Volume 232, number 1, 17–23 FEB 05800 May 1988

Two isoforms of protein phosphatase 1 may be produced from the same gene

Patricia T.W. Cohen

Department of Biochemistry, Medical Science Institute, The University, Dundee DD1 4HN, Scotland

Received 27 February 1988

A cDNA clone encoding a second type-1 protein phosphatase catalytic subunit (1α) was isolated from a rabbit skeletal muscle cDNA library constructed in λgt10. The deduced protein sequence (330 residues, 37.5 kDa) was 19 residues longer at its N-terminus than protein phosphatase 1β (311 residues, 35.4 kDa). The amino acid sequences of protein phosphatases 1α and 1β were identical after residue 33 of protein phosphatase 1α. The results suggest that the different N-terminal sequences of protein phosphatases 1α and 1β are likely to be generated by differential transcription or processing of the mRNA transcribed from a single gene. Southern blotting of rabbit DNA was consistent with this interpretation.

Protein phosphorylation; cDNA cloning; Nucleotide sequence; Amino acid sequence; Alternative splicing; Isozyme

1. INTRODUCTION

In mammalian tissues, four serine/threonine-specific protein phosphatase catalytic subunits have been identified on the basis of their sensitivity to inhibitor proteins and their substrate specificity (review [1,2]). Type-1 protein phosphatase is inhibited by nanomolar concentrations of two thermostable proteins, inhibitor-1 and inhibitor-2, while the type-2 enzymes (2A, 2B and 2C), are unaffected [3].

We have recently employed recombinant DNA techniques to investigate the structures and physiological roles of the protein phosphatases. Clones containing the entire coding regions of a type-1 [4] and two type-2A protein phosphatase catalytic subunits [5,6] have been isolated from a rabbit skeletal muscle cDNA library constructed in phage λgt10 [7]. The amino acid sequence identity between the type-1 and type-2A protein phosphatases is 43%, showing that the protein phosphatases, like the protein kinases, are members of a homologous gene family. The cDNA clones coding for the two type-2A protein phosphatases (2Aα and 2Aβ) were 82% identical at the nucleotide level in the coding region while their 3′-untranslated regions were completely different [6]. These findings established that the phosphatase subunits 2Aα and 2Aβ are the products of different genes, although they code for proteins that are 97% identical in amino acid sequence [6].

In this paper, a further clone, isolated from the same cDNA library, has been shown to encode a type-1 protein phosphatase catalytic subunit. However, in contrast to the situation with phosphatase 2A isozymes, the data suggest that two isoforms of type-1 protein phosphatases are derived from a single gene. A preliminary account of part of this work was presented at the 1st International Conference on Post-Translational Modifications of Proteins and Ageing [8].

Correspondence address: P.T.W. Cohen, Department of Biochemistry, Medical Science Institute, The University, Dundee DD1 4HN, Scotland

2. MATERIALS AND METHODS

2.1. *Subcloning and cDNA sequencing*

The 1.4 kb insert of the cDNA clone that was positive with oligonucleotide probes for the catalytic subunit of protein phosphatase 1 [4] was subcloned into the *Eco*RI site of M13 tg

Published by Elsevier Science Publishers B.V. (Biomedical Division)
00145793/88/$3.50

131 [9]. Single-stranded DNA was prepared and sequenced using the dideoxy chain termination procedure [10], [α-^{35}S]-ATPαS (Amersham, Bucks, England) and buffer gradient gels [11]. Sequencing reactions were initiated with oligonucleotides complementary to the coding sequence or with an M13 primer. To overcome the ambiguities often encountered with GC rich DNA, most reactions were also performed in the presence of 7-deaza-2'-dGTP in place of dGTP [12]. Since all subclones obtained contained the 1.4 kb insert in the same orientation, it was also subcloned into Bluescript pkS-M13$^+$ (Stratagene Cloning Systems, San Diego, USA) in order to sequence the opposite DNA strand. The double-stranded DNA was sequenced using oligonucleotide or Bluescript primers.

2.2. *Southern blot analysis*

Genomic DNA was isolated [13], digested with *Eco*RI, and fractionated on 0.6% agarose gels. Blotting of DNA was performed onto Hybond N (Amersham). Hybridisation conditions were as described in [7], except that the probes were labelled by random hexanucleotide priming [14].

