Volume 312, number 1, 10–14 FEBS 11689 November 1992

The 5′ untranslated region of the human cellular glutathione peroxidase gene is indispensable for its expression in COS-7 cells

Hiroshi Kurata[a], Keiichi Kamoshita[b], Eiko Kawai[b], Yoshikazu Sukenaga[b] and Takaharu Mizutani[a]

[a]*Faculty of Pharmaceutical Sciences, Nagoya City University, Mizuho-ku, Nagoya 467* and [b]*Research Laboratories Pharmaceuticals Group, Nippon Kayaku Co., Kita-ku, Tokyo 114, Japan*

Received 11 August 1992

We studied the expression of the human cellular glutathione peroxidase (GPx) gene, from which a key enzyme containing selenocysteine (Scy) at the active site is produced. Expression of some human GPx gene mutants in COS-7 cells revealed that the 5′ untranslated region (utr) was necessary for expression of the GPx gene, since mutant genes having 10 base pairs (bps) at the 5′utr (the complete had 311 bps) expressed GPx at very low levels. The genes with 311 or 408 bps at the 5′utr were better expressed than those having 257 bps. The GPx gene having 133 bps at the 3′utr (80 bps shorter than the entire length) was highly expressed. This deletion did not influence expression. We constructed some mutants in which 3 bases were altered at the upstream region of the Scy UGA codon in the frame of the GPx gene, by site-directed mutagenesis. GPx expression decreased but the expression was restored. Therefore, the upstream region of the in-frame Scy codon was not essential in the Scy decoding mechanisms. Finally, the 5′utr was essential for the expression of GPx gene. However, the deletion of a part of the 3′utr and the site-directed mutation upstream of the Scy codon did not show drastic effects on the expression.

Glutathione peroxidase; Selenium; Selenocysteine; Site-directed mutagenesis; 5′-Untranslated region

1. INTRODUCTION

Selenium is an essential trace element in humans; a Se deficiency causes Keshan disease which is endemic in China [1]. The syndrome is relieved by administration of selenious acid. Selenium presents as selenocysteine (Scy) in some seleno-proteins, such as cellular glutathione peroxidase (GPx) [2], plasma glutathione peroxidase [3], 5′-iodothyronine monodeiodinase (5′DI) [4], plasma seleno-protein (SeP) [5], and phospholipid hydroperoxide glutathione peroxidase (PHGPx) [6]. We previously reported the sequences of human cellular GPx cDNA and genomic DNA [7,8]. The Scy codon on mRNAs of these proteins is UGA which also functions as a stop codon. Half of all stop codons in mammals is UGA [9]. Thus, UGA has two roles, namely termination and Scy. Several years previously, a serine tRNA corresponding to UGA codon was discovered and referred to as natural opal suppressor tRNA [10]. A minor serine tRNA having NCA of anticodon has been sequenced [11]. Recently, it has been reported that this natural suppressor tRNA binds Scy [12] and that the carbon skeleton comes from serine [13]. We have reported the mechanisms of in vitro enzymatic synthesis of Scy-tRNA from Ser-tRNA in mammals [14,15]. This mechanism is similar to that in *Escherichia coli*, in which the biosynthetic mechanism of formate dehydrogenase (FDH) has been clarified [16].

The discrimination mechanisms of the Scy UGA codon from the many UGA codons on mRNAs are of interest. It is possible that a particular sequence in mRNA of seleno-proteins can decode Scy on ribosomes. Some models of the secondary structure of mRNAs have been reported [17–19]. The model of human GPx mRNA postulated by Zinoni et al. [18] did not fit with the mRNAs of mammalian other seleno-proteins. This model was derived from the stem-loop model of the FDH in *E. coli*. This discrepancy suggests that there is no common secondary structure between mRNAs of seleno-proteins from eukaryotes and prokaryotes for the Scy decoding system. The stem-loop model at the 3′ untranslated region (utr) in 5′ DI mRNA postulated by Berry [19] is currently the most appropriate, but it does not completely fit all mRNAs of seleno-proteins, such as that of PHGPx. Our model for GPx, which contained a stem-loop structure surrounding the Scy UGA codon [17], also did not fit with all mRNAs of seleno-proteins. However, it is possible that a key structure for recognition of the Scy UGA codon presents in the frame of mRNAs of mammalian seleno-proteins and is located near the Scy UGA codon. In

Correspondence address: T. Mizutani, Faculty of Pharmaceutical Sciences, Nagoya City University, Tanabe-dori, Mizuho-ku, Nagoya 467, Japan. Fax: (81) (52) 834-9309.

