

Isolation and sequence analysis of a cDNA clone encoding a type-1 protein phosphatase catalytic subunit: homology with protein phosphatase 2A

Norbert Berndt, David G. Campbell, F. Barry Caudwell, Philip Cohen, Edgar F. da Cruz e Silva, Odete B. da Cruz e Silva and Patricia T.W. Cohen

Department of Biochemistry, University of Dundee, Dundee DD1 4HN, Scotland

Received 27 August 1987

A 1.5 kb clone containing the full-length coding sequence of a type-1 protein phosphatase catalytic subunit has been isolated from a rabbit skeletal muscle cDNA library constructed in λ gt10. The protein sequence deduced from the cDNA contains 311 residues and has a molecular mass of 35.4 kDa. A single mRNA species at 1.6 kb was visualized by Northern blotting. The type-1 protein phosphatase was strikingly homologous to protein phosphatase 2A, 49% of the amino acids between residues 11 and 280 being identical. The first 10 and last 31 residues were dissimilar. Residues 1–101 of the type-1 protein phosphatase also showed 21% sequence identity with a region of mammalian alkaline phosphatases.

Protein phosphatase; cDNA cloning; Nucleotide sequence; Amino acid sequence homology

1. INTRODUCTION

Protein phosphatases 1 and 2A, two of the four principal threonine/serine-specific protein phosphatase catalytic subunits in the cytosol of mammalian cells (reviewed in [1,2]), can be distinguished in several ways. Type-1 protein phosphatase is inhibited by nanomolar concentrations of two thermostable proteins inhibitor-1 and inhibitor-2 [3] and dephosphorylates the β -subunit of phosphorylase kinase [3], whereas type-2A protein phosphatase is at least several hundred-fold less sensitive to these inhibitors [3,4] and dephosphorylates the α -subunit of phosphorylase

kinase preferentially [3]. Furthermore, peptide mapping studies have shown that the two phosphatases are the products of distinct genes [5,6].

Nevertheless, the catalytic (C) subunits of protein phosphatases 1 and 2A resemble one another in several ways. Their apparent molecular masses determined by SDS-polyacrylamide gel electrophoresis are very similar (37 kDa for protein phosphatase 1 and 36 kDa for protein phosphatase 2A [5]). They have broad and somewhat similar substrate specificities [3], are both active in the absence of divalent cations towards most substrates (in contrast to protein phosphatases 2B and 2C), and are the only two enzymes in mammalian tissues with significant phosphorylase phosphatase activity [3,7]. In addition, it has been reported that a number of monoclonal antibodies prepared against the C-subunit of protein phosphatase 2A cross-react with protein phosphatase 1 [4,8] and vice versa [8].

Recently, we isolated and sequenced a cDNA

Correspondence address: N. Berndt, Department of Biochemistry, Medical Sciences Institute, The University, Dundee DD1 4HN, Scotland

The nucleotide sequence presented here has been submitted to the EMBL/GenBank database under the accession number Y00701

clone encoding a type-2A protein phosphatase C-subunit from rabbit skeletal muscle [9]. Here, we describe the isolation of a further cDNA clone encoding a type-1 protein phosphatase C-subunit from the same tissue, and compare its sequence with that of protein phosphatase 2A. A preliminary account of part of this work was presented at the 1st International Meeting on Post-translational Modifications of Proteins and Ageing [10].

2. MATERIALS AND METHODS

2.1. Isolation of peptides from a type-1 protein phosphatase C-subunit

The glycogen-bound form of protein phosphatase 1 (4 mg) was purified from rabbit skeletal muscle as in [11] and subjected to SDS-polyacrylamide gel electrophoresis to separate the C-subunit from the glycogen-binding (G)-subunit. The gel was stained with Coomassie blue, destained in 43% methanol/7% acetic acid, washed with water, and the band containing about 0.5 mg of C-subunit was excised and broken into small pieces. The gel pieces were dried and suspended in 2.5 ml of 0.2 M *N*-methylmorpholine acetate, pH 8.1. Trypsin or chymotrypsin (5 μ g) was added and the suspension incubated at 37°C. Two further additions of proteinase (3.5 μ g) were made after 16 and 24 h. After 36 h, the supernatant was removed, and the gel washed with methanol (2 ml). The methanol extract and first supernatant were combined, dried, redissolved in 1.0 ml of 0.1% (v/v) trifluoroacetic acid, centrifuged for 15 min at $13000 \times g$ and the supernatant removed. The free C-subunit of protein phosphatase 1 [5] was also cleaved with CNBr in 70% formic acid. All peptides were separated on a Vydac C₁₈ column and analysed as in [12].

