

Prominent expression of the selenoprotein thioredoxin reductase in the medullary rays of the rat kidney and thioredoxin reductase mRNA variants differing at the 5' untranslated region

Anna-Klara RUNDLÖF*, Mattias CARLSTEN*, MaiBritt M. J. GIACOBINI† and Elias S. J. ARNÉR*¹

*Medical Nobel Institute for Biochemistry, Department of Medical Biochemistry and Biophysics, Karolinska institutet, S-171 77 Stockholm, Sweden, and †Department of Neuroscience, Karolinska institutet, S-171 77 Stockholm, Sweden

The mammalian selenoprotein thioredoxin reductase is a central enzyme in protection against oxidative damage or the redox control of cell function. Previously a neuroblastoma-cell-derived 2193 bp cDNA for rat thioredoxin reductase 1 (TrxR1) was characterized [Zhong, Arnér, Ljung, Åslund and Holmgren (1998) *J. Biol. Chem.* **273**, 8581–8591]. Here, the major rat TrxR1 mRNA was determined as 3.5 kb by Northern blotting. A corresponding full-length 3360 bp liver-derived cDNA was cloned and sequenced, being extended in the 3' untranslated region (3' UTR) compared with the previous clone. Among tissues examined, lowest TrxR1 mRNA levels were found in spleen and testis and highest in liver and kidney. High expression in kidney was unexpected and *in situ* hybridization of adult rat kidney was performed. This revealed a highly structured expression pattern with the mRNA being prominently synthesized in

the proximal tubules of the medullary rays. Analysing rat kidney cDNA, a 5' UTR domain of TrxR1 was found that was different from that in liver- or neuroblastoma-derived cDNA clones. The kidney-derived 5' UTR mRNA domain was instead highly similar to kidney-derived cDNA variants of murine apolipoprotein E. By sequence determination of the rat genomic sequence upstream of the open reading frame for TrxR1, an exon was encountered that encoded a third alternative 5' UTR domain that could also be expressed, as confirmed by its presence in a mouse skin TrxR1 cDNA clone. It can therefore be concluded that TrxR1 mRNA is expressed in at least three different variants that differ at their 5' UTRs.

Key words: alternative splicing, full-length cDNA, *in situ* hybridization, mammalian.

INTRODUCTION

The thioredoxin system [thioredoxin (Trx), thioredoxin reductase (TrxR) and NADPH] is a ubiquitous redox system found in both prokaryotic and eukaryotic cells. Trx is a well-characterized 12 kDa protein that is highly conserved in bacterial, plant and vertebrate organisms and contains an active site with two redox-active cysteine residues, reversibly oxidizable to a disulphide [1]. Trx participates in many redox reactions, including the synthesis of deoxyribonucleotides for DNA synthesis [2] and the redox control of transcription factors [3]. In addition, human Trx functions as an extracellular cytokine for lymphoblastic cells [4,5] and a truncated form is synthesized in macrophages and increases the cytotoxicity of eosinophilic cells [6,7]. The active-site disulphide in oxidized Trx is reduced by TrxR in an NADPH-dependent reaction. TrxR from *Escherichia coli* was purified 30 years ago and was found to be a homodimer composed of two 35 kDa subunits [8]. *E. coli* TrxR has a very narrow substrate specificity, in principle reducing only *E. coli* Trx. TrxR forms from plants or yeast have similar principal characteristics to those of the bacterial enzyme.

Mammalian TrxR has been purified from several sources, including calf thymus and liver [9], rat liver [10] and human placenta [11], and comprises subunits of approx. 55 kDa in comparison with the smaller 35 kDa subunits of TrxR in lower organisms. Furthermore, in contrast with the bacterial enzyme, mammalian TrxR has a remarkably broad substrate specificity

and reduces not only protein disulphides but also low-molecular-mass disulphide compounds such as 5,5'-dithiobis-(2-nitrobenzoic acid) ('DTNB') [9] and lipoic acid [12], low-molecular-mass non-disulphide compounds such as selenite [13] or alloxan [14] and also functions as a peroxidase with H₂O₂ or lipid hydroperoxides [15]. Recently the mammalian enzyme was discovered to be a selenoprotein with a penultimate catalytically active selenocysteine residue [16–18] and within the past year two isoenzymes have been discovered, one mitochondrial [19,20] and one localized predominantly in testis [21]. A human cDNA clone for the classically purified enzyme, TrxR1, was published in 1995 [22] and we characterized a rat cDNA clone for this enzyme [18]. The sequence similarity to prokaryotic TrxR is vague, and mammalian TrxR is more similar to glutathione reductase but with an additional 16-residue C-terminal elongation carrying the selenocysteine-containing motif. The selenocysteine residue in the C-terminal motif Gly-Cys-Sec-Gly (in which Sec represents selenocysteine) is conserved in the human, rat and bovine enzymes and is essential for its catalytic activity [18]. The presence of this presumably easily accessible redox-active motif should explain the wide substrate specificity. As judged from its enzymic activity, mammalian TrxR most probably carries antioxidative or redox regulatory functions as such, in addition to its role in Trx-coupled reactions. Previous immunohistochemical studies have shown that TrxR and Trx have considerably different expression patterns in cells of different tissues [23]. With the availability of cDNA clones it has now become possible to study the tissue-

Abbreviations used: EST, expressed sequence tag; ORF, open reading frame; RACE, rapid amplification of cDNA ends; Trx, thioredoxin; TrxR, thioredoxin reductase; UTR, untranslated region.

¹ To whom correspondence should be addressed (e-mail elias.arnér@mbb.ki.se).

specific expression of TrxR also on the mRNA level. Here the previously characterized rat TrxR cDNA [18] was used for probing a Northern blot with mRNA from eight different rat tissues, to determine the size of the rat mRNA and to initiate studies of tissue-specific mRNA expression patterns. The result led to the cloning of a full-length rat cDNA clone and to the discovery of a surprisingly structured and prominent expression of the enzyme in the proximal tubules of the medullary rays in the kidney. Furthermore, mRNA variants were also identified that differed in their 5' untranslated regions (UTRs).

