Biochem. J. (1992) 285, 985–992 (Printed in Great Britain)

Mouse UDP-GlcNAc: dolichyl-phosphate *N*-acetylglucosaminephosphotransferase

Molecular cloning of the cDNA, generation of anti-peptide antibodies and chromosomal localization

Bhanu RAJPUT,* Jie MA,* Nagaraja MUNIAPPA,* Laura SCHANTZ,† Susan L. NAYLOR,† Peter A. LALLEY‡ and Inder K. VIJAY*§

* Department of Animal Sciences and the Center for Agricultural Biotechnology, University of Maryland, College Park, MD 20742, U.S.A., † Department of Cellular and Structural Biology, University of Texas Health Science Center, San Antonio, TX 78284-7762, U.S.A., and ‡ Institute for Medical Research, Bennington, VT 05201, U.S.A.

A cDNA encoding UDP-GlcNAc–dolichyl-phosphate *N*-acetylglucosaminephosphotransferase (GPT; EC 2.7.8.15), an enzyme that catalyses the first step in the synthesis of dolichol-linked oligosaccharides, was isolated from mRNA prepared from mouse mammary glands. The cDNA contains an open reading frame that codes for a protein of 410 amino acids with a predicted molecular mass of 46.472 kDa. Mouse GPT has two copies of a putative dolichol-recognition sequence that has so far been identified in all eukaryotic enzymes which interact with dolichol, and four consensus sites for asparagine-linked glycosylation. It shows a high degree of conservation with yeast and hamster GPTs at the amino acid level. The mouse GPT cDNA recognized a single mRNA species of about 2 kb in mouse mammary glands when used as a probe in Northern blot analysis. An antiserum raised against a 15-residue peptide, derived from the predicted amino acid sequence of the cloned mouse cDNA, specifically precipitated the activity of GPT from solubilized mouse mammary gland microsomes, and detected a protein of about 48 kDa on Western blot. This size is in good agreement with that predicted from the cDNA sequence, and also with that (46 and 50 kDa) of purified bovine GPT. With the use of a panel of mouse/hamster somatic-cell hybrids and a specific probe derived from the 3′-non-coding region of the mouse cDNA, the GPT gene was mapped to mouse chromosome 17.

INTRODUCTION

The biosynthesis of asparagine (N)-linked glycoproteins in eukaryotic cells begins with the assembly of the dolichol(Dol)-linked oligosaccharide precursor, $Glc_3Man_9GlcNAc_2$-*PP*-Dol, in the endoplasmic reticulum (ER) (Kornfeld & Kornfeld, 1985). The oligosaccharide moiety is then transferred *en bloc* to the asparagine residue in the sequence motif Asn-Xaa-Ser/Thr (where Xaa is any amino acid except proline) of the nascent polypeptide by an oligosaccharyltransferase in the lumen of the ER (Kornfeld & Kornfeld, 1985). The synthesis of the oligosaccharide precursor occurs in a highly ordered sequence of elongation reactions involving as many as 16 different glycosyltransferases (Kornfeld & Kornfeld, 1985). The enzyme, UDP-GlcNAc–dolichyl-phosphate *N*-acetylglucosaminephosphotransferase (GPT; EC 2.7.8.15), catalyses the first step in this multistep process, i.e. the transfer of GlcNAc-1-P from UDP-GlcNAc to Dol-*P* to form GlcNAc-*PP*-Dol, and hence is a potential target for the regulation of oligosaccharide assembly and therefore of *N*-linked glycosylation.

The GPT enzyme has been purified to apparent homogeneity from bovine mammary gland in our laboratory (Shailubhai *et al.*, 1988). The purified preparation comprised two proteins of 46 and 50 kDa as determined by SDS/PAGE. Antibodies raised against either protein band inhibited GPT activity in solubilized bovine microsomes, and recognized the same-sized bands on Western blot containing purified protein, and an additional polypeptide of about 70 kDa with solubilized bovine microsomes. Recent studies on the topography of the early reactions of the oligosaccharide-biosynthetic pathway suggested that the active site of GPT faced the cytoplasmic side of the ER membrane (Abeijon & Hirschberg, 1990; Kean, 1991).

The gene for GPT has been cloned from yeast (Rine *et al.*, 1983); it encodes a protein of 448 amino acids (Hartog & Bishop, 1987) and is essential for cell growth (Kukuruzinska & Robbins, 1987). By using a fragment of the yeast GPT gene as probe, a segment of genomic DNA was isolated from a tunicamycin-resistant Chinese-hamster ovary (CHO) cell line, that appeared to have a 40–50-fold amplification of the GPT gene (Lehrman *et al.*, 1988). The cloned hamster DNA revealed a 24-amino acid stretch that was 92% conserved with the corresponding yeast sequence.

Despite the recent progress in elucidating the potential structure and topology of GPT, very little is known about the regulation of this enzyme. An early study from our laboratory had shown that three glycosyltransferases, namely GPT, GDP-Man–Dol-*P* mannosyltransferase and UDP-Glc–Dol-P glucosyltransferase, underwent differentiation-related activation during ontogeny of the mouse mammary gland, and that the lactogenic hormone, prolactin, may be involved in this regulation (Vijay & Oka, 1986). The same three enzymes were also found to be developmentally regulated in sea-urchin (Welply *et al.*, 1985), and mouse (Armant *et al.*, 1986) embryos. Oestrogen-induced differentiation of immature chick oviduct also caused the elevation of a number of glycosyltransferases involved in dolichol-linked oligosaccharide biosynthesis (Hayes & Lucas, 1983; Starr & Lucas, 1988). Our long-term goal is to study developmental and hormonal regulation of GPT in the mouse mammary gland

Abbreviations used: GPT, UDP-GlcNAc–dolichyl-phosphate *N*-acetylglucosaminephosphotransferase; ER, endoplasmic reticulum; Dol, dolichol; CHO, Chinese-hamster ovary; Caps, 3-cyclohexylamino-1-propanesulphonic acid; PBS, phosphate-buffered saline (20 mM-sodium/potassium phosphate/0.15 M-NaCl, pH 7.4); poly(A)$^+$, polyadenylated.

§ To whom correspondence and reprint requests should be addressed.

The nucleotide sequence data reported will appear in the EMBL Nucleotide Sequence Database under the accession no. X65603.

at the molecular level. A cDNA encoding mouse GPT, and anti-mouse GPT antibodies, would greatly facilitate the molecular studies at the level of RNA and protein respectively. The present paper describes the isolation of a cDNA encoding mouse GPT, preparation of anti-peptide antibodies against mouse GPT, and the localization of GPT gene on mouse chromosome. While this work was in progress, two reports on the cloning of GPT cDNA from a tunicamycin-resistant CHO cell line were published (Zhu & Lehrman, 1990; Scocca & Krag, 1990).