Abbreviations: GPx, cellular glutathione peroxidase; Scy, selenocysteine; utr, untranslated region; dxm, dexamethasone; 5′DI, 5′-iodothyronine monodeiodinase; SeP, plasma selenoprotein; PHGPx, phospholipid hydroperoxide glutathione peroxidase; FDH, formate dehydrogenase; bps, base pairs.

Published by Elsevier Science Publishers B.V.

order to understand these mechanisms, we constructed some deletion and site-specific mutants of the GPx gene, and investigated the influence of those mutations on the expression of GPx in COS-7.

2. EXPERIMENTAL

2.1. *Construction of expression vector*

Genetic procedures and recombinant techniques were carried out according to the literature [20]. Human GPx genes used were the cDNA and genomic DNA previously reported [7,8]. Genomic DNA and cDNA of GPx were subcloned in pcD [21] or pSV [22] which contain the SV40 ori and early promoter. Plasmid pSVM has the *dhfr* gene deleted and contains the mouse mammary tumor virus (MMTV) promoter which was induced with dexamethasone (dxm).

2.2. *Site-directed mutagenesis*

New constructs containing mutations were prepared using the oligonucleotide-directed in vitro mutagenesis system Ver. 2 (Amersham). A restriction fragment from pcDcGPxdA with *Hin*dIII and *Eco*RI was subcloned into M13mp19 and single-stranded DNA was prepared as a template for mutagenesis. Oligonucleotides synthesized using the ABI model 382A were as follows (the positions of mutations are underlined).

5'-GCCTCAGAGAGACGCCAC-3' (for cGPxbp5)
5'-GTGCCTCACAAAGACGCC-3' (for cGPxbp3)

Mutation was confirmed by sequencing with Sequenase Ver. 2.0 (USB). The mutated restriction fragments from replicative form DNAs of recombinant M13 were fused into the fragment made from pcDcGPxdA.

2.3. *Expression of vectors in COS-7*

COS-7 cells were cultured in Dulbecco's modified Eagle's Medium (Flow Lab.) containing 10% fetal calf serum (Irvine Sci.). Expression vectors were introduced into COS-7 by the DEAE-dextran method. After transfection, the medium was replaced with that containing 7.5 μM sodium selenite, and the transfected cells were cultured for 48 h. The expression level in COS-7 was checked with pSV-β-galactosidase (Promega).

2.4. *Western-blot analysis*

Expression of GPx was analyzed by Western blotting. Cells collected were disrupted by sonication and cellular extracts were separated by SDS-PAGE (12% acrylamide). Proteins in gels were electroblotted onto Immobilon membranes (Millipore). GPx was detected by immuno-staining with rabbit serum immunized with bovine GPx (Sigma), which clearly reacted with human GPx, and subsequently with anti-rabbit IgG goat serum conjugated horseradish peroxidase (MBL). Bands were visualized using an immuno-staining kit (Konika).

3. RESULTS

We investigated the role of the 5'utr and 3'utr of the human GPx gene on GPx expression. Fig. 1 shows a map of GPx gene in some expression vectors used in this study. These vectors were composed of genomic GPx and cDNA in pcD, pSV and pSVM based upon the SV40 promoter. In pcDcGPxdA, the 5'utr of GPx gene has all 311 base pairs (bps) but the 3'utr is 133 bps, deleted by 80 bps. Plasmid pcDcGPx10 has a complete 3'utr (213 bps) but the 5'utr is only 10 bps. The vector containing a genomic DNA with a 278 bps intron has a different sized 5'utr. In comparison with the 311 bps at the 5'utr of cDNA, pSVgGPx and pcDgGPx have 408 and 257 bps at the 5'utr, respectively. These vectors were transfected into COS-7 by the DEAE-dextran method.

