FEBS Letters 356 (1994) 51–55

FEBS 14844

Molecular cloning of cDNA encoding the 110 kDa and 21 kDa regulatory subunits of smooth muscle protein phosphatase 1M

Yu Hua Chen[a], Mao Xiang Chen[a], Dario R. Alessi[a], David G. Campbell[a], Catherine Shanahan[b], Philip Cohen[a], Patricia T.W. Cohen[a,*]

[a]*Medical Research Council Protein Phosphorylation Unit, Department of Biochemistry, University of Dundee, Dundee DD1 4HN, Scotland, UK*
[b]*Department of Medicine, Addenbrooke's Hospital, Cambridge, UK*

Received 1 November 1994

Abstract The structures of the M_{110} and M_{21} regulatory subunits of protein phosphatase-1M, the major enzyme which dephosphorylates myosin in smooth muscle, have been deduced from cloned cDNAs. The N-terminus of the M_{110} subunit from rat aorta contains seven ankyrin repeats, while the C-terminus of the M_{21} subunit from chicken gizzard contains a leucine zipper motif. The M_{110} subunit is expressed in two different forms which differ in their C-terminal sequences. One of these is highly homologous to the whole of the M_{21} subunit.

Key words: Smooth muscle; Protein phosphatase; cDNA sequence; Myosin; Ankyrin repeat; Leucine zipper

1. Introduction

Protein phosphatase-1 (PP1), one of the major protein serine/threonine phosphatases of eukaryotic cells, has been implicated in the regulation of a variety of cellular processes (reviewed in [1,2]). The structure of the 37 kDa catalytic subunit has been determined [3] and three genes encoding it have been identified in mammalian tissues [4]. In recent years, evidence has accumulated that the PP1 catalytic subunit is directed to particular subcellular locations by 'targetting' subunits which also modify the substrate specificity of the enzyme (reviewed in [5]). For example, striated muscles contain a form of PP1, termed PP1G, which is complexed to a G-subunit that contains domains for association with the sarcoplasmic reticulum and with glycogen, and which enhances PP1 activity towards the enzymes of glycogen metabolism. The G-component also plays an important role in the hormonal control of glycogen metabolism because its phosphorylation at distinct sites in response to insulin and adrenalin activate and inhibit PP1G, respectively [5]. The 124 kDa G-subunit has been cloned from rabbit and human skeletal muscle [6,7].

We have recently described further forms of PP1 that are associated with the myofibrils of striated [8] and smooth [9] muscles. These enzymes, termed PP1M, comprise the catalytic subunit associated with 'M-complexes' which enhance the rate at which PP1 dephosphorylates myosin while supressing activity towards the enzymes of glycogen metabolism. The complex from smooth muscle enhances the dephosphorylation of smooth muscle myosin, but not skeletal muscle myosin, whereas the 'M-complex' from skeletal muscle enhances the dephosphorylation of skeletal muscle myosin at least 30-fold [9]. Thus the M-complexes from skeletal and smooth muscles appear to be distinct proteins.

The M-complex from chicken gizzard smooth muscle is a heterodimer composed of two proteins whose apparent molecular masses on SDS/polyacrylamide gels are 130 kDa and 20 kDa respectively [9]. The 130 kDa subunit is the component which interacts with PP1 [9]. In order to identify the regions on the 130 kDa and 20 kDa subunits which interact with each other and with myosin, and the regions on the 130 kDa subunit which bind to PP1 and modulate its substrate specificity, it is essential to first determine the amino acid sequence of these proteins. Here, we present the amino acid sequences of the 130 kDa and 20 kDa subunits from rat aorta and chicken gizzard which show several interesting features.

2. Materials and methods

Smooth muscle PP1M was purified from chicken gizzard and the 130 kDa and 20 kDa subunits separated by chromatography on a C18 column as described [9]. The subunits were then cleaved with CNBr or digested with trypsin, chymotrypsin or a proteinase from the fungus *A. mellea* which cleaves on the N-terminal side of lysine residues, termed N-Lys proteinase [10]. The digests were applied to a Vydac C18 column (Hesperia, CA) equilibrated in 0.1% (v/v) trifluoroacetic acid and peptides separated using a gradient of increasing acetonitrile (0.33% per min). Sequence analysis was performed on an Applied Biosystems 470/120 gas-phase sequencer.

