Volume 221, number 2, 415–422 FEB 05090 September 1987

Isolation and sequence analysis of a cDNA clone encoding the entire catalytic subunit of a type-2A protein phosphatase

Odete B. da Cruz e Silva, Susana Alemany, David G. Campbell and Patricia T.W. Cohen

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

Received 18 July 1987

A 2.5 kb clone containing the full-length coding sequence of a type-2A protein phosphatase catalytic subunit has been isolated from a rabbit skeletal muscle cDNA library constructed in λgt10. The sequence of the protein deduced from the cDNA contains 309 residues (35.58 kDa). A major mRNA species at 2.0 kb and a minor component at 2.8 kb were visualized by Northern blotting in both skeletal muscle and liver. The type-2A enzyme showed weak homology with mammalian alkaline phosphatases between residues 55 and 95. The protein sequence of the type-2A phosphatase from rabbit skeletal muscle differs from that reported for the bovine adrenal enzyme in three regions.

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

1. INTRODUCTION

Four principal types of protein phosphatase catalytic subunit have been identified in the cytosol of mammalian cells which are responsible for dephosphorylating phosphoseryl and phosphothreonyl residues in proteins [1–3]. Other protein phosphatases exist, but are either located in mitochondria [4] or dephosphorylate tyrosine, rather than serine or threonine residues [5]. The four cytosolic serine/threonine-specific protein phosphatases have been subdivided into two classes, type-1 and type-2, depending on whether they dephosphorylate the β-subunit of phosphorylase kinase and are inhibited by two thermostable proteins inhibitor-1 and inhibitor-2 (type-1 protein phosphatases) or whether they dephosphorylate the α-subunit of phosphorylase kinase and are insensitive to inhibitors-1 and -2 (protein phosphatases 2A, 2B and 2C) [1–3].

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

However, it is now clear that the structure, substrate specificity and regulation of the three type-2 protein phosphatases are distinct. For example, protein phosphatase 2A is active in the absence of divalent cations, whereas protein phosphatases 2B and 2C have an absolute requirement for Ca^{2+} and Mg^{2+}, respectively.

Protein phosphatases 1, 2A and 2C have broad and overlapping substrate specificities and account for nearly all the phosphatase activity in skeletal muscle and liver extracts towards a variety of enzymes that regulate the major biodegradative and biosynthetic pathways in these tissues (review [1]). However, the phosphatases are present at similar concentrations in other tissues such as brain, where the protein substrates present in skeletal muscle and liver do not exist [6]. It is therefore likely that these phosphatases play even wider roles. For example, in brain they may regulate the phosphorylation states of proteins that control neurotransmitter synthesis and release. It is also established that the free catalytic units of protein phosphatases 1, 2A and 2B do not exist in vivo, but are complexed to other subunits that either have a

Published by Elsevier Science Publishers B.V. (Biomedical Division)
00145793/87/$3.50

regulatory function or target the phosphatases to distinct subcellular locations (review [1]).

In order to elucidate the structural relationships between each protein phosphatase catalytic subunit, and to facilitate identification of the regions which interact with regulatory macromolecules, we have undertaken the cDNA cloning of the catalytic subunits of these enzymes. In this paper, we present the first phase of this study, namely the isolation of a clone containing the entire coding region of the catalytic subunit of a type-2A protein phosphatase and the determination of its sequence. A preliminary account of part of this work was presented at the 1st International Conference on Post-Translational Modifications and Ageing in May 1987 [7].

2. MATERIALS AND METHODS

2.1. *Isolation and sequence analysis of peptides*

The free catalytic subunit of protein phosphatase 2A was purified as in [8,9] and 0.5 mg aliquots precipitated by addition of 0.1 vol. of 100% (w/v) trichloroacetic acid. Following centrifugation for 5 min at $13000 \times g$, the supernatant was discarded and the precipitate washed twice with ether to remove trichloroacetic acid. Aliquots were then either redissolved in 0.1 ml of 50% (w/v) formic acid and incubated with a crystal of CNBr for 20 h at 4°C, or resuspended in 0.1 M NH_4HCO_3 and digested for 20 h at 37°C with 0.02 mg trypsin. The digests were dried, redissolved in 1 ml of 0.1% (v/v) trifluoroacetic acid and chromatographed on a Vydac C_{18} column as in [10]. Peptides were then analysed on a Beckman 890C sequencer [10] to determine their primary structures (table 1).

