Volume 244, number 1, 47–50 FEB 06777 February 1989

Molecular cloning and sequencing of human cDNA for phosphoribosyl pyrophosphate synthetase subunit II

Taizo Iizasa, Masanori Taira, Hideaki Shimada, Sumio Ishijima and Masamiti Tatibana

Department of Biochemistry, Chiba University School of Medicine, Inohana, Chiba 280, Japan

Received 6 December 1988

cDNA clones for human phosphoribosyl pyrophosphate synthetase subunit II (PRS II) were isolated. The five overlapping clones contained 2457 base pairs (bp) covering a 954-bp complete coding region for 318 amino acid residues. Homologies between human and rat PRS II were 99% of the amino acids and 88% of the nucleotides in the coding region. This amino acid homology seems to be the highest so far reported for enzymes involved in nucleotide metabolism and glycolysis. The highly conserved structure may be required for unique catalysis and rigid regulation of this enzyme.

Ribose-phosphate pyrophosphokinase; Enzyme subunit; cDNA cloning; Nucleotide sequence; Amino acid sequence; (Human)

1. INTRODUCTION

5-Phosphoribosyl 1-pyrophosphate (*PP*Rib*P*), an essential substrate and a critical regulator in the pathways of purine, pyrimidine [1] and pyridine nucleotide production, is synthesized from MgATP and ribose 5-phosphate, by the catalysis of *PP*Rib*P* synthetase (EC 2.7.6.1). This enzyme exists in many active forms of various molecular masses and has many effectors: Mg^{2+} and inorganic phosphate (P_i) as activators; and ADP, 2,3-bisphosphoglycerate or GDP as competitive or non-competitive inhibitors [2]. Recently, we showed that rat *PP*Rib*P* synthetase had two distinct types of subunits, referred to as PRS I and PRS II, determined by cDNA cloning [3] and that human PRS I and PRS II genes (PRPS1 and PRPS2, respectively) are located on different regions of the same X chromosome [4]. We now report the cDNA sequence for human PRS II. The high amino acid homology of the human and rat PRS II was compared with those of other human and

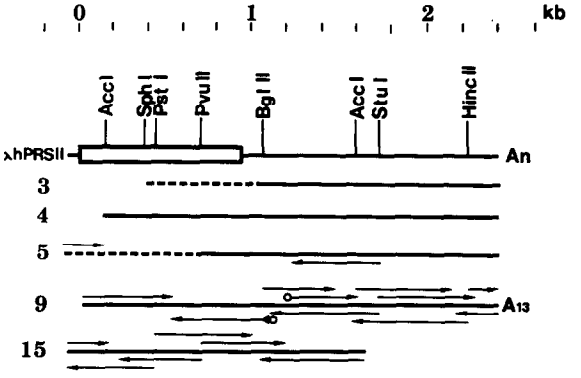

Fig.1. Restriction map and sequence strategies for human *PP*Rib*P* synthetase subunit II cDNA. The box represents the coding region, the bar denoting the non-coding region. Numbers on the left indicate clone numbers. Arrows show the direction and extent of nucleotide sequencing. Sites of *Sau*3AI (○) were used for the sequencing. The broken lines indicate regions with different restriction maps. These regions may have been unrelated fragments, accidentally ligated during construction of the cDNA library.

Correspondence address: T. Iizasa, Department of Biochemistry, Chiba University School of Medicine, Inohana, Chiba 280, Japan

Abbreviations: *PP*Rib*P*, 5-phosphoribosyl 1-pyrophosphate; PRS I (II), phosphoribosyl pyrophosphate synthetase subunit I (II); kb, kilbase(s); bp, base pairs; dT, deoxythymidylic acid; poly(A), polyadenylic acid

The nucleotide sequence presented here has been submitted to the EMBL/GenBank database under the accession no. Y00971

Published by Elsevier Science Publishers B.V. (Biomedical Division)
00145793/89/$3.50

murine enzymes. Additionally, knowledge of the human cDNA paves the way for elucidation of the molecular basis of human X-linked disease: *PP*Rib*P* synthetase superactivity [5–8].