The λgt11 chicken gizzard library was from Clontech (Palo Alto, CA) and the λZap rat aorta cDNA library is described in [11]. Degenerate oligonucleotides containing inosine (I) were synthesised by Oswel DNA Service, University of Edinburgh. Oligonucleotide primers for sequencing were synthesised by Miss A. Gough, University of Dundee. pT7Blue(R) was obtained from Novagen (Oxon, UK) and Bluescript M13+ from Stratagene (San Diego, CA).

Polymerase chain reactions were performed using either oligonucleotides 1 and 2 or 4 and 5 (0.2 μM each), 2–10 × 10^6 bacteriophage (2 μl), 1.5 mM $MgCl_2$ and AmpliTaq DNA polymerase as described by Perkin Elmer Cetus (Bucks, UK) in 20 or 100 μl reactions with the following cycling protocol: 94°C 5 min, 1 cycle; 55°C 1 min, 72°C 1 min, 94°C 1 min, 42 cycles; 55°C 5 min; 72°C 10 min. The PCR products were analysed by gel electrophoresis, followed by Southern blotting and hybridisation at 30°C with either oligonucleotide 3 or 6 which had been 5′ end-labelled with [γ-^{32}P]dATP. PCR fragments positive with the appropriate oligonucleotide were purified from the gel using Prepagene (NBL, Cramlington, UK), subcloned into pT7 Blue vector (R) and sequenced.

The chicken gizzard cDNA library was screened at high stringency with the PCR fragments labelled with [α-^{32}P]dATP according to [12] using 0.2 μM each of oligonucleotides 1 and 2 or oligonucleotides 4 and 5 in place of random hexanucleotides. Screening of the rat aorta cDNA library was carried out at 55°C with washes in 30 mM NaCl, 3 mM sodium citrate pH 7.0, 0.1% SDS. DNA sequencing was performed in

*Corresponding author. Fax: (44) (382) 23778.

0014-5793/94/$7.00 © 1994 Federation of European Biochemical Societies. All rights reserved.
SSDI 0014-5793(94)01231-8

both directions on double-stranded plasmid DNA using an Applied Biosystems 373A automated DNA sequencer and Taq dye terminator cycle sequencing according to the manufacturer's instructions.

3. Results

3.1. Cloning of cDNA encoding an M_{21} subunit from chicken gizzard

A sequence of 35 consecutive amino acids was determined by analysing several overlapping peptides obtained after cleavage of the M_{21} subunit with CNBr and N-Lys proteinase (Fig. 1a). PCR using a chicken gizzard cDNA library and oligonucleotide primers 1 and 2 derived from the peptide sequence yielded a 105 bp fragment (Fig. 1a) whose sequence encoded the same 35 amino acid peptide. Using this DNA fragment as a probe, two positive clones were identified from a chicken gizzard cDNA library, both of which included an identical open reading frame of 558 nucleotides (Fig. 2a). All the peptides that were isolated and sequenced (>100 residues) were found in the deduced sequence (Fig. 3). In the non-coding region, there were 11 single base differences and one deletion/insertion which may arise from a polymorphism (Figs. 2 and 3).

3.2. The M_{21} subunit has a leucine zipper motif and shows weak similarity to several structural proteins

The M_{21} subunit comprises 186 amino acids and its deduced molecular mass of 21 kDa is similar to that estimated previously by SDS/PAGE. The C-terminus of the protein contains a leucine zipper motif (reviewed in [13]) (Fig. 4) but the overall stucture is not closely related to any protein in the databases that we have searched. Nevertheless, there is some similarity to structural proteins, such as the myosin heavy chain (21% identity over 91 amino acids), paramyosin (23% identity over 96 amino acids), lamin LIII (22% identity over 175 amino acids), neurofilament triplet M protein (22% identity over 183 amino acids) and kinesin heavy chain (27% identity over 86 amino acids). The region of similarity in myosin lies at the junction of the rod and head domains.

