Volume 306, number 1, 54–58 FEBS 11236 July 1992

Polymerase chain reactions using *Saccharomyces, Drosophila* and human DNA predict a large family of protein serine/threonine phosphatases

Mao Xiang Chen, Yu Hua Chen and Patricia T.W. Cohen

Medical Research Council Protein Phosphorylation Unit, Department of Biochemistry, The University, Dundee, DD1 4HN, Scotland, UK

Received 27 April 1992

Using the polymerase chain reaction (PCR) to examine the protein serine/threonine phosphatase (PP) family which includes PP1, PP2A and PP2B, we have identified two, seven, and four novel protein phosphatase genes in *Saccharomyces cerevisiae, Drosophila melanogaster* and *Homo sapiens*, respectively. Consequently, the genes in the PP1/PP2A/PP2B family now number 11, 15 and 12 in these species respectively, and the data predicts still more unidentified phosphatases in higher eukaryotes. The PCR analyses also point to the presence in *Drosophila* and mammals of three or more different genes encoding PP2B, the enzyme recently identified as the target of certain immunosuppressant drugs.

Protein phosphatase; Gene family; Polymerase chain reaction; *Saccharomyces*; *Drosophila*; human

1. INTRODUCTION

Protein phosphorylation plays a crucial role in the regulation of a wide array of proteins that are involved in many cellular processes. The phosphatases which dephosphorylate serine and threonine residues in proteins have been classified into four major types, PP1, PP2A, PP2B and PP2C, based on their substrate specificities and their sensitivity to certain activators and inhibitors [1]. DNA cloning has elucidated the amino acid sequences of the catalytic subunits of PP1, PP2A and PP2B, and shown that they share 40–50% identity in the catalytic domain, while PP2C has a completely different structure [2]. Molecular cloning in lower eukaryotes has demonstrated that the amino acid sequences of PP1 and PP2A have been more highly conserved during evolution than any other enzyme [3]. In addition, from screening of DNA libraries it is clear that, not only do these phosphatases exist as several isoforms sharing approximately 90% amino acid sequence identity, but also that other phosphatases exist which cannot readily be classified as PP1, PP2A, or PP2B by sequence similarity [2].

A member of the same protein phosphatase family (λORF221) is encoded in the genome of bacteriophage λ. This enzyme was initially identified as a protein phosphatase by a data base search [4] and subsequently by enzyme assay [5]. Comparison of its structure with those of mammalian protein phosphatases, identified regions that proved to be conserved across all known protein phosphatases in this family, indicative of a critical role in catalytic activity [2,5]. In order to ascertain how many novel protein phosphatases exist in eukaryotes and to begin to study their functions, we have carried out polymerase chain reactions within the most highly conserved region using the same two oligonucleotides with DNA from *Saccharomyces cerevisiae, Drosophila melanogaster* and *Homo sapiens*.

Correspondence address: M. Chen, MRC Protein Phosphorylation Unit, Department of Biochemistry, The University, Dundee, DD1 4HN, Scotland, UK. Fax: (44) (382) 23778.

2. METHODS

Oligonucleotide primers 1 and 2 (Fig. 1) were designed to recognise two of the most conserved regions of the PP1/PP2A/PP2B family [5,6]. The primers were phosphorylated at the 5'-end according to Maniatis et al. (1982). PCR was carried out using AmpliTaq DNA polymerase (Perkin Elmer Cetus, Bucks, UK) in 20 μl reactions. The procedure was as described by Perkin Elmer Cetus, with the modifications detailed in [6]. Samples were treated with UV light of 300 nm wavelength, for 15 min after addition of all components, except the template DNA. PCR reactions were carried out as follows: 94°C for 3 min; 49°C (50°C for human DNA) for 1 min, 72°C for 1 min, 94°C for 1 min, 35 cycles; followed by 49°C (or 50°C) for 5 min and then 72°C for 10 min.