The free C-subunit of protein phosphatase 1 isolated as in [5] was cleaved with CNBr as described for protein phosphatase 2A [9], chromatographed on the Vydac C₁₈ column and analysed as described above.

2.2. Isolation and sequence analysis of cDNA clones coding for a type-1 protein phosphatase

Oligonucleotides were constructed as in [9]. Screening of the rabbit skeletal muscle λ gt10

cDNA library, isolation of positive clones and purification of their DNA was carried out as described in [13]. The DNA from positive clones was digested with *Eco*RI and the largest insert ligated into the *Eco*RI site of Bluescript pKS-M13⁺ (Stratagene Cloning Systems, San Diego, USA). Following transformation of *E. coli* JM109 and alkaline lysis of the cells, the recombinant plasmid was purified on caesium chloride gradients. A set of 'nested' deletions was produced using exonuclease III and mung bean nuclease [14,15], and DNA containing appropriate deletions was prepared by a rapid 'small scale' boiling method [16]. The double-stranded DNA was sequenced using the dideoxy chain termination procedure [17], [α -³⁵S]dATP α S and buffer gradient gels [18]. The sequence of the complementary strand was determined by removing restriction fragments and sequencing the religated recombinants, or by using synthetic oligonucleotides as primers. To overcome the ambiguities often encountered with GC-rich DNA, most of the reactions were also performed in the presence of 7-deaza-2'-dGTP in place of dGTP [19]. Additionally, some gels contained 40% formamide and 7 M urea to minimize the formation of secondary structures in the reaction products [20].

3. RESULTS

3.1. Sequences of tryptic, chymotryptic and cyanogen bromide peptides

Peptide sequences comprising 145 residues of the C-subunit of protein phosphatase 1_G were determined and are given in table 1. Five of these were used to construct the oligonucleotide probes shown in table 2.

3.2. Screening of the cDNA library

Screening of 6×10^4 recombinants yielded four clones that were positive with the 29 base oligonucleotide 4 (table 2). Electrophoresis of restriction digests showed the cDNA insert sizes to be 0.6 kb, 1.0 kb, 1.4 kb and 1.5 kb. The two largest clones were also positive with oligonucleotides 1, 2, 3 and 5. The 1.5 kb clone was selected for sequence analysis and subcloned into Bluescript pKS-M13⁺ for DNA sequencing.

Table 1

Amino acid sequences of tryptic (T), chymotryptic (C), and cyanogen bromide (CB) peptides isolated from the C-subunit of protein phosphatase 1 and the acetonitrile concentrations at which they were eluted from the Vydac C₁₈ column

Peptide	Sequence	Acetonitrile (%)	Molecular mass (M + H) ⁺		Position in sequence
			Expected	Measured	
T1	E I F L S Q P I L L E L E A P L K	35			25– 41
T2	I C G D I H G Q Y T D L L R	24	1604	1604	42– 55
T3	L F E Y G G F P P E A N Y L F L G D Y V D R	33	2584	2584	56– 77
T4	T F T D C F N C L P I A A I V D E K	32			132–149
T5	I F C C H G G L S P D L Q S M E Q I R	27			150–168
T6	I M R P T D V X D T G L L X D L L	34			170–186
T7	D V Q G W G E N D R	14			193–202
T8	A H Q V V E D G Y E F F A K	25			228–241
C1	G E N D R G V S F	17			198–206
C2	T F G A D V V S K F	24	1070	1070	207–216
CB1	E Q I R R I M	14			165–171

X, denotes residues that could not be positively identified. Amino acids that differ from those predicted from the cDNA sequence are underlined. Sequences were determined on an Applied Biosystems 470A gas-phase sequencer and molecular masses on a VG-250SE fast-atom bombardment mass spectrometer