MATERIALS AND METHODS

Materials

Mid-lactating mouse mammary glands were purchased from Hilltop Laboratory Animals (Scottdale, PA, U.S.A.). mRNA purification kit and CNBr-activated Sepharose 4B were obtained from Pharmacia. Random-primed DNA-labelling kit was bought from Boehringer-Mannheim. [α-^{32}P]dCTP, [α-[^{35}S]thiol]dATP, ^{125}I-labelled donkey anti-rabbit antibodies, and Rainbow molecular-mass marker proteins were supplied by Amersham Corp. UDP-*N*-acetyl[^{3}H]glucosamine was from du Pont–New England Nuclear. Restriction endonucleases, T_4 DNA ligase and kinase, bacterial alkaline phosphatase, vectors pUC18 and M13mp18, DNA and RNA standards were purchased from Bethesda Research Laboratories. *Hin*dIII linker, λgt11 forward and reverse primers and anti-rabbit IgG–peroxidase conjugate were obtained from Promega Biotech (Madison, WI, U.S.A.). DNA-amplification reagent and Sequenase DNA-sequencing kits were bought from Perkin–Elmer/Cetus (Branchburg, NJ, U.S.A.) and United States Biochemical (Cleveland, OH, U.S.A.) respectively. Ampicillin, isopropyl β-D-thiogalactopyranoside, 5-bromo-4-chloro-3-indolyl β-D-galactopyranoside, 2,2-azinodi-3-ethylbenzothiazolinsulphonate and 3-cyclohexylamino-1-propanesulphonic acid (Caps) were supplied by Sigma. Zetaprobe and nitrocellulose membranes and protein A–agarose were from Bio-Rad. Microtitre plates (Immulon) were the product of Dynatech Laboratories (Chantilly, VA, U.S.A.) and Centriflo membrane cones that of Amicon.

Preparation and screening of cDNA library

Total RNA was isolated from mid-lactating mouse mammary glands by the method of Chomczynski & Sacchi (1987). Polyadenylated [Poly(A)$^+$]RNA was prepared on oligo(dT)–cellulose spin columns supplied in the mRNA-purification kit (Pharmacia) and sent to Clontech (Palo Alto, CA, U.S.A.) for the synthesis of a cDNA library in λgt11 (catalogue no. ML10386). The amplified library contained 2×10^6 recombinant phages.

Initial screening of the library was performed using a partial cDNA clone encoding the putative GPT in rat liver (kindly provided by M. Lehrman, University of Texas Southwestern Medical Center, Dallas, TX, U.S.A.). An *Eco*RI insert containing the rat liver GPT cDNA was labelled with [α-^{32}P]dCTP by the random-priming method (Feinberg & Vogelstein, 1983) and was used to screen the library by the '*in situ*' plaque-hybridization technique of Benton & Davis (1977) as described (Degen *et al.*, 1986). Subsequent screenings were performed with probes derived from the 5′ end of the clone, GPT 1.4 (200 bp *Eco*RI–*Bam*HI fragment) and the 3′ end of the clone, GPT 1.5 (350 bp *Eco*RI fragment) (Fig. 1). The final 5′ end of the mouse cDNA was obtained by amplification of the cDNA library with a 5′ primer (nucleotides 89–108), derived from the 5′ non-coding region of the published sequence of hamster GPT cDNA (Zhu & Lehrman, 1990), and a 3′ primer (reverse complement of nucleotides 194–213), derived from the 5′ end of the clone, GPT 1.6 (Fig. 2) using the DNA thermal cycler and the DNA-amplification reagent kit from Perkin–Elmer/Cetus.

Subcloning

Positive clones were plaque purified through three rounds of plating and screening, and DNA was prepared from purified phage by the plate lysis method (Steffens & Gross, 1989), and digested with *Eco*RI to determine the insert size. All of the positive clones (GPT 1.0, 1.5, 1.4 and 1.6) (Fig. 1) gave rise to two insert fragments, suggesting the presence of a naturally occurring internal *Eco*RI site. The full length of each insert was determined by PCR using λgt11 forward and reverse primers and the respective purified phage plaque. The *Eco*RI fragments were subcloned into vectors, pUC18 and M13mp18, for further analysis. The PCR-amplified products were ligated into the *Hin*dIII sites of the above vectors after the addition of *Hin*dIII linkers. A cDNA containing the entire coding region was obtained by ligating GPT 0.2 and GPT 1.6 at the unique *Bgl*I site (Fig. 1) and will be referred to as mGPT.

DNA sequence analysis

DNA was sequenced by the dideoxy chain-termination method of Sanger *et al.* (1977) using Sequenase and universal and synthetic oligonucleotide primers. All the cDNA fragments shown in Fig. 1, and the full-length cDNA (mGPT) were sequenced in their entirety on both strands. DNA sequence analyses were performed with PCGENE software from Intelligenetics (Mountain View, CA, U.S.A.), and the Genetics Computer Group software package from University of Wisconsin, Madison, WI, U.S.A.

Northern-blot analysis

Poly(A)$^+$ RNA from mid-lactating mouse mammary glands was fractionated on a 1% formaldehyde–agarose gel, transferred to Zetaprobe membrane (Bio-Rad), and hybridized to a radiolabelled mGPT probe as described (Botteri *et al.*, 1990).

Preparation of solubilized enzyme and GPT assay

Preparation of microsomes, solubilization and stabilization of enzyme from mid-lactating mouse mammary glands and the assay for GPT activity were performed essentially as described for bovine mammary tissue (Shailubhai *et al.*, 1988), except that the final concentration of Nonidet P40 used for solubilization of enzyme from microsomal membranes was 0.4%.

Generation of anti-peptide antibodies

To raise antiserum against mouse GPT, a peptide corresponding to amino acids 303–317 of the deduced amino acid sequence (Fig. 2) was synthesized (Peninsula Laboratories, Belmont, CA, U.S.A.). Peptide (1 mg) was dissolved in 1 ml of immunogenic enhancement buffer (L. Tamarkin, unpublished work), mixed with 1 ml of complete Freund's adjuvant and injected subcutaneously into a New Zealand White rabbit. Multiple secondary injections of the above mixture in 1 ml of incomplete Freund's adjuvant were given at 2–3 week intervals and the blood was collected 10–14 days after the final booster. The rabbit antibodies against the peptide were purified by passing the crude antiserum through a column packed with Sepharose 4B covalently coupled to the peptide. Specifically bound antibodies were eluted with 50 mM-glycine, pH 2.0, and immediately neutralized with 1 M-Tris, pH 8.0. The eluates were concentrated by ultrafiltration with Centriflo membrane cones and dialysed against 20 mM-sodium/potassium phosphate buffer/0.15 M-NaCl, pH 7.4 (PBS) at 4 °C for 36 h.