2.2. *Synthesis of oligonucleotides*

Oligonucleotides complementary to two peptides synthesized by Dr Gerald Zon, Center for Drug and Biologies, Division of Biochemistry and Biophysics, Bethesda, MD) were constructed on the basis of most frequent codon usage [11], with the use of GT pairing [12] or insertion of inosine and cytosine [13] in the third position of certain codons (table 2).

2.3. *Screening of the cDNA library and identification of positive clones*

A rabbit skeletal muscle cDNA library constructed in λgt10 [14] was screened without amplification on duplicate filters using [γ-^{32}P]ATP 5′-end-labelled oligonucleotides. Putative positive clones were plaque purified and their DNA isolated from plate lysates by formamide extraction following caesium chloride density gradient purification of the phage [15].

2.4. *Subcloning and sequencing of the cDNA insert*

Phage DNA was digested with *Eco*RI, extracted with phenol and used without further purification for ligation of the insert into the *Eco*RI site of Bluescript pKS-M13$^+$ [16]. Following transformation of *Escherichia coli* JM109, recombinant plasmid DNA was purified by caesium chloride density gradient centrifugation, after alkaline lysis of the cells [17].

A set of 'nested' deletions was produced using exonuclease III and mung bean nuclease [16]. DNA from the plasmids containing the deletions was prepared as above, or by a rapid 'small-scale' boiling method [18]. DNA sequencing was performed directly on the double-stranded, supercoiled plasmid DNA [19], using the dideoxy chain termination procedure [20], [α-^{35}S]dATPαS and buffer gradient gels [21]. Gaps and ambiguities were clarified by removing restriction fragments and sequencing the religated, truncated recombinant, or by using synthetic oligonucleotides as sequencing primers.

2.5. *Northern blot analysis*

Poly(A)$^+$ RNA was prepared from rabbit skeletal muscle and rabbit liver as described in [14], and blotting according to [17]. The probe was the full-length cDNA insert in Bluescript and was labelled by nick-translation [17].

3. RESULTS

3.1. *Screening of the cDNA library*

Screening of 1.2×10^5 recombinants yielded two clones that were positive with the 36-base oligonucleotide and five that were positive with the 38-base oligonucleotide (table 2). However, no clones were positive with both oligonucleotides. Restriction digests showed the cDNA insert sizes to range from 2.3 to 2.7 kb, sufficient to contain the entire coding region. Since the 36-base

Table 1

Amino acid sequences of cyanogen bromide (CB) and tryptic (T) peptides isolated from the catalytic subunit of protein phosphatase 2A

Peptide	Sequence	Position in sequence
CB1	E L F R I G G K S P D T N Y	67– 80
CB2	E L D D T L K Y S F L Q F D P A P R X G	277–296
A1	P A P R R G E P	291–298
T1	E L D Q W I E Q L N E C K	9– 21
T2	G A G Y T F G Q D I S E T F N X A N G L T L V S	215–238
T3	Y S F L Q F D P A P R	284–294
T4	T P D Y F	304–308

The peptide A1 was generated by cleavage of an Asp–Pro bond, as described in the text. X denotes residues that could not be positively identified

oligonucleotide was constructed from a peptide whose sequence lacked amino acids with high codon degeneracy, such as serine, the two positives with this oligonucleotide were investigated first. One positive clone had an internal *Eco*RI site, but had lost one of the terminal *Eco*RI sites. Thus digestion with *Eco*RI yielded a 600 base pair insert, the rest of the insert remaining attached to the λgt10 arms. The second positive clone had a 2.5 kb insert with no internal *Eco*RI site, and was therefore subcloned into Bluescript for sequence investigation.