The results of the transient expression of those plasmids containing cDNA and genomic DNA are shown in Fig. 2A and B, respectively. Lane 1 in Fig. 2A shows the profile of the authentic human GPx enzyme. GPx in pcDcGPxdA was not expressed in the absence of Se

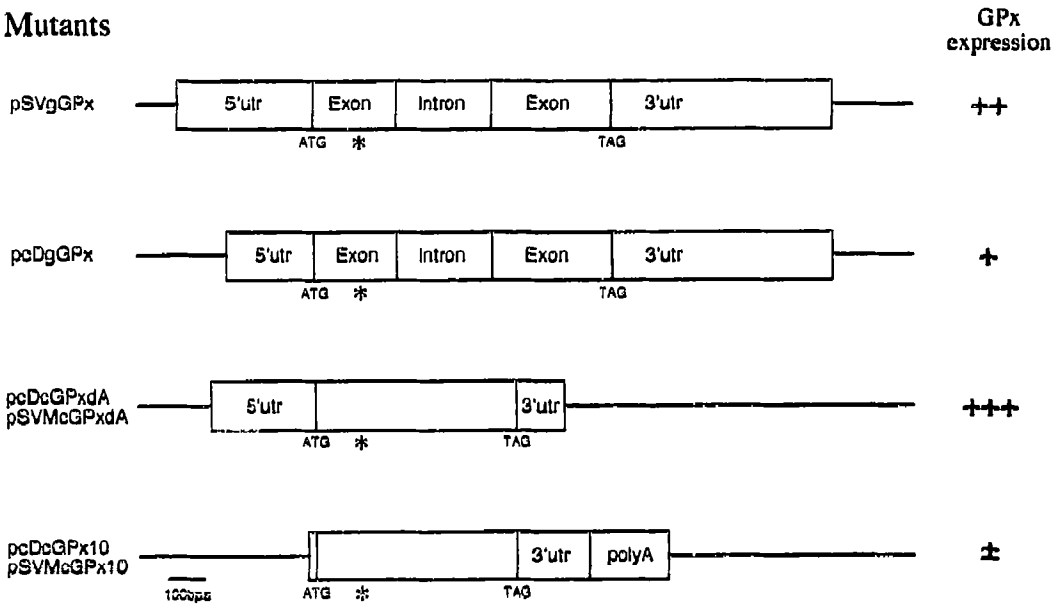

Fig. 1. Some deletion mutants of GPx gene expressed in COS-7. The expression levels of GPx are indicated on the right side. Asterisks indicate the positions of the Sey UGA codon.

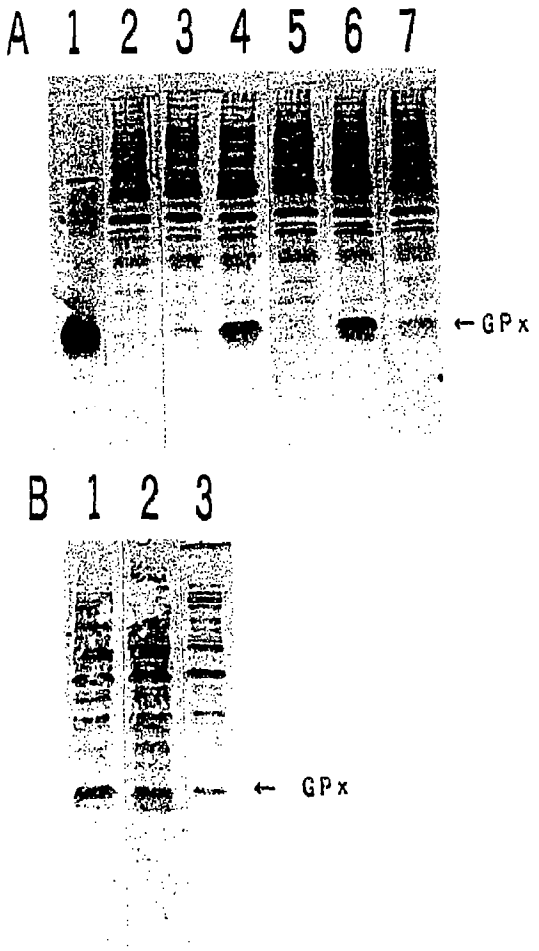

Fig. 2. Profiles of GPx expressed in COS-7 visualized by Western-blotting. Panels A and B contain cDNA and genomic DNA, respectively. Panel A: (1) human GPx protein 0.1 μg; (2) pSG5 control plasmid containing β-globin gene; (3 and 4) pcDcGPxdA; (5) pcDcGPx10; (6 and 7) pSVMcGPxdA. Lane 3 is in the absence of Se in medium. Lane 6 is induced with dxm but 7 is not. Panel B: (1) positive control pcDcGPxdA; (2) pSVgGPx; (3) pcDgGPx. Arrows indicate the position of the GPx enzyme. The levels of the expression of each plasmid are shown in Fig. 2 by +++, ++, or ±.

