,** 75S (1984).
3. Tatemoto, K. & Mutt, V. *Proc. natn. Acad. Sci. U.S.A.* **75,** 4115–4119 (1978).
4. Tatemoto, K. *Proc. natn. Acad. Sci. U.S.A.* **79,** 2514–2518 (1982).
5. Tatemoto, K. & Mutt, V. *Proc. natn. Acad. Sci. U.S.A.* **78,** 6603–6607 (1981).
6. Tatemoto, K., Rökaeus, Å, Jörnvall, H., McDonald, T. J. & Mutt, V. *FEBS Lett.* **164,** 124–128 (1983).
7. Mutt, V. in *Gastrointestinal Hormones* (ed. Glass, G. B. J.) 169–180 (Raven, New York, 1980).
8. Jorpes, J. E., Mutt, V, Magnusson, S. & Steele, B. B. *Biochem. biophys. Res. Commun.* **9,** 275–278 (1962).
9. Said, S. I. & Mutt, V. *Eur. J. Biochem.* **28,** 199–204 (1972).
10. Brown, J. C., Cook, M. A. & Oryburgh, J. R. *Gastroenterology* **62,** 401–404 (1972).
11. Rökaeus, Å & Brownstein, M. J. *Proc. natn. Acad. Sci.U.S.A.* **83,** 6287–6291 (1986).
12. Grunstein, M. & Hogness, D. *Proc. natn. Acad. Sci. U.S.A.* **72,** 3961–3965 (1975).
13. Young, W. S. III, Mezey, E. & Siegel, R. E. *Neurosci. Lett.* **70,** 198–203 (1986).
14. Von Heijne, G. *Nucleic Acids Res.* **14,** 4683–4690 (1986).
15. Hopp, T. P. & Woods, K. R. *Proc. natn. Acad. Sci. U.S.A.* **78,** 3824–3828 (1981).
16. Melton, D. A. *et al. Nucleic Acids Res.* **12,** 7035–7056 (1984).
17. Pelham, H. R. B. & Jackson, R. J. *Eur. J. Biochem.* **67,** 247–256 (1976).
18. Kozak, M. *Cell* **44,** 283–292 (1986).
19. Schally, A. V. *et al. Biochem. biophys. Res. Comm.* **82,** 582–588 (1978).
20. Kiso, Y. *et al. FEBS Lett.* **155,** 281–284 (1983).
21. Maniatis, T., Fritsch, E. F. & Sambrook, J. *Molecular Cloning; A Laboratory Manual* (Cold Spring Harbor Laboratory, New York, 1982).
22. Sanger, F., Nicklen, S. & Coulson, A. R. *Proc. natn. Acad. Sci. U.S.A.* **74,** 5463–5467 (1977).
23. Chirgwin, J. M., Przybyla, A. E., MacDonald, R. J. & Rutter, W. J. *Biochemistry* **18,** 5294–5299 (1979).
24. Feinberg, A. P. & Vogelstein, B. *Analyt. Biochem.* **137,** 266–267 (1984).
25. O'Farrell, P. *J. biol. Chem.* **250,** 4007–4021 (1975).

Cloning of decay-accelerating factor suggests novel use of splicing to generate two proteins

Ingrid W. Caras, Michael A. Davitz*†, Lucy Rhee, Greg Weddell, David W. Martin, Jr & Victor Nussenzweig*

Genentech, Inc., 460 Point San Bruno Boulevard, South San Francisco, California 94080, USA
* Departments of Pathology and †Environmental Medicine, New York University School of Medicine, New York, New York 10016, USA

Decay-accelerating factor (DAF), a glycoprotein that is anchored to the cell membrane by phosphatidylinositol[1,2], binds activated complement fragments C3b and C4b, thereby inhibiting amplification of the complement cascade on host cell membranes[3–5]. Here, we report the molecular cloning of human DAF from HeLa cells. Analysis of DAF complementary DNAs revealed two classes of DAF messenger RNA, one apparently derived from the other by a splicing event that causes a coding frameshift near the C terminus. The apparent 'intron' sequence contains an *Alu* family member and encodes contiguous protein sequence. Two DAF proteins are therefore possible, having divergent C-terminal domains which differ in their hydrophobicity. Both mRNAs are found on polysomes, suggesting that both are translated. We propose that the major (90%) spliced DAF mRNA encodes membrane-bound DAF whereas the minor (10%) unspliced DAF mRNA may encode secreted DAF and we present expression data supporting this. The deduced DAF sequence contains four repeating units homologous to a consensus repeat found in a recently described family of complement proteins[6].