3.3. DNA sequencing

This was performed on both strands of the 1.5 kb insert in Bluescript pKS-M13⁺ as shown in fig.1, and the nucleotide and deduced amino acid sequence is presented in fig.2. The cDNA clone consists of 1492 base pairs, and has 240 base pairs

of 5'-non-coding region, an open reading frame of 933 base pairs and 319 base pairs of 3'-non-coding region, assuming that the initiating ATG codon is at the position indicated (see section 4). The coding region is terminated by a TAG stop codon. At position 1469 there is the AATAAA signal for

Table 2
Oligonucleotides used to screen the cDNA library

Peptide T2	I C G D I H G Q Y T D L L R
Oligonucleotide-1 3'	TAG ACA CCG CTG TAG GTA CCG GTC ATG TGT CTG GA 5'
Peptide T3	L F E Y G G F P P E A N Y L F L G D Y V D R
Oligonucleotide-2 3'	AAA CTC ATA CCG CCG AAG GGG GGA CTC CGG TTG ATG GAC AAG GAC CCG CTG ATA CAC CTG 5'
Peptide T4	T F T D C F N C L P I A A I V D E K
Oligonucleotide-3 3'	AAG TGT CTG ACG AAG TTG ACG GAC GGG TAA CGA CGG TAA CAC CTA CTC TCC 5'
Peptide T7	D V Q G W G E N D R
Oligonucleotide-4 3'	CTG CAC GTC CCC ACC CCC CTC TTG CTG GC 5'
	I I I
Peptide T8	A H Q V V E D G Y E F F A K
Oligonucleotide-5 3'	CGG GTA GTC CAC CAC CTC CTA CCG ATA CTC AAG AAA CGG TTC 5'

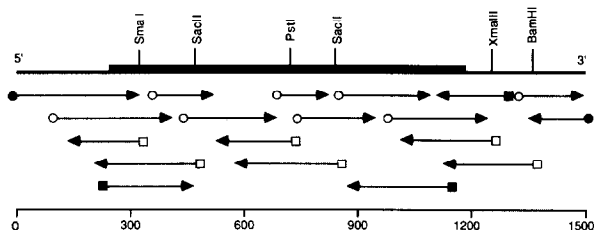


Fig.1. Partial restriction map and strategy used to sequence the cDNA clone. The scale indicates the nucleotide position in bases from the 5'-end of the cDNA insert. The arrows indicate the direction and length of the DNA sequences obtained. All sequences were determined at least twice. Sequences were initiated with Bluescript primers for the intact insert (●), 'nested' deletions (○) and restriction fragments (□), and with specific oligonucleotide primers (■).

cleavage of the messenger RNA and addition of the poly(A)⁺ tail. The use of 7-deaza-2'-dGTP in sequencing reactions and formamide gels eliminated most of the difficulties in reading GC rich regions. However, it should be noted that a section near the N-terminus of the protein (bases 263–272) was still difficult to interpret.

The 145 residues of peptide sequence determined by primary structure analysis were all found in the cDNA sequence (fig.2) establishing that the clone encodes a type-1 protein phosphatase C-subunit. Tryptic peptides were preceded by lysine or arginine, chymotryptic peptides by phenylalanine or tryptophan, and the cyanogen bromide peptide by methionine, as expected. However, five differences were found between the peptide sequences and the primary structure deduced from the