in the medium (lane 3). However, GPx was expressed in the presence of Se (lane 4) at the position shown by the arrow on the right side. Selenium is necessary for the expression of GPx gene in COS-7. Lane 2 is the control plasmid containing β-globin gene and GPx was not expressed. Vector pcDcGPx10, which contained only 10 bps at the 5'utr, was hardly expressed GPx in COS-7 as shown in lane 5. Vector pSVMcGPxdA, which contained a promoter induced by dxm, was expressed GPx at the same level as that of pcDcGPxdA in the presence of an inducer (lane 6) but at low levels in the absence of dxm (lane 7). Plasmid pSVMcGPx10, which contained the dxm promoter and 10 bps at the 5'utr, expressed little GPx in the presence of inducer (data not shown). Thus, both pcDcGPx10 and pSVMcGPx10 were expressed GPx at very low levels in COS-7. These results showed that the 5'utr was necessary for the expression of human GPx. Plasmids pcDcGPxdA and pSVMcGPxdA, which harbour an 80 bps deletion at the 3'utr, were highly expressed GPx. This suggests that part (80 bps) of the 3'utr is not necessary for the GPx expression.

Fig. 2B shows the expression of genomic DNA in the presence of Se. The expression of pSVgGPx (lane 2), which contained 408 bps at the 5'utr, was the same level as that of pcDcGPxdA (lane 1), which contained 311 bps at the 5'utr. The expression of pcDgGPx, which contained 257 bps at the 5'utr, was at a low level (lane 3). The possibility cannot be excluded that the mutational alteration affected transcription of mRNA stability. Overall, however, we found that the 5'utr was important for the expression of GPx. Fig. 2B also showed that an intron in genomic GPx genes did not influence expression. These results are summarized in Fig. 1. We confirmed the expression of GPx not only by Western blotting but also by the direct measurement of GPx activity in cell extracts. From this, we estimated the amount of GPx enzyme expressed with pcDcGPxdA to be about 1 μg/ml medium and 3.1 μg/10^6 cells. This is the same level as that of other proteins expressed in cultured cells. This showed that the translation level of GPx as a seleno-protein was reserved by the turnover of Scy-tRNA, which was synthesized by a system containing several enzymes [14,15], although $tRNA^{Scy}$ was at low levels in the cytosol.

Next, we investigated the influences of some base changes in the GPx frame, especially the region upstream of the Scy UGA codon. We constructed two GPx genes by site-directed mutagenesis. The mutated positions are shown in Fig. 3. The strategy of the mu-

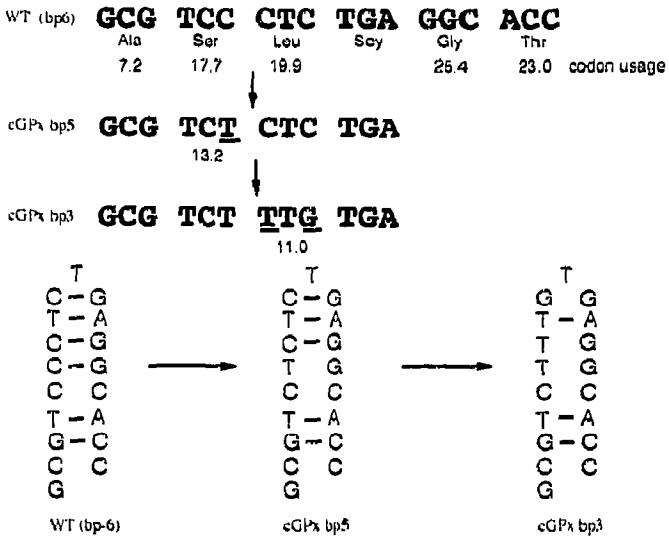

Fig. 3. Design for site-directed mutagenesis. The upper area shows the positions of mutation indicated by underlines. The lower area shows the stem-loop model of mutants (cGPxbp5 and cGPxbp3) and the original GPx gene (pcDcGPxdA indicated by WTbp6). The numbers under codons show codon usage.