3.3. Cloning of cDNA encoding part of an M_{110} subunit from chicken gizzard

A sequence of 35 amino acids of the M_{110} subunit was determined by analysis of several overlapping peptides obtained by cleavage with CNBr and N-Lys proteinase (Fig. 1b). PCR using a chicken gizzard cDNA library and oligonucleotide primers 4 and 5 derived from the peptide sequence yielded a DNA fragment of 87 base pairs whose sequence encoded part of the 35 amino acid segment (Fig. 1b). Using this DNA fragment as a probe, four positive clones were isolated from a chicken gizzard cDNA library all of which contained different N-terminal regions interrupted by stop codons. This indicated that these N-termini are likely to be artifacts arising in the cDNA library preparation. Nevertheless, the C-termini of the clones were identical, and contained several of the peptides isolated by digestion of the M_{110} subunit (Fig. 5).

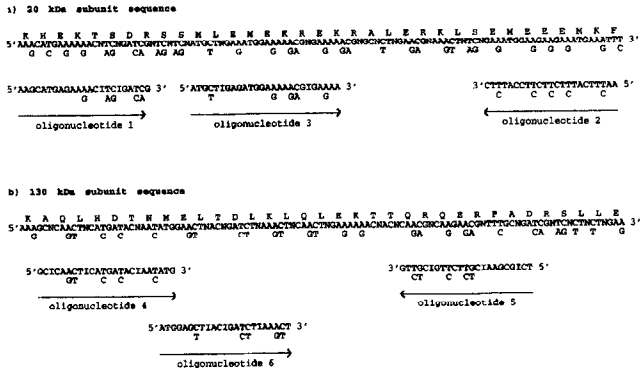

Fig. 1. Peptide sequences of (a) the 20 kDa and (b) the 130 kDa subunit of chicken gizzard smooth muscle PP1M, and the corresponding oligonucleotides used to isolate the cDNA encoding them. The sequences were obtained from overlapping peptides generated by cleavage with CNBr and N-Lys proteinase.

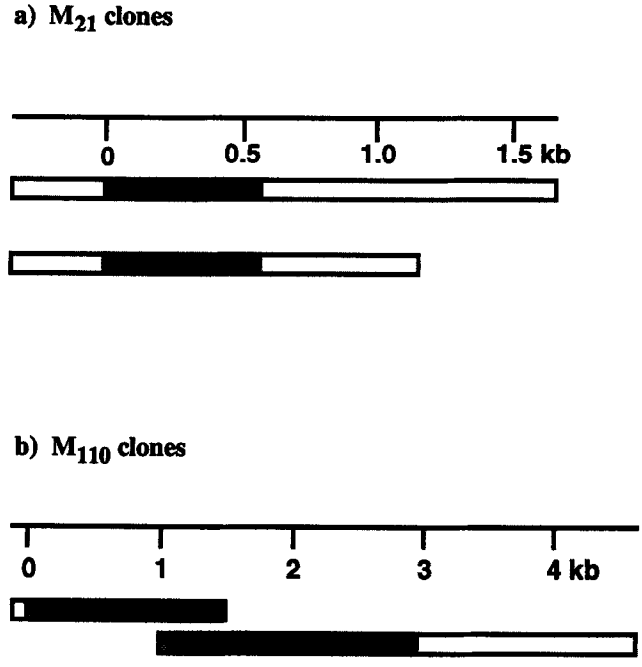

Fig. 2. Schematic representation of the cDNA clones isolated. Open bars indicate non-coding regions and filled bars the coding regions. (a) M_{21} clones from chicken gizzard. Both clones were sequenced over the 1.6 kb overlapping region and the shorter clone had the following differences from the longer clone (Fig. 3) in the non-coding regions. Nucleotides −295 to −270 were deleted and nucleotide substitutions at −240 (G to A), −237 (T to C), −236 (A to G), −121 (C to T), 579 (T to C), 750 (C to T), 763 (C to T), 850 (T to G), 973 (T to A), 1124 (G to C), 1132 (G to A) occurred in the longer clone. (b) M_{110} clones from rat aorta. The full sequence was determined from the two overlapping clones.

