Molecular cloning of a novel N-terminal variant of annexin II from rat basophilic leukaemia cells

A. Louise UPTON and Stephen E. MOSS*

* Department of Physiology, University College London, Gower Street, London WC1E 6BT, U.K.

Rat annexin II cDNA clones were isolated from a rat basophilic leukaemia cell plasmid library by cross-species hybridization with a mouse probe, and fully sequenced using the dideoxy-chain-termination method. Alignment of the derived amino-acid sequence with those of other mammalian annexin II species revealed a high level of conservation, characteristic of the annexin family of proteins. One of the cDNAs isolated contained an additional six nucleotides close to the N-terminus, lying in-frame and at a point corresponding to an intron/exon boundary in the human annexin II gene. As the two rat cDNAs were identical apart from the six nucleotide insert, it is likely that these represent alternatively spliced transcripts of a single gene, rather than the products of two separate genes. The six nucleotides encode serine–glutamine and therefore introduce an additional potential phosphorylation site into a region already containing one tyrosine and two serine phosphorylation sites. The discovery of this novel annexin II variant may have important implications both for p11 binding and for regulation of annexin II function by phosphorylation.

INTRODUCTION

Annexin II (also known as calpactin 1) is a 36 kDa calcium-dependent phospholipid-binding protein of the annexin supergene family. The annexin family comprises at least ten mammalian genes, with a further three annexins identified in lower eukaryotes (for reviews see Crompton et al., 1988; Klee, 1988; Creutz, 1992; Gerke, 1992; Moss, 1992). The apparent ubiquitous expression of annexins, both phylogenetically and within cells and tissues, suggests that they perform functions of fundamental importance. Accordingly, numerous functions have been proposed for the annexins, including phospholipase A_2 inhibition (Russo-Marie, 1992), anticoagulation (Tait et al., 1988), calcium-channel regulation (Huber et al., 1990; Diaz-Munoz et al., 1990), exocytosis and endocytosis (Ali et al., 1989; Lin et al., 1992; Smythe et al., 1994), inositol phosphate metabolism (Ross et al., 1990) and protein kinase C inhibition (Schlaepfer et al., 1992). Most of these suggested functions reflect *in vitro* properties of annexins, and the true function of any annexin awaits unequivocal proof.

Annexin II is one of the most extensively studied of the annexins, and is widely considered to participate in membrane fusion events, such as those that occur during exocytosis (Drust and Creutz, 1988; Ali et al., 1989; Nakata et al., 1990; Creutz, 1992). Recently, annexin II has also been shown to be associated with components of the endocytic pathway (Emans et al., 1993; Harder and Gerke, 1993), suggesting that it may have a bi-directional function. Annexin II was originally discovered as a major cellular substrate for phosphorylation on tyrosine by the transforming-gene product of the Rous sarcoma virus ($pp60^{v\text{-}src}$) (Erikson and Erikson, 1980; Gerke and Weber, 1984). Annexin II is also a substrate *in vivo* for phosphorylation on serine by protein kinase C (Glenney and Tack, 1985; Gould et al., 1986). Although the effects of phosphorylation on the membrane fusogenic properties of annexin II have not been fully characterized, it has been shown that phosphorylation of Tyr-23 leads to a greater calcium requirement for phospholipid binding (Powell and Glenney, 1987). Similarly, phosphorylation of the annexin II tetramer by protein kinase C was reported to cause a diminished ability to fuse lipid vesicles (Johnstone et al., 1992). However, the ability of annexin II to prevent exocytotic run-down in permeabilized chromaffin cells was shown to be markedly potentiated by prior phosphorylation by protein kinase C (Sarafian et al., 1991).

Annexin II conforms to the protein structure that characterizes the annexin family. The protein core comprises four conserved, repeated domains, each approx. 70 amino-acid residues in length. The N-terminus distinguishes annexin II from other members of the family, and it is here that the phosphorylation sites and p11-binding site are located (Gerke, 1992). Annexin II has been cloned from a variety of species including human (Huang et al., 1986), bovine (Kristensen et al., 1986), pig (Gerke, 1992), mouse (Saris et al., 1986), chicken (Gerke and Koch, 1990) and *Xenopus laevis* (Gerke et al., 1991; Izant and Bryson, 1991). In this paper we report the molecular cloning of rat annexin II and describe the first mammalian splice variant of this protein. The splice sequence lies within the N-terminus and introduces a potential phosphorylation site into this important domain.