3.2. *DNA sequence analysis*

DNA from each deletion of the 2.5 kb insert in Bluescript was electrophoresed and blotted onto nitrocellulose. Hybridisation revealed that the oligonucleotide-binding site was present in most of the clones, but absent from those with the largest deletions. The smallest clone that hybridised to the oligonucleotide was therefore sequenced and shown to contain a cDNA sequence that encoded peptide T1 (table 2). The entire cDNA insert was subsequently sequenced and the results are shown in figs 1 and 2.

The cDNA clone is 2557 base pairs, and it has 894 base pairs of 5′-non-coding region, an ATG codon at the start of an open reading frame of 927 base pairs, and 736 base pairs of 3′-non-coding region. The coding region is terminated by two consecutive stop codons. At position 2480 there is the AATAAA signal for cleavage of the messenger RNA and addition of the poly(A)$^+$ tail. The molecular mass of the protein calculated from the sequence is 35582 Da.

The peptide sequences determined by primary

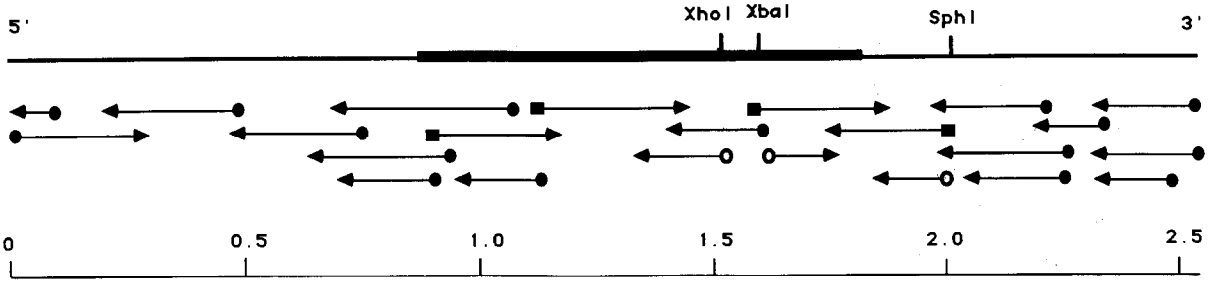

Fig.1. Restriction endonuclease map and strategy used to sequence the cDNA clone. The scale indicates the nucleotide position in kilobases from the 5′-end of the cDNA insert. The arrows indicate the direction and length of the DNA sequences determined. Sequences were initiated with Bluescript primers for the nested deletions (●), Bluescript primers for the restriction fragments (○), and specific oligonucleotide primers (■).