The PCR products were subjected to agarose gel electrophoresis, excised from the gel and subcloned into the *Eco*RV site of Bluescript (Stratagene, La Jolla, CA) [6]. Sequencing of both strands was carried out on double stranded DNA with Sequenase (United States Biochemical Corp. Cleveland, OH) using T3 and T7 primers.

3. RESULTS

Electrophoretic analysis of the PCR fragments produced with chromosomal DNA from *S. cerevisiae*, *D. melanogaster*, and *Homo sapiens* is shown in Fig. 2. With each species, a prominent band of the expected size (120 pb) was visible. In addition, a second, much weaker, band of higher molecular weight (about 250 bp)

Published by Elsevier Science Publishers B.V.

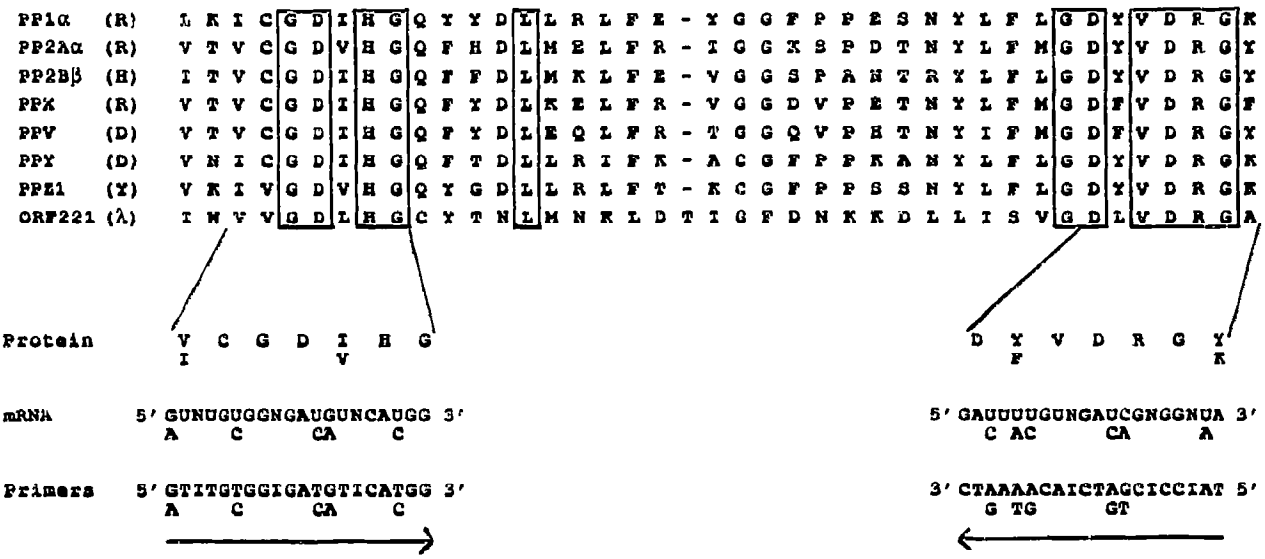

Fig. 1. The conserved region of protein serine/threonine phosphatase used for designing primers for the polymerase chain reaction. The section shown is amino acids 59–98 of rabbit PP1α [2,7] and the corresponding region of PP2Aα [8], PP2Bβ [9], PPX [2], PPV [2], PPY [10], PPZ1 [2], ORF221 [4,5]. Invariant amino acids are boxed. R, rabbit; H, human; D, *Drosophila melanogaster*; Y, *Saccharomyces cerevisiae*; λ, bacteriophage λ. The coding sequences for the most conserved sections and the nucleotide sequences of the primers are shown below. A, adenine; C, cytosine; G, guanine; T, thymine; I, inosine; N=A, C, G and T.

was also amplified. This larger DNA fragment was later shown by cloning and sequencing to have arisen from spurious PCR products.