EXPERIMENTAL

Isolation of rat annexin II cDNA clones

Mouse annexin II cDNA (a generous gift from Dr. Tony Hunter, The Salk Institute, CA, U.S.A.), was ^{32}P-labelled by random-priming and used to screen a plasmid cDNA library (60000 colonies) prepared from RBL-2H3 cells (kindly provided by Dr. Ernie Peralta, Harvard, MA, U.S.A.), according to Maniatis et al. (1982). The library was constructed in the plasmid pcDNA I (British Biotechnology Ltd.). Several positive clones were isolated after two rounds of screening, which on the basis of restriction mapping and DNA sequencing, comprised the entire rat annexin II coding and untranslated sequences.

* To whom correspondence should be addressed.

```
                ggaggctctctgcaataggtgcccggcccagctttttttttcaaa  44
Start
ATGTCTACTGTCCACGAAATCCTGTGCAAGCTCAGCTTGGAGGGTGATTCTCAGCATTCT  104
M  S  T  V  H  E  I  L  C  K  L  S  L  E  G  D  S  Q  H  S    20

ACACCCCCAAGTGCCTATGGGTCGGTCAAACCCTACACCAACTTCGACGCTGAGAGGGAT  164
T  P  P  S  A  Y  G  S  V  K  P  Y  T  N  F  D  A  E  R  D    40

GCTTTGAACATTGAAACAGCAATCAAGACCAAAGGCGTGGACGAGGTCACCATTGTCAAC  224
A  L  N  I  E  T  A  I  K  T  K  G  V  D  E  V  T  I  V  N    60

ATTCTGACTAACCGCAGCAATGCACAGAGGCAGGACATTGCCTTCGCCTACCAGAGGAGG  284
I  L  T  N  R  S  N  A  Q  R  Q  D  I  A  F  A  Y  Q  R  R    80

ACCAAAAAGGAACTGCCATCGGCGATGAAGTCGGCCTTGTCTGGTCACCTGGAGACCGTG  344
T  K  K  E  L  P  S  A  M  K  S  A  L  S  G  H  L  E  T  V   100

ATGTTAGGCCTGTTCAAGACACCTGCTCAGTACGATGCCTCTGAGCTCAAAGCCTCCATG  404
M  L  G  L  L  K  T  P  A  Q  Y  D  A  S  E  L  K  A  S  M   120

AAGGGCCTGGGGACTGATGAGGACTCCCTCATCGAGATCATCTGCTCAAGAACCAACCAG  464
K  G  L  G  T  D  E  D  S  L  I  E  I  I  C  S  R  T  N  Q   140

GAGCTGCAGGAGATTAACCGAGTGTATAAGGAAATGTACAAGACCGATCTGGAGAAGGAC  524
E  L  Q  E  I  N  R  V  Y  K  E  M  Y  K  T  D  L  E  K  D   160

ATCATCTCTGACACATCTGGAGAATTCCGAAAGCTGTTGGTCGCCCTTGCAAAGGGTAAA  584
I  I  S  D  T  S  G  E  F  R  K  L  L  V  A  L  A  K  G  K   180

CGGGCAGAGGATGGTTCTGTTATTGACTACGAGCTGATTGACCAGGATGCCCGGGAGCTC  644
R  A  E  D  G  S  V  I  D  Y  E  L  I  D  Q  D  A  R  E  L   200

TATGATGCTGGGGTGAAGAGGAAAGGAACCGATGTCCCCAAGTGGATCAGCATCATGACT  704
Y  D  A  G  V  K  R  K  G  T  D  V  P  K  W  I  S  I  M  T   220

GAGCGCAGTGTGTGCCACCTCCAGAAAGTGTTCGAAAGGTACAAGAGCTACAGTCCTTAT  764
E  R  S  V  C  H  L  Q  K  V  F  E  R  Y  K  S  Y  S  P  Y   240

GACATGCTGGAGAGCATCAGGAAAGAGGTCAAAGGAGACCTGGAGAACGCCTTCCTGAAC  824
D  M  L  E  S  I  R  K  E  V  K  G  D  L  E  N  A  F  L  N   260

CTGGTTCAGTGCATTCAGAACAAGCCCCTGTACTTTGCTGACCGGCTGTATGACTCCATG  884
L  V  Q  C  I  Q  N  K  P  L  Y  F  A  D  R  L  Y  D  S  M   280

AAGGGCAAGGGGACTCGAGACAAGGTCCTGATTAGAATCATGGTCTCTCGCAGTGAAGTG  944
K  G  K  G  T  R  D  K  V  L  I  R  I  M  V  S  R  S  E  V   300

GACATGTTGAAAATCAGATCTGAATTCAAGAGGAAATATGGCAAATCCCTGTACTACTTC 1004
D  M  L  K  I  R  E  S  F  K  R  K  Y  G  K  S  L  Y  Y  F   320

ATCCAGCAATACACTAAGGGTGACTACCAGAAGGCGCTGCTGTACCTGTGTGGTGGGGAC 1064
I  Q  Q  D  T  K  G  D  Y  Q  K  A  L  L  Y  L  C  G  G  D   340

GACTGAagggcttggcatggtggattgcccagaagtggccctacctgtgccccaacctaa 1124
D  Stop                                                       341

tgttctagagaatcagcctgccactaatggacccctgaactcctccctgtgaagatgacg 1184

acagagctgccgacccatcccccatcttagctgcctttgcctggctttcccctcattctc 1244

tcctttatgccaaagaaatgaacattddagggagttggacgtaccgtctgtgacatgaga 1304

cacttcctcatatgtgtcgtgaataaaccatttttacttaaaaaaaaaaaaaaaaaaa   1362
```