Immunoadsorption of GPT activity

Various amounts of either preimmune or anti-peptide antiserum, made up to a total volume of 100 μl with PBS containing 1 mg

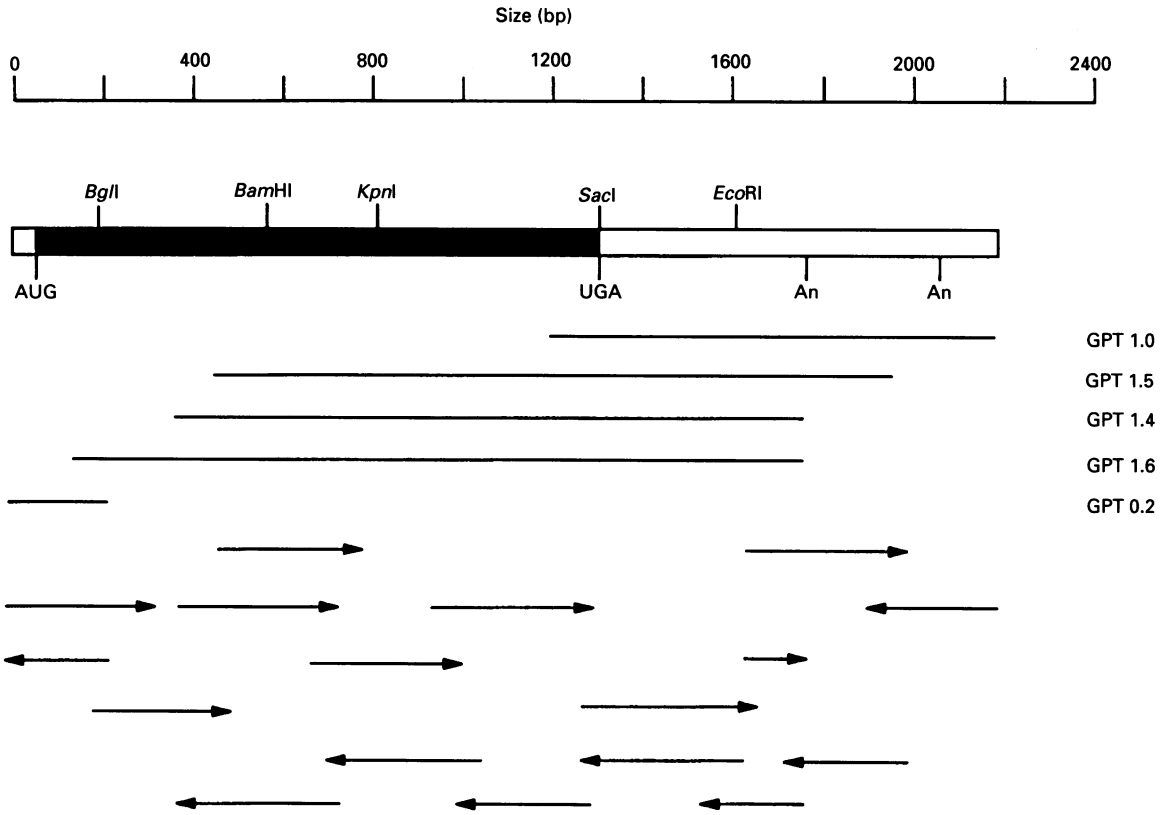

Fig. 1. Schematic representation of GPT cDNA clones and sequencing strategy

The isolation of the overlapping clones, GPT 1.0, 1.5, 1.4, 1.6 and 0.2, is detailed in the text. The scale in bp, cleavage sites of several restriction endonucleases, the open reading frame (solid bar), the translation initiation (AUG) and termination (UGA), as well as two polyadenylation (An) sites are shown. The arrows indicate the direction and the extent of individual sequence runs.

of BSA/ml (PBS/BSA), were mixed with 200 μl of protein A–Sepharose (equilibrated 1:1 with PBS/BSA) by end-over-end rotation at 4 °C overnight. The antibody–protein A–Sepharose complexes were recovered by centrifugation and incubated with a constant amount of solubilized enzyme from mouse mammary-gland microsomes for 1 h at 4 °C. GPT activity remaining in the supernatant was determined as described above, after removal of antigen–antibody complexes by centrifugation.

Western blot analysis

Portions of solubilized mouse mammary-gland microsomes (approx. 75 μg of total protein) or purified bovine GPT (Shailubhai *et al.*, 1988) were electrophoresed on a 10% (w/v) polyacrylamide gel under reducing conditions (Laemmli, 1970), and the proteins were electrophoretically blotted on to nitrocellulose in 10 mM-Caps/10% (v/v) methanol, pH 11.0, for 45 min at 0.5 A (Matsudaira, 1987). After electrotransfer, the blots were treated essentially as described (Pain *et al.*, 1988), except that the primary antibodies employed were affinity-purified preimmune serum (1:100), affinity-purified anti-peptide antibodies (1:100) or anti-(bovine GPT) antibodies (1:1000) (Shailubhai *et al.*, 1988), and the secondary reagent used was ^{125}I-labelled anti-rabbit antibodies.

Chromosomal mapping

The construction and characterization of the mouse/hamster hybrid cell lines used in this study have been described [Rajput *et al.* (1987) and references therein]. The mouse chromosome content of the hybrids was determined by karyotyping and analysis of enzyme markers of known chromosomal locations, at the same cell passage as that used for DNA isolation. The DNAs from mouse and hamster parental cell lines as well as from 14 hybrid cell lines were subjected to PCR analysis using a primer pair derived from the 3′-non-coding region of mGPT cDNA (5′ primer, nucleotides 1421–1440; 3′ primer, reverse complement of nucleotides 1595–1614) (Fig. 2) as described (Theune *et al.*, 1991).