Sequencing of the 120 bp PCR fragments cloned in DH5α cells (64 from yeast, 84 from *Drosophila* and 53 from man) identified seven protein phosphatase sequences from *S. cerevisiae* and two from *D. melanogaster* that were already known (see discussion). However, the other cloned 120 bp PCR fragments encoded two protein phosphatase like sequences from *S. cerevisiae*, seven from *D. melanogaster* and four from man. The nucleotide and deduced amino acid sequences of these novel putative protein phosphatase fragments are shown in Fig. 3. and their amino acid sequences are compared with those of known phosphatases in Fig. 4.

Table I shows the level of amino acid identity of the PCR fragments to the corresponding region of known phosphatases in the same species, with the exception that human PP2Bβ was used for comparison with the *Drosophila* sequences, because the sequence of PP2B in *Drosophila* has not yet been published. From Table I it is clear that the two yeast sequences (Y25 and Y58) do not encode isoforms of known phosphatases. Y25 shows only 46–58% identity to known yeast phosphatases, while Y58 is even more distantly related to the known phosphatases, with amino acid identity of 40–52%. Y58 also contains an extra lysine residue at position 11.

Of the seven novel *Drosophila* PCR sequences, D3 is not closely related to any other phosphatase, sharing only 44–56% identity with known sequences. It also contains an extra proline residue at position 20. In contrast, D19 is nearly identical (96%) to that of the male specific gene PPY [10], with only a single amino acid difference. However as there are 12 changes in the

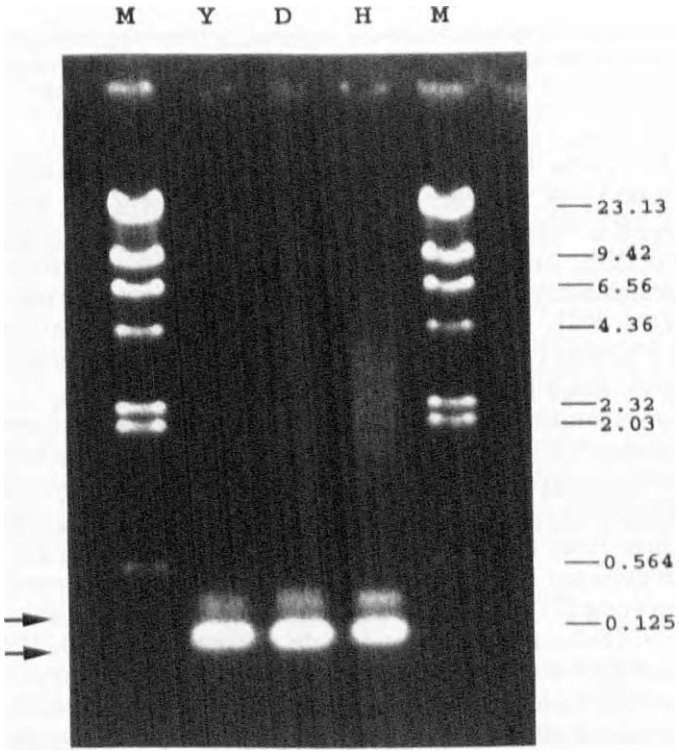

Fig. 2. Gel electrophoresis of the polymerase chain reaction amplification products from genomic DNA using the protein serine/threonine phosphatase primers shown in Fig. 1. (Y) *Saccharomyces cerevisiae* genomic DNA. (D) *Drosophila melanogaster* genomic DNA (H) Human genomic DNA. (M) λ *Hin*dIII marker DNA. The sizes of the DNA fragments are shown in kilobase pairs. Arrows indicate the PCR amplification products.