Figure 1 Complete cDNA sequence and derived amino-acid sequence of rat annexin II

The nucleotide sequence is shown above the amino-acid sequence. Non-coding sequence at the 5′- and 3′-ends is shown in lower-case letters and the coding sequence as upper-case letters. The amino-acid sequence is represented as single letter code. The initiation methionine (Start) and termination codon (Stop) are indicated. The position of the six-nucleotide insert is underlined and the polyadenylation signal is shown in bold.

Nucleic acid sequencing

Sequencing of double-stranded plasmid clones was performed according to the dideoxyribonucleotide-chain-termination method of Sanger et al. (1977), using Sequenase Version 2 (USB) according to the manufacturer's instructions. The library plasmid, pcDNA I, contains SP6 and T7 promoters on either side of the polylinker region. Therefore several hundred bases at the 5′ and 3′ ends of the cDNA in the library plasmid were sequenced, and visual alignment with the published sequence of mouse annexin II confirmed the identity of the clones. Restriction enzyme fragments of the cDNA clones were subcloned into pBluescript (Stratagene), to complete the sequencing. The sequence obtained (Figure 1) was 96% identical with the mouse annexin II cDNA and was derived from the composite data of two non-identical rat cDNA clones.

RESULTS AND DISCUSSION

Screening of an RBL-2H3-cell-derived cDNA library with a mouse annexin II probe yielded a number of positive clones which were subcloned into proprietary vectors and sequenced. The resulting nucleotide sequence, which combines data from several overlapping clones, is shown in Figure 1 with the derived amino-acid sequence. The cDNA contains an open reading frame of identical length to the other published mammalian annexin II sequences, and has 5′ and 3′ untranslated sequences also of similar length and composition to human and mouse annexin II. The high level of sequence conservation typical of the annexins is preserved in rat annexin II, alignment of the rat, human and mouse annexin II protein sequences reveals virtual sequence identity between the three species (Figure 2). Of the two rat annexin II clones isolated, one contained a six nucleotide insert located close to the 5′ end (Figure 3), at a position corresponding to the boundary of the first and second coding exons in the mouse annexin II gene (Figure 4) (Amiguet et al., 1990). Given the high degree of conservation of annexin intron/exon boundary positions between species, the location of the insert at an intron/exon boundary strongly suggests that the two isoforms arise as a consequence of alternative splicing (Breitbart et al., 1987). If the six nucleotides constitute a complete cassette exon, then it equals in size the smallest known exon (in the cardiac troponin C gene). However, the sequence of the insert conforms closely to the consensus for a splice donor site, which favours a retained intron. Outside the proposed alternative splice site the two forms of annexin II were found to be identical.