MATERIALS AND METHODS

Northern blot hybridization

A commercial filter (Clontech, Palo Alto, CA, U.S.A.) with rat mRNA from eight different tissues (heart, brain, lung, liver, spleen, skeletal muscle, kidney and testis) was probed with [α - 32 P]dCTP-labelled rat TrxR cDNA [18], using a β -actin probe included with the filter as an internal control for mRNA amount. For α - 32 P-labelling of the probes, the Prime-a-gene kit (Promega) was used. The hybridizations were performed with 2.6×10^6 c.p.m./ml probe (the specific radioactivity of the TrxR cDNA probe was 3.2×10^9 c.p.m./ μ g; for the actin probe it was 2.21×10^9 c.p.m./ μ g) in Expresshyb Hybridization solution (Clontech) in accordance with the protocol provided by the manufacturer.

In situ hybridization

Adult Sprague–Dawley rats were killed and kidney tissues were rapidly dissected, frozen and sectioned as described [24]. Two anti-sense oligonucleotide probes based on nt 181–230 and 1011–1060 of the rat TR cDNA sequence [18] were used for *in situ* hybridization. In addition to these two TrxR-specific probes, a random probe with the sequence 5'-GGA GTT TTC TTG GCC CCT CGG TAA TTG CGG ATG ATT GGT CTC GCA TGA CT-3' [25] was used as one negative control and a GDNF probe based on nt 540–589 [26] was used as a second negative control for a protein known to be down-regulated in the adult rat kidney [27]. Throughout the kidney, the random probe and the GDNF probe gave very low non-specific signals (results not shown). The oligonucleotides were labelled at the 3' end with adenosine 5'-[α - 35 S]thio]triphosphate ([α - 35 S]dATP) by using terminal deoxynucleotidyltransferase and purified with a QIAquick Nucleotide Removal kit (Qiagen).

Sections of adult rat kidney were hybridized for 17 h at 42 °C in solution containing 50% (v/v) formamide, 4 \times SSC (SSC is 0.15 M NaCl/0.015 M sodium citrate), 0.02% polyvinylpyrrolidone, 0.02% Ficoll, 0.02% BSA, 10% (w/v) dextran sulphate, 1% (w/v) sarcosyl, 0.02 M Na₃PO₄, 0.5 mg/ml salmon sperm DNA, 50 mM dithiothreitol and 1.15×10^7 c.p.m./ml [α - 35 S]dATP-labelled probe (specific radioactivity 3.8×10^8 c.p.m./ μ g), as described [28]. The slides were washed at 60 °C for 2 h with four changes of SSC, followed by dehydration in ethanol. Autoradiography was performed with Kodak NTB-2 liquid emulsion. After development, the sections were stained with Cresyl Violet, mounted with Entellan (Merck) and photographed. Artifacts such as dust particles and loosened tissue fragments were retouched with the Adobe Photoshop program.

Generation and sequence determination of a full-length cDNA clone of rat TrxR

To generate a full-length cDNA clone, adult Sprague–Dawley rat liver Marathon-Ready cDNA Amplification (Clontech) was

used. For rapid amplification of the 5' cDNA end (5' RACE) a gene-specific primer complementary to nt 1257–1279 of the rat TrxR cDNA clone [18] was used, whereas for the 3' RACE a primer complementary to nt 1029–1051 was used, thereby also producing an overlapping sequence to confirm the integrity of the amplification. The products obtained (2363 bp 3' RACE, 1279 bp 5' RACE) were analysed on a 1% (w/v) agarose gel, isolated and cloned into pGEM-T vectors (Promega). The plasmids were propagated in the DH5 α *E. coli* strain (Clontech) and purified with the Wizard mini-prep kit (Promega). The ends of the inserts were confirmed by using standard forward and reverse primers (Pharmacia Biotech) in an automated laser fluorescent sequencer (Pharmacia). The AutoRead Sequencing Kit (Pharmacia Biotech) was used for the dideoxy sequence reactions.

For generation of the full-length cDNA by PCR, two oligonucleotides based on the sequence information obtained from the 5' RACE and the 3' RACE were used (5'-GCC TTG CAG CGG CAC GGG CAA ACG GA-3' and 5'-ACG ACG AGT GTA CTC CTT CCA TGG GG-3'). Touchdown PCR (94 °C for 3 s, 72 °C for 4 min, 5 cycles; 94 °C for 3 s, 70 °C for 4 min, 5 cycles; 94 °C for 3 s, 68 °C for 4 min, 27 cycles) was performed with 1.8 units of Expand Long Template PCR system enzyme (Boehringer Mannheim) and 5 μ l of adult Sprague–Dawley rat liver cDNA provided in the Marathon-Ready cDNA Amplification kit (Clontech) in a buffer containing 50 mM Tris/HCl, pH 9.2, 16 mM (NH₄)₂SO₄, 2.25 mM MgCl₂, 20% (v/v) DMSO and 1% Tween 20. The PCR product (3392 bp) was recovered from a 1% (w/v) agarose gel and ligated into a pGEM-T vector, propagated in the DH5 α *E. coli* strain (Clontech), purified with plasmid DNA and sequenced as described above and then with subsequent primers based on the sequence information obtained.

5' RACE and 3' RACE of TrxR from a rat kidney cDNA library

To amplify the 5' and 3' ends of TrxR in kidney, an adult Sprague–Dawley rat kidney Marathon-Ready cDNA library (Clontech) was used together with the same gene-specific primers as used for the generation of a full-length rat liver cDNA clone (as described above). From the 5' RACE this generated a 1.4 kb product and from the 3' RACE a 2.4 kb product; these were cloned and sequenced as described in the generation of a full-length clone in rat liver.

Mapping of the genomic upstream region

To map the upstream region of rat TrxR, a PromoterFinder DNA walking kit (Clontech) was used. Two nested TrxR-specific primers (5'-GCC TCC CGA GCC TCC TCC AAT GAT GGA T-3' and 5'-TCA TAG GAC TTA GGG GCA TCT TTA GAG-3') were designed on the basis of the first part of the rat cDNA sequence, excluding the first 80 nt, which were not similar to the human sequence [18]. For the primary PCR, a touchdown PCR was performed (94 °C for 2 s, 72 °C for 4 min, 5 cycles; 94 °C for 2 s, 70 °C for 4 min, 5 cycles; 94 °C for 2 s, 67 °C for 4 min, 37 cycles) with adapter-specific primers and the TrxR-specific primers with the PromoterFinder DNA walking kit.

