Biochem. J. (1988) 250, 547–555 (Printed in Great Britain)

Rat liver β-glucuronidase

cDNA cloning, sequence comparisons and expression of a chimeric protein in COS cells

Penelope P. POWELL, John W. KYLE, Raymond D. MILLER, John PANTANO, Jeffrey H. GRUBB and William S. SLY*

E.A. Doisy Department of Biochemistry, St. Louis University Medical Center, 1402 S. Grand Blvd., St. Louis, MO 63104, U.S.A.

A cDNA for rat liver β-glucuronidase was isolated, its sequence determined and its expression after transfection into COS cells studied. The deduced amino acid sequence of the rat liver clone showed 77% homology with that from the cDNA for human placental β-glucuronidase and 47% homology with that deduced from the cDNA for *Escherichia coli* β-glucuronidase. Several differences were found between the cDNA from rat liver and that previously reported from rat preputial gland. Only one change leads to an amino acid difference in the mature enzyme. A chimeric clone was constructed by using a fragment encoding the first 18 amino acid residues of the signal sequence from the human placental cDNA clone and a fragment from the rat clone encoding four amino acid residues of the signal sequence, all 626 amino acid residues of the mature rat enzyme, and all of the 3′ untranslated region. After transfection into COS cells the chimeric clone expressed β-glucuronidase activity that was specifically immunoprecipitated by antibody to rat β-glucuronidase. The M_r value of 76000 of the expressed gene product was characteristic of the glycosylated rat enzyme. It was proteolytically processed in COS cells to M_r 75000 6 h after metabolic labelling. At least 50% of the expressed enzyme was secreted at 60 h post-transfection, but the secreted enzyme did not undergo proteolytic processing. These results provide evidence that the partial cDNA isolated from a rat liver library contains the complete coding sequence for the mature rat liver enzyme and that the chimeric signal sequence allows normal biosynthesis and processing of the transfected rat liver enzyme in COS cells.

INTRODUCTION

The lysosomal acid hydrolase β-glucuronidase has been studied extensively, both as a model of enzyme sorting and transport (Hasilik & Neufeld, 1980; Natowicz *et al.*, 1982; Creek & Sly, 1984) and as an example of a hormonally and developmentally regulated enzyme (Pfister *et al.*, 1984; Watson *et al.*, 1985). Studies on mucopolysaccharidosis type VII, the human genetic disorder resulting from deficiency of β-glucuronidase, have demonstrated the important role of this enzyme in glycosaminoglycan breakdown. In the mouse kidney β-glucuronidase is subject to induction by androgens, which can elevate mRNA concentrations 120-fold (Palmer *et al.*, 1983). In female rat preputial glands β-glucuronidase is regulated by oestrogen, and may comprise as much as 7% of the total protein in the gland (Levy *et al.*, 1958). The extralysosomal roles of β-glucuronidase, suggested by the presence of an egasyn-bound form in the endoplasmic reticulum of mouse and rat tissue (Medda & Swank, 1985) and its abundance in secretions from rat preputial gland, have not yet been elucidated.

The enzyme is a tetramer of identical subunits formed from a single gene product. The native enzyme in rat liver has an M_r of about 280000 (Delvin & Granetto, 1970). The sequences of cDNA clones for β-glucuronidase from three different species have been recently reported. Nishimura *et al.* (1986) reported a cDNA for rat preputial-gland β-glucuronidase. The protein synthesized from this cDNA in a translation system *in vitro* had an M_r of 70000, and was co-translationally inserted into microsomal membranes. Jefferson *et al.* (1986) cloned β-glucuronidase from *Escherichia coli* and determined the M_r to be 68000. They suggested that β-glucuronidase could be used as a reporter enzyme in a gene fusion system. We have recently cloned, sequenced and expressed the human placental β-glucuronidase (Oshima *et al.*, 1987). On transfection into COS cells (from African-green-monkey kidney) the expressed human cDNA showed an immunoprecipitable protein of M_r 74000 after removal of carbohydrate with endoglycosidase F.

In the present paper we provide data on the cloning and the sequence of a cDNA for β-glucuronidase from rat liver. The amino acid sequence deduced from this cDNA is compared with those deduced for the three enzymes reported previously. An expressing clone of the rat liver cDNA for β-glucuronidase was constructed. Coding sequence from the human enzyme provided 18 amino acid residues of the 22-amino-acid-residue signal sequence. The rest of the signal sequence and the complete sequence for the mature rat enzyme were derived from the rat liver cDNA clone. After transfection of COS cells, the chimeric cDNA clone produced active enzyme that was targeted to lysosomes and underwent

Abbreviations used: SSC, standard saline citrate; DMEM, Dulbecco's modified Eagle's medium; SV40, simian virus 40.

* To whom all correspondence should be addressed.

These sequence data have been submitted to the EMBL/GenBank Data Libraries under the accession number Y00717.

proteolytic processing similar to that seen for endogenous enzyme in hepatocytes (Rosenfeld *et al.*, 1982).

EXPERIMENTAL

Materials

All reagents were analytical grade and, unless otherwise specified, from Sigma Chemical Co. or Fisher Scientific Co. DNA restriction and modification enzymes were from Bethesda Research Laboratories or New England Biolabs. Nitrocellulose was from Schleicher and Schuell. Oligonucleotides were synthesized by the phosphoramidite method on an Applied Biosystems model 380A DNA synthesizer. Oligo(dT)–cellulose was from Collaborative Research. ^{35}S- and ^{32}P-labelled nucleotides were purchased from Amersham–Searle, New England Nuclear Corp. or International Chemical and Nuclear Corp. Endoglycosidase F (grade II) was from Boehringer Mannheim. Acrylamide was from Bio-Rad Laboratories. IgGsorb was from the Enzyme Centre. COS-7 cells, derived from African-green-monkey kidney tissue, were obtained from Dr. Maurice Green, Institute of Molecular Virology, St. Louis University Medical Center. Culture media were obtained from Mediatech (Herndon, VA, U.S.A.).