RESULTS AND DISCUSSION

Isolation and characterization of GPT cDNA from mouse mammary glands

Mid-lactating mouse mammary glands were considered a good source for the isolation of mRNA for the construction of a cDNA library, since an earlier study from our laboratory had shown that GPT activity was highest at this stage of development of the mouse mammary gland (Vijay & Oka, 1986). A probe prepared from a partial rat liver GPT cDNA (see the Materials and methods section) hybridized to a unique mRNA species from mouse mammary glands under stringent hybridization conditions (results not shown), and thus was used for preliminary screening of the cDNA library. Seven positive plaques were found on screening 1.3×10^6 phage; two clones with the longest inserts, 1.5 kb (GPT 1.5) and 1.4 kb (GPT 1.4) are shown in Fig. 1. Sequencing of these two DNA inserts indicated a single open reading frame with a termination codon but without a functional initiator methionine. Hence the library was rescreened with probes derived from the 5′ end of the clone, GPT 1.4, and 3′ end of the clone, GPT 1.5. The secondary screen yielded two new

```
1       TGTGCCGGGTTGCTAGTTAGCTGAGTAGGCGGCGGGGCGGCGGCCACCGGAGGGTCACC   59

60      ATGTGGGCCTTCCCGGAGTTGCCCCTGCCGCTGCCGCTGCTGGTGAATTTGATCGGCTCG  119
1       MetTrpAlaPheProGluLeuProLeuProLeuProLeuLeuValAsnLeuIleGlySer   20

120     CTGTTGGGATTCGTGGCTACAGTCACCCTCATCCCTGCCTTCCGTAGCCACTTTATCGCC  179
21      LeuLeuGlyPheValAlaThrValThrLeuIleProAlaPheArgSerHisPheIleAla   40

180     GCGCGCCTCTGTGGCCAGGACCTCAACAAGCTCAGCCAGCAGCAGATCCCAGAGTCCCAG  239
41      AlaArgLeuCysGlyGlnAspLeuAsnLysLeuSerGlnGlnGlnIleProGluSerGln   60

240     GGAGTGATCAGCGGTGCTGTTTTCCTTATCATCCTCTTCTACTTCATCCCTTTCCCCTTC  299
61      GlyValIleSerGlyAlaValPheLeuIleIleLeuPheTyrPheIleProPheProPhe   80

300     CTGAACTGCTTCGTGGAGGAGCAGTGTAAGGCATTCCCCCACCATGAATTTGTGGCCCTA  359
81      LeuAsnCysPheValGluGluGlnCysLysAlaPheProHisHisGluPheValAlaLeu  100

360     ATAGGTGCCCTCCTTGCCATCTGCTGCATGATCTTCCTGGGGTTTGCTGATGATGTCCTC  419
101     IleGlyAlaLeuLeuAlaIleCysCysMetIlePheLeuGlyPheAlaAspAspValLeu  120

420     AATCTCCGCTGGCGCCACAAGCTGCTGCTGCCCACAGCTGCCTCACTACCTCTCCTCATG  479
121     AsnLeuArgTrpArgHisLysLeuLeuLeuProThrAlaAlaSerLeuProLeuLeuMet  140

480     GTCTACTTCACAAACTTTGGCAATACAACCATCGTGGTGCCCAAGCCCTTCCGCTGGATT  539
141     ValTyrPheThrAsnPheGlyAsnThrThrIleValValProLysProPheArgTrpIle  160

540     CTGGGCCTGCATTTGGACTTGGGGATCCTGTACTACGTCTACATGGGGCTGCTTGCAGTG  599
161     LeuGlyLeuHisLeuAspLeuGlyIleLeuTyrTyrValTyrMetGlyLeuLeuAlaVal  180

600     TTCTGTACCAATGCCATCAACATCCTGGCGGGCATTAATGGCCTAGAGGCCGGTCAGTCA  659
181     PheCysThrAsnAlaIleAsnIleLeuAlaGlyIleAsnGlyLeuGluAlaGlyGlnSer  200

660     CTAGTCATCTCTGCTTCTATCATTGTCTTCAACCTGGTGGAACTGGAAGGTGATTATCGA  719
201     LeuValIleSerAlaSerIleIleValPheAsnLeuValGluLeuGluGlyAspTyrArg  220

720     GATGATCATATCTTTTCCCTTTACTTCATGATACCATTTTTCTTTACCACCTTGGGACTG  779
221     AspAspHisIlePheSerLeuTyrPheMetIleProPhePhePheThrThrLeuGlyLeu  240

780     CTTTACCACAACTGGTACCCGTCCCGCGTGTTTGTGGGAGACACCTTCTGTTACTTTGCG  839
241     LeuTyrHisAsnTrpTyrProSerArgValPheValGlyAspThrPheCysTyrPheAla  260

840     GGCATGACTTTTGCCGTGGTGGGGATCTTGGGACACTTCAGCAAGACCATGCTGCTCTTC  899
261     GlyMetThrPheAlaValValGlyIleLeuGlyHisPheSerLysThrMetLeuLeuPhe  280

900     TTTATGCCACAAGTATTCAATTTCCTCTACTCACTGCCTCAGCTCTTCCATATCATCCCC  959
281     PheMetProGlnValPheAsnPheLeuTyrSerLeuProGlnLeuPheHisIleIlePro  300

960     TGCCCTCGACACCGGATGCCCAGACTCAACGCAAAGACAGGCAAACTGGAAATGAGCTAT 1019
301     CysProArgHisArgMetProArgLeuAsnAlaLysThrGlyLysLeuGluMetSerTyr  320

1020    TCCAAGTTCAAGACCAAGAACCTCTCTTTCCTGGGCACCTTTATTTTAAAGGTAGCAGAG 1079
321     SerLysPheLysThrLysAsnLeuSerPheLeuGlyThrPheIleLeuLysValAlaGlu  340

1080    AACCTCCGGTTAGTGACAGTTCACCAAGGTGAGAGTGAGGACGGTGCCTTCACTGAGTGT 1139
341     AsnLeuArgLeuValThrValHisGlnGlyGluSerGluAspGlyAlaPheThrGluCys  360

1140    AACAACATGACCCTCATCAACTTGCTACTTAAAGTCTTTGGGCCTATACATGAGAGAAAC 1199
361     AsnAsnMetThrLeuIleAsnLeuLeuLeuLysValPheGlyProIleHisGluArgAsn  380

1200    CTCACCCTGCTCCTGCTGCTCCTGCAGGTCCTAAGCAGCGCCGCCACCTTCTCCATTCGT 1259
381     LeuThrLeuLeuLeuLeuLeuLeuGlnValLeuSerSerAlaAlaThrPheSerIleArg  400

1260    TACCAACTCGTCCGACTCTTCTATGATGTCTGAGCTCCCTGACAGCTGCCCTTTACCTCA 1319
401     TyrGlnLeuValArgLeuPheTyrAspValEnd                              410

1320    CAGTCTCCATTGGACCTCAGCCAGGACCAGCCTCTGTCTGGTCCGAGATGACCCTCTGGT 1379

1380    CCAGGCCTCGCTGACACTTTTGTTCTCAGCTTCTGCCATCTGTGACTACTGATATCCTGG 1439

1440    ATGGACACCTTGCTGGACTTGAAGTCCGCTAGTTGGACTTTGCCTAGGGCTTTCATCTTG 1499

1500    CCTTGCCCTCCCTTTCTGTCCCATCTGCAGCCTCACCAGGTGGGCTTGTAGCCTCTATTA 1559

1560    TGCAAATATTCGTAGCTCAGCTTTCAGAGCGCTAACTCTAAAGGAATTCACCTGAGCCTT 1619

1620    GAGAGAGAACCTGGGCTAGGGCTAGAGTTAGGGCTACATACTCCAAGGTGACCTCACATT 1679

1680    TGACTATCAAATGAAGTGTTGTGATTGGGAAGCGTAGAGGCAGGGCATGTGCTCAGAACG 1739

1740    GTGACAATAAAGGACTGCCTTTTAC                                      1764
```