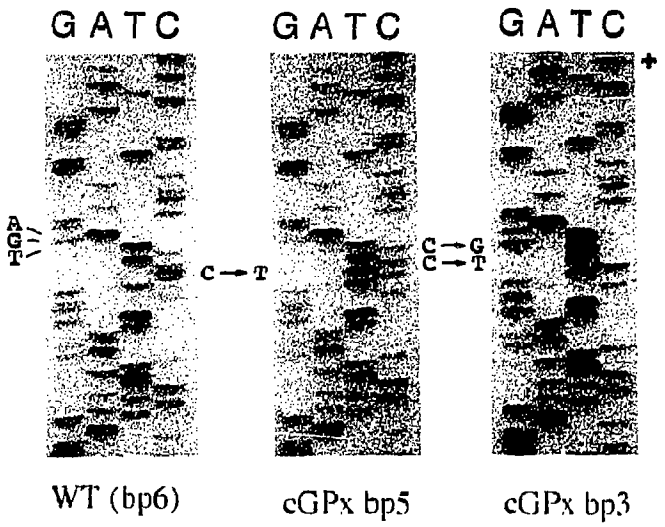

Fig. 4. Sequence for confirmation of base changes in mutants by site-directed mutagenesis. (Left) WTbp6 (pcDcGPxdA); (center) cGPxbp5; (right) cGPxbp3. On the right side, the position of Scy TGA is indicated and the mutated bases are also indicated between the photos.

tagenesis was silent mutation (changes of bases and no change of amino acids), and similar codon usage. In the original cGPx gene (pcDcGPxdA) which highly expressed GPx, a stem between bases after and before the Scy UGA codon can be formed as shown in Fig. 3A. The mutant cGPxbp5 has a base change from C to T at the 4th base upstream of the Scy UGA codon and the number of bps is reduced from 6 to 5. The mutant cGPxbp3, made from cGPxbp5, has base changes from C to G and from C to T at the 1st and 3rd positions upstream of the Scy UGA codon, respectively. The mutation was confirmed by standard dideoxy sequencing as shown in Fig. 4.

The results of expression of these mutants in COS-7 are shown in Fig. 5. The expression of cGPxbp5 (lane 4) was half that of the original pcDcGPxdA (lane 3). The expression level of cGPxbp3 (lane 5) was the same level as that of cGPxbp5. These base changes did not cause drastic effects in the expression of the GPx gene because clear expression was found in both mutants. Thus, we considered that the upstream region of the Scy codon did not have an essential role in effective GPx expression. Codon usage of UCU at the mutated position in cGPxbp5 and cGPxbp3 was 13.2, which was less than the 17.7 of UCC in the original pcDcGPxdA [9]. Codon usage of UUG in cGPxbp3 was 11.0, which was less than the 19.9 of CUC in pcDcGPxdA and cGPxbp5. Thus, these values represent low codon usage [9]. Therefore, the low expression of GPx by cGPxbp5 and cGPxbp3 might be dependent upon low codon usage at the upstream region of the Scy UGA codon.

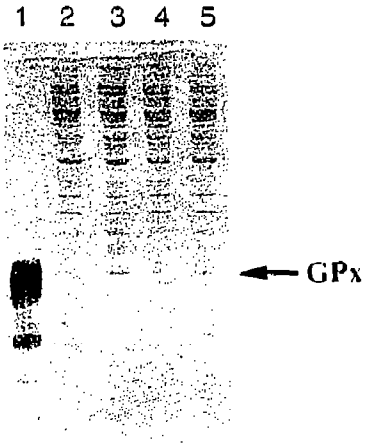

Fig. 5. Profiles of GPx expressed in COS-7 with mutants cGPxbp5 and cGPxbp3 by Western blotting. (Lane 1) Human GPx enzyme 0.1 μg; (2) blank; (3) WTbp6 (pcDcGPxdA); (4) cGPxbp5; (5) cGPxbp3. Arrow indicates the position of GPx enzyme. The levels of GPx expression are indicated below the blot.