A cDNA library was constructed by Michael McPhaul (McPhaul & Berg, 1986, 1987) by using the pcD vector (Okayama & Berg, 1982) and polyadenylated RNA isolated from the liver of a female Sprague–Dawley rat. After construction, he digested samples from the library with *Sal*I, *Pvu*I or *Cla*I restriction endonuclease, each of which cuts the pcD vector once. The samples were pooled, sized by agarose-gel electrophoresis, re-ligated by size class, and introduced into bacteria to form a series of sub-libraries.

Screening of the cDNA library

The cDNA sub-library containing inserts between 2.2 and 2.8 kb was screened by colony hybridization (Grunstein & Hogness, 1975). Twenty plates (10 cm × 10 cm) containing 10000 colonies each were screened. The probe used was a 1.95 kb human β-glucuronidase partial cDNA (Guise *et al.*, 1985). Pre-hybridization and hybridization were performed at 65 °C in a solution containing 6 × SSC, 0.5% SDS, 5 × Denhardt solution and 100 μg of single-stranded DNA/ml (1 × SSC is 0.15 M-NaCl/0.15 M-sodium citrate buffer, pH 7.0; 50 × Denhardt solution contains 10 mg of bovine serum albumin/ml, 10 mg of Ficoll/ml and 10 mg of polyvinylpyrrolidone/ml). For hybridization, the nick-translated probe (10^7 c.p.m./μg of DNA) was added to the solution such that there were 10^6 c.p.m./ml. Filters were washed for 5 min in 2 × SSC/0.1% SDS at room temperature, followed by three successive 15 min washes in 0.2 × SSC/0.1% SDS at 68 °C. Positive plaques were detected by autoradiography at −70 °C with Kodak XAR film with intensifying screens.

DNA sequencing

Fragments from the longest cDNA clone, pRLG5, were subcloned into the M13 vectors mp18 and mp19 and used to transform *E. coli* JM109. Single-stranded templates were prepared and sequenced by the dideoxy chain-termination procedure (Messing, 1983). The sequencing reaction was primed with either the universal primer or specific oligonucleotides complementary to sequence already determined (Strauss *et al.*, 1986).

Transfections

Cloned cDNAs for human and rat β-glucuronidases in the SV40 expression vector pJC119 were transfected into COS-7 cells by a DEAE-dextran procedure (Oshima *et al.*, 1987). Cells, seeded at 2×10^5 cells per 60 mm dish the preceding day, were washed with DMEM without fetal bovine serum. Plasmid (10 μg) was added in 2 ml of DEAE-dextran/DMEM/50 mM-Tris/HCl buffer, pH 7.4. After 10 h the cells were washed with DMEM containing 10% (v/v) fetal bovine serum and treated for 3 h with 100 μM-chloroquine in 2 ml of DMEM/fetal bovine serum. Chloroquine was removed and the cells were allowed to grow for 36–72 h. Cell lysates were assayed for enzyme activity or analysed by gel electrophoresis. β-Glucuronidase activity was assayed fluorimetrically with 4-methylumbilliferyl β-glucuronide (Glaser & Sly, 1973).

Pulse–chase and immunoprecipitation

At 48–72 h post-transfection, COS cells were pre-incubated for 2 h with methionine-free DMEM. They were then grown for 30 min in the presence of 50 μCi of [^{35}S]methionine in 1 ml of methionine-free DMEM containing 5% fetal bovine serum. The cells were washed and placed in normal media supplemented with 2 mM-methionine for a chase of 0, 1, 3, 6 or 12 h. The cells were harvested and solubilized in 0.2 ml of solubilization buffer (10 mM-Tris/HCl buffer, pH 8.5, containing 1 mM-$MgCl_2$, 0.5% sodium deoxycholate, 1% Nonidet P-40, 10 μg of aprotinin/ml and 0.1% SDS). The cell extracts and media were pre-cleared with pre-immune serum, followed by addition of 100 μl of 20% (w/v) IgGsorb. This was incubated for 1 h at 4 °C, and spun at approx. 14000 g for 10 min. After centrifugation, goat anti-(rat preputial-gland β-glucuronidase) antibody was added to the supernatant and samples were incubated for 6 h at 4 °C. Immunoprecipitated material was collected by addition of 100 μl of 20% IgGsorb, incubation at 4 °C for 1 h and centrifugation. The pellets were washed five times with 150 mM-Tris/HCl buffer, pH 7.4, containing 0.57 M-NaCl, 1% sodium deoxycholate and 0.4% Nonidet P-40. The pellets were washed a final time with 10 mM-Tris/HCl buffer, pH 7.4.

β-Glucuronidase was liberated from the pellet by boiling for 5 min in 40 μl of 1% (w/v) SDS in 50 mM-Tris/HCl buffer, pH 6.8. Samples were split into two 20 μl portions. To each was added 20 μl 0.15 M-sodium citrate buffer, pH 5.3. One part was treated with 1 munit of endoglycosidase F for 20 h at 37 °C. Samples were analysed by electrophoresis on 7.5% polyacrylamide gels. ^{35}S was detected by fluorography.