3. RESULTS

3.1. *Isolation of the cDNA clone*

Berndt et al. [4] described the isolation of four cDNA clones from a rabbit skeletal muscle cDNA library in λgt10 [7] that were positive with the 29 base oligonucleotides

3'-CTG CAC GTC CCC ACC CCC CTC TTG CTG GC-5'
I I I

complementary to tryptic peptide T7 of protein phosphatase 1 DVQGWGENDR. The insert sizes of these clones were 0.6, 1.0, 1.4 and 1.5 kb and the two largest clones were also positive with four other oligonucleotides complementary to sequences coding for tryptic peptides T2, T3, T4 and T8. The 1.5 kb clone was isolated and sequenced to give the complete primary structure of a type-1 protein phosphatase [4], now termed 1β. The analysis of the 1.4 kb clone, designated 1α, is described below.

3.2. *Sequence analysis of cDNA coding for protein phosphatase 1α*

The sequencing strategy is shown in fig.1 and the nucleotide and deduced amino acid sequences are presented in fig.2. The cDNA clone has 36 nucleotides preceding the putative initiating ATG codon, an open reading frame of 990 base pairs terminated by a TAG stop codon and 337 nucleotides of 3'-untranslated region which includes the start of the poly(A) tail. The proposed initiating ATG codon is preceded by a G/C rich region which conforms with the consensus eukaryotic protein synthesis initiation sequence [15]. The poly(A) tail of 17 bases was preceded 22 bases upstream by a polyadenylation signal (AATAAA). The molecular mass of the protein calculated from the sequence is 37479 Da, very similar to the apparent molecular mass (37 kDa) estimated by SDS-polyacrylamide gel electrophoresis (review [1]).

3.3. *Comparison of the sequences of protein phosphatases 1α and 1β*

Protein phosphatases 1α and 1β are identical from amino acid 34 of 1α to the C-terminus (residue 330 of 1α, fig.3). However, the N-terminal 33 amino acids of 1α are completely different from the N-terminal section of 1β, which is 19 residues shorter (fig.3). The nucleotide sequences of the 5'-non-coding regions and the N-terminal coding sections (nucleotides 1–133 of 1α and 1–280 of 1β) are likewise completely different (fig.3). In contrast, the nucleotide sequences of the rest of the coding regions and the 3'-non-coding region