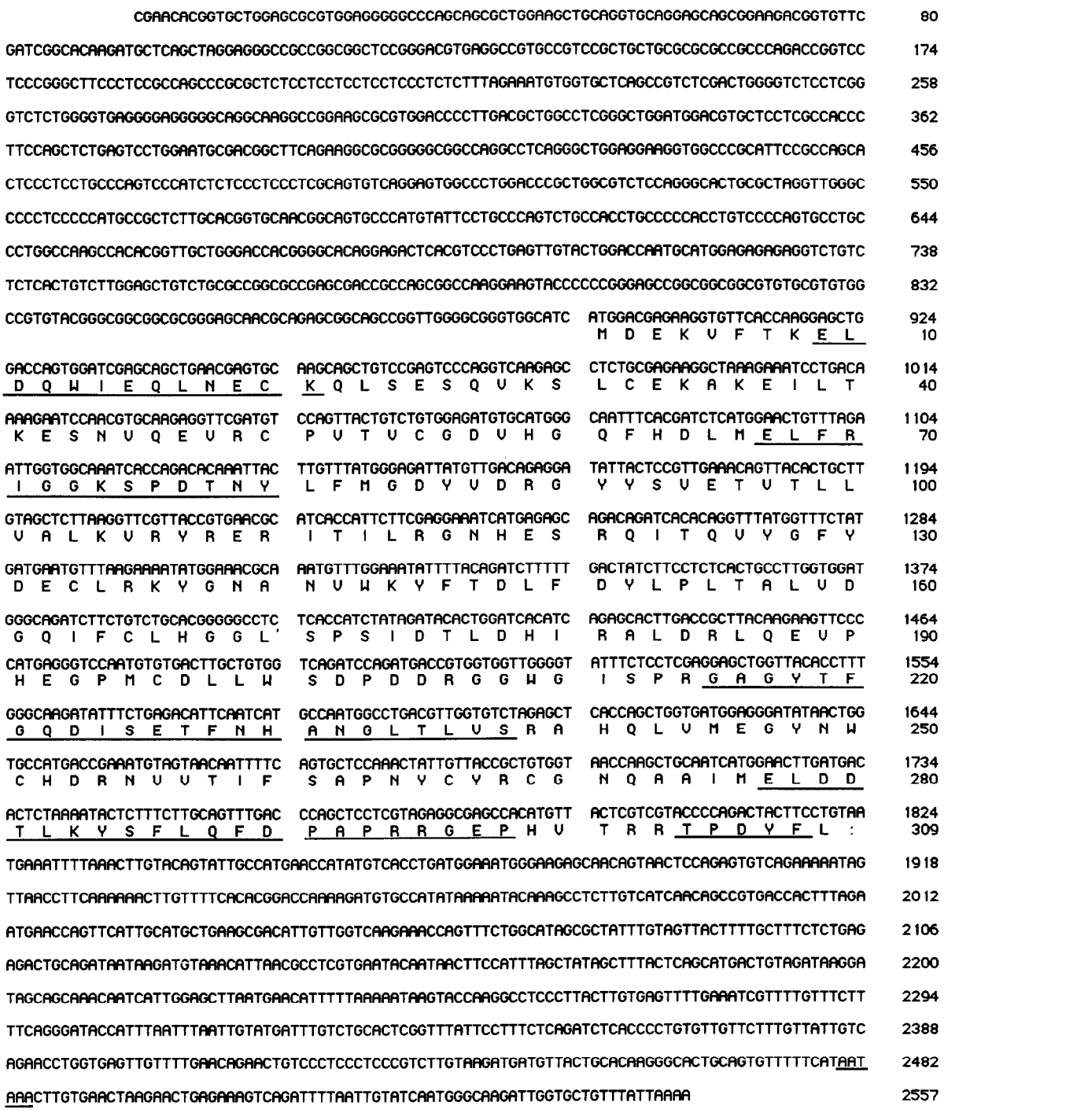

Fig.2. The cDNA and translated protein sequence of the type-2A protein phosphatase catalytic subunit from rabbit skeletal muscle. Information obtained by amino acid sequencing and the consensus sequence for addition of the poly(A)$^+$ tail are underlined.

Table 2

Oligonucleotides used to screen the cDNA library

(1) Peptide T1			E	L	D	Q	W	I	E	Q	L	N	E	C	K	
mRNA	5'	GAA	CUN	GAU	CAA	UGG	AUA	GAA	CAA	CUN	AAU	GAA	UGU	AAA	3'	
		G	UUA	C	G		U	G	G	UUA	C	G	C	G		
			G				C			G						
Oligo nucleotide (36 bases)	3'	CTT	GAG	CTG	GTT	ACC	TAG	CTT	GTT	GAG	TTG	CTT	ACG		5'	
			C							C						
(2) Peptide CB2			E	L	D	D	T	L	K	Y	S	F	L	Q	F	D	P	
mRNA	5'	GAA	CUN	GAU	GAU	ACN	CUN	AAA	UAU	UCN	UUU	CUN	CAA	UUU	GAU	CCN	3'	
		G	UUA	C	C		UUA	G	C	AGU	C	UUA	G	C	C			
			G				G			C		G						
Oligo-nucleotide (38 bases)	3'			CTG	CTG	TGI	GAG	TTT	ATG	TCG	AAG	GAG	GTT	AAG	CTG	GG	5'	
						C	C					C						

N indicates that any of the four bases can be present

structure analysis (table 1) were all found in the cDNA sequence (fig.2), establishing that the clone encodes a type-2A protein phosphatase catalytic subunit. The peptides isolated from tryptic digests were preceded by either lysine or arginine, and the CNBr peptides by methionine, apart from the peptide A1 with the N-terminal sequence PAPRRGEP (table 1), which was generated by acid cleavage of an Asp–Pro bond.