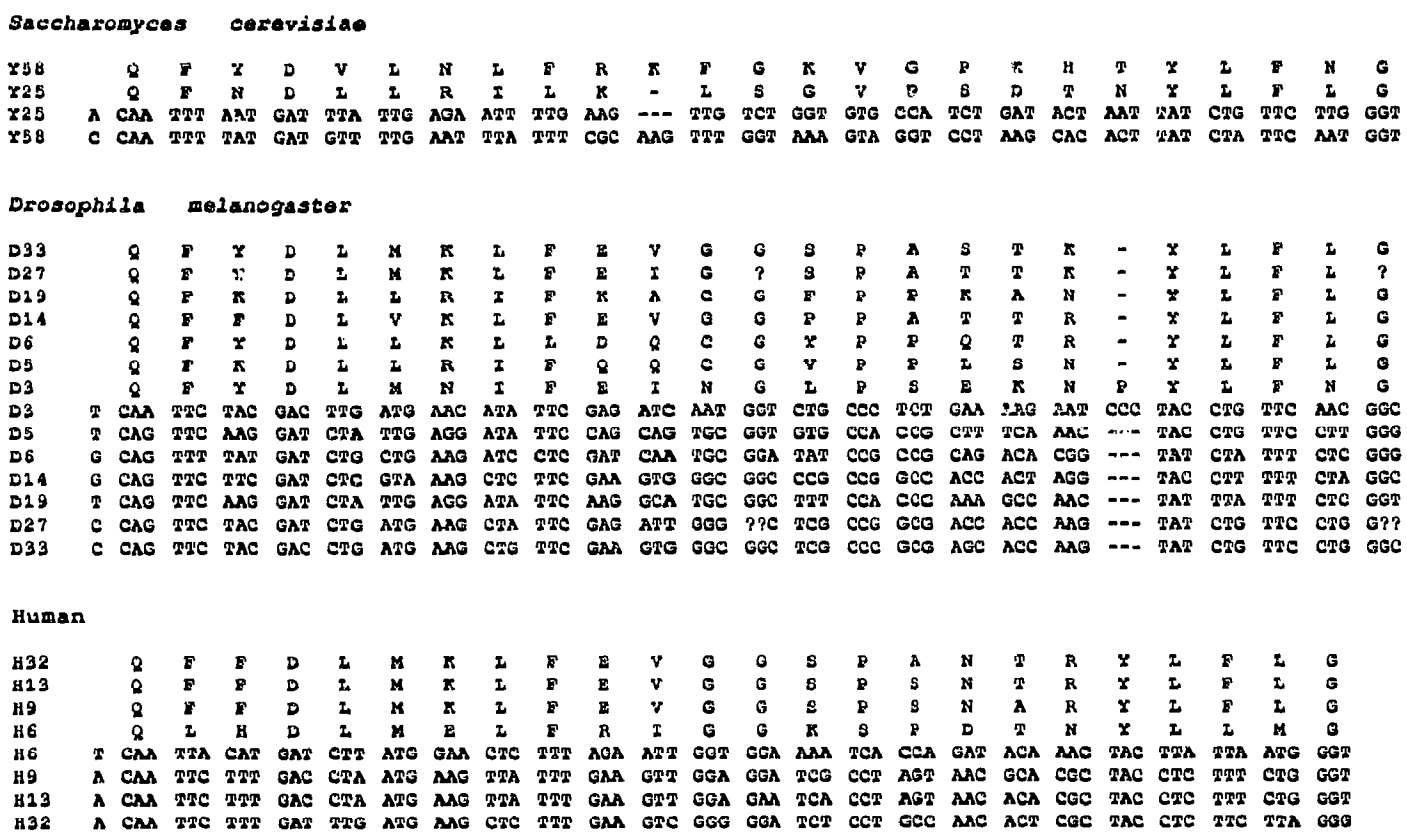

Fig. 3. Novel protein phosphatase genes and their encoded amino acid sequence identified by PCR. Y, yeast; D, *Drosophila*; H, human. The numbers refer to the order of the PCR fragments isolated. The question marks indicate that the base or amino acid could not be unambiguously assigned, even though the nucleotide sequence was read in both directions.