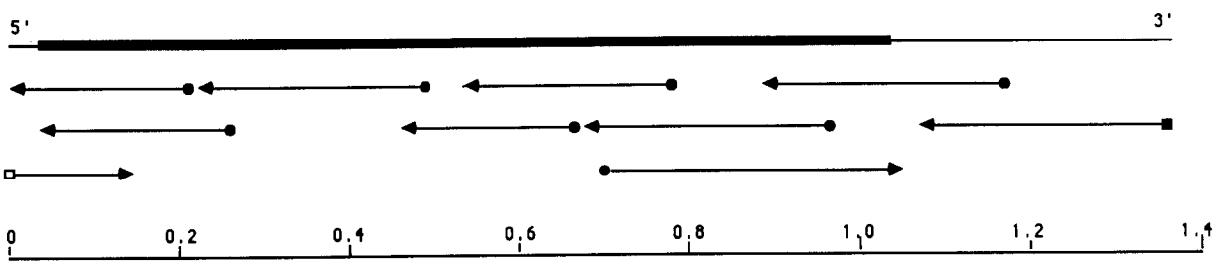


Fig.1. Strategy used to sequence the cDNA clone coding for protein phosphatase 1α. The scale indicates the nucleotide position in kilobases from the 5'-end of the cDNA insert. The arrows indicate the direction and length of DNA sequences obtained. All sequences were determined at least twice. Sequences were initiated with an M13 primer (■), specific oligonucleotide primers (●) or a Bluescript primer (□).

```
                                                                5' - GCCCGGAGCTGCTGGGCCCCGAGCGGCGGCGCCGCC      36

ATGTCCGACAGCGAGAAGCTCAACCTGGAC  TCTATCATCGGGCGCCTGCTGGAAGTGCAG  GGCTCGCGGCCCGGAAAGAATGTGCAGCTG  126
 M  S  D  S  E  K  L  N  L  D    S  I  I  G  R  L  L  E  V  Q    G  S  R  P  G  K  N  V  Q  L   30

ACGGAGAACGAGATCCGTGGTCTGTGCCTC  AAATCCCGGGAGATCTTCCTGAGCCAGCCC  ATTCTGCTGGAGCTGGAGGCGCCCCTCAAG  216
 T  E  N  E  I  R  G  L  C  L    K  S  R  E  I  F  L  S  Q  P    I  L  L  E  L  E  A  P  L  K   60

ATCTGCGGTGACATCCACGGCCAGTACTAC  GACCTGCTGCGGCTGTTCGAGTACGGCGGC  TTCCCCCCAGAGAGCAACTACCTGTTCCTG  306
 I  C  G  D  I  H  G  Q  Y  Y    D  L  L  R  L  F  E  Y  G  G    F  P  P  E  S  N  Y  L  F  L   90

GGTGACTACGTGGACCGCGGCAAGCAGTCC  CTGGAGACCATCTGCCTGCTGCTGGCCTAC  AAGATCAAGTACCCCGAGAACTTCTTCCTG  396
 G  D  Y  V  D  R  G  K  Q  S    L  E  T  I  C  L  L  L  A  Y    K  I  K  Y  P  E  N  F  F  L  120

CTGCGCGGGAACCACGAGTGCGCCAGCATC  AACCGCATCTACGGCTTCTACGACGAGTGC  AAGAGACGCTACAACATCAAGCTGTGGAAG  486
 L  R  G  N  H  E  C  A  S  I    N  R  I  Y  G  F  Y  D  E  C    K  R  R  Y  N  I  K  L  W  K  150

ACGTTCACCGACTGCTTCAACTGCCTGCCC  ATCGCGGCCATTGTGGACGAGAAGATATTC  TGCTGCCATGGCGGCCTCTCCCCCGACCTG  576
 T  F  T  D  C  F  N  C  L  P    I  A  A  I  V  D  E  K  I  F    C  C  H  G  G  L  S  P  D  L  180

CAGTCCATGGAGCAGATCCGGCGCATCATG  CGGCCCACGGACGTGCCCGACCAGGGCCTG  CTGTGTGACCTGCTGTGGTCTGACCCCGAC  666
 Q  S  M  E  Q  I  R  R  I  M    R  P  T  D  V  P  D  Q  G  L    L  C  D  L  L  W  S  D  P  D  210

AAGGACGTGCAGGGCTGGGGCGAGAACGAC  CGCGGCGTCTCCTTCACCTTCGGCGCGGAG  GTGGTGGCCAAGTTCCTGCACAAGCATGAC  756
 K  D  V  Q  G  W  G  E  N  D    R  G  V  S  F  T  F  G  A  E    V  V  A  K  F  L  H  K  H  D  240

CTGGACCTCATCTGCCGGGCGCACCAGGTG  GTGGAGGACGGCTATGAGTTCTTTGCCAAG  CGGCAGCTGGTGACACTTTTCTCAGCCCCC  846
 L  D  L  I  C  R  A  H  Q  V    V  E  D  G  Y  E  F  F  A  K    R  Q  L  V  T  L  F  S  A  P  270

AACTACTGTGGCGAGTTCGACAACGCCGGA  GCCATGATGAGCGTGGACGAGACCCTCATG  TGCTCCTTCCAGATCCTTAAGCCGGCCGAC  936
 N  Y  C  G  E  F  D  N  A  G    A  M  M  S  V  D  E  T  L  M    C  S  F  Q  I  L  K  P  A  D  300

AAGAACAAGGGCAAGTACGGGCAGTTAAGT  GGCCTGAACCCTGGAGGCCGACCCATCACC  CCACCCCGCAACTCTGCCAAAGCCAAGAAA  1026
 K  N  K  G  K  Y  G  Q  L  S    G  L  N  P  G  G  R  P  I  T    P  P  R  N  S  A  K  A  K  K  330

TAGCCCCCGGGTGCGGCCCCCTGCCCAGACAGGGGCTGACTGCACAGGAACCATGCTGCCACGGCCGTGCCGCCTCGGCCGCCACCACGGGGAC  1120

ACGGGCCTCGGTGGATCTGTTTTTTAATGAATCACTAGCAGCACCCACCCCCCAGGCTCCCTCCAGCCTGCACCTGCGCGGCTGCAAGCAGGAT  1214

CCTGGGGCCAAGGCTGCAGCTCAGGGCGACGGCGGCCAGACTGGGTCTCGTCTCGTGCGGCCGCAGGGCTGGCAGCCGGGTCCCAGGGCAGCCC  1308

GTCTGGTCTCCTGAATAAAGGTCAAAGCTGGATTCTGAAACAAAAAAAAAAAAAAAAA - 3'                                1366
```