3.3. *Northern blot analysis*

Analysis of poly(A)$^+$ RNA from rabbit skeletal muscle and rabbit liver by Northern blotting showed that the protein phosphatase 2A clone recognized a major mRNA species at 2.0 kb and a minor band at 2.8 kb in both muscle and liver (fig.3). In addition, a minor 1.3 kb band was also evident.

4. DISCUSSION

The molecular mass of the type-2A catalytic subunit deduced from the cDNA (35.58 kDa) is in very close agreement with the apparent molecular mass of 36 kDa estimated by SDS-polyacrylamide gel electrophoresis [22]. In the proposed sequence, the initiator methionine is followed by an aspartyl residue and it has been noted that such Met–Asp bonds are not cleaved after synthesis [23]. Since

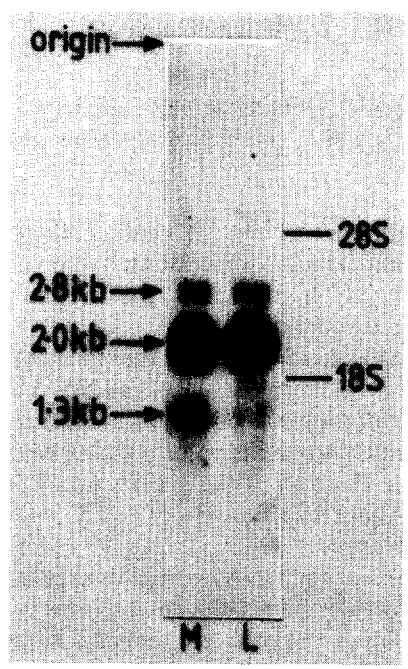

Fig.3. Northern blot of 5 μg poly(A)$^+$ RNA from muscle (M) and liver (L). The probe used was the full-length cDNA clone and the final wash was 0.2 × SSC at 65°C for 15 min (1 × SSC: 0.15 M NaCl, 0.015 M sodium citrate, pH 7). The positions of the 28 S and 18 S ribosomal RNAs and the origin are indicated.

Fig.4. Comparison of protein and cDNA sequences of rabbit skeletal muscle (this paper) and bovine adrenal [27] type-2A protein phosphatases from residues 32–40 and 292–309. The differences between the sequences are underlined.

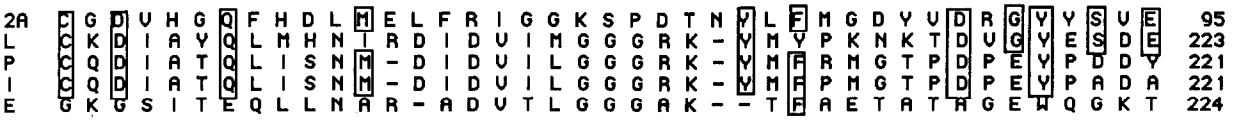

Fig.5. Homology between protein phosphatase 2A (2A) and alkaline phosphatases from human liver (L), placenta (P), intestine (I) and *E. coli* (E). Identities are boxed, and the number of the last amino acid in each sequence is given.

the N-terminus of the protein is blocked (unpublished) the initiator methionine is most likely to be acetylated. In this case, the molecular mass would be 35623 Da.

An unexpected finding was that the cDNA clone was 2.5 kb, whereas the major mRNA detected by Northern blotting was only 2.0 kb. However, since a minor mRNA band is present at 2.8 kb (fig.3), it is possible that this is the species which has been cloned. Alternatively, some of the 5′-end may be a cloning artefact.

A feature of the DNA coding for skeletal muscle proteins is their high GC content. This is striking in the coding regions, particularly in the third codon position where G or C is found 80–90% of the time [14,24]. In contrast, the cDNAs of liver proteins have about 50% G or C in the third codon position [24]. It is therefore of interest that the coding region of protein phosphatase 2A from rabbit skeletal muscle only contains 45% G or C in the third codon position, suggesting that it might not be a muscle-specific protein. This conclusion was supported by Northern blotting experiments which showed that the muscle cDNA clone hybridised equally efficiently to liver and muscle mRNA, yielding bands of identical sizes (fig.3). If protein phosphatase 2A is the same gene product in muscle and liver, this would be similar to protein phosphatase 2C, where the same two isoforms ($2C_1$ and $2C_2$) are found in both skeletal muscle and liver, as judged by peptide mapping [25].