Fig. 2. Nucleotide sequence and translation of mouse GPT cDNA

The nucleotide sequence shown here was derived from the full-length cDNA, mGPT (described in the Materials and methods section). Its translation is shown below the DNA sequence. The consensus Dol-recognition sites are underlined, the potential *N*-linked glycosylation sites appear in boldface type and the polyadenylation signal is shown in boldface italics.

positive clones, GPT 1.6 and GPT 1.0 (Fig. 1). DNA sequence analysis of the latter showed that it coded for more sequence in the 3′ non-coding region including an additional polyadenylation site, whereas that of the former contained about 250 bp more of the coding sequence at the 5′ end, but still lacked a functional initiator methionine. While this work was in progress, the sequence of hamster GPT cDNA was published (Zhu & Lehrman, 1990), and a comparison of the two cDNA sequences indicated a very high degree of conservation (88 % identity at the nucleotide level) (results not shown). Therefore this information was utilized to devise a primer pair (see the Materials and methods section for detail) to obtain the missing 5′ end of the mouse cDNA by PCR. A DNA fragment of the expected size (approx. 200 bp) was amplified (GPT 0.2) (Fig. 1) and its sequence showed the presence of an initiator methionine (see below).

There is an internal *Eco*RI site in the mouse GPT cDNA (Fig. 1). That the two segments of DNA flanking the *Eco*RI site were truly contiguous, i.e. derived from a single mRNA species, and not from the ligation of two unrelated cDNA fragments, was ascertained by the following observations: (1) the internal *Eco*RI site did not have an artificial *Eco*RI linker sequence, (2) the sequences flanking the *Eco*RI site were the same in all four clones and (3) the two segments flanking the *Eco*RI site hybridized to a similar-sized mRNA on a Northern blot when used individually as probes (results not shown).

The DNA sequence (with its translation) of mouse GPT cDNA (mGPT) is shown in Fig. 2. The methionine codon (nucleotides 60–62) designated as initiator methionine was in a favourable position for translation initiation as determined by Kozak (1987). The sequence revealed a single open reading frame that coded for a protein of 410 amino acids with a predicted molecular mass of 46.472 kDa, which is in good agreement with the size (46 and 50 kDa doublet) shown for purified bovine GPT (Shailubhai *et al.*, 1988). Like the hamster GPT (Zhu & Lehrman, 1990), the mouse sequence also contained two copies of a putative dolichol-recognition sequence (underlined) which has so far been identified in all eukaryotic enzymes that interact with dolichol phosphate or its derivatives (Albright *et al.*, 1989). The first copy (residues 69–81) was identical with that of hamster GPT, except for a tyrosine instead of a cysteine at position 6 (see Fig. 3) of the 13-amino acid consensus. Similarly, the second copy (residues 224–236) differed from that of hamster GPT only at position 1, with an isoleucine instead of valine. Both of these changes were in non-conserved positions according to the refined consensus proposed by Zhu & Lehrman (1990). There are four potential *N*-linked glycosylation sites (shown in bold), the significance of which is not very clear, since neither the bovine (Shailubhai *et al.*, 1988) nor the hamster (Zhu & Lehrman, 1990) GPT appears to bind to concanavalin A–agarose.

GPT, like the other enzymes of the dolichol cycle, is known to be membrane bound, and a resident of the ER (Kornfeld & Kornfeld, 1985; Shailubhai *et al.*, 1988; Kaushal & Elbein, 1985). Hydropathy analyses by the methods of Rao & Argos (1986) and Kyte & Doolittle (1982) showed the protein to be very hydrophobic with a potential for forming several membrane-spanning segments (results not shown), consistent with its tight membrane association. The mouse sequence was also searched for the presence of an *N*-terminal signal sequence (von Heijne, 1985), a *C*-terminal sequence, Lys-Asp-Glu-Leu (KDEL), typical of soluble luminal ER proteins (Munro & Pelham, 1987), and yet another *C*-terminal motif, shown to be required for the retention of type-I transmembrane proteins in the ER (Jackson *et al.*, 1990). The search was negative for all three motifs. Maybe the signal essential for the retention in the ER of transmembrane proteins with topologies other than type I, as appears to be the case with GPT with several transmembrane domains (above; Zhu & Lehrman, 1990; Scocca & Krag, 1990), is different (Jackson *et al.*, 1990).

A search of the SWISS-PROT (release 17.0) database with the deduced amino acid sequence of mGPT showed 85 % identity between residues 250–284 of mGPT and residues 283–317 of yeast ALG7 gene product (Hartog & Bishop, 1987). This region included the 24-amino acid segment shown to be 92 % conserved with a hamster GPT gene fragment (Lehrman *et al.*, 1988). The yeast ALG7 gene product shares an overall 41–43 % identity with the mouse (results not shown) and hamster (Zhu & Lehrman, 1990; Scocca & Krag, 1990) GPTs. A check against the GENBANK (release 67.0)/DMBL (release 26.0) databases revealed significant sequence similarity only to the hamster GPT (Zhu & Lehrman, 1990; Scocca & Krag, 1990). Alignment of the two sequences is shown in Fig. 3. As can be seen, there is a very high degree of conservation (96 % identity) between these two rodent species.

The various overlapping clones shown in Fig. 1 added up to a total of about 2.2 kb. In addition to the 1230 bp of the coding sequence, there are 59 bp of 5′-non-coding and either 475 or 775 bp of 3′-non-coding sequences (Figs. 1 and 2). Northern blot analysis using poly(A)$^+$ RNA from mid-lactating mouse mammary glands and mGPT probe showed a single hybridizing band of about 2 kb (Fig. 4). This suggested that the first polyadenylation site was probably the major one, since utilization of the second polyadenylation site would produce a much larger transcript. It is interesting to note that only a single major GPT mRNA species was observed in mouse mammary glands, whereas a family of RNAs coding for GPT have been found in wild-type and tunicamycin-resistant CHO (1.5, 1.8, 2.0 and 2.2 kb) (Lehrman *et al.*, 1988; Scocca & Krag, 1990) and yeast (1.4 and 1.6 kb) (Kukuruzinska & Robbins, 1987) cells. In both instances, the size heterogeneity could be attributed to the presence of multiple transcription termination sites (Kukuruzinska & Robbins, 1987; Zhu & Lehrman, 1990). In this regard there appears to be a tighter regulation of the usage of the two polyadenylation sites in mouse mammary glands.