2. MATERIALS AND METHODS

We reported the cloning of human PRS II cDNA [4]. Briefly, a cDNA library was constructed from adult human testis poly(A)$^+$ RNA in λgt10 phage vector, as described [3]. Filters were hybridized with a nick-translated fragment of rat PRS II cDNA (*Bst*EII/*Hin*cII, 1.50 kb) [3,4] as a probe, and washed in 0.3 M NaCl, 0.03 M sodium citrate and 0.1% SDS at 68°C. Nucleotide sequencing of cDNA obtained was performed as described [3].

3. RESULTS AND DISCUSSION

Fourteen clones were selected from 1350000

```
                                  -60 CGCTGTTGCCTCCGCCACCTCCTCCGCCGCCGCGCGCCCCTCGGAGTTCCGCGCCCCACC   -1
    1
 M  P  N  I  V  L  F  S  G  S  S  H  Q  D  L  S  Q  R  V  A  D  R  L  G  L  E  L  G  K  V    29
ATGCCCAACATCGTGCTGTTCAGCGGCAGCTCGCATCAGGACCTATCCCAGCGCGTGGCCGACCGCCTGGGCCTGGAGCTGGGCAAGGTG    90

 V  T  K  K  F  S  N  Q  E  T  S  V  E  I  G  E  S  V  R  G  E  D  V  Y  I  I  Q  S  G  C    59
GTCACGAAGAAGTTCAGCAACCAGGAGACCAGCGTGGAGATTGGTGAAAGCGTGAGAGGGGAAGATGTCTACATCATCCAGAGCGGCTGC   180

 G  E  I  N  D  N  L  M  E  L  L  I  M  I  N  A  C  K  I  A  S  S  S  R  V  T  A  V  I  P    89
GGGGAAATTAACGACAACCTGATGGAACTCCTCATCATGATCAATGCCTGCAAGATTGCGTCATCATCCAGAGTAACTGCCGTGATCCCG   270

 C  F  P  Y  A  R  Q  D  K  K  D  K  S  R  A  P  I  S  A  K  L  V  A  N  M  L  S  V  A  G   119
TGTTTCCCATACGCCCGACAAGATAAAAAGGACAAGAGTCGTGCCCCAATTTCTGCAAAACTTGTGGCCAATATGCTGTCGGTGGCTGGG   360

 A  D  H  I  I  T  M  D  L  H  A  S  Q  I  Q  G  F  F  D  I  P  V  D  N  L  Y  A  E  P  A   149
GCGGATCACATCATCACCATGGACCTGCATGCTTCTCAGATACAGGGATTCTTTGATATTCCTGTGGATAATTTGTATGCGGAGCCCGCA   450

 V  L  Q  W  I  R  E  N  I [A] E  W [K] N  C  I  I  V  S  P  D  A  G  G  A  K  R  V  T  S   179
GTCCTGCAGTGGATTCGGGAAAACATTGCCGAGTGGAAGAACTGTATCATTGTTTCACCTGACGCAGGGGGAGCCAAAAGGGTTACATCA   540

 I  A  D  R  L  N  V  E  F  A  L  I  H  K  E  R  K  K  A  N  E  V  D  R  M  V  L  V  G  D   209
ATTGCAGACAGGTTGAATGTGGAATTTGCTTTGATCCACAAAGAGAGGAAGAAGGCGAATGAAGTGGACCGGATGGTCCTGGTGGGCGAC   630

 V  K  D  R  V  A  I  L  V  D  D  M  A  D  T  C  G  T  I  C  H  A  A  D  K  L  L  S  A  G   239
GTGAAGGACCGTGTGGCCATCCTCGTGGATGACATGGCTGACACTTGCGGCACCATCTGCCATGCTGCGGACAAGCTGCTGTCAGCTGGA   720