While this paper was in preparation we received a preprint of the work of Mumby and co-workers [26] who have isolated and sequenced cDNA coding for a type-2A catalytic subunit from bovine adrenal cortex. Their sequence differs from the rabbit skeletal muscle protein in three regions. Firstly, there is a single amino acid change at residue 55, which is cysteine in the skeletal muscle sequence and arginine in the bovine adrenal protein, resulting from a single base change. Secondly, the sequence from residue 32 to 40 is completely

different (fig.4). Thirdly, the bovine adrenal sequence is completely different from residue 294 and has 16 additional residues at the C-terminus, terminating at residue 325. However, inspection of nucleotide sequences shows that the differences between residue 32 and 40 are explained by the absence of a single G in the bovine adrenal sequence at nucleotide residue 991 and insertion of a G after 1011 (fig.4). Similarly, at the C-terminus, the differences are explained by the presence of an additional T at nucleotide 1773 of the rabbit skeletal muscle cDNA clone (fig.4). It is unknown whether these discrepancies are explained by species differences or sequencing errors. However, the C-terminus of the skeletal muscle enzyme is correct, since three peptides (T3, T4 and CB2) corresponding to this region have been sequenced (table 1). We are also confident that the nucleotide sequence encoding residues 32–40 is correct since the sequences of both cDNA strands agreed in this region. If the differences are explained by sequencing errors, then there would only be a single amino acid difference between the rabbit skeletal muscle and bovine adrenal enzymes.

There are 39 differences in the third codon position and a single difference in the first codon position between the rabbit skeletal muscle and bovine adrenal cDNAs that do not change the deduced amino acid sequence. There is a further single base change in the first codon position, that accounts for the single amino acid difference at residue 55, previously discussed.

A short section of the type-2A protein phosphatase from skeletal muscle (residues 55–95) shows weak homology with the alkaline phosphatases from human placenta [27,28], liver [29] and intestine [30,31] (fig.5). The type-2A protein phosphatase and all the mammalian alkaline phosphatases are identical at six sites. At a further two, the type-2A protein phosphatase is identical to the placental and intestinal alkaline phosphatases, while at three other sites it is identical to the liver alkaline phosphatase. The overall identity between the type-2A protein phosphatase and alkaline phosphatases in this region is 27%. To obtain this homology it was necessary to introduce a single amino acid deletion in the sequences of the alkaline phosphatases with respect to protein phosphatase 2A. However, this is at a position where *E. coli* alkaline phosphatase has already undergone a deletion compared to the mammalian alkaline phosphatases [32]. The metal ligand sites and residues that interact with phosphate are known for *E. coli* alkaline phosphatase, and do not lie in this section of the protein. It is therefore unclear whether the conserved amino acids are essential for the function of mammalian phosphatases or whether they merely reflect a distant evolutionary relationship.

ACKNOWLEDGEMENTS

This work was supported by Grants from the Wellcome Trust and the University of Dundee, a Group Support Grant from the Medical Research Council, London (to P.T.W.C.) and a Programme Grant from the Medical Research Council (to Professor Philip Cohen). O.B.C.S. is the recipient of a Postgraduate Research Studentship from the Medical Research Council.