Immunological characterization of mouse GPT

In order to show that the mouse cDNA isolated here indeed coded for the catalytic moiety of GPT, antibodies were raised against a segment of the deduced amino acid sequence. The selection of the peptide (residues 303–317, Fig. 2), present in the large cytoplasmic loop between the ninth and the tenth membrane-spanning segments of the model of GPT structure proposed by Zhu & Lehrman (1990), was based on the studies by Abeijon & Hirschberg (1990) and Kean (1991), which suggested that the active site of GPT faced the cytoplasmic side of the ER. Incubation of increasing amounts of anti-peptide antiserum with a constant amount of solubilized enzyme from mouse mammary-gland microsomes progressively precipitated out the GPT activity (Fig. 5). Control experiments with preimmune serum did not show significant removal of enzyme activity. Although the above data clearly showed that the anti-peptide antibodies depleted solubilized mouse microsomes of GPT activity, it was not certain whether this was due to binding of the antibodies to GPT *per se* or due to an indirect effect on GPT activity by complexing with some modifier protein. To address the question of binding more directly, affinity-purified preimmune and immune sera were used to probe Western blots (Fig. 6). The preimmune control failed to bind to any protein in solubilized mouse microsomes (lane 1), whereas the affinity-purified anti-peptide antibodies recognized a protein of about 48 kDa (lane 2). This size was in good agreement with that predicted from the cDNA sequence (46.472 kDa), and also with that of the two polypeptides (46 and 50 kDa) of purified and immunoreactive (lane 3) bovine GPT (Shailubhai *et*

```
M-  MWAFPELPLPLPLLVNLIGSLLGFVATVTLIPAFRSHFIAARLCGQDLNKLSQQQ -55
    |||||||||| ||||| ||||||||||||||||||||||||||||||||| ||
H-  MWAFPELPLP--LLVNLFGSLLGFVATVTLIPAFRSHFIAARLCGQDLNKLSRQQ -53

M-  IPESQGVISGAVFLIILFYFIPFPFLNCFVEEQCKAFPHHEFVALIGALLAICCM -110
    |||||||| ||||||||| ||||||||||||||||||||||||||||||||||||
H-  IPESQGVICGAVFLIILFCFIPFPFLNCFVEEQCKAFPHHEFVALIGALLAICCM -108

M-  IFLGFADDVLNLRWRHKLLLPTAASLPLLMVYFTNFGNTTIVVPKPFRWILGLHL -165
    |||||||||||| ||||||||||||||||||||||||||||||||||||||||||
H-  IFLGFADDVLNLPWRHKLLLPTAASLPLLMVYFTNFGNTTIVVPKPFRWILGLHL -163

M-  DLGILYYVYMGLLAVFCTNAINILAGINGLEAGQSLVISASIIVFNLVELEGDYR -220
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||
H-  DLGILYYVYMGLLAVFCTNAINILAGINGLEAGQSLVISASIIVFNLVELEGDYR -218

M-  DDHIFSLYFMIPFFFTTLGLLYHNWYPSRVFVGDTFCYFAGMTFAVVGILGHFSK -275
    |||.|||||||||||||||||||||| ||||||||||||||||||||||||||||
H-  DDHVFSLYFMIPFFFTTLGLLYHNWYPSQVFVGDTFCYFAGMTFAVVGILGHFSK -273

M-  TMLLFFMPQVFNFLYSLPQLFHIIPCPRHRMPRLNAKTGKLEMSYSKFKTKNLSF -330
    ||||||.|||||||||||||| | |||||||.|||| ||||||||||||||||||
H-  TMLLFFIPQVFNFLYSLPQLLHAIPCPRHRIPRLNPKTGKLEMSYSKFKTKNLSF -328

M-  LGTFILKVAENLRLVTVHQGESEDGAFTECNNMTLINLLLKVFGPIHERNLTLLL -385
    |||||||||| | ||||| ||||||||||||||||||||.|||||||||||||
H-  LGTFILKVAERLQLVTVHRGESEDGAFTECNNMTLINLLLKIFGPIHERNLTLLL -383

M-  LLLQVLSSAATFSIRYQLVRLFYDV -410
    ||||.|||| ||||||||||||||
H-  LLLQILSSAVTFSIRYQLVRLFYDV -408
```

Fig. 3. Comparison of the deduced amino acid sequences of mouse and hamster GPT cDNAs

The deduced amino acid sequences of mouse (M) and hamster (H) were aligned using PCGENE software. The numbering of hamster sequence is according to Zhu & Lehrman (1990). Identical residues are shown by a vertical line and similar residues by a single dot. Dashes indicate gaps introduced to produce maximum alignment. The consensus dolichol-recognition sites are shown in boldface type and are overlined.

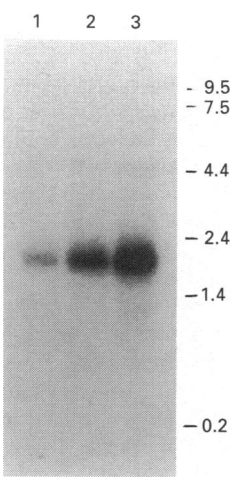

Fig. 4. Northern blot analysis

Poly(A)$^+$ RNA from mid-lactating mouse mammary glands was isolated, fractionated on agarose gel, transferred to Zetaprobe membrane and hybridized to mGPT probe, as described in the Materials and methods section. Lanes 1, 2 and 3 contain 1.0, 2.5 and 5.0 μg of poly(A)$^+$ RNA respectively. The positions of the RNA molecular-mass standards (in kb) are shown.