5'-GGCACCTTGGACCACACGAGGCTCTAGGGGTGGCAGAGGGGCTCTGCACATAG	56
CAGAGAGGGCGTGGTTCCTCCGTTGGAGCTCGGGGCAATCGTGGGGCAATGTAGGAGGCAAGAGAGGCGAGCGGTACACACACCTGATTC	148
CAGCGTCGTCAGACCAACACCAACAGAACTGGGGTTCACACCATCGCTCAGCAGCGCTTCAGCAGGTGGTGACTATTCCGGTAACT	240
ATGGTTACAAATATGACACACAGTGAATAT TTATCAGGATACGAGATCCGTGGTCTGTGC CTCAAATCCGGGAGATCTTCTGAGCCAG	330
M U T I M T T S E Y L S G Y E I R G L C L K S R E I F L S Q	30
CCCATCTGCTGGAGCTGGAGGCGCCCTC AAGATCTGCGGTGACATCCAGGCCAGTAC TACGACCTGCTGCGGTGTTCGAGTACGGC	420
P I L L E L E A P L K I C G D I H G Q Y V D L L R L F E Y G	60
GGCTTCCCCCGGAGAGCACTACCTGTTCT GGGGTGACTACGTGGACCGGGCAGGAG TCCCTGGAGACCATGCTGCTGCTGCTGGCC	510
G F P P E S N V L F L G D V U D R G K Q S L E T I C L L L A	90
TACAGATCAAGTACCCGAGAACTTCTTC CTGCTGCGGGGAGACAGAGTGGCCAGC ATCAACCGCATCTACGGCTTCTACGACGAG	600
Y K I K Y P E N F F L L R G N H E C A S I N R I Y G F Y D E	120
TGCAGAGACGCTACACATCAGCTGTGG AAGACGTTACCGACTGCTTCACTGCCTG CCCATCGCGGCATTGTGGACGAGAGATA	690
C K R R Y N I K L W K T F T D C F N C L P I A A I U D E K I	150
TTCTGCTGCCATGGCGGCTCTCCCCGAC CTGACGTCCATGGAGCAGATCCGGCGCATC ATGCGGCCACGGACGTGCCGACGAGGGC	780
F C C H G G L S P D L Q C S M E Q I R R I M R P T D U P D Q G	180
CTGCTGTGTGACCTGCTGGTCTGACCCC GACAGGACGTGCAGGGCTGGGGCGAGAAC GACCGCGCGTCTCCTTCACTTCGGCGCG	870
L L C D L L W S D P D K D U Q G W G E N D R G U S F T F G A	210
GAGGTGGTGGCCAGTTCTCTGCACAGCAT GACCTGGACCTCATCTGCCGGGCGACCAAG GTGGTGGAGGACGGCTATGAGTTCTTTGCC	960
E U U A K F L H K H D L D L I C R A H Q V V E D G Y E F F A	240
AAGCGGACGCTGGTACACTTTTCTCAGCC CCCAACTACTGTGGCGAGTTCGACACGCC GGAGCCATGATGAGCGTGGACGAGACCCCTC	1050
K R Q L V T L F S A P N Y C G E F D N A G A M M S U D E T L	270
ATGTGCTCCTTCAGATCCTTAAGCCGGCC GACAGAGACAGGGCAGTACGGGCAAGTAA AGTGGCTGAACCTGGAGGCGGACCCATC	1140
M C S F Q I L K P A D K N K G K Y G Q L S G L N P G G R P I	300
ACCCACCCCGCACTGTGCCAAGCCAGG AATAGCCCCGGGTGCGGCCCTGCCCCAGACAGGGGCTGACTGCACAGGAACCATGCTG	1231
T P P R N S A K A K K :	311
CCACGGCCGTGCCCTCGGCCGCCACCGGGACACGGGCTCGGTGGATCTGTTTTTATGATCACTAGCAGTACCCACCCCCACGG	1323
CTCCCTCAGCCTGCACCTGCGCGGCTGCAAGCAGGATCCTGGGGCAAGGCTGCAGCTCAGGGCGACGGCGGCCAGACTGGGTCTCGTCTC	1415
GTGCGGCCGAGGGCTGGCAGCCGGGTCCAGGGGAGCCCGTCTGGTCTCTCTGAATAAGGTCAGAGCTGGATTCTG-3'	1492

Fig.2. The cDNA and translated protein sequence of the type-1 protein phosphatase C-subunit from rabbit skeletal muscle. Information obtained by amino acid sequencing and the consensus sequence for the addition of the poly(A)⁺ tail are underlined.