The PCR products from the primary PCRs were diluted 1:50 and 1 μ l aliquots of the dilutions were subjected to a secondary PCR. The secondary PCR was performed as described for the primary PCR but with the nested adapter- and TrxR-specific primers. The amplified PCR products were analysed on a 1% (w/v) agarose gel and a 1.5 kb fragment was recovered and cloned into a pGEM-T vector. The plasmid was propagated in the DH5 α *E. coli* strain (Stratagene) and the plasmid DNA was

purified with the Wizard mini-prep kit (Promega). The insert was sequenced, with standard forward and reverse primers, in an automated laser fluorescent sequencer. The AutoRead Sequencing Kit was used for the dideoxy sequence reactions.

To clone and determine the sequence of further upstream regions of the genomic DNA, two new specific nested primers (5'-GGG TAC ATG GGA CTC TAT AAG GCT CAC-3' and 5'-CCT GTG GTG GAA AGT CCT GTC TAG TGT-3') were based on the first 5' part of the previously cloned genomic sequence. Primary and secondary PCRs were performed with the same procedure as that described above. To detect correct bands among the several products of the secondary PCR, an oligonucleotide based on the upstream part of the sequence obtained previously (5'-AAG AAT TAT GGA TGG ACA GAA ATT TCA GAA-3') was used as a probe for Southern hybridization, performed with standard protocols. Two positive products, 0.8 and 3.6 kb, were cloned into pGEM-T vectors (Promega) and their sequences were determined by using the same methods as those described above. In total, 4619 nt in the genomic sequence upstream of nt 81 in the previously reported cDNA clone [18] were determined.

RESULTS

Northern blot analysis

To determine the size of the native rat TrxR mRNA and to study tissue distribution, a Northern blot filter with mRNA from eight

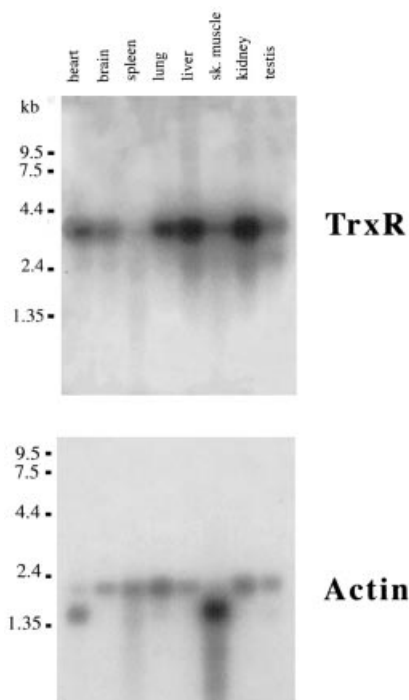


Figure 1 Tissue expression of rat TrxR mRNA

A multiple tissue Northern blot (Clontech) with mRNA from the different rat tissues indicated was probed with rat TrxR cDNA (upper panel) or the control β -actin probe supplied by the manufacturer (lower panel). The size of the TrxR1 mRNA was determined as 3.5 kb; there was a minor band at 2.6 kb. The intensities of the two forms of β -actin, with a shorter form predominantly expressed in muscle tissue, were combined for quantification and analysis of the relative levels of TrxR1 mRNA in comparison with those of β -actin, as given in Table 1. Abbreviation: sk. muscle, skeletal muscle.

Table 1 Relative TrxR mRNA levels in rat tissues

The 3.6 kb TrxR mRNA in the Northern blot and the control probe for β -actin mRNA with the same filter, shown in Figure 1, were quantified by densitometry (with the signal of both forms of β -actin combined). The pattern of expression of the TrxR mRNA between the tissues was then determined; the results are shown. The second column shows the 3.6 kb mRNA level in liver set to 1.00; the third column shows the TrxR mRNA signal divided by the actin signal, with ratios in absolute values; the fourth column shows the ratio between TrxR mRNA and the control actin probe, with the ratio in liver set to 1.00.

Tissue	TrxR	TrxR/actin (absolute ratios)	TrxR/actin (relative ratios)
Testis	0.34	0.71	0.38
Kidney	1.26	3.37	1.82
Skeletal muscle	0.38	0.54	0.29
Liver	1.00	1.85	1.00
Lung	0.58	1.43	0.77
Spleen	0.16	0.12	0.06
Brain	0.21	0.42	0.23
Heart	0.31	1.24	0.67

rat tissues was probed with an [α - 32 P]dCTP-labelled rat TrxR cDNA probe (Figure 1). The major mRNA form of rat TrxR was found to be 3.5 kb in all tissues with a possible isoform of 2.6 kb, which was significantly larger than the previously characterized 2.19 kb rat TrxR cDNA clone derived from nerve growth factor (NGF)-treated neuroblastoma cells and obtained from the Institute for Genomic Research (Rockville, MD, U.S.A.) [18]. The expression of the TrxR mRNA in the different tissues varied between 0.1-fold and 3.4-fold the signal of the actin mRNA in absolute values; the relative ratio between these mRNAs varied between 0.06-fold and 1.8-fold (with the ratio in liver set to 1.0) (Table 1). The highest expression of TrxR1 was clearly found in kidney and liver; the weakest signal was seen in spleen.

In situ hybridization

The high levels of TrxR mRNA in the adult rat kidney were unexpected because this tissue had previously been reported to contain little TrxR protein in comparison with other tissues [23]. *In situ* hybridization studies were therefore performed to determine any cell-type-specific expression in the kidney. Sections of adult rat kidney were probed with two different 35 S-labelled anti-sense oligonucleotides, corresponding to nt 230–181 and 1060–1011 within the open reading frame (ORF) of the previously published [18] rat TrxR cDNA clone. Both these probes gave identical patterns, with TrxR mRNA prominently localized to the medullary rays in a pattern typical of the proximal tubules (Figure 2).