3.4. Cloning of cDNA encoding an M_{110} subunit from rat aorta

cDNA encoding the C-terminal region of the M_{110} subunit from chicken gizzard was used to screen a rat aorta cDNA library. Several positive clones were identified. Two overlapping clones (Fig. 2b) encoded a protein of 976 amino acids, with a molecular mass of 110 kDa (Fig. 5). The N-terminal section of the protein (residues 39–296) contains seven 33 amino acid ankyrin repeats [14] (Fig. 6), while a leucine zipper motif, consisting of four heptad repeats where leucine is the seventh residue [13], is present at the C-terminus (Fig. 4). An acidic region, rather than the basic DNA binding region found in transcription factors, precedes the leucine zipper. Between the ankyrin repeats and leucine zipper motif the protein is remarkably hydrophilic. 62% of the the amino acids between residues 400 and 800 are Asp/Glu, Asn/Gln, Ser/Thr or Lys/Arg (Fig. 5).

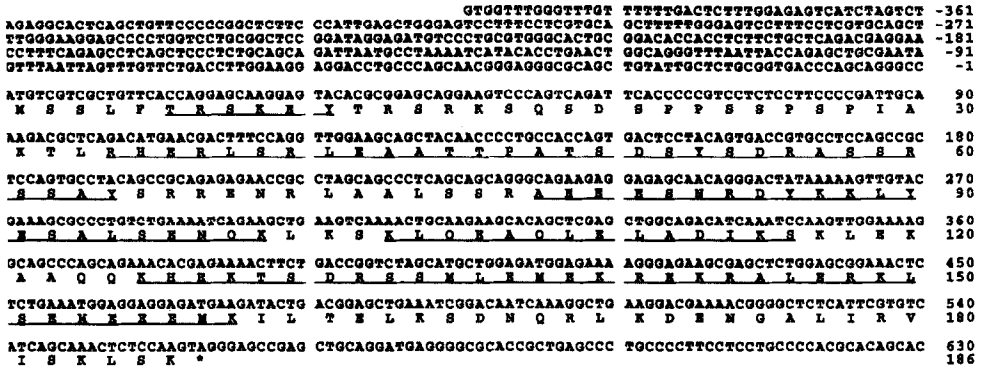

Fig. 3. Complementary DNA and predicted protein sequence of the M_{21} subunit from chicken gizzard. Peptide sequences determined from the isolated subunit are underlined and were obtained from a variety of digests with trypsin, chymotrypsin, N-Lys proteinase and CNBr.

3.5. Comparison of the M_{110} subunits from rat aorta and chicken gizzard

All the peptides isolated from the chicken M_{110} subunit (123 residues) were highly homologous to peptide sequences present in the rat aorta M_{110} protein (96% identity) and the C-terminal 359 amino acids of the chicken M_{110} subunit are 75% identical to the homologous C-terminal section of the rat aorta protein (Fig. 5). However, the chicken protein terminates 35 residues before the rat protein and does not contain a leucine zipper motif.

4. Discussion

The finding that PP1 is complexed to many different subunits has raised the important question of how the correct proportion of PP1 is targetted to different subcellular locations. Several possible solutions to this problem could be envisaged. Firstly, as several isoforms of the catalytic subunit of PP1 have been identified in mammalian tissues which are 90–94% identical in amino acid sequence [4], one potential mechanism would involve each targetting subunit binding to a different PP1 isoform. However, this does not appear to be the explanation, at least in muscle, because the M_{110} subunit and the G-subunit are both complexed to the same PP1 isoform (PP1β) [8,9]. Moreover, all three expressed isoform of PP1 are capable of binding to either the G-subunit or the M_{110} subunits in vitro

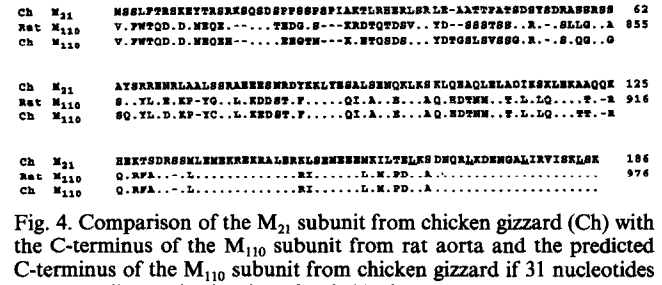

Fig. 4. Comparison of the M_{21} subunit from chicken gizzard (Ch) with the C-terminus of the M_{110} subunit from rat aorta and the predicted C-terminus of the M_{110} subunit from chicken gizzard if 31 nucleotides corresponding to the deletion after 2793 of the rat sequence are omitted. Residues in the M_{110} sequences that are identical to the M_{21} sequence are indicated by dots and deletions by dashes. The conserved leucines in the four leucine zipper heptad repeats are underlined.