Fig.2. The cDNA and translated protein sequence of the catalytic subunit of protein phosphatase 1α.

(nucleotides 134–1345 of 1α, and 281–1492 of 1β) are identical except for two nucleotides at positions 285 and 1162 (1α) which are A and C, respectively, in 1α, but G and T, respectively, in 1β.

3.4. *Genomic sequences hybridising to protein phosphatase 1α and 1β catalytic subunit*

The full-length cDNA coding for the 1α catalytic subunit hybridised with a 28 kb *Eco*RI fragment of

```
1βcDNA      5'-GGCACCTTGGACCACCACGAGGCTCTAGGGGTGGCAGAGGGGGTCCTGCACAATAGCAGAGAGGCCGTGGTTCCCGTGGAGCTCGGGGC   93

1βcDNA      AATCGTGGGGGCAATGTAGGAGGCAAGAGAAAGGCAGACGGGTACAACCAACCTGATTCCAAGCGTCGTCAGACCAACAACCAACAGAACTGGG  183
1αcDNA                                                                  5'-GCCCGGAGCTGCTGGGCCCCGAGCGGCGGCGCCGCC   36

1βcDNA      GTTCCCAACCCATCGCTCAGCAGCCGCTTC  AGCAAGGTGGTGACTATTCCGGTAACTATG  GTTACAATAATGACAACCAGTGAATATTTA  273
1αcDNA      ATGTCCGACAGCGAGAAGCTCAACCTGGAC  TCTATCATCGGGCGCCTGCTGGAAGTGCAG  GGCTCGCGGCCCGGAAAGAATGTGCAGCTG  126
1αProtein   M  S  D  S  E  K  L  N  L  D    S  I  I  G  R  L  L  E  V  Q    G  S  R  P  G  K  N  V  Q  L   30
1βProtein                                                                   M  V  T  I  M  T  T  S  E  Y  L   11

1βcDNA      TCAGGATACGAGATCCGTGGTCTGTGCCTC  ---
1αcDNA      ACGGAGAACGAGATCCGTGGTCTGTGCCTC  ---
1αProtein   T  E  N  E  I  R  G  L  C  L    -
1βProtein   S  G  Y  E  I  R  G  L  C  L    -
```

Fig.3. Comparison of the N-terminal protein sequence and 5'-non-coding and coding regions of protein phosphatases 1α and 1β.

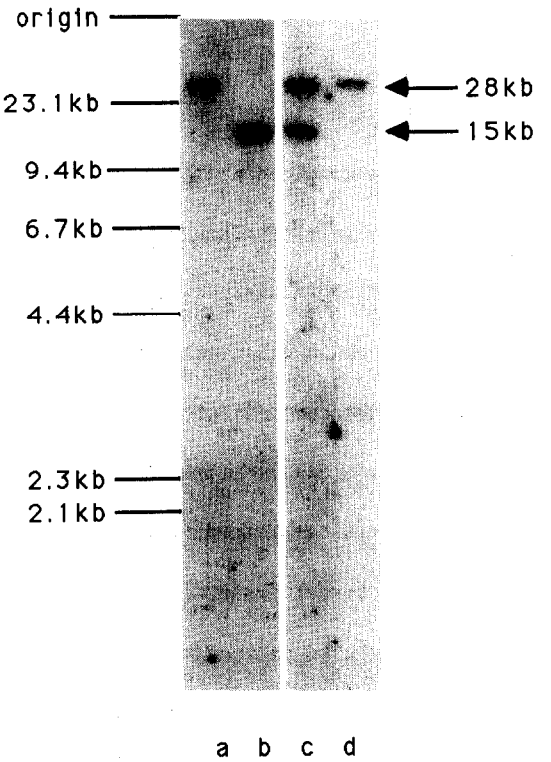


Fig.4. Southern blot analysis of rabbit genomic DNA digested with *Eco*RI, 10 μg per lane. ^{32}P-labelled probes used for hybridisation were: a, the full-length 1α cDNA insert; b, nucleotides 1–310 of the 1β cDNA; c, the full-length 1β cDNA; d, nucleotides 311–1345 of the 1β cDNA.