Full-length and variant sequences of rat TrxR cDNA

The highly structured expression of mRNA in the kidney could possibly be a reflection of cell-type-specific transcription, which could be studied at a cDNA level. First, to obtain a full-length rat TrxR cDNA clone that would correspond to the size of the mRNA (Figure 1), two primers, based on the coding region of the known sequence, were used for overlapping 5' and 3' RACE. With rat liver cDNA this generated a sequence of 3360 bp including the coding region and an extended 3' UTR, highly similar to the previously published human cDNA sequence [22]. However, no 5'-end sequence extending beyond the first 80 nt of

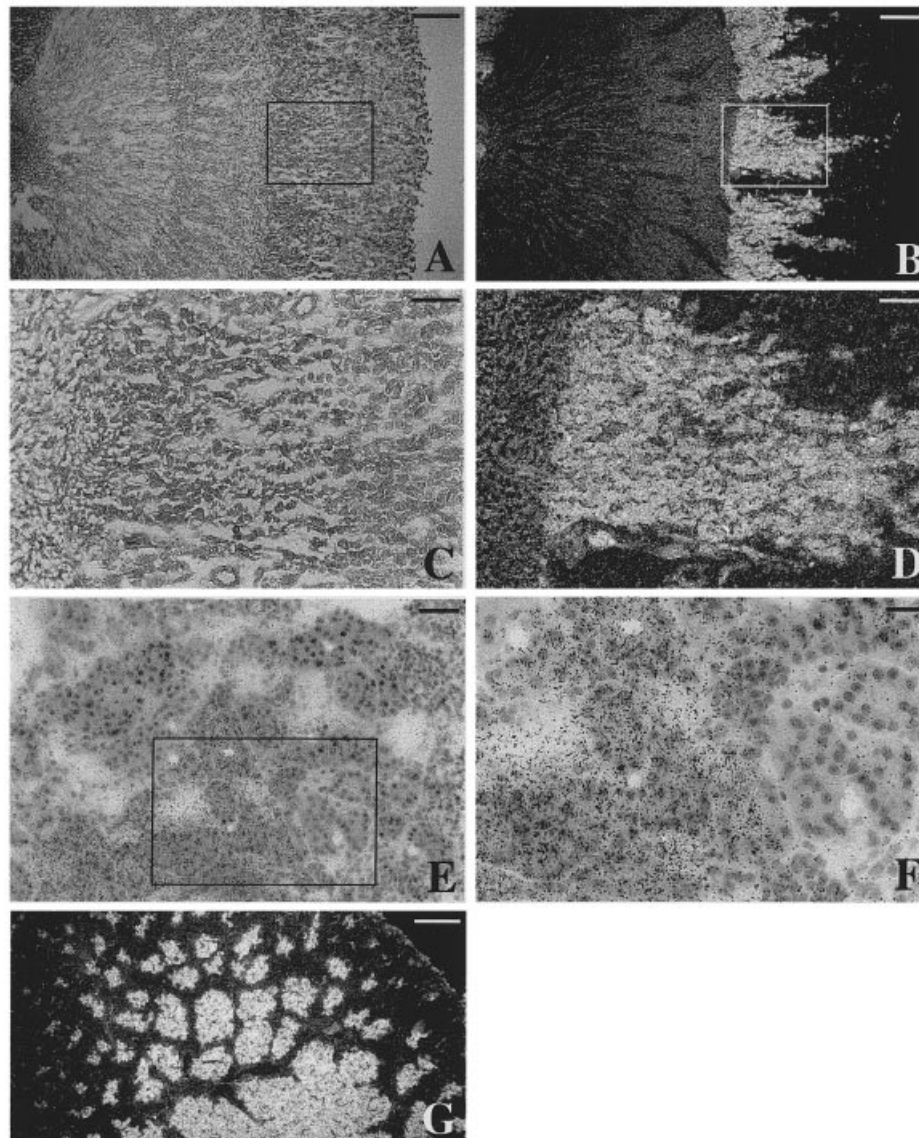


Figure 2 Expression of TrxR mRNA in the medullary rays of the kidney

Adult rat kidney sections hybridized with two 50-mers complementary to different parts of the TrxR ORF both showed the same pattern of the highly structured expression seen in this Figure. (A–F) Photographs from the same section showing both medulla and cortex; (G) sagittal section through the cortex. (A, B) The same section under bright-field (A) and dark-field (B) illumination. The squares in (A) and (B) indicate the positions of (C) and (D), which were photographed at a higher magnification. The position of (F), which has the highest magnification, is indicated with a square in (E). Scale bars: (A), (B), (G), 2 mm; (C), (D), 0.5 mm; (E), 0.1 mm; (F), 50 μ m.

the previously characterized rat cDNA clone was found; instead, the original 5'-end sequence was confirmed, although this sequence lacked similarity to the corresponding part of the human cDNA.

Interestingly, with rat kidney cDNA another 5'-end domain was identified, whereas the 5' end of liver cDNA was not found. This 5'-end domain was highly similar to cDNA clones for apolipoprotein E of mouse kidney and a kidney-derived mouse TrxR expressed sequence tag (EST) clone, the possible relevance of which will be further discussed below. No in-frame alternative upstream initiation ATG could be identified in either of these 5'-end domains. The longer 3' UTR found in the liver cDNA was also present in the kidney-derived cDNA; the mRNA species

differing at the 5' UTR could not be distinguished in Northern blots owing to the small differences in size.

With the different 5'-end variants of TrxR cDNA, we initiated sequence determination of the genomic region flanking the 5' side of the ORF. Within a 4.6 kb genomic fragment flanking the ORF, nt 76–160 in the original cDNA clone [18] were identical; however, the first 73 nt of the cDNA could not be found. Splice site prediction [29] also suggested an intron with a splice site corresponding to position 76 of the cDNA. Upstream of this exon, a typical intron with an Alu repetitive sequence motif of the Alu-J subtype was located. Unexpectedly, we found a novel exon further upstream flanked by two predicted splice sites. This exon was also present in the 5' UTR of a mouse TrxR EST