[15]. The simplest alternative mechanism would involve each targetting subunit binding to the same or overlapping sites on PP1, such that the binding of one subunit precludes the binding of another. In this case, the amount of PP1 directed to any particular location would be determined by the levels of expression of each targetting subunit and by their relative affinities for the PP1 catalytic subunit. However, as yet, no clear sequence identities that might provide a binding site for PP1 can be discerned among the regulatory subunits that have been sequenced, including the mammalian M_{110} subunit (Fig. 4), the G-subunit [6,7], inhibitor-1 [10] and inhibitor-2 [16,16a] and the yeast subunits GAC1 [17] and sds22 [18]. Nevertheless, it has been shown previously that the PP1 binding site is located in the N-terminal third of the G-subunit [19] and the preceeding paper shows that this is also the case for the skeletal muscle M_{110} subunit [20]. This suggests that the ankyrin repeat (Fig. 6) is likely to contain the PP1 binding site, since many proteins that contain ankyrin repeats interact specifically with other proteins via this domain.

A most interesting finding made in the present study was that the C-terminal region of the M_{110} subunit from rat aorta was homologous to the whole of the M_{21} subunit from chicken gizzard, similarity being particular striking in the C-terminal two thirds of the M_{21} subunit which terminates in the leucine zipper motif (Fig. 4). Since such structures are believed to interact with other leucine zipper proteins, it is possible that these structures in the M_{21} and M_{110} subunits are either involved in mediating interactions between these two subunits or in binding to myosin which also contains many of these motifs (originally called the coiled-coil structure).

In contrast to the rat aorta M_{110} subunit, the chicken gizzard M_{110} subunit lacks the C-terminal leucine zipper sequence and this result was confirmed by isolation of the predicted C-terminal peptide from the chicken gizzard protein (Fig. 5). However, inspection of the 3′ non-coding region of the chicken M_{110} cDNA reveals that it encodes a structure identical to the C-terminal sequence of the rat aorta M_{110} subunit is present in one reading frame (Fig. 4). Moreover, we have shown by PCR analysis that a rat uterus library does contain DNA which encodes an M_{110} subunit which terminates at the same place as