RESULTS

Cloning and DNA sequence of the rat liver β-glucuronidase

cDNA clones for β-glucuronidase were identified in the rat liver cDNA library as described in the Experimental section. Fig. 1 gives the restriction map of the longest clone, pRLG5, and the sequencing strategy. The restriction map of the insert is nearly identical with that previously published for the rat preputial-gland β-glucuronidase (Nishimura *et al.*, 1986). Fig. 2 shows the nucleotide sequence from the cDNA clone. On comparison with the sequences previously published for human

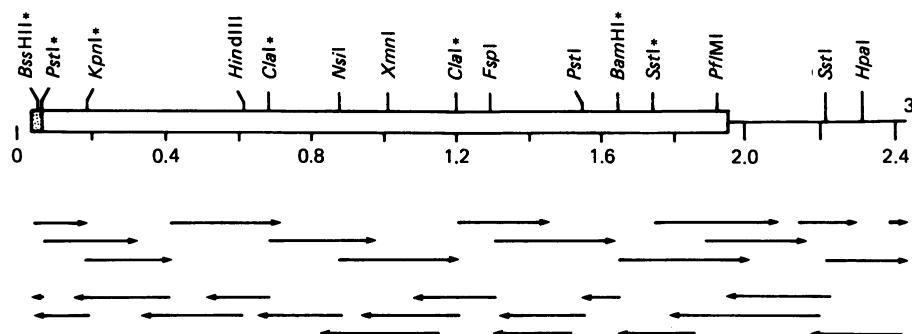

Fig. 1. Restriction map and sequencing strategy for the longest rat liver β-glucuronidase clone, pRLG5

The large bar represents the coding region of the gene; the stippled portion of the large bar represents the region coding for the signal sequence; the line represents the 3′ untranslated region. The numbers represent length in kilobases from the ATG initiation codon. Asterisks indicate restriction sites in common with human β-glucuronidase cDNA (Oshima *et al.*, 1987). The arrows show the relative position, extent and direction of DNA sequence determination by the dideoxy chain-termination procedure (Messing, 1983).

placental (Oshima *et al.*, 1987) and rat preputial-gland β-glucuronidase (Nishimura *et al.*, 1986), several points can be made. (1) The cDNA from rat liver contains the entire coding sequence for the mature rat enzyme plus the sequence that would encode nine residues of the 22-amino-acid-residue *N*-terminal signal sequence. (2) Seven nucleotide differences were noted between rat liver cDNA and that from rat preputial gland. Four of the seven differences were in the 5′ sequence that would translate to the *N*-terminal signal sequence, one was in the coding sequence for the mature enzyme, and two were in the 3′ untranslated region. (3) One of the differences between rat liver and rat preputial-gland cDNAs provides a unique *Bss*HII site near the 5′ end, which is also present in the human cDNA. This site was utilized in the construction of the chimeric cDNA described in Fig. 4. (4) The rat liver and rat preputial-gland cDNAs also differ at the 3′ end. Both rat cDNA clones contain two AATAAA 3′ cleavage signals. The sequence reported for the 3′ untranslated region of the cDNA from rat liver agrees with that reported for the cDNA from rat preputial gland up to the poly(A) stretch, but the cDNA from rat preputial gland contained 17 additional residues after the 18-base poly(A) segment. This unusual feature at the 3′ end of the cDNA for β-glucuronidase from rat preputial gland led Nishimura *et al.* (1986) to speculate that the 3′ end of this cDNA may be incomplete. (5) Comparison of the nucleotide sequence of the rat liver cDNA with that of the human placental cDNA (Oshima *et al.*, 1987) shows substantial conservation (79% overall homology for the coding sequence). However, the 3′ untranslated region of the rat liver cDNA is 276 nucleotide residues longer than that in the cDNA from human placenta, and the 3′ untranslated sequences are sufficiently divergent that they cannot be reliably aligned.

Fig. 3 presents a comparison of the deduced amino acid sequence of rat liver β-glucuronidase with those of human placental, rat preputial-gland and *E. coli* β-glucuronidases. The amino acid sequence deduced from the rat liver cDNA shows three differences from that reported for the rat preputial-gland enzyme. Two differences, amino acid changes from glutamine to glutamate and from valine to leucine, occur in the signal sequence. The third difference represents an amino acid change from methionine to leucine at amino acid 510 (in the alignment shown in Fig. 3). Comparison of sequence deduced for the rat liver enzyme with that for the human placental enzyme shows the rat enzyme to have one additional residue at the *C*-terminal end, but to be missing four internal residues (numbers 313–315 and 166). The overall amino acid sequence homology between rat liver and human placental β-glucuronidase, excluding the signal sequences, is 77%. The least conserved regions of homology involve the signal sequence, where the homology is only 42%, and the last 18 amino acid residues, where the homology is only 40%. Both the human and rat sequences have four potential *N*-linked glycosylation sites, three of which are homologous. Three glycosylated peptides have been characterized in a mouse macrophage cell line (Goldberg & Kornfeld, 1981).

The publication of the sequence of *E. coli* β-glucuronidase (Jefferson *et al.*, 1986) permits a sequence comparison of the mammalian lysosomal enzyme with its prokaryotic counterpart (Fig. 3). The amino acid sequence of β-glucuronidase from rat liver is 47% homologous to the enzyme from *E. coli*. A weaker overall homology had previously been recognized between the sequences for β-glucuronidases from rat preputial gland (Nishimura *et al.*, 1986) and human placenta (Oshima *et al.*, 1987) and β-galactosidase from *E. coli*.