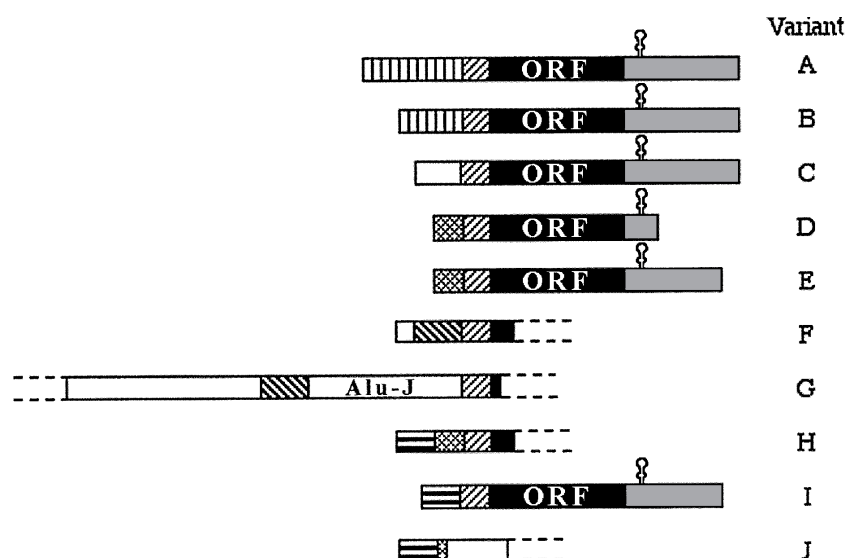


Figure 3 Schematic alignment of TrxR1 cDNA variants

The different mammalian TrxR1 cDNA variants are schematically aligned. The black boxes shows the highly homologous ORF while white boxes indicate regions with no similarities between the sequences. The different sequence domains showing more than 95% identity between cDNA of different sources are shown with the same filling patterns. Although of different lengths, the 3' UTRs of all variants are highly similar and all contain the selenocysteine insertion sequence (SECIS) element needed for selenocysteine insertion [18,40], as indicated by the 'hairpin-like' structures above the grey-shaded boxes. Note that the 5' UTR sequences of the human and bovine cDNA clones are different from the murine 5' UTRs, except for the first 60 bp flanking the 5' side of the ORF, which is identical in all TrxR cDNA sequences. Three distinct murine 5'-end domains were found; these display a possible tissue-specific expression; i.e. that of a skin EST clone (with the sequence confirmed as an exon in the rat genomic sequence), that of liver or neuroblastoma cells (found in three different cDNA clones) and that of kidney (with the same domain found in a mouse TrxR EST clone, a rat TrxR 5' RACE determination and three kidney-derived EST clones of mouse apolipoprotein E). The cDNA variants listed as A–J in the figure are summarized in the following Table and the murine sequences are given in detail in Figure 4.

Variant	Type of sequence	Organism	Tissue/cell origin	Reference	Database accession no.
A	TrxR1, cDNA	Human	Placenta	[22]	NM-003330
B	TrxR1, cDNA	Human	Bone marrow	[34]	D88687
C	TrxR1, cDNA	Bovine	Small intestine	[42]	AF053984
D	TrxR1, cDNA	Rat	NGF-treated neuroblastoma cell-line	[18]	U63923
E	TrxR1, cDNA	Rat	Liver	[32]; the present work	AF108213 AF220760
F	TrxR1 EST clone*	Mouse	Skin	M. Marra and others	AA168412
G	Genomic DNA	Rat		The present work	AF189711
H	TrxR1 EST clone*	Mouse	Kidney	M. Marra and others	AI956288
I	TrxR1, cDNA	Rat	Kidney	The present work	AF220761
J	Apolipoprotein E EST clones*	Mouse	Kidney	M. Marra and others	AI226627 AI527732 AI526517

*These sequences were directly deposited into Genbank (The WashU-NCI Mouse EST Project) in 1996 (TrxR1 EST clone, skin), 1999 (TrxR1 EST clone, kidney) or 1998/99 (apolipoprotein E EST clones).

cDNA clone, originating from skin [30]. That EST cDNA sequence was similar to nt 93–240 of the previously characterized rat neuroblastoma TrxR cDNA and to nt 2735–3078 of the 4.6 kb upstream genomic sequence, which thereby confirmed the exon as functional.

Taken together, these results indicate the finding of three different 5' UTR variants of rat TrxR cDNA, all dissimilar to the 5'-end of the human TrxR cDNA [22]. The 5' UTR domains of the rat sequences could possibly be expressed in a tissue-dependent manner, with one variant detected in the genomic upstream sequence and (mouse) skin-derived cDNA, one found in kidney-derived cDNA and one found in liver or neuroblastoma cells. By using either the 5' RACE determinations in this study

or by extensive database searches we have not found any two of the different 5' UTR variants simultaneously from the same tissue or cell type. The different cDNA domains of TrxR1 are described schematically in Figure 3 and the murine 5' UTR domains characterized here are shown in detail and aligned in Figure 4. [While this paper was under review, a mouse cDNA sequence derived from thymic lymphoma cells highly homologous to the 5' UTR variant of the rat kidney-derived cDNA found here (variant I, Figures 3 and 4), determined by H. Kawai and M. Tatsuka (unpublished work) was deposited in GenBank® (accession number AB027565). This confirms further that the 5' UTR variants shown here are expressed in diverse murine tissues.]

Murine TrxR1 5'-UTR domain variant nr. 1

cDNA variant	from nt	sequence	to nt
Variant F (mouse skin EST)	1	CAGGCTCCACAGTGTCTTCGCAGACCTCAGAGCCCTGGCGGCTGGCCCTCATAAACAGCCGTGGATGGACACTCTACTAA	80
Variant G (rat genomic sequence)	2656	CAGTACACTGGCTAACTGACAATGTTTACCTGCTACTGTGATCTGAGTGGCTGTTTTTACCAGCTTGTFTCCCCCCACA	2735
Variant F (mouse skin EST)	81	GTGCCCTGCATTGAAGGAGAAGCCCTGGTACCATGCCAGTTGATGACTGCTGGCTGTACTTCCAGCTTCTCGAGGTAG	160
Variant G (rat genomic sequence)	2736	GTGCCCTGCATTGCGAAGAAGCCCTGGTACATGCCAGTTGATGGTGGCTGGCTGCTTCCCGACTTC-CGAGGTAG ↑	2814
Variant F (mouse skin EST)	161	AACCTTTGTGCAGA-CTGTCTGGGTGGCACCCACTTGCCTCAACTGTGTGGTTCAGGTTTTTCCCTCCAGTCCCC	239
Variant G (rat genomic sequence)	2815	GGCTTTGGCGAGACTTGTCTGGGTGACGCGCACTTGCCTCAACTGTGTGGTTCAGGTTTTTCCCTCCAGTCCCC	2895
Variant F (mouse skin EST)	240	CGGCCACCACATGTGCCCGTCTGTCTGCTGAGGGGCCCTCGTGGGTAGTGTCTTCCCTT-CACGTCCCTCCAAGACAC	319
Variant G (rat genomic sequence)	2896	CGGCAACAACAGTGTGCCCTGCGTSCCGCCGAGGGGCTCTGTGGGGTGTACTTCCCTTACACATCCCTCCAAGGGAC	2976
Variant F (mouse skin EST)	320	TCCCTCCTCATCCCAGACGCCCTGTCTACTGACCCCTGTATCTGCCCTCCACCCTCCACACCTGATAGTAGGCAGGAA	400
Variant G (rat genomic sequence)	2977	TCCCTCCTCATCCCAGACGCCCTGTCTACTGACCCCTGTATCTGCCCTCCACCCTCCACACCTGATAGTAGGCAGGAA	3057
Variant F (mouse skin EST)	401	AAAAATACGCAATCTGAGCTG... /continued with TrxR consensus domain/	419
Variant G (rat genomic sequence)	3058	AAAAATACGCAATCTGAGCTG... /continued with TrxR consensus domain/ ↑	3115