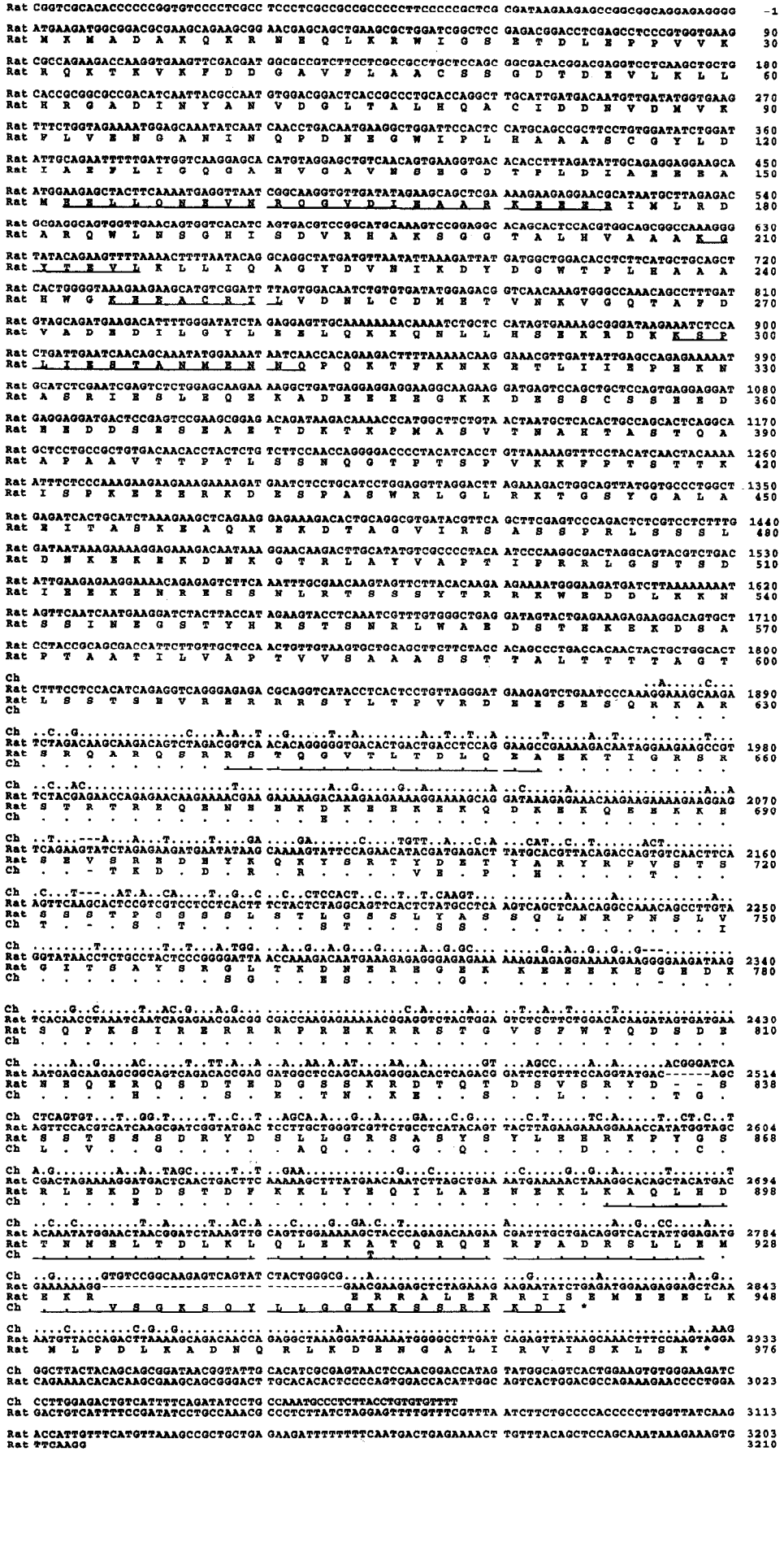

Fig. 5. Complementary DNA and predicted protein sequence of the M_{110} subunits from rat aorta and chicken gizzard (Ch). Residues in the chicken subunit that are identical to the rat aorta sequence are indicated by dots and deletions by dashes. Peptide sequences from the isolated chicken gizzard M_{110} subunit are underlined and were obtained by digestion of the M_{110} subunit with CNBr and N-Lys proteinase. They were identical to the rat sequence except for amino acids 308, 309 and 913 (M, E and A in the rat) which were L, D and T, respectively, in the chicken.
←

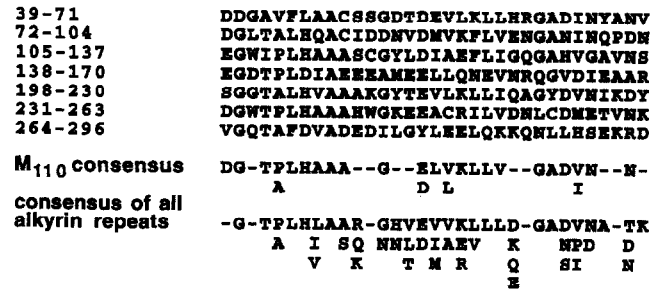

Fig. 6. The ankyrin repeat structure at the N-terminus of the M_{110} subunit from rat aorta. The consensus sequence from other known ankyrin repeats is taken from [14].

the chicken gizzard M_{110} sequence and hence lacks the leucine zipper motif (Y.H. Chen, M.X. Chen and P.T.W. Cohen, unpublished work). These results demonstrate that the rat aorta M_{110} subunit is the mammalian homologue of the chicken M_{110} subunit.