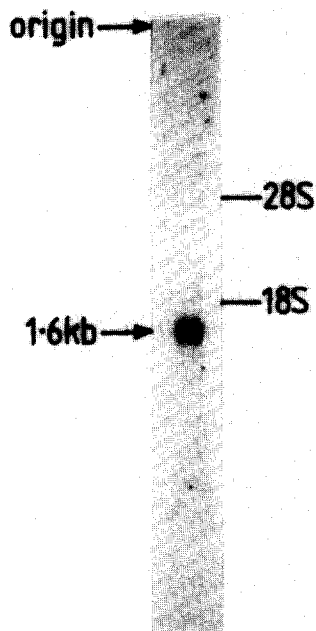


Fig.3. Northern blot of 5 μ g of poly(A)⁺ RNA from muscle. This was carried out as in [13], except that the probe was labelled by random hexanucleotide priming [21]. The probe used was the full-length cDNA clone and the final wash was $0.2 \times$ SSC at 65°C for 15 min ($1 \times$ SSC is 0.15 M sodium chloride, 0.015 M sodium citrate, pH 7). The positions of the 28 S and 18 S ribosomal RNAs and the origin are indicated.

cDNA. These differences, at positions 51, 66, 179, 211 and 214 of the amino acid sequence (table 1, fig.2) do not appear to be due to sequencing errors, and are considered further in section 4.

3.4. Northern blotting

Poly(A)⁺ RNA from rabbit skeletal muscle was shown to contain a messenger RNA species of 1.6 kb that hybridised to the 1.5 kb phosphatase 1 cDNA probe (fig.3). The difference of 0.1 kb may be accounted for by the missing poly(A)⁺ tail.

4. DISCUSSION

The molecular mass of the type-1 C-subunit deduced from the cDNA would be 35432 Da if the initiating codon is at position 241–243 of the nucleotide sequence (fig.2) or 34987 Da if the initiating ATG is the methionine at residue 5. These

values are similar to the apparent molecular mass of 37 kDa estimated by polyacrylamide gel electrophoresis [5]. The cDNA clone (1.5 kb) and messenger RNA species of 1.6 kb (fig.3) demonstrate that the type-1 protein phosphatase 1 C-subunit is not synthesized as a 70 kDa protein, as suggested in [22].

Five amino acid differences were found between the peptide sequences determined by primary structure analysis and the protein sequence deduced from the cDNA (fig.2). Since the same sequence was obtained from both strands of the DNA and the gas-phase sequencer results were unambiguous and confirmed by mass spectrometry (table 1), the differences do not appear to be explained by sequencing errors. The cDNA library was prepared from 11-day-old New Zealand White rabbits, while protein phosphatase 1 was purified from five adult rabbits obtained from a different supplier. The differences in primary structure can therefore be explained in two ways. Firstly, there could be allelic genes in the rabbit population, as found for other proteins [23–25]. Secondly, there may be more than one type-1 protein phosphatase 1 C-subunit in skeletal muscle, or the type-1 C-subunit may alter during muscle development.

The primary structure of the type-1 protein phosphatase is strikingly homologous to protein phosphatase 2A (fig.4). Homology extends from residues 11–280 of the coding region of the type-1 C-subunit, and only a single deletion in protein phosphatase 1 and one in protein phosphatase 2A is necessary to maximise homology (fig.4). The overall sequence identity is 43% (50% between residues 11 and 280), and homology rises to 59% (67% between residues 11 and 280) if conservative substitutions are included. These findings provide a structural basis for previous observations that monoclonal antibodies to protein phosphatase 2A cross-react with protein phosphatase 1 and vice versa (see section 1). One might expect that homologous regions are concerned with catalytic activity, while dissimilar sequences may be those involved in interaction with regulatory proteins [2]. The type-1 C-subunit is 12 residues shorter than the type-2A enzyme at its N-terminus, but 14 residues longer at its C-terminus.

The overall G+C content of the protein phosphatase 1 clone (62%) is high in both the coding and non-coding regions. This is accounted