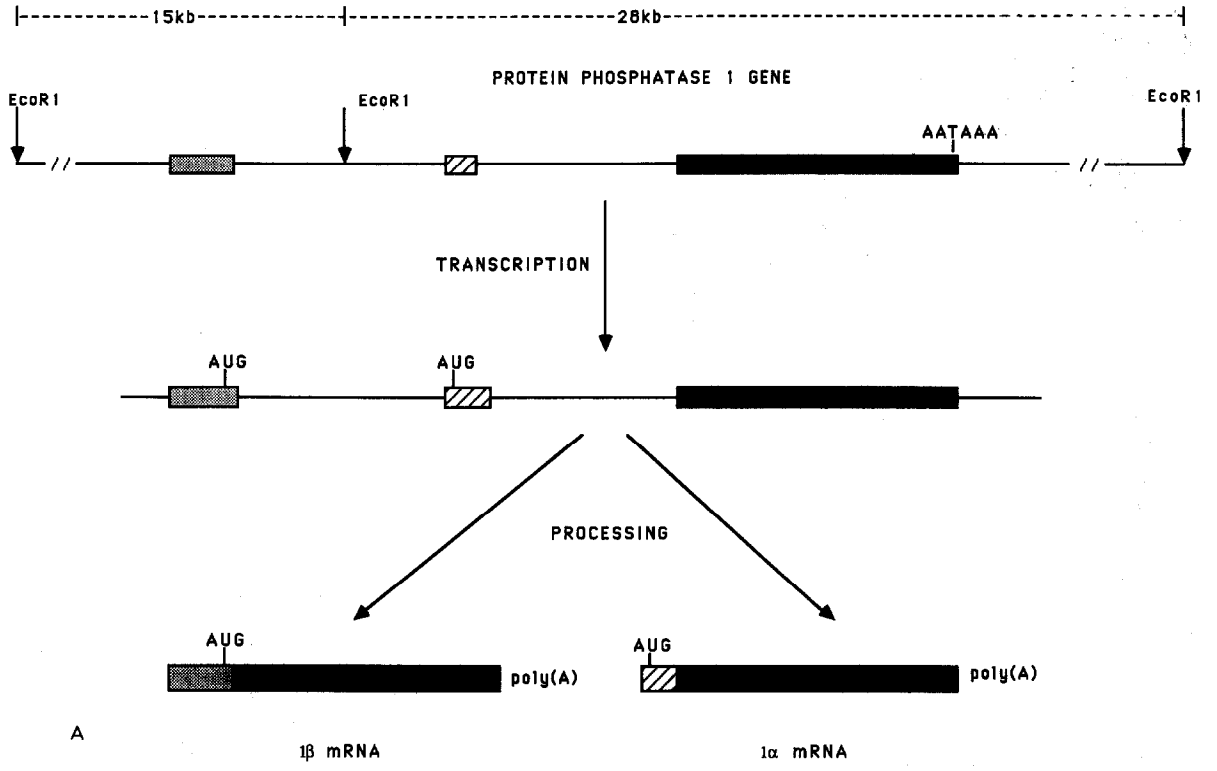
rabbit DNA in Southern blotting experiments (fig.4). Since no other hybrising bands were visualized, the entire gene for the catalytic subunit must be contained within this 28 kb gene fragment. In contrast, the full-length 1β cDNA hybridised with a 15 kb fragment, as well as the 28 kb fragment. Hybridisation with a *Sma*I/*Eco*RI fragment of 1β (bases 311–1492), encoding the identical C-terminal sections of 1α and 1β detected only the 28 kb band, indicating that the genomic sequences for the C-terminus of both 1α and 1β lie within the 28 kb fragment. Hybridisation with the 5′-non-coding region and sequences encoding the N-terminal 23 amino acids of 1β (an *Eco*RI/*Sma*I fragment comprising nucleotides 1–310) detected the 15 kb band only (fig.4).