Construction of an expressing rat liver β-glucuronidase cDNA

Prior studies on the cloning of human placental β-glucuronidase, where cDNAs corresponding to two different mRNAs were isolated, only one of which specifies a functional enzyme, indicated the need for expression studies to verify that the cDNA is functionally full length (Oshima *et al.*, 1987). Since the rat liver cDNA lacked some of the 5′ nucleotide sequence specifying the signal sequence, we constructed a chimeric cDNA that would contain the first 18 amino acid residues from the signal sequence of the human enzyme and the last four amino acid residues (including the cleavage site) of the signal sequence for the rat enzyme. This construction took advantage of a common *Bss*HII restriction site at a point of homology between the two relatively divergent

```
39                             --------
A GAG CTG CTG TGC AGC TGC GCG CTG GCG▼CTG CAG GGC GGG ATG CTG TTC CCG AAG GAG ACC CCT TCG CGG GAA CTC AAG GTG CTG GAC GGC CTG  132
C C                           C   G
TGG AGC TTC CGT GCC GAT TAC TCG AAC AAT CGG TTG CAG GGC TTC GAG AAG CAG TGG TAC CGG CAG CCG CTA CGG GAG TCG GGC CCA ACC TTG GAC  228
ATG CCG GTC CCT TCC AGC TTC AAT GAC ATC ACC CAA GAA GCA GAG CTT CGG AAC TTC ATT GGC TGG GTG TGG TAT GAA CGG GAA GCA GTG CTT CCA  324
CAG CGA TGG ACC CAG GAC ACC GAC AGG AGA GTG GTG TTG AGA ATC AAC AGC GCC CAT TAC TAT GCA GTT GTG TGG GTG AAT GGG ATT CAT GTG GTG  420
GAA CAT GAG GGA GGT CAC CTC CCC TTT GAG GCT GAC ATC ACC AAG CTG GTC CAG AGT GGG CCC CTG ACC ACC TTC CGG GTC ACC ATC GCC ATC AAC  516
AAC ACA CTG ACC CCT TAT ACC CTT CCA CCG GGG ACC ATT GTC TAC AAG ACT GAT CCT TCC ATG TAT CCC AAG GGT TAC TTC GTC CAG GAC ATA AGC  612
TTC GAC TTC TTC AAC TAT GCG GGG CTG CAC CGG TCT GTG GTC CTG TAC ACC ACC CCT ACC ACC TAT ATC GAT GAT ATC ACT GTG ACC ACT GAC GTG  708
GAC CGG GAC GTT GGG CTG GTG AAC TAC TGG ATT TCT GTC CAG GGC AGT GAC CAT TTC CAG CTA GAA GTG CGT CTT CTG GAT GAG GAT GGC AAA ATT  804
GTG GCC CGT GGA ACA GGG AAT GAG GGT CAA CTT AAG GTG CCG CGT GCC CAC CTC TGG TGG CCT TAC CTG ATG CAT GAG CAT CCA GCC TAC CTG TAC  900
TCC TTG GAG GTG ACG ATG ACA ACA CCT GAG TCT GTG TCT GAC TTC TAT ACC CTC CCT GTC GGG ATT CGA ACA GTG GCT GTC ACA AAG AGC AAG TTC  996
CTC ATA AAT GGG AAG CCT TTC TAC TTC CAA GGC GTC AAC AAG CAT GAG GAT TCA GAT ATC CGA GGG AGA GGC TTC GAC TGG CCT CTG CTG ATA AAG 1092
GAT TTC AAC CTC CTC CGT TGG CTC GGG GCA AAT TCC TTT CGT ACC AGC CAC TAT CCC TAC TCG GAG GAG GTA CTT CAG CTC TGT GAC CGA TAT GGA 1188
ATT GTG GTC ATC GAT GAG TGT CCC GGT GTG GGC ATC GTG CTG CCC CAG AGT TTT GGC AAC GTG TCT CTT CGG CAC CAC CTA GAG GTG ATG GAC GAG 1284
CTG GTG CGC AGG GAC AAA AAT CAC CCT GCG GTC GTG ATG TGG TCT GTG GCC AAT GAG CCT GTC TCT TCT CTG AAA CCT GCC GGA TAT TAC TTC AAG 1380
ACG CTG ATC GCC CAC ACC AAA GCC CTG GAC CCC ACC CGT CCT GTG ACC TTT GTG AGC AAT ACC AGA TAT GAC GCA GAC CTG GGG GCC CCG TAC GTG 1476
                                                                                                          A
GAC GTG ATT TGT GTG AAC AGT TAC TTA TCC TGG TAT CAT GAC TAC GGG CAT CTG GAG GTG ATT CAG CTG CAG CTG ACT AGC CAG TTT GAG AAC TGG 1572
TAT AAG ATG TAC CAG AAG CCA ATT ATC CAG AGC GAG TAT GGA GCA GAC GCC GTC TCG GGG CTT CAT GAG GAT CCA CCT CGC ATG TTC AGT GAG GAG 1668
TAC CAG ACA GCT CTC CTG GAG AAT TAT CAT TTG ATC CTG GAT GAG AAA CGA AAA GAA TAT GTC ATC GGA GAG CTC ATC TGG AAT TTT GCT GAC TTC 1764
ATG ACG AAC CAG TCA CCA CTG AGA GTA ACA GGA AAC AAG AAG GGG ATC TTC ACT CGA CAG AGA AAC CCC AAG ATG GCA GCC TTC ATT TTG CGA GAG 1860
AGA TAC TGG AGG ATT GCC AAC GAA ACC AGA GGT TAC GGT TCA GTG CCG AGG ACC CAG TGT ATG GGA AGC AGA CCG TTC ACA TTC TAA AGCTACGACTA 1958
                                                                                                                  *
CCTCACCGCAGACAGGGCCTGATGAAGTGAAAGGTATTCTCTCAGCAGAAGAGTTTTTCCTCTGGAGTTTGTGAACCAACACAAGGGGGACTTTAGAAGTGACATGGAAGTTTTCCTTGTGTGGGGA 2085
                                                                                     A
TAATGACTTGCAGTGCCTTAGGCTCAGTAATGTGTCCTGCTGAGAAGTGTCCCAGGTCATTTCTGTCAGGTTTTTTTTGGCCTTGGCTTTGTGTACTCTTGAAAACCTACTGTGTGAATACTCACAC 2212
TAAGAGCTCCTGGTTACCGGCTGAGGGACGGCTCACAATTTAAGAATACTGACTGTTCTACCAGGGGTTCCAGATTCGATTCCTAGCACCCACGTGGTCGTTAACTATGAGCTGTTGCCCCATTTCT 2339
                      A
CAGGTGATGTACGAACATACATGTGGGGGAGGGGAATAGTCATACGTGTAAAATGAAAATAAATCTAAATAAAACTGTTTTCTTTCATT                                    2429
```