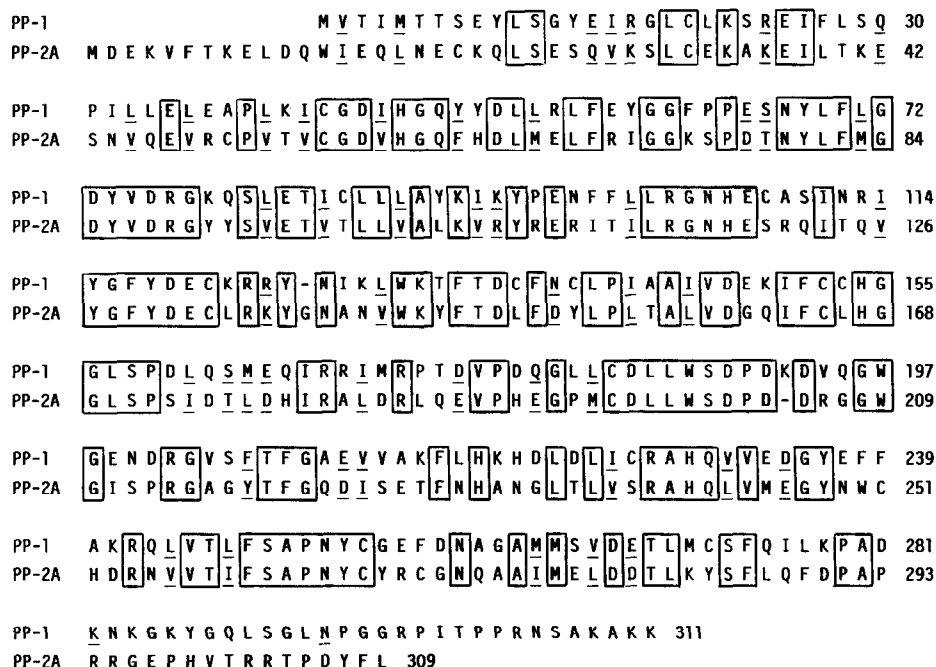


Fig.4. Comparison of the primary structures of the type-1 and type-2A protein phosphatase C-subunits from rabbit skeletal muscle. Identities are boxed and conservative replacements are underlined.

for in the coding region by the almost exclusive use of G or C (87%) in the third codon position. A similar phenomenon has been observed for other skeletal muscle proteins [13,26]. However, the G + C content of the cDNA of protein phosphatase 2A was only 45%, perhaps because it is not a muscle-specific protein [9]. For this reason, the nucleotide sequences of protein phosphatases 1

and 2A showed only 53% identity in the region coding for amino acid residues 11–280. No homology was evident in the 5'- and 3'-non-coding regions.

We reported in [9] that residues 55–95 of protein phosphatase 2A showed homology with mammalian alkaline phosphatases. Homology between the type-1 protein phosphatase and alkaline

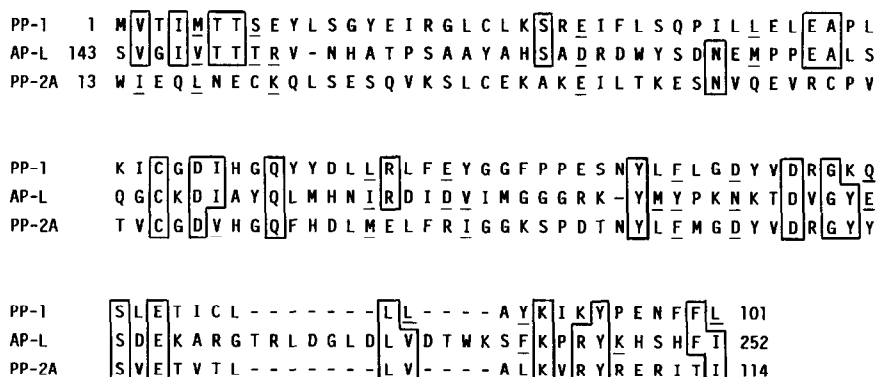


Fig.5. Homology between protein phosphatase 1 (PP-1) and protein phosphatase 2A (PP-2A) from rabbit skeletal muscle and alkaline phosphatase from human liver (APL) [27]. Identities are boxed and conservative replacements underlined. The number of the first and last amino acid in each sequence is also given.

phosphatases extends from the N-terminus to residue 101, if the sequences are aligned to maximise similarity (fig.5). In the first 101 residues there are 21 identities and a further 13 conservative replacements. However, this is not the region that contains the phosphate binding site and some of the metal ligand sites of the alkaline phosphatases.

ACKNOWLEDGEMENTS

The work was supported by grants from the Wellcome Trust, University of Dundee and Medical Research Council. O.B.C.S. is a Postgraduate Research Student of the MRC. We thank Mrs Valerie Kirk for expert technical assistance, Dr Charles Holmes for advice on peptide purification, and Mr Alastair I.H. Murchie (Dundee) and Dr Michael Brownstein (Laboratory of Cell Biology, Mental Health Institute, Bethesda, USA) for making the synthetic oligonucleotides.