Our data also leads us to speculate that the M_{21} subunit from chicken gizzard may be transcribed from a second M_{110} subunit gene. If this were the case then each of the two precursor mRNAs might be spliced to generate three different products, namely two M_{110} subunits with and without the C- terminal leucine zipper and an M_{21} subunit. These six products may be expressed to varying extents in different cells and might account, at least in part, for the distinct contractile behaviour of different smooth muscles. Although at this stage we cannot exclude the possibility that the M_{21} subunit is expressed from a completely separate gene, its generation by alternative splicing of an M_{110} subunit would be analogous to the situation in smooth muscle myosin light chain kinase (MLCK). The C-terminal 155 residues of MLCK are expressed in some smooth muscles as a separate protein, termed telokin, which is transcribed from a promoter located within an intron of the MLCK gene [21]. Since telokin is expressed at a much higher concentration than MLCK in turkey gizzard [22], it will be interesting to see whether the M_{21} subunit of PP1M is synthesized in a molar excess over the M_{110} subunit.

Acknowledgements: We thank the UK Medical Research Council for financial support and for providing D.R.A. with a postdoctoral training fellowship. This work has benefitted from the use of the SERC funded SEQNET facility at Daresbury, UK.

References

[1] Cohen, P. (1989) Annu. Rev. Biochem. 58, 453–508.
[2] Cohen, P and Cohen, P.T.W. (1989) J. Biol. Chem. 264, 21435–21438.
[3] Berndt, N., Campbell, D.G., Caudwell, F.B., Cohen, P., da Cruz e. Silva, E.F., da Cruz e. Silva, O.B. and Cohen, P.T.W. (1987) FEBS Lett. 223, 340–346.
[4] Barker, H.M., Brewis, N.D., Street, A.J., Spurr, N.K. and Cohen, P.T.W. (1994) Biochim. Biophys. Acta 1220, 212–218.
[5] Hubbard, M.J. and Cohen, P. (1993) Trends Biochem. Sci. 18, 172–176.
[6] Tang, P.M., Bondor, J.A., Swiderek, K.M. and DePoali-Roach, P. (1991) J. Biol. Chem. 266, 15782–15789.
[7] Chen, Y.H., Hansen L., Chen, M.X., Bjørbaek, C, Vestergaard H., Hansen, T., Cohen, P.T.W. and Pedersen, O. (1994) Diabetes 43, 1234–1241.
[8] Dent, P., MacDougall, L.K., MacKintosh, C., Campbell, D.G. and Cohen, P. (1992) Eur. J Biochem. 210, 1037–1044.
[9] Alessi, D., MacDougall, L.K., Sola, M.M., Ikebe, M. and Cohen, P. (1992) Eur. J. Biochem 210, 1023–1035.
[10] Aitken, A., Bilham, T. and Cohen, P (1982) Eur J. Biochem. 126, 235–246.
[11] Shanahan, C.M., Weissberg, P.L. and Metcalfe, J.C. (1993) Circ. Res. 73, 193–204.
[12] Feinberg, A.P. and Vogelstein, B. (1984) Anal. Biochem. 137, 266–267.
[13] Hu, J.C. and Sauer, R.T. (1993) Nucleic Acids Mol. Biol. 6, 83–101.
[14] Michaely, P. and Bennett, P. (1992) Trends Cell Biol. 2, 127–129.
[15] Alessi, D.R., Street, A.J., Cohen, P. and Cohen, P.T.W. (1993) Eur. J. Biochem. 213, 1055–1066.
[16] Holmes, C.F.B., Campbell, D.G., Caudwell, F.B., Aitken, A. and Cohen, P. (1986) Eur. J. Biochem. 155, 173–182.
[16a] Helps, N.R., Street, A.J., Elledge, S.J. and Cohen, P.T.W. (1994) FEBS Lett. 340, 93–98.
[17] Francois, J.M., Thompson-Jaeger, S., Skrotch, J., Zellenka, U., Spevak, W. and Tatchell, K. (1992) EMBO J. 11, 87–96.
[18] Ohkura, H. and Yanagida, M. (1991) Cell 64, 149–157.
[19] Hiraga, A., Kemp, B.E. and Cohen, P. (1987) Eur. J. Biochem. 163, 253–258.
[20] Moorhead, G., MacKintosh, R.W., Morrice, N., Gallagher, T. and MacKintosh, C. (1994) FEBS Lett., 356, 46–50 (preceeding paper).
[21] Gallagher, P.J. and Herring, B.P. (1991) J. Biol. Chem. 23945–23952.
[22] Ito, M., Dabrowska, R., Guerriero, V. and Hartshorne, D.J. (1989) J. Biol. Chem. 264, 13971–13974.