A 69-base synthetic oligonucleotide 'long-probe'[7,8] based on the N-terminal sequence of DAF[9] was used to screen a λgt10 cDNA library (1×10^6 recombinants) derived from HeLa cell poly $(A)^+$ RNA. Of several hybridizing clones, a single 1,395-base-pair (bp) clone was shown by DNA sequence analysis to encode the known amino-acid sequence. Northern analysis of DAF mRNA revealed two major mRNA bands of ~1.5 and ~2.2 kilobases (kb). We therefore rescreened the library and identified several longer clones. The complete sequence of the cDNA and deduced protein are shown in Fig. 1*a*. To confirm that this cDNA encodes DAF, we raised anti-peptide antibodies against the deduced peptides 1, 2, and 3 (Fig. 1*a*), and showed by immunoblotting that all three antibodies specifically recognize the relative molecular mass 70,000 (M_r 70 K) DAF protein purified from human erythrocytes (Fig. 1*b*).

Two classes of DAF cDNA clones were identified by DNA sequence analysis. These differed by an apparent internal deletion of 118 bp which probably represents a splicing event (see below). The translation initiation codon was assigned to the first in-frame ATG which is flanked by sequences that conform to Kozak's rules[10]. This predicts a hydrophobic signal peptide of 34 amino acids preceding the known N terminus of DAF. The deduced mature DAF contains four contiguous repeating units of ~63 amino acids (Fig. 1*c*) that conform to a consensus repeat sequence found in the complement proteins Factor B, C2, Factor H, C4b-binding protein, CR1, and Clr (all of which bind C3 or C4), and in a few non-complement proteins including the IL-2 (interleukin-2) receptor and a subunit of clotting factor XIII[6]. Outside the general consensus framework, the first and second repeats of DAF show significant homology (~38%), as do the third and fourth (~40%), but the homology of repeats 1 or 2 compared with 3 or 4 is weaker. This suggests that these repeats arose by sequential internal duplications. The deduced protein contains one site for *N*-linked glycosylation (Asn, residue 61). Several clusters of threonine and serine residues flank the repeats and may be sites of *O*-linked glycosylation[11,12].

Three out of seven clones sequenced contained a precise 118-bp deletion within the C-terminal portion of the coding region. The sequences bordering this deletion match the 5′ and 3′ consensus sequences for RNA splicing[13,14], suggesting that the deleted segment represents an unspliced intron. Interestingly, an *Alu* family sequence displaying an 80% homology with the *Alu* consensus[15] lies within this apparent intron (Fig. 1*a*). The intron contains no stop codons so that the unspliced DAF mRNA could encode a protein of 440 amino acids (including signal peptide) of which 39 residues (328–366) near the C terminus would be encoded by the intron. Removal of the intron by splicing causes a frameshift in the C-terminal coding region, resulting in a shorter protein of 381 amino acids with a different C terminus (Figs 1*a* and 2*a*). The hydropathy profiles of the two DAF proteins encoded by the spliced or unspliced DAF mRNAs differ markedly in the divergent C-terminal domains, being hydrophobic in the case of DAF deduced from the spliced mRNA (short protein), or hydrophilic in DAF from the unspliced mRNA (long protein) (Fig. 2*b*, *c*). This difference may have implications for the cellular localization of DAF.

Each class of DAF cDNA displays heterogeneity at the 3′ end. Four polyadenylation signals, AATAAA[16], are present in the 3′ noncoding region of the sequence. DNA sequence analysis of several independent clones revealed that at least the first and last of these sites are used, producing 3′ noncoding regions differing in size by 788 bp, resulting in the two major mRNA species (~2.2 and ~1.5 kb) identified by Northern analysis (Fig. 3*a*, lane 1). To confirm this, we rescreened the Northern blot