REFERENCES

- [1] Ingebritsen, T.S. and Cohen, P. (1983) *Science* 221, 331–338.
- [2] Cohen, P. (1985) *Eur. J. Biochem.* 151, 439–448.
- [3] Ingebritsen, T.S. and Cohen, P. (1983) *Eur. J. Biochem.* 132, 255–261.
- [4] Brautigan, D.L., Gruppuso, P.A. and Mumby, M. (1986) *J. Biol. Chem.* 261, 14924–14928.
- [5] Tung, H.Y.L., Hemmings, B.A., Resink, T.J., Shenolikar, S. and Cohen, P. (1984) *Eur. J. Biochem.* 138, 635–641.
- [6] Silberman, S.R., Speth, M., Nemani, R., Ganapathi, M.K., Dombradi, V., Paris, H. and Lee, E.Y.C. (1984) *J. Biol. Chem.* 259, 2913–2921.
- [7] Ingebritsen, T.S., Foulkes, J.G. and Cohen, P. (1983) *Eur. J. Biochem.* 132, 263–274.
- [8] Speth, M., Alejandro, R. and Lee, E.Y.C. (1984) *J. Biol. Chem.* 259, 3475–3481.
- [9] Da Cruz e Silva, O.B., Alemany, S., Campbell, D.G. and Cohen, P.T.W. (1987) *FEBS Lett.* 221, 415–422.
- [10] Cohen, P.T.W., Berndt, N., Da Cruz e Silva, O.B. and Campbell, D.G. (1987) in: 1st Int. Meet. on Post-translational Modifications of Proteins and Ageing, Ischia, Italy, May 1987 (Zappia, V. ed.) Plenum, London, in press.
- [11] Strålfors, P., Hiraga, A. and Cohen, P. (1985) *Eur. J. Biochem.* 149, 295–303.
- [12] Holmes, C.F.B., Campbell, D.G., Caudwell, F.B., Aitken, A. and Cohen, P. (1986) *Eur. J. Biochem.* 155, 173–182.
- [13] Da Cruz e Silva, E.F. and Cohen, P.T.W. (1987) *FEBS Lett.* 220, 36–42.
- [14] Henikoff, S. (1984) *Gene* 28, 351–359.
- [15] Laskowski, M.S. (1980) *Methods Enzymol.* 65, 263–275.
- [16] Holmes, D.S. and Quigley, M. (1981) *Anal. Biochem.* 114, 193–197.
- [17] Sanger, F., Nicklen, S. and Coulson, A.R. (1977) *Proc. Natl. Acad. Sci. USA* 74, 5463–5467.
- [18] Biggin, M.D., Gibson, T.J. and Hong, G.F. (1983) *Proc. Natl. Acad. Sci. USA* 80, 3963–3965.
- [19] Mizusawa, S., Nishimura, S. and Seela, F. (1986) *Nucleic Acids Res.* 14, 1319–1324.
- [20] Martin, R. (1986) *Focus* 9, 8–10.
- [21] Feinberg, A.P. and Vogelstein, B. (1983) *Anal. Biochem.* 132, 6–13 and addendum (1984) 137, 266–267.
- [22] Brautigan, D.L., Shriner, C.L. and Gruppuso, P.A. (1985) *J. Biol. Chem.* 260, 4295–4302.
- [23] Hunter, T. and Munro, A. (1969) *Nature* 223, 1270–1272.
- [24] Galizzi, A. (1970) *Eur. J. Biochem.* 17, 49–55.
- [25] Cohen, P.T.W., Burchell, A. and Cohen, P. (1976) *Eur. J. Biochem.* 66, 347–356.
- [26] Newgard, C.B., Nakano, K., Hwang, P.K. and Fletterick, R.J. (1986) *Proc. Natl. Acad. Sci. USA* 83, 8132–8136.
- [27] Weiss, M.J., Henthorn, P.S., Lafferty, M.A., Slaughter, C., Raducha, M. and Harris, H. (1986) *Proc. Natl. Acad. Sci. USA* 83, 7182–7186.