Interestingly, N-terminal variants of annexin II have also been identified in *Xenopus* oocytes (Izant and Bryson, 1991) although the differences were more varied and suggested the presence of two distinct genes. These findings, together with those reported here, suggest that species-specific microheterogeneity may be a common phenomenon in the annexin family. Similar findings were reported for annexin I isoforms cloned from pigeon cropsac (Horseman, 1989; Hitti and Horseman, 1991). Nevertheless, it is important to make the distinction between closely related genes, such as those for annexin II found in *Xenopus* oocytes and for annexin I in pigeon cropsac, and alternative splice variants, which almost certainly explain the rat annexin II isoforms described here. Alternative splicing has also been demonstrated for other members of the annexin family, including annexins VI (Moss and Crumpton, 1990), VII (Magendzo et al., 1991) and XI (Towle et al., 1992). With the exception of annexin VI, in which the splice site lies within the repeated conserved core of the protein, all the annexin splice variants (including that described here) arise as a consequence of differences in their N-termini. Given the importance of the N-terminus in defining the properties of annexins, it is clearly possible that the different isoforms of annexins II, VII and XI have different cellular functions.

The six-nucleotide insert reported here is interesting and may be functionally important because it encodes serine–glutamine, thus introducing an additional potential phosphorylation site into a region of the protein already known to have several phosphorylation sites. It is known that phosphorylation of the N-terminus of annexin II and binding to its cellular ligand, p11, are likely to be mechanistically interdependent. However, if the alternatively spliced serine residue is not a phosphorylation site (it does not lie within an obvious consensus sequence for known serine kinases), the importance of the two amino-acid insert may depend on whether or not it induces changes in the folding of the annexin II N-terminus.

Figure 2 Alignment of the amino-acid sequences of rat, human and mouse annexin II

The amino-acid sequences of rat, mouse and human annexin II are shown as single-letter code, and arranged to emphasize the internally repetitive structure of proteins of the annexin family. Hyphens indicate amino-acid identity and all substitutions are indicated. Breaks in the mouse and human sequences (represented by dots) have been introduced to maintain sequence similarity.

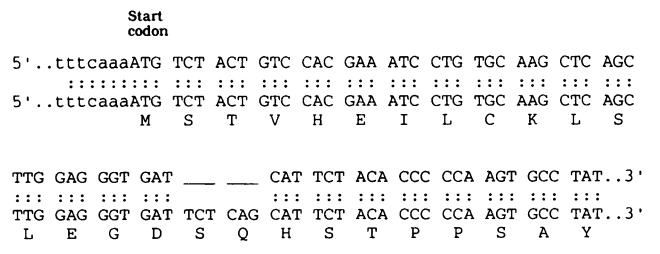

Figure 3 Alignment of the alternative 5′ ends of rat annexin II

The nucleotide and amino-acid sequences at the 5′ ends of the two annexin II isoforms are shown. The two sequences are identical except for the insertion of six nucleotides encoding serine–glutamine.

We are grateful to Paul Smith for technical advice and to Bastien Gomperts for helpful and interesting discussions. A.L.U. holds a Prize Studentship from the Wellcome Trust.

REFERENCES

Ali, S. M., Geisow, M. J. and Burgoyne, R. D. (1989) Nature (London) **340**, 313–315
Amiguet, P., D'Eustachio, P., Kristensen, T., Wetsel, R. A., Saris, C. J. M., Hunter, T., Chaplin, D. D. and Tack, B. F. (1990) Biochemistry **29**, 1226–1232
Breitbart, R. E., Andreadis, A. and Nadal-Ginard, B. (1987) Annu, Rev. Biochem **56**, 467–495
Creutz, C. E. (1992) Science **258**, 924–931
Crompton, M. R., Moss, S. E. and Crumpton, M. J. (1988) Cell **55**, 1–3
Diaz-Munoz, M., Hamilton, S. L., Kaetzel, M. A., Hazarika, P. and Dedman, J. R. (1990) J. Biol. Chem. **265**, 15894–15899
Drust, D. S. and Creutz, C. E. (1988) Nature (London) **331**, 88–91
Emans, N., Gorvel, J.-P., Walter, C., Gerke, V., Kellner, R., Griffiths, G. and Gruenberg, J. (1993) J. Cell Biol. **120**, 1357–1369
Erikson, E. and Erikson, R. L. (1980) Cell **21**, 829–836
Gerke, V. (1992) in The Annexins (Moss, S. E., ed.), pp. 47–59, Portland Press, London and Chapel Hill
Gerke, V. and Koch, W. (1990) Nucleic Acids Res. **18**, 4246
Gerke, V. and Weber, K. (1984) EMBO J. **3**, 227–233
Gerke, V., Koch, W. and Thiel, C. (1991) Gene **104**, 259–264
Glenney, J. R., Jr. and Tack, B. F. (1985) Proc. Natl. Acad. Sci. U.S.A. **82**, 7884–7888
Gould, K., Woodgett, J. R., Isacke, C. M. and Hunter, T. (1986) Mol. Cell. Biol. **6**, 2738–2744
Harder, T. and Gerke, V. (1993) J. Cell Biol. **123**, 1119–1132
Hitti, Y. S. and Horseman, N. D. (1991) Gene **103**, 185–192
Horseman, N. D. (1989) Mol. Endocrinol. **3**, 773–779
Huang, K.-S., Wallner, B. P., Mattaliano, R. J., Tizard, R., Burne, C., Frey, A., Hession, C., McGray, P., Sinclair, L. K., Chow, E. P. et al. (1986) Cell **46**, 191–199
Huber, R., Romisch, J. and Paques, E.-P. (1990) EMBO J. **9**, 3867–3874
Izant, J. G. and Bryson, L. J. (1991) J. Biol. Chem. **266**, 18560–18566
Johnstone, S. A., Hubaishy, I. and Waisman, D. M. (1992) J. Biol. Chem. **267**, 25976–25981
Klee, C. B. (1988) Biochemistry **27**, 6645–6653
Kristensen, T., Saris, C. J. M., Hunter, T., Hicks, L. J., Noonan, D. J., Glenney, J. R., Jr. and Tack, B. F. (1986) Biochemistry **25**, 4497–4503