Fig. 2. Nucleotide sequence of the rat liver β-glucuronidase cDNA

The nucleotides are numbered from the initiation of translation. The arrow indicates where the derived protein sequence is cleaved by signal peptidase. The *Bss*HII site used in the construction of an expressing clone is indicated with a dashed overline. The two potential 3′ cleavage signals (AATAAA) are underlined. The stop codon is marked with an asterisk. Nucleotide differences in the rat preputial-gland cDNA are shown below the rat liver sequence.

```
                                ↓
RL              ELLCSCALALQGGMLFPKETPSRELKVLDGLWSFRADYSNNRLQGFEKQWYRQPLRESGPTLDMPVPSSFNDITQEAELRNFIGWVWYEREAVLPQRWTQDTDRRVV
RPG MSPRRSVCWFVLGQ......V.................................................................................................
HP  MARGSAVAWAALGP..WG...G......Y.Q.S....C.E..........F.D..RR...E....R..W.....V...........S.DWR..H.V........VI..E.....LRT...  120
EC                            ..R.V...T..I.K.....A.SL.RE.CGIDQRWWESAL.ES.AIA------..G....QFAD.DI..YA.N...Q..VFI.KG.AGQ---.I.

RL  LRINSAHYYAVVWVNGIHVVEHEGGHLPFEADITKLVQSGPLTTF-RVTIAI[NNT]LTPYTLPPGTIVYKTDPSMYPKGYFVQDISFDFFNYAGLHRSVVLYTTPTTYIDDITVTTDVDRD
RPG .............................................-......[...]..................................................................
HP  ...G...S..I.....VDTL.....Y.......SN...V...PSRL.I.....[...]...T.......Q.L..T.K........NTY.........Q...L...............S.EQ.  240
EC  ..FDAVTH.GK...NNQE.M..Q..YT.....V.PY.IA.KSV---.I.VCV.[.E.]NWQ.I...MVITDEN------.KKK.SYFH.......I....M.....N.WV.....V.H.AQ.

RL  VGLVNYWISVQGSDHFQLEVRLLDEDGKIVARGTGNEGQLKVPRAHLWWPYLMHEHPAYLYSLEVTMTTPES---VSDFYTLPVGIRTVAVTKSKFLINGKPFYFQGVNKHEDSDIRGRG
RPG ......................................................................---...............................................
HP  S.....Q...K..NL.K.......AEN.V..[N..].TQ......GVS.........R.........QL.AQT.LGP..................Q..........H.......A....K.  360
EC  CNHASVDWQ.VANGDVSV.L.--.A.QQV..[T.Q].TS.T.Q.VNP...Q.GEG-----...E.C..AKSQTE----C.I.P.R....S...KGEQ....H.....T.FGR...A.L..K.

RL  FDWPLLIKDFNLLRWLGANSFRTSHYPYSEEVLQLCDRYGIVVIDECPGVGIVLPQSFG[NVS]LRH-----------------HLEVMDELVRRDKNHPAVVMWSVANEPVSSLKPAGYYF
RPG ...........................................................[...]...-----------------......................................
HP  ......V............A........A...M.M................LA...F.[N..].H.-----------------.MQ..E.V.................A.H.ES....L  480
EC  ..NV.MVH.HA.MD.I....Y.......A..M.DWA.EH.......TAA..FN.SLGI.[FEA]GNKPKELYSEEAVNGETQQA..QAIK..IA......S.....I....DTRPQVH.NIS

RL  KTLIAHTKALDPTRPVTFVSNTRYDA--DLGAPYVDVICVNSYLSWYHDYGHLEVIQLQLTSQFENWYKMYQKPIIQSEYGADAVSGLHEDPPRMFSEEYQTALLENYHLILDEKRKEYV
RPG ..........................--.M..........................................................................................
HP  .MV.....S...S........SN.A.--.K.........L...Y..........L.....AT.......K............ETIA.F.Q...L..T....KS...Q...G..Q..RK..  600
EC  PLAE.TR.-.......I.C.NVMFC..HT.TISDLF..L.L.R.YG..VQS.D..TAEKV.EKELLA.QEKLHQ...IT...V.TLA...SMYTD.W.....C.W.DM..RVF.--.VSA.

RL  IGELIWNFADFMT[NQS]PLRVTGNKKGIFTRQRNPKMAAFILRERYWRIA[NET]RGYGSVPRTQCMGSRPFTF
RPG .............[...]....................................[...]....................
HP  V............[E..].T..L...........Q..S...L......K..[...].YPH..AKS..LENS...   670
EC  V..QV.......A.S.GI...G.........D.K..S...L.QK.WTGMNFGEKPQQGGKQ
```

Fig. 3. Comparisons of the amino acid sequences of β-glucuronidases from various sources

The amino acid sequence for β-glucuronidase from rat liver (RL), deduced from the nucleotide sequence (Fig. 2), was aligned with sequences previously reported from cDNA clones from rat preputial gland (RPG) (Nishimura *et al.*, 1986), from human placenta (HP) (Oshima *et al.*, 1987) and from *E. coli* (EC) (Jefferson *et al.*, 1986). Alignment was carried out by using the FASTP program from the Protein Identification Resource of the National Biomedical Research Foundation (Lipman & Pearson, 1985) with the use of a gap penalty of 4. The dots indicate identity with the rat liver sequence. The dashes indicate absences of amino acid residues from the respective sequence. The arrow indicates the site of the signal sequence cleavage. The boxes indicate sites for potential *N*-glycosylation in the mammalian enzymes. Amino acids are represented by the single-letter code.