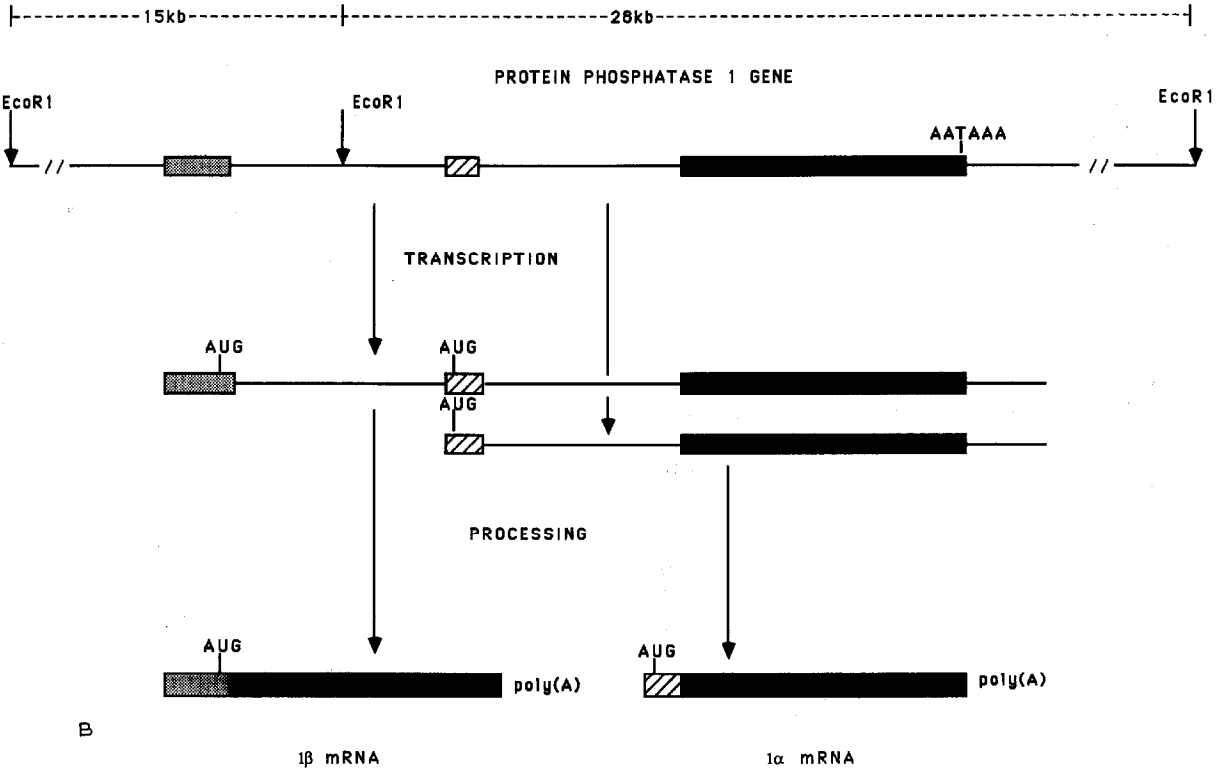
4. DISCUSSION

The nucleotide and deduced amino acid sequence of protein phosphatases 1α and 1β were completely different in the 5′-non-translated region and the N-termini of the proteins (fig.3). Protein phosphatase 1α is 19 amino acid residues longer than phosphatase 1β and residues 20–33 of 1α show no homology with amino acids 1–14 of the phosphatase 1β. This contrasts with the virtual identity in nucleotide and amino acid sequence after this position. The identity in the 3′-non-coding regions suggests that phosphatases 1α and 1β are the products of the same gene, and that the divergent N-terminal sequences are either generated by differential transcription or processing. The two nucleotide differences that were seen at positions 285 and 1162 are not sequencing errors, since the sequencing gels were very clear in these regions. They are probably explained by the presence of allelic genes in the rabbit population, as found for other proteins [16–18]. The cDNA library [7] was prepared from pooled skeletal muscle from a family of 11-day-old New Zealand white rabbits, allowing for a maximum of 4 allelic genes to be present.