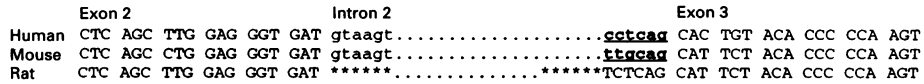

Figure 4 Relationship between the rat annexin II splice variants and the genomic structures of mouse and human annexin II

The mouse and human annexin II genomic sequences surrounding the boundary between exons 2 and 3 are aligned to the six-nucleotide rat annexin II insert. The rat insert is placed in alignment with the 3′ end of intron 2. Intronic sequence is shown in lower-case letters, asterisks indicate (as yet uncharacterized) rat intronic sequence. The 3′ ends of human and mouse intron 2 are in bold and underlined.

Lin, H. C., Sudhof, T. C. and Anderson, R. G. W. (1992) Cell **70**, 283–291
Magendzo, K., Shirvan, A., Cultraro, C., Srivastava, M. and Pollard, H. B. (1991) J. Biol. Chem. **266**, 3228–3232
Maniatis, T., Fritsch, E. F. and Sambrook, J. (1982) Molecular Cloning. A Laboratory Manual, Cold Spring Harbor, NY
Moss, S. E. (1992) in The Annexins (Moss, S. E., ed.), pp. 1–9, Portland Press, London and Chapel Hill
Moss, S. E. and Crumpton, M. J. (1990) FEBS Lett. **261**, 299–302
Nakata, T., Sobue, K. and Hirokawa, N. (1990) J. Cell Biol. **110**, 13–25
Powell, M. A. and Glenney, J. R. (1987) Biochem. J. **247**, 321–328
Ross, T. S., Tait, J. F. and Majerus, P. W. (1990) Science **248**, 605–607
Russo-Marie, F. (1992) in The Annexins (Moss, S. E., ed.), pp. 35–46, Portland Press, London and Chapel Hill
Sanger, F., Nicklen, S. and Coulson, A. R. (1977) Proc. Natl. Acad. Sci. U.S.A. **74**, 5463–5467
Sarafian, T., Pradel, L.-A., Henry, J. P., Aunis, D. and Bader, M. F. (1991) J. Cell Biol. **114**, 1135–1147
Saris, C. J. M., Tack, B. F., Kristensen, T., Glenney, J. R. Jr. and Hunter, T. (1986) Cell **46**, 201–212
Schlaepfer, D. D., Jones, J. and Haigler, H. T. (1992) Biochemistry **31**, 1886–1891
Smythe, E., Smith, P. D., Jacob, S. M., Theobald, J. and Moss, S. E. (1994) J. Cell Biol. **124**, 301–306
Tait, J. F., Sakata, M., McMullen, B. A., Miao, C. H., Funakoshi, T., Hendrickson, L. E. and Fujikawa, K. (1988) Biochemistry **27**, 6268–6276
Towle, C. A., Weissbach, L. and Treadwell, B. V. (1992) Biochim. Biophys. Acta **1131**, 223–226

Received 10 February 1994/16 March 1994; accepted 22 March 1994