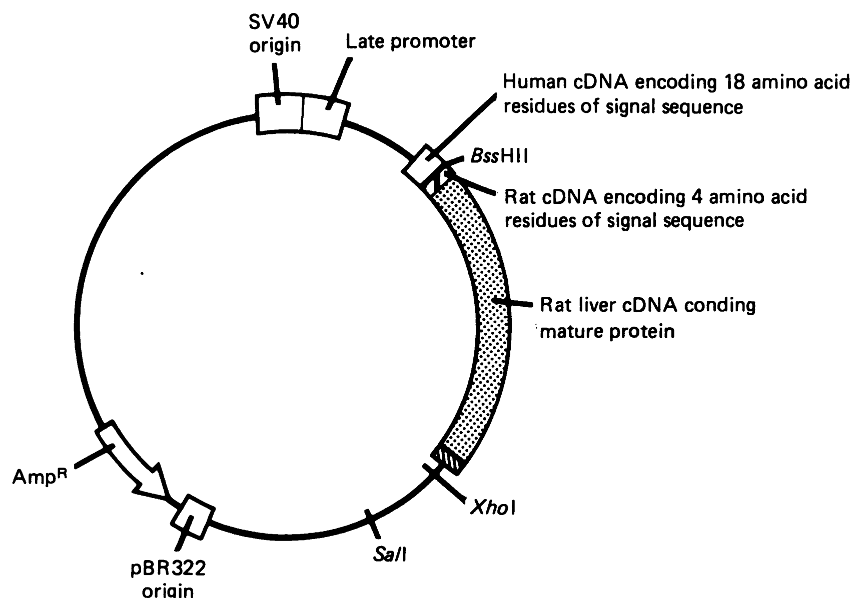

Fig. 4. Construction of a full-length-expressing rat liver β-glucuronidase cDNA

Part of the human β-glucuronidase cDNA clone encoding the signal sequence was ligated to the 5′ end of the longest rat liver clone, pRLG5, by use of a homologous *Bss*HII site, briefly as follows. The rat liver clone was excised from the SV40 early vector pcD in which the library had been constructed and ligated into pJC119, the SV40 late vector, at the unique *Xho*I site. This plasmid was restricted with *Bss*HII and *Sal*I to remove a 3.0 kb fragment containing the entire rat coding sequence with four amino acid residues of the signal sequence, and part of the vector DNA. The full-length-expressing human β-glucuronidase clone, pSVL(JC)-HUGP13, was digested with *Bss*HII and *Sal*I, and the human β-glucuronidase cDNA, excluding nucleotide sequence encoding the first 18 amino acid residues of the signal sequence, was removed. A directional ligation of the rat cDNA into this vector completed the signal sequence at the 5′ end. This vector is denoted pHRLG5-SVL.

Table 1. β-Glucuronidase activity of protein expressed from cloned rat and human cDNAs in transfected COS cells

COS-7 cells were transfected with vector containing no insert (pJC119) or with the expressing rat liver cDNA construct (pHRLG5-SVL) or with the expressing human placental cDNA clone [pSVL(JC)-HUGP13] as described in the Experimental section. Cell extracts and cell medium were assayed for β-glucuronidase activity 60 h post-transfection.

	β-Glucuronidase activity			
	In cell extract		In medium	
Plasmid	(total units)	(units/mg of protein)	(total units)	(units/mg of protein)
pJC119	14	85	17	0.7
pHRLG5-SVL	403	2240	1100	47
pSVL(JC)-HUGP13	540	3260	216	9

signal sequences. The details of the construction are illustrated in the legend to Fig. 4. The resultant chimeric cDNA, pHRLG5-SVL, should encode a protein specified entirely by rat cDNA sequence once the chimeric signal peptide has been cleaved by signal peptidase in the endoplasmic reticulum.

Expression of rat liver β-glucuronidase in COS cells

Table 1 presents data on the expression of β-glucuronidase in COS cells after transfection with the chimeric human–rat cDNA construct in the SV40 late promoter expression vector, pJC119 (Fig. 4) (Sprague *et al.*, 1983). Transfection with vector containing the cDNA led to a 27-fold increase in intracellular β-glucuronidase activity over that seen after transfection with vector containing no insert. The increase was comparable with that seen when COS cells were transfected with vector containing the human cDNA for β-glucuronidase.

The enzyme expressed in cells transfected with the chimeric human–rat cDNA construct was not precipitated by a goat antibody raised to human placental β-glucuronidase that precipitates human enzyme but does not precipitate rat enzyme. It was precipitated by goat antibody raised to purified rat enzyme. Expression of the rat enzyme in COS cells after transfection supports the conclusion from sequence comparisons with the human cDNA clone, namely that the cDNA from rat liver contains the entire coding sequence for the mature rat β-glucuronidase enzyme.

Biosynthesis and processing of metabolically labelled rat enzyme

Fig. 5 presents data from metabolic labelling experiments in which expressed enzyme was examined by fluorography after immunoprecipitation. The control for

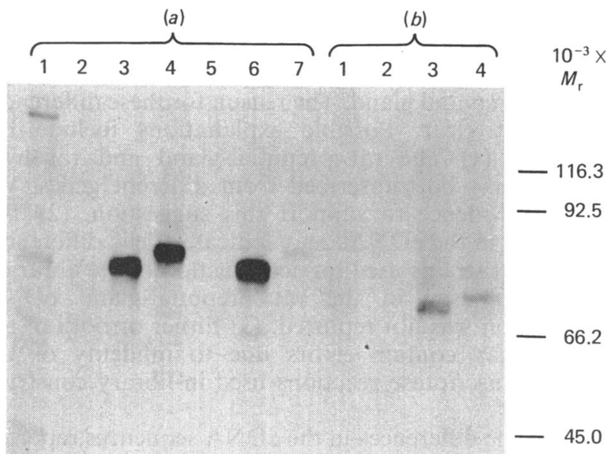

Fig. 5. Expression of rat liver and human placental β-glucuronidases after transfection into COS cells