 A  T  K  V  Y  A  I  L  T  H  G  I  F  S  G  P  A  I  S  R  I  N  N  A  A  F  E  A  V  V   269
GCCACCAAAGTGTATGCTATCCTTACCCATGGGATCTTCTCTGGACCAGCTATTTCCAGAATAAATAATGCCGCCTTTGAGGCTGTTGTC   810

 V  T  N  T  I  P  Q  E  D  K  M  K  H  C [T] K  I  Q  V  I  D  I  S  M  I  L  A  E  A  I   299
GTCACAAACACAATTCCGCAAGAGGACAAAATGAAACACTGCACCAAGATTCAGGTCATTGACATTTCCATGATCTTGGCCGAAGCAATC   900

 R  R  T  H  N  G  E  S  V  S  Y  L  F  S  H  V  P  L  *                                    317
CGAAGGACACACAATGGGGAATCCGTGTCCTACCTGTTCAGCCATGTCCCGCTATAAATCCAGAATGGGAAGTGTCCAGCAAGCCTACTC   990
TGACTTCTGACTTGTTTTTGTTTTCTGGATTTTTAGCTGTAGGTATTCAGCAATGATAGGTTAATCACTGGCAAAAGCATCAGATCTTTG  1080
TATATGCTAAGATTTATTGTTTCCCCTTCTAAAGCTCAAGATCATTTCTTTCCAGTTTTTGGGGAAATGGTGGTGGTTATTTGGTCTTTA  1170
AGTGAACTGTCTTAAATGAGAAACGTTTTTGTCATTTTGACTTTTAACAGGTACAGGTGATCTCTTCCTTTGTTCTTTCAGTACTTTGAG  1260
GCGACAACTTTCAAGTATATAATTTCATTGTGGAAGTCATAGTTTATATATTTCGAGGTTGCCAAAGGTGACTTCACATTAAAGCCTTCT  1350
GTGTAAATATATACTGATAATGCCTATGGACATTTGGGTAAAACCCTGTATAGAATTAATTATCCTTTTACTTTGGAGTGAACCTTGGAA  1440
AATTTATAATTATAATACCATGGATTTTGAATTTTCCTTTTTTTTTTTTTTTTGGATAACTCAGTTTCAGATAAACCATCTTGGTTAC  1530
TGTGCTTAATTTGGACCAAATTTTATTTAGCTTAATATGGACACTGACACATTTTGGGGGGTATACATTAGACATATCAGAGCAGTGTAT  1620
TTCTGGATCATTTTTTAAATGACCTCTTCTAAAACATAACTGTCACTTACCTGAAATGCTGCATCCTAAAATTCCAAAATTATATTGAGC  1710
AATCGCCAAGGCCTAAAGCCAACTGACTTAAAGGTAATCATTTCAGCTAAGATTAAATTTAAAGCCTAAGAATGTATAGAGCTAGTTTTA  1800
AAATAATGATCTCAGATTTTTAAAAAGGATATAGGAACCTGCATTGTCATTCTCTGAATTAAGAACTGATGGTTTCTATCATTATTTAGC  1890
CCCACCTTTGTATTTTAAAATCCTTCAGAATACATTTATGAACCAATGCGACTGGACTTAGCCACACACAATGGAAATTCAGACCTTGAC  1980
TATTTGGTGTTTCCAGTTCACAAAGGTGATGAAGACTGTCTTGGGAGCAGCTTAATCCCAAAATTTGTACATTTCTTGCTGCTCCTGGCG  2070
TGGAAACTTAAGTGAGACCACCAAATACATTGGTCCTGTCCAATTCTACTGAATGGGGGTGGACCTGGCATTTATCTGGCCAAAAACAGG  2160
AGCCAGAGAAATATGAATATACCAAAGTTGTTTGTTTAGCCTCCAACTTAAATTACATTAGTCAACTTATAGATACTCATATGATCACTT  2250
TTCTTTTTAGATACTACATCAACTAGATTCAGGAGTATATCATTTGCAGTGCTTGTATTGGTTTAAAATGTAAGATTTTAAGATCCTCTA  2340
ACACTGTACTAAAACATTTCAATAAAATCATTCTGACTGCGTTCAAAAAAAAAAAAA                                    2397
```