4. DISCUSSION

In mammalian 5'DI, it has been reported that a large stem-loop structure, located near 1000 nucleotides downstream from the peptide termination codon, is essential for expression of 5'DI [19]. This model is the best to date. The similar secondary structure model is also found at the GPx 5'utr which is essential for the expression of GPx, as shown in Fig. 6. This model at the 3'utr of 5'DI mRNA may fit some mammalian seleno-proteins. However, the 3'utr in many mRNAs contains more A and U, and the base pairing is generally related to the stability of the mRNAs [23]. Therefore, the base pairing at the 3'utr in the model by Berry may not be related to discrimination of the Scy UGA codon but with the stability of 5'DI mRNA. It is possible that the mutant mRNA with a deletion at the 3'utr is labile and cannot produce 5'DI activity.

In this study, we showed that the 5'utr of the GPx

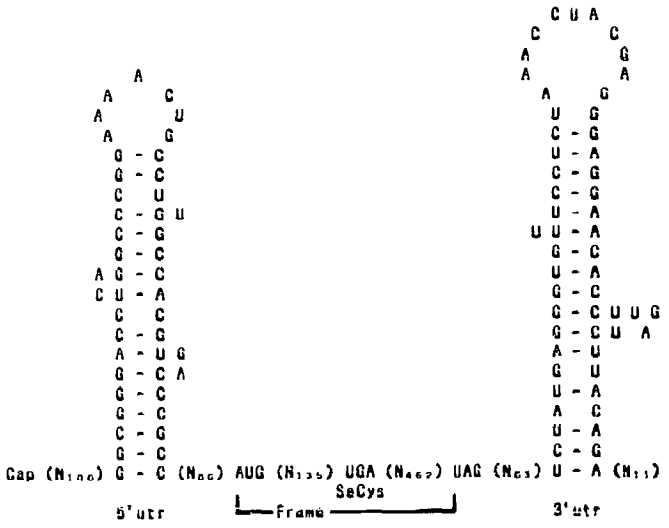

Fig. 6. A secondary structure model at the 5'utr and 3'utr on the mRNA of pcDcGPxdA, according to the literature [19].

Table I

Chain lenghts of the 5' and 3' untranslated regions in mRNAs of some mammalian seleno-proteins

Seleno-proteins	5'utr	3'utr	Refs.
Human GPx	319	213	[7]
Murine GPx	37	214	[2]
5'DI	6[a]	1,313	[4]
SeP	34[a]	1,598	[5]
Plasma GPx	53[a]	871[a]	[3]
PHGPx	–	206	[6]

[a] These are from cDNA published in the papers and may be shorter than that from those mature mRNAs.

gene was necessary for expression of the GPx gene. We summarized the sizes of the 5'- and 3'utr of cDNA of seleno-proteins in Table I. Those are estimated from cDNA and may not be complete mature mRNAs. The size of the 5'utr of human cGPx is longer than other seleno-protein mRNAs as shown in Table I. As shown in Table I, the sizes of mammalian seleno-protein mRNAs are random and it is difficult to obtain a general profile for decoding the Scy UGA codon from the lengths of the 5'- and 3'utrs.

A stem-loop structure downstream of the Scy UGA codon on FDH mRNA in *E. coli* can decode Scy [18]. Recently, a new mechanism for UGA termination events in *E. coli* has been proposed as follows. The 5'-UCAUCA-3' motif in 16 S rRNA initially interacts with the stop UGA codon, then RF-2 recognizes this complex [24]. However, this sequence was not found at a similar region in mammalian 18 S rRNA. The primary structure of RF-2 of *E. coli* differs from mammalian eRF. We made an expression vector in *E. coli*, which contained the runaway ori and Tac promoter, and an expression vector in *Saccharomyces cerevisiae*, which contained the glyceraldehyde-3-phosphate dehydrogenase promoter. However, these two vectors did not express GPx in *E. coli* and *S. cerevisiae*, respectively. This suggests that the UGA decoding systems of both stop and Scy differ between mammals and bacteria.

Figs. 2 and 5 do not show any small peptides terminated at the Scy UGA codon and stained with the polyclonal antiserum for GPx at the electrophoretic front. Meanwhile, the polyclonal antibody for 5'DI reacted well with the small peptide terminated at the Scy UGA codon [19]. In the case of expression of 5'DI, the main product was this small peptide terminated at the Scy UGA codon. Meanwhile, this study with mutants of GPx gene suggests that the decoding information of the Scy codon is stored after mutagenesis.