Normal rat kidney cells and COS cells that were 48 h post-transfection were starved for 2 h in methionine-free medium and labelled for 1 h with 50 μCi of [^{35}S]-methionine followed by a 3 h chase with unlabelled methionine. The cell extracts and medium were pre-cleared with pre-immune serum, and proteins were immunoprecipitated with goat anti-(rat β-glucuronidase) antibody. Panel (*a*) shows proteins not treated with endoglycosidase F. Panel (*b*) shows proteins treated with 1 munit of endoglycosidase F for 20 h. Lane 1, normal rat kidney cells; lane 2, COS cells transfected with vector (pJC119) alone; lane 3, cells transfected with rat liver cDNA, pHRLG5-SVL; lane 4, human placenta cDNA pSVL(JC)-HUG13 (Oshima *et al.*, 1987); lane 5, medium from cells transfected with vector alone; lane 6, medium from cells transfected with pHRLG5-SVL; lane 7, medium from cells transfected with pSVL(JC)-HUGP13. The positions of protein M_r markers are indicated.

intracellular rat enzyme (Fig. 5*a*, lane 1) was provided by metabolically labelled normal rat kidney cells, which showed a band at apparent M_r 75000–76000. Fig. 5(*a*) shows immunoprecipitation of intracellular β-glucuronidase produced after transfection with the vectors containing no insert (lane 2), the chimeric human–rat cDNA (lane 3) and the human cDNA (lane 4). Transfection with the rat construct produced an intense band of intracellular protein at M_r 75000–76000. Transfection with the human cDNA produced an intense band at M_r 82000. Fig. 5(*b*) shows the effects of prior treatment of the immunoprecipitated enzymes with endoglycosidase F. The apparent M_r of both the rat and human enzymes decreased about 6000. The difference between the M_r values of human and rat enzymes persists, indicating that it was not due to carbohydrate.

Lanes 6 and 7 in Fig. 5(*a*) show the enzyme immunoprecipitated from the medium 3 h after labelling the cells transfected with the constructs containing rat and human cDNAs respectively. The rat enzyme was secreted at a much higher rate than was the human enzyme. This is consistent with the data in Table 1, where activity in the medium 60 h post-transfection with the cDNA for rat liver β-glucuronidase was 5 times that seen after transfection with the cDNA for human β-glucuronidase.

The processing and fate of metabolically labelled enzyme expressed after transfection were studied in a pulse–chase experiment presented in Fig. 6. Transfected COS cells were metabolically labelled for 30 min, and then isolated after periods of chase up to 12 h. The major band present intracellularly at 30 min was at M_r 76000. By 6 h a band at M_r 75000 appeared, and this increased to become the predominant band by 12 h of chase (Fig. 6*a*). Endoglycosidase F treatment decreased the M_r by roughly 6000 at all time points, suggesting that the late processing is due to loss of polypeptide (proteolytic cleavage) and not due to carbohydrate processing. The secreted enzyme was prominent in the medium by 3 h (as in Fig. 5) and there was no change in M_r in the first 12 h. Endoglycosidase F decreases the M_r of the monomers of the secreted enzyme by about 8000.

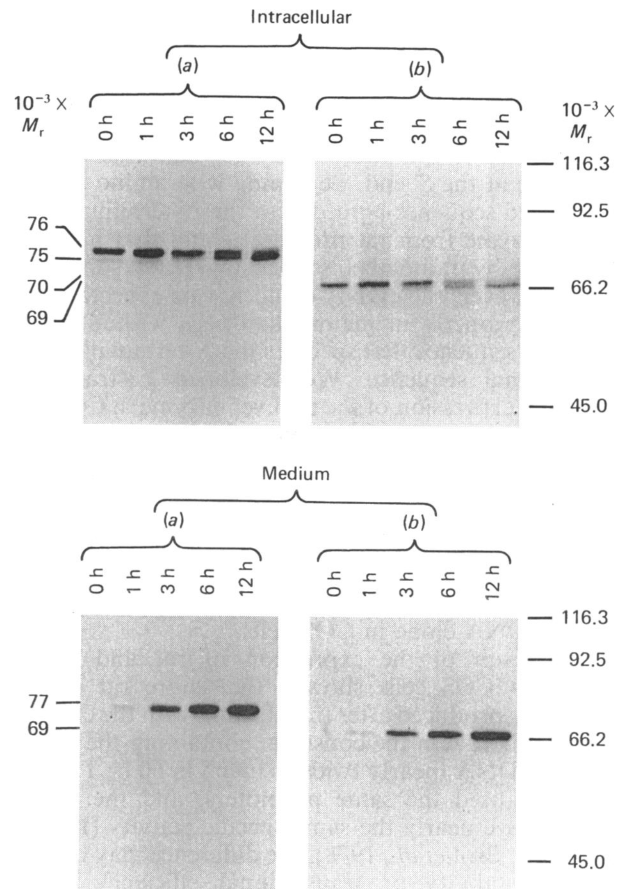

Fig. 6. Post-translational events in the synthesis of β-glucuronidase

COS-7 cells transfected with the rat liver cDNA clone in the late expression vector, pHRLG5-SVL, were labelled with [^{35}S]methionine for 30 min and chased for various times with an excess of the unlabelled amino acid. Labelled proteins in the cells or the media were immunoprecipitated with goat anti-(rat β-glucuronidase) antibody, and the products were isolated by SDS/polyacrylamide-gel electrophoresis. The first lane (0 h) shows immunoprecipitation immediately after the 30 min pulse, the next lanes after a chase of 1, 3, 6 or 12 h, as indicated. Panel (*a*) shows no endoglycosidase F treatment. Panel (*b*) shows extracts treated with 1 munit of endoglycosidase F for 20 h at 37 °C. ^{35}S was detected by fluorography. The positions of protein M_r markers are indicated.