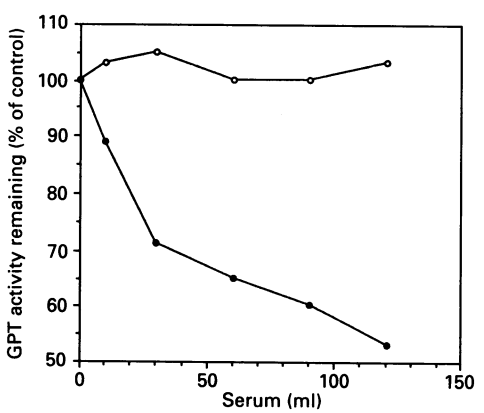

Fig. 5. Precipitation of GPT activity with anti-peptide antibodies

An antibody was raised against a 15-amino-acid peptide corresponding to amino acids 303–317 of mouse GPT. Increasing amounts of either preimmune (○) or immune (●) serum, prebound to protein A–Sepharose, were incubated with a constant amount of solubilized enzyme from mouse mammary-gland microsomes. GPT activity remaining in the supernatant, after removal of antigen–antibody complexes, was measured and compared with a control sample which received no serum (100% point). The data shown are averages for two experiments.

al., 1988). Thus the anti-peptide antibodies did appear to bind specifically to a mouse protein with a molecular mass expected of GPT from available data. Therefore, together with the observation that the mouse and the hamster enzymes were virtually identical (96% identity), the immunoprecipitation and the immunoblotting results provided additional supportive evidence

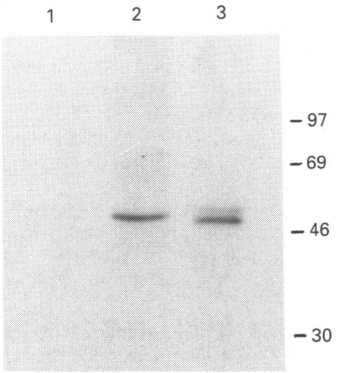

Fig. 6. Western blot analysis

Samples were electrophoresed on an SDS/polyacrylamide gel (10%, w/v, acrylamide, transferred to nitrocellulose, incubated with the indicated primary antibodies, followed by binding to ^{125}I-labelled anti-rabbit IgG. Lanes 1 and 2, solubilized enzyme from mouse mammary-gland microsomes probed with affinity-purified pre-immune serum and affinity-purified anti-peptide antibodies respectively. Lane 3, purified bovine GPT probed with anti-bovine GPT antibodies. The positions of the protein molecular-mass standards (in kDa) are shown.

for the conclusion that the mouse cDNA studied here coded for GPT.

Chromosomal localization of mouse GPT gene

The GPT gene was assigned to mouse chromosome 17 with a panel of 14 mouse/hamster somatic-cell hybrids that retained overlapping subsets of mouse chromosomes. Instead of the use of the standard Southern hybridization technique to analyse the cell lines, PCR was used for rapid analysis. PCR has been used for characterization of somatic-cell hybrids and for gene mapping, and comparison of the results with that obtained by the Southern hybridization method showed that PCR could be used reliably for these purposes (Theune *et al.*, 1991). Using a primer pair designed from the 3′-non-coding region of the mouse cDNA, a specific 194 bp fragment was amplified from the mouse DNA (Fig. 7, lane 8) but not from the hamster DNA (Fig. 7, lane 7).

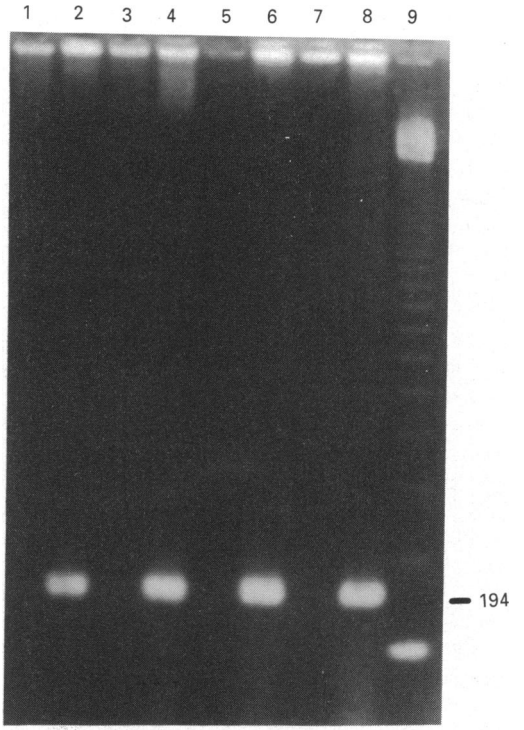

Fig. 7. Chromosomal mapping/PCR analysis of parent and representative somatic-cell hybrid DNAs

A primer pair derived from the 3′-non-coding region of the mouse GPT cDNA was used to amplify DNA from mouse (lane 8) and hamster (lane 7) parent cell lines, as well as DNAs from various mouse/hamster hybrid cell lines. Lanes 1–6, hybrids EBS-15, EBS-17, EBS-18, EBS-51, EBS-63 and EBS-71 respectively. Lane 9 contains a 123 bp DNA ladder. The position of the 194 bp band is shown.

All hybrids (except EBS-15) that contained chromosome 17 showed the presence of the 194 bp mouse-specific band, whereas hybrids that lacked chromosome 17 were negative for this band

Table 1. Segregation of mouse GPT with mouse chromosomes in mouse/hamster somatic-cell hybrids

The presence of mouse GPT gene was determined by scoring for the presence of the 194 bp mouse-specific band, as described in the legend to Fig. 7. The mouse chromosome content of the hybrids was determined as detailed in the Materials and methods section. Female mouse parental cells were used, thus eliminating the Y chromosome. Discordance indicates the presence of the GPT gene but the absence of a specific chromosome, or vice versa.

	Presence of mouse GPT gene																				
Hybrid	GPT	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	X
EBS-1	+	+	+	+	+	−	+	+	+	+	+	−	+	+	+	+	+	+	+	+	+
EBS-2	+	+	+	+	−	−	+	+	+	+	+	−	+	+	−	+	+	+	−	+	+
EBS-3	+	+	+	−	−	+	+	+	+	−	−	−	+	+	−	+	+	+	−	+	+
EBS-4	+	−	−	−	−	−	−	−	−	−	−	−	−	+	−	+	−	+	−	−	+
EBS-5	+	+	+	+	+	+	+	+	+	−	+	−	+	+	+	+	+	+	−	+	+
EBS-9	+	−	+	+	+	−	+	+	+	+	−	−	+	+	−	+	+	+	+	+	+
EBS-10	−	−	+	+	−	−	−	+	−	−	+	−	+	+	+	+	+	−	+	−	+
EBS-11	−	+	−	−	−	−	−	+	−	−	+	−	+	−	−	+	+	−	−	−	+
EBS-15	−	−	+	+	+	−	+	+	+	−	+	−	+	+	−	+	−	+	−	+	+
EBS-17	+	+	+	−	−	+	−	+	−	+	−	−	−	+	−	+	−	+	−	−	+
EBS-18	−	−	+	−	−	−	−	−	−	−	−	−	−	−	−	+	−	−	−	−	+
EBS-51	+	−	+	+	−	+	+	+	+	+	−	−	+	−	−	+	+	+	+	+	+
EBS-63	−	+	−	−	−	−	−	+	−	−	−	−	−	−	−	−	−	−	−	−	+
EBS-71	+	+	+	+	−	+	−	+	−	+	−	−	+	−	−	+	+	+	+	−	+
Discordance (%)...		36	29	36	50	29	29	36	29	21	64	64	36	29	57	29	29	7	43	29	36

(some representative hybrids are shown in lanes 1–6 of Fig. 7). The EBS-15 hybrid was positive for chromosome 17 markers, C3 and Glo I (Lalley & McKusick, 1985), but since both are present on the same region of chromosome 17, other breaks cannot be ruled out. With this single exception, chromosome 17 showed a discordance of 7%, while all other chromosomes had 21% or greater discordance for GPT (Table 1).