Murine TrxR1 5'-UTR domain variant nr. 2

cDNA variant	from nt	sequence	to nt
Variant J (mouse kidney apoE EST's)	1	GTGTGCTTCCGTGAGCCCTCGCGTCCACGCGGGAGGTGCGGGACGC	46
Variant H (mouse kidney TrxR)	1	AATGCGCAGTCGCTCCCGGTACTTGTCTCCGCCAGGCCCTCGCGTCCACGCGGGAGGTGCGGGACGC	67
Variant I (rat kidney TrxR)	1	CGCATCCACGCGGGAGGTGCGGGACGC	27
Variant J (mouse kidney apoE)	47	CGACACCGCGGGCGGAGAAGAGCTGGTGGTTTCACCTTCCCTTGTTCAT... /continued with domain nr. 3/	94
Variant H (mouse kidney TrxR)	68	CGACACCGCGGGCGGAGAAGAGCTGGTGGTTTCACCTTCCCTTGTTCAT... /continued with domain nr. 3/	115
Variant I (rat kidney TrxR)	28	CGACACCGCGGGCGGAGAAGAGCTGGTGGTTTCACCTTCCCTTGTTCAT... /continued with TrxR consensus domain/	75

Murine TrxR1 5'-UTR domain variant nr. 3

cDNA variant	from nt	sequence	to nt
Variant J (mouse kidney apoE EST's)	95	...AGGGCGCGGGCCCTTGCAGCGCGGGCGGAGCGGAAAGGCGGGAGGCGGGAGCCAGCGGAA	161
Variant H (mouse kidney TrxR)	116	...AGGGCGGTGGGGCTTGCAGTGGCGGGCGGAGCGGAAAGGCGGGAGGCGGGAGCCAGCGGAA	182
Variant E (rat liver TrxR)	1	GCCTTGCAGCGGGCGGCAACGGAGGGCCCGGGAGGCGGAAAGCGCGGAA	56
Variant D (rat neuroblastoma TrxR)	1	AATTCGCGCAGCAACGGAGGGCCCGGGAGGCGGAAAGCGCGGAA	51
Variant J (mouse kidney apoE EST's)	162	GGTG... /continued with apoE-specific sequence/	
Variant H (mouse kidney TrxR)	183	GGTGCAGCCG-GCGGAGGGCGGCCATGGTCCAGCCCTGAAG... /continued with TrxR consensus domain/	222
Variant E (rat liver TrxR)	57	GCGAGGGAGAGCGGAGGGCGGCCATGGTCCAGCCCTGAAG... /continued with TrxR consensus domain/	97
Variant D (rat neuroblastoma TrxR)	52	GGCAGGGAGAGCGGAGGGCGGCCATGGTCCAGCCCTGAAG... /continued with TrxR consensus domain/	92

Consensus murine TrxR1 5'-UTR domain and beginning of ORF

cDNA variant	from nt	sequence	to nt
Variant G (rat genomic sequence)	4496	... /Alu-J containing intron/ ...GGGCATGTGCTTTTCTGTTAGAGTCTTTTCATTTATGTTTTCCTG	4545
Variant F (mouse skin EST)	420	...CCGAACAAAAAGGCCAAGCTTCAAAGCTGCCAACAATGAATGGCTCCAA-GATCCCCCTGGTCCCTATGA...	489
Variant H (mouse kidney TrxR)	223	...CCGAACAAAAAGGCCAAGCTTCAAAGCTGCCAACAATGAATGGCTCCAAAGATCCCCCTGGTCCCTATGA...	293
Variant E (rat liver TrxR)	98	...CCGAACAAAAAGGCCAAGCTTCAAAGCTGCCAACAATGAATGGCTCCAAAGATGCCCTTAAGTCCCTATGA...	168
Variant D (rat neuroblastoma TrxR)	93	...CCGAACAAAAAGGCCAAGCTTCAAAGCTGCCAACAATGAATGGCTCCAAAGATGCCCTTAAGTCCCTATGA...	163
Variant I (rat kidney TrxR)	76	...CCGAACAAAAAGGCCAAGCTTCAAAGCTGCCAACAATGAATGGCTCCAAAGATGCCCTTAAGTCCCTATGA...	146
Variant G (rat genomic sequence)	4546	...CAGCCAAACAAAAAGGCCAAGCTTCAAAGCTGCCAACAATGAATGGCTCCAAAGATGCCCTTAAGTCCCTATGA... ↑	4619

ORF

Figure 4 Alignment of 5' UTR domain variants of murine TrxR1

The different murine 5' UTRs described schematically in Figure 3 are shown with identical nucleotides indicated by black dots. Species-specific differences between the variants are indicated with vertical bars and nucleotides found in three out of four sequences are marked with plus signs. The 5' UTR domain variant 1 was found in a mouse EST clone originating from skin and confirmed as an exon in the rat genomic sequence. The highly similar part of this sequence was flanked by a predicted acceptor (simple arrow) and donor (double arrow) splice site. The genomic sequence continued with an Alu-J motif-containing 1.5 kb intron before the murine consensus 5' UTR domain of TrxR1 and the beginning of the ORF, also flanked by an acceptor splice site (arrow). The kidney-specific 5' UTR sequence found in two murine TrxR cDNA clones and in three mouse EST clones of apolipoprotein E (apoE) are shown as 5' UTR domain variant 2. The apolipoprotein E sequence and the TrxR sequence from mouse kidney continued with domain variant 3, also found in the rat liver and neuroblastoma TrxR cDNA sequences, but lacking from the TrxR sequence from the rat kidney-derived cDNA. The murine apolipoprotein E sequence continued with a sequence specific for apolipoprotein E, exon 4, which showed no similarity to any other sequence. It should be emphasized that no other entries in GenBank[®] showed significant similarity to any of the 5' UTR domain variants given in the Figure, using the BLAST algorithm [41] at the time of submission of this paper.