Acknowledgements: We thank Dr. Osamu Koiwai of the Cancer Research Institute of Aichi Prefectural Cancer Center for the generous gift of COS-7 and for many helpful suggestions, and Mr. K. Kitagawa of Nagoya University for much technical advice. This work was supported by a grant-in-aid for scientific research on priority areas from the Ministry of Education, Science and Culture of Japan, and by a research grant from the Ishida Foundation.

REFERENCES

[1] Keshan Disease Group (1979) Chin. Med. J. 92, 477–482.
[2] Chambers, I., Frampton, J., Goldfarb, P., Affara, N., McBain, W. and Harrison, P.R. (1986) EMBO J. 5, 1221–1224.
[3] Takahashi, K., Akasaka, M., Yamamoto, Y., Kobayashi, C., Mizoguchi, J. and Koyama, J. (1990) J. Biochem. 108, 145–148.
[4] Berry, M.J., Banu, L. and Larson, P.R. (1991) Nature 349, 438–440.
[5] Hill, K.E., Lloyd, R.S., Yang, J.-G., Read, R. and Burk, F. (1991) J. Biol. Chem. 266, 10050–10053.
[6] Schuckelt, R., Brigelius-Flohe, R., Maiorino, M., Roveri, A., Reumkens, J., Strassburger, W., Ursini, F., Wolf, B. and Flohe, L. (1991) Free Radical. Res. Commun. 14, 343–361.
[7] Sukenaga, Y., Ishida, K., Takeda, T. and Takagi, K. (1987) Nucleic Acids Res. 15, 7178.
[8] Ishida, K., Morino, T., Takagi, K. and Sukenaga, Y. (1987) Nucleic Acids Res. 15, 10051.
[9] Wada, K., Aota, S., Tsuchiya, R., Ishibashi, F., Gojobori, T. and Ikemura, T. (1990) Nucleic Acids Res. 18, 2367–2411.
[10] Hatfield, D. and Portugal, F.H. (1970) Proc. Natl. Acad. Sci. USA 67, 1200–1206.
[11] Kato, N., Hoshino, H. and Harada, F. (1983) Biochem. Int. 7, 635–645.
[12] Lee, B.J., Worland, P.J., Davis, J.N., Stadtman, T.C. and Hatfield, D. (1989) J. Biol. Chem. 264, 9724–9727.
[13] Sunde, R.A. and Evenson, J.K. (1987) J. Biol. Chem. 262, 933–937.
[14] Mizutani, T., Kurata, H. and Yamada, K. (1991) FEBS Lett. 289, 59–63.
[15] Mizutani, T., Kurata, H., Yamada, K. and Totsuka, T. (1992) Biochem. J. 284, 827–834.
[16] Bock, A., Forchhammer, K., Heider, J. and Baron, C. (1991) Trends Biochem. Sci. 16, 463–467.
[17] Mizutani, T. and Hitaka, T. (1988) FEBS Lett. 232, 243–248.
[18] Zinoni, F., Heider, J. and Bock, A. (1990) Proc. Natl. Acad. Sci. USA 87, 4660–4664.
[19] Berry, M.J., Banu, L., Chen, Y., Mandel, S.J., Kieffer, J.D., Harney, J.W. and Larsen, P.R. (1991) Nature 353, 273–276.
[20] Sambrook, J., Fritch, E.F. and Maniatis, T. (1989) in: Molecular Cloning: A Laboratory Manual, (Cold Spring Harbor Lab., Cold Spring Harbor, NY), 2nd edn., J. Virol. 61, 3543–3549.
[21] Okayama, H. and Berg, P. (1983) Mol. Cel. Biol. 3, 280–289.
[22] Lee, F., Mulligan, R., Berg, P. and Ringold, G. (1981) Nature 294, 228–232.
[23] Shaw, G. and Kamen, R. (1986) Cell 46, 659–667.
[24] Goringer, H.U., Hijazi, K.A., Murgola, E.J. and Dahlberg, A.E. (1991) Proc. Natl. Acad. Sci. USA 88, 6603–6607.