REFERENCES

[1] Cohen, P. (1985) Eur. J. Biochem. 151, 439–448.
[2] Ingebritsen, T.S. and Cohen, P. (1983) Science 221, 331–338.
[3] Ingebritsen, T.S. and Cohen, P. (1983) Eur. J. Biochem. 132, 255–261.
[4] Bradford, A.P. and Yeaman, S.J. (1986) Advances in Protein Phosphatases III (Merlevede, W. and DiSalvo, J. eds) pp.73–106, Leuven, University Press, Belgium.
[5] Foulkes, J.G. (1983) Curr. Top. Microbiol. Immunol. 107, 163–180.
[6] Ingebritsen, T.S., Stewart, A.A. and Cohen, P. (1983) Eur. J. Biochem. 132, 297–307.
[7] Cohen, P.T.W., Berndt, N., Da Cruz e Silva, O.B. and Campbell, D.G. (1987) 1st International Meeting on Post-translational Modifications and Ageing, Ischia, Italy, May 1987 (Zappia, z. ed.), in press.
[8] Resink, T.J., Hemmings, B.A., Tung, H.Y.L. and Cohen, P. (1983) Eur. J. Biochem. 133, 455–461.
[9] Tung, H.Y.L., Resink, T.J., Hemmings, B.A., Shenolikar, S. and Cohen, P. (1983) Eur. J. Biochem. 138, 635–641.
[10] Holmes, C.F.B., Campbell, D.G., Caudwell, F.B., Aitken, A. and Cohen, P. (1986) Eur. J. Biochem. 155, 173–182.
[11] Lathe, R. (1985) J. Mol. Biol. 183, 1–12.
[12] Jaye, M., H. de la Salle, Schamber, F., Ballard, A., Kohto, V., Findeli, A., Tolstoshev, P. and Lecocq, J.P. (1983) Nucleic Acids Res. 11, 2325–2335.

[13] Takahashi, Y., Kato, K., Hayashizaki, Y., Wakabayashi, T., Ohtsuka, E., Matsuki, S., Ikehara, M. and Matsubara, K. (1985) Proc. Natl. Acad. Sci. USA 82, 1931–1935.
[14] Da Cruz e Silva, E.F. and Cohen, P.T.W. (1987) FEBS Lett. 220, 36–42.
[15] Davis, R.W., Botstein, D. and Roth, J.R. (1986) in: Advances in Bacterial Genetics: A Manual for Genetic Engineering, Cold Spring Harbor Laboratory, NY.
[16] Bluescript Exo/Mung DNA Sequencing System (1986) Stratagene, San Diego, USA.
[17] Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) in: Molecular Cloning: A Laboratory Handbook, Cold Spring Harbor Laboratory, NY.
[18] Holmes, D.S. and Quigley, M. (1981) Anal. Biochem. 114, 193–197.
[19] Chen, E.Y. and Seeburg, P.H. (1985) DNA 4, 165–170.
[20] Sanger, F., Nicklen, S. and Coulson, A.R. (1977) Proc. Natl. Acad. Sci. USA 74, 5463–5467.
[21] M13 Cloning and Sequencing Handbook (1984) Amersham International, Bucks, England.
[22] Tung, H.Y.L., Alemany, S. and Cohen, P. (1985) Eur. J. Biochem. 148, 253–263.
[23] Persson, B., Flinta, C., Von Heijne, G. and Jornvall, H. (1985) Eur. J. Biochem. 152, 523–527.
[24] Newgard, C.B., Nakano, K., Hwang, P.K. and Fletterick, R.J. (1986) Proc. Natl. Acad. Sci. USA 83, 8132–8136.
[25] McGowan, C.H. and Cohen, P. (1987) Eur. J. Biochem., in press.
[26] Green, P.D., Sung-Il, Y. and Mumby, M.C. (1987) Proc. Natl. Acad. Sci. USA, 84, 4880–4884.
[27] Kam, W., Clauser, E., Kim, Y.S., Kan, W.Y. and Rutter, W.J. (1985) Proc. Natl. Acad. Sci. USA 82, 8715–8719.
[28] Millan, J.L. (1986) J. Biol. Chem. 261, 3112–3115.
[29] Weiss, M.J., Henthorn, P.S., Lafferty, M.A., Slaughter, M.R. and Harris, H. (1986) Proc. Natl. Acad. Sci. USA 83, 7182–7186.
[30] Berger, J., Garattini, E., Hua, J. and Udenfriend, S. (1987) Proc. Natl. Acad. Sci. USA 84, 695–698.
[31] Henthorn, P.S., Raducha, M., Edwards, Y.H., Weiss, M.J., Slaughter, C., Lafferty, M.A. and Harris, H. (1987) Proc. Natl. Acad. Sci. USA 84, 1234–1238.
[32] Kikuchi, Y., Yoda, K., Yamasaki, M. and Tamura, G. (1981) Nucleic Acids Res. 9, 5671–5678.