DNA, the sequence cannot have arisen from the PPY gene by PCR error. Consequently, it is likely that D19 is an isoform of PPY. D5 and D6 are also most similar to PPY, but since the identity is only 75% and 63%, respectively, they may be functionally distinct enzymes. D14, D27, and D33 are 88%, 83%, and 87% identical to human PP2Bβ [9,26] respectively and may be *Drosophila* homologues of human PP2B.

Of the four novel human sequences, H6 differs from human PP2Aα [23,24] and PP2Aβ [23,25] at only two amino acids (92% identity). The nucleotide sequence of H6 shows nine differences from that of PP2Aα, but only three from the same region of PP2Aβ. Since one pseudogene for PP2Aβ has been identified by chromosomal blotting [27], it is possible that H6 is this pseudogene. H9, H13, and H32 may encode isoforms of PP2B. H9 and H13 are very similar to that of human PP2Bβ (83 and 96% identical, respectively), while H32 has exactly the same amino acid sequence. However the nucleotide sequence of H32 shows 17 changes from human PP2Bβ, demonstrating that it must be part of a distinct gene. H9 and H13 show 17 and 15 nucleotide differences from human PP2Bβ, respectively. The amino acid sequence of H13 is identical to that of a PP2β isoform cloned from murine testis [28].

4. DISCUSSION

We have identified two, seven, and four novel protein serine/threonine phosphatase genes in yeast, *Drosophila*, and man, respectively. The use of genomic DNA rather than cDNA for PCR means that in theory all protein phosphatases with the conserved sequence representative of the family could be identified from a single reaction. Although the possibility that some of these are pseudogenes rather than expressed phosphatase genes cannot be ruled out, particularly in man, it is likely that most, if not all, of those in *Drosophila* and yeast are functional genes. In fact, we have now shown that the yeast PCR fragments are part of long open reading frames (M.X. Chen, Y.H. Chen and P.T.W. Cohen, unpublished data). The *Drosophila* PCR fragments, including D3, D5, D6, D19, D27 and D33 have been used to probe head and eye imaginal disc cDNA libararies under high stringency, and, except for D6, positive clones were identified in all cases. These probes have been hybridized under identical conditions to *Drosophila* genomic DNA cleaved with several restriction enzymes. In all cases only a single band was seen, except for the putative PP2Bβ homologues D27 and D33, which both hybridised to the same two bands

(data not shown). This clearly indicates that the cDNA clones detected at the same stringencies are almost certain to contain the PCR probe sequences. Indeed sequencing of one of the cDNAs has shown this is correct (L. Brown, M. Chen and P.T.W. Cohen, unpublished data). In addition, D14 sequence has been found in a cDNA encoding a *Drosophila* PP2B (Drs. D. Guerini and O.B. Klee, personal communication).

A total of 13 new phosphatase genes have been identified in this study by PCR, bringing the number of protein Ser/Thr phosphatase genes in the PP1/PP2A/PP2B family to 11, 15 and 12 in *S. cerevisiae*, *D. melanogaster* and *H. sapiens*, respectively (Fig. 4). Seven of the nine known yeast genes including DIS2S1 (PP1) [11], PPH21 (PP2A), PPH22 (PP2A) [13], SIT4 [14], CMP1 (PP2B) [15,16], CMP2 (PP2Bw) [12,15–17] and PPZ2 [12] were identified during our analysis, suggesting that the majority of yeast phosphatase genes in this family may have been found. In contrast, although seven new sequences were amplified in *Drosophila*, only two of the seven genes that were known already were identified, suggesting that much larger numbers of phosphatase genes may exist in *Drosophila*. The human species is likely to have homologues of all of the *Drosophila* phosphatases. The fact that less novel genes were found in man, than in yeast and *Drosophila* partly reflects the smaller number of clones analysed (see results). It is also clear that phosphatase genes with introns in this region would have escaped cloning. Recent sequencing has shown that the *Drosophila* PPV gene does indeed have a small intron in this section (D.J. Mann and P.T.W. Cohen, unpublished data). Although it may be possible to obtain sequences with small introns by cloning DNA fragments of larger size from the PCR products, it is unlikely that sequences with large introns would be amplified effectively by PCR. In addition, it was evident from our experiments with human and *Drosophila* DNA, that many phosphatase sequences amplified by PCR were not cloned because 'rearrangement' of clones (which produced clones with an unexpected structure) appeared to be common with DNA from the higher organisms (see data in ref. 6). Such rearrangement was low with yeast DNA. An increase in the number of phosphatase genes through evolution is not unexpected in order to deal with the increased complexity of cellular regulation by reversible protein phosphorylation in higher eukaryotes.