DISCUSSION

The highly structured pattern of TrxR mRNA in the adult rat kidney with localization to the medullary rays, was interesting and unexpected. The medullary rays harbour both proximal and distal tubuli as well as collecting ducts, but the pattern with a distinct and clear border in the outer medulla is typical of the straight part of the proximal tubules. Interestingly, the same expression pattern in kidney has been found for another selenoprotein, plasma glutathione peroxidase, which is secreted to the plasma by the proximal tubular epithelial cells [30]. Co-localization of glutathione peroxidase and TrxR could be indications of a functional intracellular coupling between TrxR and the synthesis or function of plasma glutathione peroxidase, particularly because plasma glutathione peroxidase is already known to be a substrate for TrxR [31] and that TrxR reduces selenite to selenide [13] (a reaction needed for the use of selenite in selenoprotein synthesis).

The full-length rat cDNA clone generated and sequenced in this study had an extended 3' UTR, but an identical ORF and 5' UTR in comparison with the shorter rat TrxR cDNA clone reported previously [18] and is essentially identical with another full-length clone published while this manuscript was in preparation [32]. The sizes of the non-polyadenylated full-length cDNA clones also agree well with the size of mRNA found on a Northern blot (Figure 1). The 3826 bp cDNA for human TrxR [22] also agrees well in size with the larger 3.9 kb mRNA seen in human tissue [33,34]. The difference in sequence and length of the 5' end between rat and human TrxR cDNA should therefore be genuine and not a cloning artifact. Interestingly, the 5' UTR of the human clones contained an alternative upstream ATG codon that resulted in an ORF for a protein of approx. 60 kDa [22,34] that was functional, at least in transfected cells [34]. In the shorter 5' UTR variants of the rat cDNA sequence, no such alternative in-frame upstream ATG codon could be found. The 5' end of the human sequence is more than 300 bp longer than the corresponding part of the rat sequence; the first 80 nt of the rat cDNA sequence are different from the human sequence, as illustrated in Figure 3. The genomic sequence revealed a splice site corresponding to nt 76 in the rat cDNA; upstream of this site an Alu-J motif-containing sequence, often found in introns, was located [35]. It was surprising that this intron was flanked by an exon that was similar not to the previously sequenced 5' end but to a novel sequence that also was present in a mouse EST cDNA clone for TrxR derived from skin. This finding supports a notion of species- or tissue-specific variants, the function of which remains to be demonstrated.

Possibly, differential splicing might yield mRNA with different regulatory elements for stability. The identical 5'-end domains of cDNA variants for the two different murine proteins TrxR and apolipoprotein E was an unexpected finding. It seems unlikely that these sequences are mere artifacts, for several reasons. First, two independent TrxR clones and three independent apolipoprotein E clones carried identical 5' ends, whereas no other sequences in GenBank® (neither genomic, non-redundant nor EST clones) showed any sequence significantly similar to this (as at September 1999). Secondly, all the cDNA clones containing this domain were murine and, most noteworthy, all came from kidney-derived cDNA. Thirdly, with the same 5'-end domain in the cDNA sequence, both apolipoprotein E and TrxR seem to have the same prominent expression in the proximal tubules of the kidney. Apolipoprotein E is expressed in a wide variety of tissues but is expressed in the kidney in the proximal tubules in a pattern highly reminiscent of that found here for TrxR [36]. The same pattern was seen when the human apolipo-

protein E gene was expressed in transgenic mice [37]. On the basis of the studies with transgenic mice expressing variants of the human gene for apolipoprotein E [37,38] or murine cells transfected with variants of the rat gene [39] it was concluded that several regulatory regions, both far downstream as well as in the 5'-proximal flanking region, directed the kidney-specific expression pattern. On the basis of the results presented here, the 5'-end domain of kidney-specific transcripts could well be an indication of a level of tissue-specific regulation whose functional properties should be studied further.

As mentioned, plasma glutathione peroxidase is also prominently expressed in the proximal tubules of the kidney [30]. For this gene no cDNA variant has yet been found containing this murine kidney-derived 5'-end domain. Both apolipoprotein E and plasma glutathione peroxidase are, however, secreted to the plasma from these cells, and cells in culture transfected with the genes for these proteins secrete the product to the medium [30,39]. No evident signal peptide could be identified for TrxR, either within the kidney-specific 5'-end domain or within the other 5' variants listed in Figure 3. Interestingly, it has nevertheless already been shown that (kidney-derived) COS-1 cells transfected with the human TrxR gene (variant B; Figure 3) secrete the protein to the medium in forms with three different molecular masses (55 to approx. 60 kDa); and a non-transfected human lung adenocarcinoma cell line also secretes the endogenous protein but only in the 55 kDa form [34]. On the basis of those results and the results presented here, it could well be proposed that a secretion also occurs of TrxR to the plasma from the proximal tubules of the kidney. This would resemble the situation with apolipoprotein E or plasma glutathione peroxidase, being secreted to plasma from the proximal tubule cells, showing the same expression pattern as that for TrxR1 in the kidney and, most surprisingly, for apolipoprotein E also sharing a 5' UTR domain in kidney-derived mRNA.

In conclusion, we have found a structured expression of TrxR1 in the proximal tubules of the kidney; as discussed above, we suggest that the enzyme might be secreted to the plasma from this location. Moreover, we have found three different 5' UTR variants of the murine TrxR1 mRNA, demonstrating the differential expression of specific 5' UTR domains in TrxR1 mRNA. In all the TrxR1 cDNA analysed, these domains were exclusively tissue-specific, i.e. no tissue or cell-type expressed more than one variant. This most probably reflects a functional regulation at the level of splicing, possibly in combination with alternative promoters and resulting in different transcription start sites.

We thank Eva Lindqvist for help with the *in situ* hybridization. This study was supported by grants from Karolinska institutet, Swedish Cancer Society (project 3775-B96-01XAB), the Swedish Medical Research Council (project 14X-03185) and the USPHS (projects NS09199 and AGO4418) and by a stipend to A.-K. R. from the Foundation for Strategic Research through the Biomedical Graduate School of Stockholm.