Fig.2. Nucleotide and deduced amino acid sequences of human PRS II cDNA. Numbers in the right margin, nucleotides from the predicted initiation codon and amino acids beginning with 1 for the first amino acid of the mature protein. Amino acids are shown in single-letter codes. (★) Stop codon. Boxed residues, positions of the different amino acids from those of rat PRS II. Bar, putative *PP*Rib*P*-binding site [14]. The polyadenylation signal is underlined. λhPRSII-9 and λhPRSII-15 covered from 30 to 2397 and from −60 to 1671, respectively. The number of dT stretch at nos 1478–1495 is 18 in the two clones λhPRSII-15 and λhPRSII-5, and 17 in the one clone λhPRSII-9. Since the PRS II gene was located on chromosome X and cDNA was made from testis containing a single X, the sequence of 17 dTs may be an artifact, not a polymorphism.

plaques of a human testis cDNA library. Five of the clones were strongly hybridized with the rat PRS II probe: the inserts were 2.5, 2.4, 2.3, 2.1 and 1.7 kb in length. A restriction map and the sequencing strategies for the selected five clones are shown in fig.1. Since the size of the human PRS II mRNA was estimated to be 2.7 kb, as determined by Northern blot analysis of human testis mRNA, using rat [9] and human (not shown) PRS II cDNA probes, the 2.4 kb (no.9) and 1.7 kb (no.15) cDNAs containing 2457 base pairs (bp) covered nearly the full length.

The nucleotide and deduced amino acid sequences of the overlapping cDNA are shown in fig.2. The ATG triplet at position 1 was determined as the initiation codon, based on a sequence similar to those of rat PRS I and II cDNAs and on Kozak's rule [10]. The calculated molecular mass was 34638 Da, without the initiation methionine, a value in accord with that of the reported human enzyme subunit (34.5 kDa) [11].

Homologies between human PRS II and rat PRS II were 88% for 954 bp coding sequence and 99% for 318 amino acid residues. Of the deduced amino acid sequence, three substitutions were observed between human and rat PRS II with Ala-159, Lys-162 and Thr-284 in human PRS II, instead of Thr, Arg and Ser in rat PRS II, respectively, as indicated by boxes in fig.2. These are conservative substitutions based on Dayhoff's mutation data [12]. The putative *PP*Rib*P*-binding site (fig.2) was identical to those of rat PRS I and II.

As shown in table 1, the amino acid homology was the highest in enzymes known to be involved in nucleotide metabolism and glycolysis. Direct comparisons between a pair of known nucleotide sequences in coding regions facilitate evaluation of the fraction of numbers of sequence differences from the total synonymous sites (synonymous difference, Ks), i.e. sites at which nucleotide substitutions do not cause amino acid substitutions [13]. Notable findings are as follows: (i) the Ks value of *PP*Rib*P* synthetase (0.495) is close to those of 10 other groups of human and murine enzymes (0.440 ± 0.068) without hypoxanthine-guanine phosphoribosyltransferase (0.255); (ii) there seems

Table 1

Comparison of enzymes involved in nucleotide metabolism and glycolysis with amino acid homology and synonymous difference between human and murine