The calcium-calmodulin regulated protein phosphatase PP2B is the target of the immunosuppressant drugs, cyclosporin and FK506, used in transplantatition surgery [29]. Our experiments indicate that there are at least three PP2B isozymes in *Drosophila* and at least four in man (assuming none of the PCR sequences are from pseudogenes). Investigation of whether mRNA corresponding to these PP2B sequences is present in human T cells may be valuable for studying the action of immunosuppressant drugs.

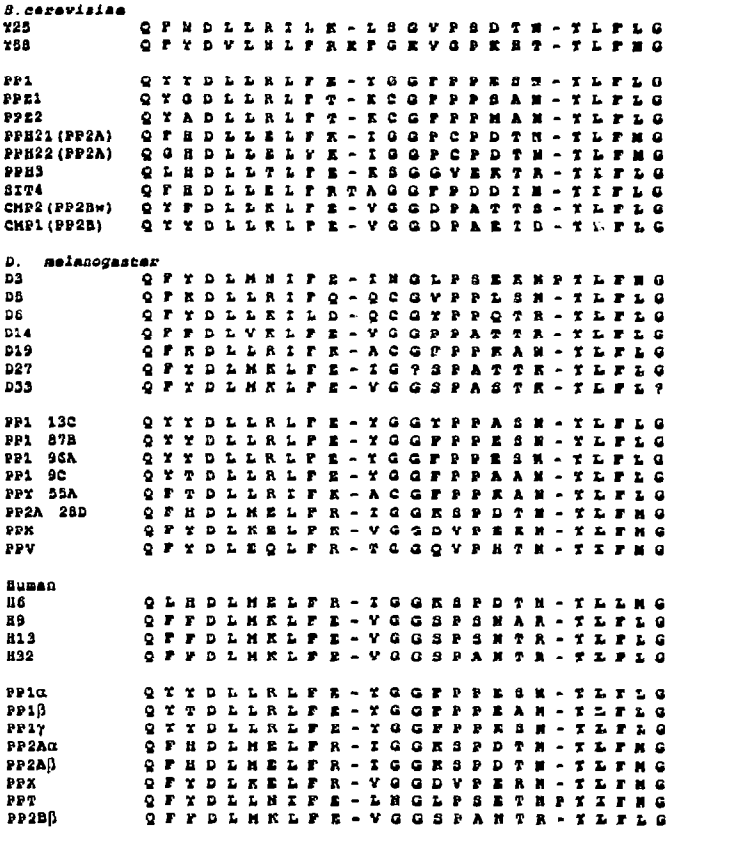

Fig. 4. Novel protein phosphatase amino acid sequences. *S. cerevisiae* PCR fragments Y25 and Y58 have not been identified previously. Seven of the known phosphatases, including DIS2S1 (PP1) [11], PPZ2 [12], PPH21 (PP2A), PPH22 (PP2A) [13], SIT4 [14], CMP1 (PP2B) [15,16], and CMP2 (PP2B$_w$) [12,15–17] were identified by PCR. The two genes that were known already and not identified in this experiment were PPZ1 [12] and PPH3 [18]. *D. melanogaster* PCR fragments D3, D5, D6, D14, D19, D27, and D33 were not previously identified. Of the known phosphatases, only two (PP1 96A [19] and PPY 55A [10]) were identified by PCR. The others including PP1 87B [20], PP1 13C (V. Dombrádi, D.J. Mann and P.T.W. Cohen, unpublished data), PP1 9C [19], PP2A 28D [21], PPX (N.D. Brewis and P.T.W. Cohen, unpublished data), and PPV [2] were not identified in this experiment. The number and letter following the phosphatase name signify its chromosomal location. (c) Human PCR fragments H6, H9, H13, and H32 have not been identified previously. Known phosphatases, including PP1α [22], PP1β, PP1γ (H.M. Barker and P.T.W. Cohen, unpublished data), PP2Aα [23,24], PP2Aβ [23,25], PPX (N.D. Brewis and P.T.W. Cohen, unpublished data), PPT (A.E. McPartlin and P.T.W. Cohen, unpublished data), and PP2Bβ [9,26], were not identified in this experiment.