REFERENCES

- Holmgren, A. (1989) *J. Biol. Chem.* **264**, 13963–13966
- Laurent, T. C., Moore, E. C. and Reichard, P. (1964) *J. Biol. Chem.* **239**, 3436–3444
- Okamoto, T., Sakurada, S., Yang, J. P. and Merin, J. P. (1997) *Curr. Top. Cell. Regul.* **35**, 149–161
- Yodoi, J. and Tursz, T. (1991) *Adv. Cancer Res.* **57**, 381–411
- Tagaya, Y., Maeda, Y., Mitsui, A., Kondo, N., Matsui, H., Hamuro, J., Brown, N., Arai, K., Yokota, T., Wakasugi, H. et al. (1989) *EMBO J.* **8**, 757–764
- Sahaf, B., Söderberg, A., Spyrou, G., Barral, A. M., Pekkarı, K., Holmgren, A. and Rosén, A. (1997) *Exp. Cell Res.* **236**, 181–192

- 7 Silberstein, D. S., McDonough, S., Minkoff, M. S. and Balcewicz Sablinska, M. K. (1993) *J. Biol. Chem.* **268**, 9138–9142
- 8 Thelander, L. (1967) *J. Biol. Chem.* **242**, 852–859
- 9 Holmgren, A. (1977) *J. Biol. Chem.* **252**, 4600–4606
- 10 Luthman, M. and Holmgren, A. (1982) *Biochemistry* **21**, 6628–6633
- 11 Oblong, J. E., Gasdaska, P. Y., Sherrill, K. and Powis, G. (1993) *Biochemistry* **32**, 7271–7277
- 12 Arnér, E. S. J., Nordberg, J. and Holmgren, A. (1996) *Biochem. Biophys. Res. Commun.* **225**, 268–274
- 13 Kumar, S., Björnstedt, M. and Holmgren, A. (1992) *Eur. J. Biochem.* **207**, 435–439
- 14 Holmgren, A. and Lyckeberg, C. (1980) *Proc. Natl. Acad. Sci. U.S.A.* **77**, 5149–5152
- 15 Björnstedt, M., Hamberg, M., Kumar, S., Xue, J. and Holmgren, A. (1995) *J. Biol. Chem.* **270**, 11761–11764
- 16 Tamura, T. and Stadtman, T. C. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 1006–1011
- 17 Gladyshev, V. N., Jeang, K.-T. and Stadtman, T. C. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 6146–6151
- 18 Zhong, L., Arnér, E. S. J., Ljung, J., Åslund, F. and Holmgren, A. (1998) *J. Biol. Chem.* **273**, 8581–8591
- 19 Lee, S. R., Kim, J. R., Kwon, K. S., Yoon, H. W., Levine, R. L., Ginsburg, A. and Rhee, S. G. (1999) *J. Biol. Chem.* **274**, 4722–4734
- 20 Miranda-Vizuete, A., Damdimopoulos, A. E., Pedrajas, J. R., Gustafsson, J. A. and Spyrou, G. (1999) *Eur. J. Biochem.* **261**, 405–412
- 21 Sun, Q. A., Wu, Y., Zappacosta, F., Jeang, K. T., Lee, B. J., Hatfield, D. L. and Gladyshev, V. N. (1999) *J. Biol. Chem.* **274**, 24522–24530
- 22 Gasdaska, P. Y., Gasdaska, J. R., Cochran, S. and Powis, G. (1995) *FEBS Lett.* **373**, 5–9
- 23 Rozell, B., Hansson, H. A., Luthman, M. and Holmgren, A. (1985) *Eur. J. Cell Biol.* **38**, 79–86
- 24 Wetmore, C. (1992) Thesis, Karolinska Institutet, Stockholm, Sweden
- 25 Nosrat, C. A. and Olson, L. (1995) *J. Comp. Neurol.* **360**, 698–704
- 26 Nosrat, C. A., Tomac, A., Lindqvist, E., Lindskog, S., Humpel, C., Stromberg, I., Ebdal, T., Hoffer, B. J. and Olson, L. (1996) *Cell. Tiss. Res.* **286**, 191–207
- 27 Nosrat, C. A., Tomac, A., Hoffer, B. J. and Olson, L. (1997) *Exp. Brain Res.* **115**, 410–422
- 28 Dagerlind, Å., Friberg, K., Bean, A. J. and Hökfelt, T. (1992) *Histochemistry* **98**, 39–49
- 29 Smith, R. F., Wiese, B. A., Wojzynski, M. K., Davison, D. B. and Worley, K. C. (1996) *Genome Res.* **6**, 454–462
- 30 Maser, R. L., Magenheimer, B. S. and Calvet, J. P. (1994) *J. Biol. Chem.* **269**, 27066–27073
- 31 Björnstedt, M., Xue, J., Huang, W., Åkesson, B. and Holmgren, A. (1994) *J. Biol. Chem.* **269**, 29382–29384
- 32 Fujiwara, N., Fujii, T., Fujii, J. and Taniguchi, N. (1999) *Biochem. J.* **340**, 439–444
- 33 Gasdaska, J. R., Gasdaska, P. Y., Gallegos, A. and Powis, G. (1996) *Genomics* **37**, 257–259
- 34 Koishi, R., Kawashima, I., Yoshimura, C., Sugawara, M. and Serizawa, N. (1997) *J. Biol. Chem.* **272**, 2570–2577
- 35 Claverie, J. M. and Makalowski, W. (1994) *Nature (London)* **371**, 752
- 36 Lin, C. T., Xu, Y. F., Wu, J. Y. and Chan, L. (1986) *J. Clin. Invest.* **78**, 947–958
- 37 Simonet, W. S., Bucay, N., Lauer, S. J. and Taylor, J. M. (1993) *J. Biol. Chem.* **268**, 8221–8229
- 38 Simonet, W. S., Bucay, N., Lauer, S. J., Wirak, D. O., Stevens, M. E., Weisgraber, K. H., Pitas, R. E. and Taylor, J. M. (1990) *J. Biol. Chem.* **265**, 10809–10812
- 39 Fung, W. P., Howlett, G. J. and Schreiber, G. (1986) *J. Biol. Chem.* **261**, 13777–13783
- 40 Low, S. C. and Berry, M. J. (1996) *Trends Biochem. Sci.* **21**, 203–208
- 41 Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W. and Lipman, D. J. (1997) *Nucleic Acids Res.* **25**, 3389–3402
- 42 Terashima, H., Gotoh, S., Yagi, K. and Mizoguchi, T. (1999) *DNA Sequence* **10**, 203–205

Received 24 November 1999/25 January 2000; accepted 16 February 2000