The isolation of unique mouse (work described here) and hamster (Zhu & Lehrman, 1990; Scocca & Krag, 1990) GPT cDNAs, and the simple hybridization pattern seen on Southern blots of differently restricted genomic DNA from mouse (B. Rajput, unpublished work) and wild-type CHO cells (Scocca & Krag, 1990), suggests that there is probably a single gene coding for GPT in mouse and hamster, and that it resides on chromosome 17 in mouse. This is the first report of a gene of the Dol cycle to be mapped. It would be of interest to see if other genes of this biosynthetic pathway map to the same chromosome, since a concerted and co-ordinated action of up to 16 glycosyltransferases may be involved in the assembly of the Dol-linked oligosaccharide precursor (Kornfeld & Kornfeld, 1985).

In summary, we have isolated a full-length cDNA coding for mouse GPT, and generated anti-mouse GPT-specific antibodies. The availability of these reagents will greatly facilitate future studies on the developmental and hormonal regulation of GPT gene expression in mouse mammary glands.

We thank Dr. Florence Botteri for her help in the preparation of RNA from mid-lactating mouse mammary glands, Dr. Mark Lehrman for supplying the partial rat liver GPT cDNA clone, Dr. Kunwar Shailubhai for providing anti-bovine GPT antibodies, Dr. Bob Leclerc for help with the use of the Wisconsin Package, Dr. Tamarkin, Jiulio Paciotti and Lisa Banghart for help in the generation of anti-peptide antibodies, Dongming Fan for preparing Fig. 1, Dr. Ian Mather for the critical reading of this manuscript and Margaret Kempf for her secretarial assistance. This work was supported by National Institutes of Health Grant, DK-19682, and a Distinguished Faculty Research Fellowship (to I. K. V.) from the Graduate School, University of Maryland. This paper is Scientific Article A6192, Contribution 8361, of the Maryland Agricultural Experiment Station.

REFERENCES

Abeijon, C. & Hirschberg, C. B. (1990) J. Biol. Chem. **265**, 14691–14695
Albright, C., Orlean, P. & Robbins, P. (1989) Proc. Natl. Acad. Sci. U.S.A. **86**, 7366–7369
Armant, D. R., Kaplan, H. A. & Lennarz, W. J. (1986) Dev. Biol. **113**, 228–237
Benton, W. D. & Davis, R. W. (1977) Science **196**, 180–182
Botteri, F. M., Ballmer-Hofer, K., Rajput, B. & Nagamine, Y. (1990) J. Biol. Chem. **265**, 13327–13334
Chomczynski, P. & Sacchi, N. (1987) Anal. Biochem. **162**, 156–159
Degen, S. J. F., Rajput, B. & Reich, E. (1986) J. Biol. Chem. **261**, 6972–6985
Feinberg, A. P. & Vogelstein, B. (1983) Anal. Biochem. **132**, 6–13
Hartog, K. & Bishop, B. (1987) Nucleic Acids Res. **15**, 3627
Hayes, G. R. & Lucas, J. J. (1983) J. Biol. Chem. **258**, 15095–15100
Jackson, M. R., Nilsson, T. & Peterson, P. A. (1990) EMBO J. **9**, 3153–3162
Kaushal, G. & Elbein, A. (1985) J. Biol. Chem. **260**, 16303–16309
Kean, E. L. (1991) J. Biol. Chem. **266**, 942–946
Kornfeld, R. & Kornfeld, S. (1985) Annu. Rev. Biochem. **54**, 631–663
Kozak, M. (1987) Nucleic Acids Res. **15**, 8125–8148
Kukuruzinska, M. A. & Robbins, P. W. (1987) Proc. Natl. Acad. Sci. U.S.A. **84**, 2145–2149
Kyte, J. & Doolittle, R. F. (1982) J. Mol. Biol. **157**, 105–132
Laemmli, U. K. (1970) Nature (London) **227**, 680–685
Lalley, P. A. & McKusick, V. A. (1985) Cytogenet. Cell Genet. **40**, 536–566
Lehrman, M., Zhu, X. & Khounlo, S. (1988) J. Biol. Chem. **263**, 19796–19803
Matsudaira, P. (1987) J. Biol. Chem. **262**, 10035–10038
Munro, S. & Pelham, H. (1987) Cell **48**, 899–907
Pain, D., Kanwar, Y. S. & Blobel, G. (1988) Nature (London) **331**, 232–237
Rajput, B., Marshall, A., Killary, A. M., Lalley, P. A., Naylor, S. L., Belin, D., Rickles, R. J. & Strickland, S. (1987) Som. Cell Mol. Genet. **13**, 581–586
Rao, M. J. K. & Argos, P. (1986) Biochim. Biophys. Acta **869**, 197–214
Rine, J., Hansen, W., Hardeman, E. & Davis, R. (1983) Proc. Natl. Acad. Sci. U.S.A. **80**, 6750–6754
Sanger, F., Nicklen, S. & Coulson, A. (1977) Proc. Natl. Acad. Sci. U.S.A. **74**, 5463–5467
Scocca, J. R. & Krag, S. S. (1990) J. Biol. Chem. **265**, 20621–20626
Shailubhai, K., Dong-Yu, B., Saxena, E. S. & Vijay, I. K. (1988) J. Biol. Chem. **263**, 15964–15972
Starr, C. M. & Lucas, J. J. (1988) Arch. Biochem. Biophys. **237**, 261–270
Steffens, D. L. & Gross, R. W. (1989) Biotechniques **7**, 674–678
Theune, S., Fung, J., Todd, S., Sakaguchi, Y. & Naylor, S. L. (1991) Genomics **9**, 511–516
Vijay, I. K. & Oka, T. (1986) Eur. J. Biochem. **154**, 57–62
von Heijne, G. (1985) J. Mol. Biol. **184**, 99–105
Welply, J. K., Lan, J. T. & Lennarz, W. J. (1985) Dev. Biol. **107**, 252–258
Zhu, X. & Lehrman, M. A. (1990) J. Biol. Chem. **265**, 14250–14255

Received 22 November 1991/22 January 1992; accepted 29 January 1992