Southern blot analysis of rabbit genomic DNA showed that while the nucleotide sequences coding for the N-termini of 1α and 1β were in two separate *Eco*RI fragments (28 and 15 kb, respectively), those coding for the identical central and C-terminal sections of 1α and 1β were in the 28 kb *Eco*RI fragment only. While it is possible that the central and C-terminal coding sequences were duplicated in the 28 kb fragment, this is unlikely for the following reason. When the cDNA coding for 1β was used as a probe, the 28 kb hybridising

Fig.5. Schematic representation of the protein phosphatase 1α/1β gene and mRNAs. A. Processing of a common precursor mRNA to yield the mature 1α and 1β mRNAs. B. Differential transcription initiation generating 1α and 1β mRNAs. Hatched boxes and stippled boxes are unique for 1α and 1β sequences, respectively. Solid boxes are common to both. Solid lines represent intervening sequences. AUG and AATAAA are the putative translation initiation codon and the polyadenylation signal, respectively.


15kb
28kb
PROTEIN PHOSPHATASE 1 GENE
EcoR1
EcoR1
EcoR1
AATAAA
TRANSCRIPTION
AUG
AUG
PROCESSING
AUG
poly(A)
AUG
poly(A)
A
1β mRNA
1α mRNA


15kb
28kb
PROTEIN PHOSPHATASE 1 GENE
EcoR1
EcoR1
EcoR1
AATAAA
TRANSCRIPTION
AUG
AUG
AUG
PROCESSING
AUG
poly(A)
AUG
poly(A)
B
1β mRNA
1α mRNA

band was roughly of equal intensity to the 15 kb band, and not double the intensity as might have been expected if duplicate genes were present. Southern blot analysis therefore suggests that the same genomic sequences encode the central and C-terminal regions of phosphatases 1α and 1β, while those encoding the N-termini are distinct, and in the case of the N-terminus of 1β are distant from the central and C-terminal coding sequences. Fig.5 illustrates schematically the most likely gene structure of protein phosphatase 1 and the ways in which 1α and 1β mRNAs could be generated by alternative splicing at the 5'-end of the common precursor message (fig.5A) or by transcription initiation at different promoters followed by differential splicing (fig.5B). Detailed analysis of the gene will be required to substantiate the proposed gene structure. Production of divergent N-termini from the same gene has now been demonstrated for myosin light chains [19], granulocyte macrophage stimulating factor [20] and tyrosine hydroxylase [21,22].

Three forms of phosphatase 1, termed protein phosphatases 1_I, 1_G and 1_M have been identified in skeletal muscle. Protein phosphatase 1_I is an inactive cytosolic form, which is a complex between a catalytic subunit and inhibitor-2 [23–25]. Protein phosphatase 1_G is an active form, which is bound to glycogen-protein particles and likely to be the species involved in the regulation of glycogen metabolism. It is composed of a catalytic subunit complexed to a 103 kDa G-subunit responsible for anchoring the phosphatase to glycogen [26,27]. Protein phosphatase 1_M is also an active species that is tightly bound to myosin and likely to be the form that dephosphorylates the myosin P-light chain. It has not yet been purified to homogeneity, but by analogy with protein phosphatase 1_G, may be composed of a catalytic subunit complexed to a myosin-binding (M) subunit [28]. Although catalytic subunits of protein phosphatases 1_I and 1_G are indistinguishable by peptide mapping [24,26], phosphatases 1α and 1β are so similar that they would not have been distinguished by this criterion. An attractive hypothesis would be that the different N-terminal sequences of 1α and 1β are the regions that interact with the glycogen- or myosin-binding components, or inhibitor-2. They may therefore play a key role in targetting protein phosphatase 1 to different subcellular locations.

Partial amino acid sequencing of a phosphatase 1 catalytic subunit (probably derived from 1_G) has shown that this preparation contains 1α [29]. Detailed structural analysis of the catalytic subunits isolated from 1_M and 1_I will be required to see whether they contain 1α or 1β.