A PCR experiment carried out with *E. coli* genomic DNA did not yield any clear PCR products. This is consistent with the finding that protein serine/threonine phosphatase activity, as measured with phosphorylated casein or phosphorylase as substrates, cannot be found in this bacterium, provided that there is no infection by phage λ or other phosphatase encoding phage [4,5]. It is probable that *E. coli* do not possess this class of protein phosphatase genes.

Table 1

Percent amino acid identity of the PCR-amplified fragment of novel protein phosphatases to the same region of known phosphatases.

S. cerevisiae

	DIS2 (PP1)	PPZ1	PPZ2	PPH21 (PP2A)	PPH22 (PP2A)	PPH3	SIT4	CMP2 (PP2Bw)	CMP1 (PP2B)
Y25	54	54	54	50	58	50	52	50	46
Y58	52	40	40	48	44	40	44	40	44

D. melanogaster

	PP1 87B	PP1 96A	PP1 9C	PP1 13C	PPY 55A	PP2A 28D	PPX 19C	PPV 6A	PP2Bβ (human)
D3	56	56	48	52	52	52	56	44	48
D5	67	67	63	67	75	50	50	46	50
D6	54	54	50	58	63	46	46	46	58
D14	58	58	58	58	50	54	54	50	**88**
D19	67	67	75	63	**96**	50	54	46	50
D27	61	61	57	61	48	57	52	52	**83**
D33	61	61	57	61	48	61	57	52	**87**

Human

	PP1α	PP1β	PP1γ	PP2Aα	PP2Aβ	PPX	PPT	PP2Bβ
H6	50	50	50	**92**	**92**	63	38	50
H9	58	63	58	54	54	54	46	**83**
H13	58	58	58	58	58	54	54	**96**
H32	58	58	58	58	58	54	50	**100**

Acknowledgement: This work was supported by a Human Frontier Science Programme grant and by the Medical research Council, London.