Human enzymes	Chromosomal location[a]	Species	Amino acid homology[b] (%)	No. of total residues[b]	Ks[e]
Pyruvate kinase (liver type)	1	rat	92	543	0.401
Orotidine-5′-monophosphate decarboxylase	3	mouse[c]	90	262[c]	0.516
Aldolase B	9	rat	95	364	0.396
Lactate dehydrogenase A	11	rat	94	332	0.511
Glyceraldehyde-3-phosphate dehydrogenase	12	rat	94	333	0.324
Adenine phosphoribosyltransferase	16	mouse	82	180	0.432
Aldolase A	16	rat	97	364	0.403
Thymidylate synthetase	18	mouse	88	307	0.413
Adenosine deaminase	20	mouse	83	352	0.548
Hypoxanthine-guanine phosphoribosyltransferase	X	mouse	96	218	0.255
Glucose-6-phosphate dehydrogenase	X	rat	94	515	0.459
Phosphoribosyl pyrophosphate synthetase subunit II	X[d]	rat	99	318	0.495

[a] From [(1985) Cytogenet. Cell Genet. 40, 1–823]
[b] From [(1988) Nucleic Acids Res. suppl. 16, r403–r408]
[c] Partial sequence
[d] From [4]
[e] Ks value, synonymous difference (from [13])

to be no distinction between the Ks values of enzymes, the genes of which are located on autosomes and X-chromosomes; and (iii) there is no significant correlation between the amino acid homology and Ks values of the 12 groups (correlation coefficient = −0.337). The amino acid sequence of *PP*Rib*P* synthetase was strikingly conserved between humans and rats. Thus, during evolution, there may have been a highly selective constraint to maintain the structure of this protein for the unique catalysis of a pyrophosphoryl transfer reaction and/or tightly controlled regulatory functions.

Acknowledgements: The authors thank Drs H. Fuse and A. Shimazaki, Chiba University School of Medicine, for the human testis obtained at surgery and M. Ohara for critical comments. This study was supported in part by grants from the Ministries of Education, Science and Culture, and of Health and Welfare, Japan.

REFERENCES

[1] Keppler, D. and Holstege, A. (1982) in: Metabolic Compartmentation (Sies, H. ed.) pp.147–203, Academic Press, London.
[2] Becker, M.A., Raivio, K.O. and Seegmiller, J.E. (1979) Adv. Enzymol. 49, 281–306.
[3] Taira, M., Ishijima, S., Kita, K., Yamada, K., Iizasa, T. and Tatibana, M. (1987) J. Biol. Chem. 262, 14867–14870.
[4] Taira, M., Kudoh, J., Minoshima, S., Iizasa, T., Shimada, H., Shimizu, Y., Tatibana, M. and Shimizu, N. (1989) Somat. Cell Mol. Genet., in press.
[5] Sperling, O., Boer, P., Persky-Brosh, S., Kanarek, E. and De Vries, A. (1972) Rev. Eur. Etud. Clin. Biol. 17, 703–706.
[6] Becker, M.A., Kostel, P.J., Meyer, L.J. and Seegmiller, J.E. (1973) Proc. Natl. Acad. Sci. USA 70, 2749–2752.
[7] Zoref, E., De Vries, A. and Sperling, O. (1977) Hum. Hered. 27, 73–80.
[8] Becker, M.A., Losman, M.J., Wilson, J. and Simmonds, H.A. (1986) Biochim. Biophys. Acta 882, 168–176.
[9] Taira, M., Iizasa, T., Yamada, K., Shimada, H. and Tatibana, M. (1989) Biochim. Biophys. Acta, in press.
[10] Kozak, M. (1986) Cell 44, 283–292.
[11] Fox, I.H. and Kelly, W.N. (1971) J. Biol. Chem. 246, 5739–5748.
[12] Dayhoff, M.O., Schwartz, R.M. and Orcutt, B.C. (1978) in: Atlas of Protein Sequence and Structure (Dayhoff, M.O. ed.) vol.5, suppl.3, pp.345–352, National Biomedical Research Foundation, Washington, DC.
[13] Miyata, T. and Yasunaga, T. (1980) J. Mol. Evol. 16, 23–36.
[14] Hove-Jensen, B., Harlow, K.W., King, C.J. and Switzer, R.L. (1986) J. Biol. Chem. 261, 6765–6771.