Acknowledgements: This work was supported by a grant from the Wellcome Trust and by Group Support from the Medical Research Council, London. I thank Mrs Valerie Kirk for expert technical assistance and Mr Alastair I.H. Murchie for synthesis of oligonucleotide primers.

REFERENCES

[1] Cohen, P. (1985) Eur. J. Biochem. 151, 439–448.
[2] Ballou, L.M. and Fischer, E.H. (1987) in: The Enzymes (Boyer, P.D. and Krebs, E.G. eds) vol.17, pp.312–361.
[3] Ingebritsen, T.S. and Cohen, P. (1983) Eur. J. Biochem. 132, 255–261.
[4] 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.
[5] Da Cruz e Silva, O.B., Alemany, S., Campbell, D.G. and Cohen, P.T.W. (1987) FEBS Lett. 221, 415–422.
[6] Da Cruz e Silva, O.B. and Cohen, P.T.W. (1987) FEBS Lett. 226, 176–178.
[7] Da Cruz e Silva, E.F. and Cohen, P.T.W. (1987) FEBS Lett. 220, 36–42.
[8] Cohen, P.T.W., Berndt, N., Da Cruz e Silva, O.B. and Campbell, D.G. (1987) 1st International Meeting on Post-translational Modifications of Proteins and Ageing, Ischia, Italy, May 1987 (Zappia, V. ed.) pp.549–557.
[9] Kieny, M.P., Lathe, R. and Lecocq, J.P. (1983) Gene 26, 91–99.
[10] Sanger, F., Nicklen, S. and Coulson, A.R. (1977) Proc. Natl. Acad. Sci. USA 74, 5463–5467.
[11] Biggin, M.D., Gibson, T.J. and Hong, G.F. (1983) Proc. Natl. Acad. Sci. USA 80, 3963–3965.
[12] Mizusawa, S., Nishimura, S. and Seela, F. (1986) Nucleic Acids Res. 14, 1319–1324.
[13] Blin, N. and Stafford, D.W. (1976) Nucleic Acids Res. 3, 2303–2306.
[14] Feinberg, A.P. and Vogelstein, B. (1983) Anal. Biochem. 132, 6–13 and addendum (1984) 137, 266–267.
[15] Kozak, M. (1984) Nucleic Acids Res. 12, 857–872.
[16] Hunter, T. and Munro, A. (1969) Nature 223, 1270–1272.
[17] Galizzi, A. (1970) Eur. J. Biochem. 17, 49–55.
[18] Cohen, P.T.W., Burchell, A. and Cohen, P. (1976) Eur. J. Biochem. 66, 347–356.
[19] Nabeshima, Y., Fujii-Kuriamya, Y., Muramatsu, M. and Ogata, K. (1984) Nature 308, 333–338.
[20] Stanley, E., Metcalf, D., Sobieszczuk, P., Gengh, N.M. and Dunn, A.R. (1985) EMBO J. 4, 2569–2573.
[21] Grima, B., Lamouroux, A., Boni, C., Julien, J., France, J. and Mallet, J. (1987) Nature 326, 707–711.

[22] O'Malley, K.L., Anhalt, M.J., Martin, B.M., Kelsoe, J.R., Winfield, S.L. and Ginns, E.I. (1987) Biochemistry 26, 6910–6914.
[23] Hemmings, B.A., Resink, T.J. and Cohen, P. (1982) FEBS Lett. 15, 319–324.
[24] Ballou, L.M., Brautigan, D.L. and Fischer, E.H. (1983) Biochemistry 22, 3393–3399.
[25] Tung, H.Y.L. and Cohen, P. (1984) Eur. J. Biochem. 145, 57–64.
[26] Stralfors, P., Hiraga, A. and Cohen, P. (1985) Eur. J. Biochem. 149, 295–303.
[27] Hiraga, A., Kemp, B.E. and Cohen, P. (1987) Eur. J. Biochem. 163, 253–258.
[28] Chisholm, A.K. and Cohen, P. (1988) Biochim. Biophys. Acta, in press.
[29] Johnson, G.L., Brautigan, D.L., Shriner, C., Jaspers, S., Arino, J., Mole, J.E., Miller, J.B. and Mumby, M.C. (1987) Mol. Endocrinol. 1, 745–748.