REFERENCES

[1] Cohen, P. (1989) Annu. Rev. Biochem. 58, 453–508.
[2] Cohen, P.T.W., Brewis, N.D., Hughes, V. and Mann, D.J. (1990) FEBS Lett. 268, 355–359.
[3] Cohen, P.T.W., in: Genetics and Human Nutrition (P.J. Randle, Ed.), J. Libbey, 1990, pp.27–40.
[4] Cohen, P.T.W., Collins, J.F., Coulson, A.F.W., Berndt, N. and da Cruz e Silva, O.B. (1988) Gene 69, 131–134.
[5] Cohen, P.T.W. and Cohen, P. (1989) Biochem. J. 260, 931–934.
[6] Chen, M.X. and Cohen, P.T.W., in: Protein Phosphorylation: a Practical Approach, D.G. Hardie, Ed., IRL Press, Oxford, 1992, in press.
[7] Berndt, N., Cambell, 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.
[8] da Cruz e Silva, O.B., Alemany, S., Cambell, D.G. and Cohen, P.T.W. (1987) FEBS Lett. 221, 415–422.
[9] Guerini, D. and Klee, C.B. (1989) Proc. Natl. Acad. Sci. USA 86, 9183–9187.
[10] Dombrádi, V., Axton, J.M., Glover, D.M. and Cohen, P.T.W. (1989) FEBS Lett. 247, 391–395.
[11] Ohkura, H., Kinoshita, N., Miyatani, S., Toda, T. and Yanagida, M. (1989) Cell 57, 997–1007.
[12] da Cruz e Silva, E.F., Hughes, V., McDonald, P., Stark, M.J.R. and Cohen, P.T.W. (1991) Biochim. Biophys. Acta 1089, 269–272.
[13] Sneddon, A.A., Cohen, P.T.W. and Stark, M.J.R. (1990) EMBO J. 9, 4339–4346.
[14] Arndt, K.T., Styles, C.A. and Fink, G.R. (1989) Cell 56, 527–537.
[15] Liu, Y., Ishii, S., Tokai, M., Tsutsumi, H., Ohki, O., Akada, R., Tanaka, K., Tsuchiya, E., Fukui, S. and Miyakawa, T. (1991) Mol. Gen. Gent 227, 52–59.
[16] Cyert, M.S., Kunisawa, R., Kaim, D. and Thorner, J. (1991) Proc. Natl. Acad. Sci. USA 88, 7376–7380.
[17] da Cruz e Silva, E.F. and Cohen, P.T.W. (1989) Biochim. Biophys. Acta 1009, 293–296.
[18] Ronne, H., Carlberg, M. Hu, G.-Z. and Nehlin, J.O. (1991) Mol. Cell. Biol. 11, 4876–4884.
[19] Dombrádi, V., Axton, J.M., Brewis, N.D., da Cruz e Silva, E.F., Alphey, L. and Cohen, P.T.W. (1990) Eur. J. Biochem. 194, 739–745.
[20] Dombrádi, V., Axton, J.M., Glover, D.M. and Cohen, P.T.W. (1989) Eur. J. Biochem. 183, 603–610.
[21] Orgad, S., Brewis, N.D., Alphey, L., Axton, J.M., Dudai, Y. and Cohen, P.T.W. (1990) FEBS Lett. 275, 44–48.
[22] Barker, H.M., Jones, T.A., da Cruz e Silva, E.F., Spurr, N.K., Sheer, D. and Cohen, P.T.W. (1990) Genomics 7, 159–166.
[23] Arino, J., Woon, C.W., Brautigan, D.L., Miller, T.B. and Johnson, G.L. (1988) Proc. Natl. Acad. Sci. USA 85, 4252–4256.
[24] Stone, S.R., Wernet, W., Mayer, R., Maurer, F., Hofsteenge, J. and Hemmings, B.A. (1988) Nucleic Acids Res. 16, 11365.
[25] Hemmings, B.A., Mayer, R., Wernet, W., Maurer, F., Hofsteenge, J. and Stone, S.R. (1988) Nucleic Acids Res. 16, 11366.
[26] McPartlin, A.E., Barker, H.M. and Cohen, P.T.W. (1991) Biochim. Biophys. Acta 1088, 308–310.
[27] Khew-Goodall, Y., Mayer, R.E., Maurer, F., Stone, S.R. and Hemmings, B.A. (1991) Biochemistry 30, 89–97.
[28] Muramatsu, T., Giri, P.R., Higuchi, S. and Kincaid, R.I. (1992) Proc. Natl. Acad. Sci. USA 89, 529–533.
[29] Liu, J., Farmer, J.D., Lane, W.S., Friedman, J., Weissman, I. and Schreiber, S.L. (1991) Cell 68, 1–20.