DISCUSSION

Before the report of the cDNA clone for rat preputial-gland β-glucuronidase (Nishimura *et al.*, 1986), we had concluded that the rat liver cDNA contained the coding sequence for the mature enzyme but that it was incomplete at the 5′ end. Beginning with amino acid 10, the deduced sequence agreed with the *N*-terminus of the mature enzyme from rat preputial gland that had been established by radiolabel sequencing (Erickson *et al.*, 1984). However, the cDNA could not be expected to be expressed properly in mammalian cells without the 5′ nucleotide sequence that specifies the *N*-terminal portion of the signal sequence. We developed a strategy to investigate expression of the rat liver enzyme in COS cells by providing the *N*-terminal end of the signal sequence from the cDNA for human placental β-glucuronidase to form a chimeric human–rat signal sequence. The expression studies reported in the present paper show that, although the human and rat signal sequences are rather divergent, the chimeric signal sequence functioned very efficiently and permitted analysis of the expression of the rat liver cDNA clone in COS cells.

Comparison of the expression of rat and human enzymes in COS cells showed that more rat enzyme activity was produced after transfection with the chimeric construct than with the construct containing the human placental cDNA (nearly twice as much in 60 h; Table 1). Since they used the same promoters, and the mature enzymes have nearly the same specific activity (Himeno *et al.*, 1975; Brot *et al.*, 1978), the difference may indicate increased stability or translational efficiency of the mRNA for the rat liver enzyme. In addition, a larger fraction of the rat enzyme was secreted, which could mean either that the human enzyme is more efficiently targeted to lysosomes in COS cells or that the more rapidly produced rat enzyme oversaturates the system for segregating enzyme to lysosomes. We have no evidence to distinguish these alternatives.

There were two differences between the secreted and intracellular expressed rat enzyme. (1) The secreted form underwent a larger change in M_r on treatment with endoglycosidase F, suggesting that it contained more carbohydrate. This could be a result of addition of extra sugar residues (especially sialic acid residues) in the processing of the secreted form, or from partial deglycosylation of the intracellular forms in lysosomes. (2) The secreted form did not undergo the proteolytic processing that the intracellular form underwent, presumably in lysosomes.

The kinetics of biosynthesis, processing and proteolytic processing of the enzyme expressed in COS cells resemble those reported by Rosenfeld *et al.* (1982) for β-glucuronidase in rat hepatocytes. In pig kidney cells the proteolytic processing event has been reported to occur near the *C*-terminus (Erickson & Blobel, 1983). We have no information on the location of the proteolytic processing event in COS cells.

The homologies between the nucleic acid sequences reported here and those reported previously demonstrate striking conservation of the β-glucuronidase sequence from *E. coli* to man. This suggests a common ancestry for the prokaryotic enzyme, which has a neutral pH optimum and is cytosolic, and the mammalian enzymes, which are glycoproteins, have acid pH optima and contain information that directs their targeting to lysosomes. On the other hand, there are several differences between the sequence we report here for the rat liver cDNA and that previously reported for the cDNA from rat preputial gland. The reason for these differences is not yet clear. Possible explanations include the following. (1) The rat preputial-gland and rat liver enzymes may be transcribed from different genes. We have no evidence to support this suggestion. (2) The differences in the cDNAs may indicate allelic differences between rat strains used for construction of the libraries. The strain used in the rat preputial-gland cDNA construction was not reported. (3) Either or both of the cDNAs may contain errors due to infidelity of the reverse transcriptase reactions used in library construction.

Could the differences in the cDNA sequences reported for rat β-glucuronidase from the two different tissue sources be the basis for the differences in targeting of the enzyme in the two tissues? The enzyme is mainly targeted to lysosomes in hepatocytes and mainly to exocrine secretory vesicles in the acinar cells of the preputial gland (Beyler & Szego, 1954). Two of the three amino acid changes involve changes from glutamine to glutamate and valine to leucine in amino acids 14 and 21 of the signal sequence. The signal sequence is probably cleaved in the endoplasmic reticulum, and therefore differences here would be unlikely to affect post-translational processing and sorting, which take place in the Golgi apparatus. The third change (methionine to leucine in residue 510) is the only one in the amino acid sequence of the mature enzyme. This difference also seems unlikely to be important in the targeting of β-glucuronidase. The alternative enzyme targeting more probably reflects tissue-specific differences in post-translational processing in the liver and preputial gland. Consistent with this conclusion is the report by Andy *et al.* (1986) that the β-glucuronidase synthesized after transfection of HeLa cells with the cDNA from rat preputial gland is targeted to lysosomes, rather than to secretory granules as in the rat preputial gland.

We thank Dr. Michael McPhaul (University of Texas Health Science Center, Dallas, TX, U.S.A.) for providing us with the rat liver cDNA library, Joseph Hoffmann for helpful advice and Liz Torno for typing this manuscript. This work was supported by grants from the U.S. Public Health Service (National Institutes of Health Grants GM34182 and AI22359) and from the National Foundation, March of Dimes.

REFERENCES

Andy, R., Rosenfeld, M., Adesnik, M. & Sabatini, D. (1986) J. Cell Biol. **103**, abstr. 1336

Beyler, A. L. & Szego, C. M. (1954) Endocrinology (Baltimore) **54**, 323–332

Brot, F. E., Bell, C. E. & Sly, W. S. (1978) Biochemistry **17**, 385–391

Creek, K. M. & Sly, W. S. (1984) in Lysosomes in Biology and Pathology (Dingle, J. T., Dean, R. T. & Sly, W. S., eds.), pp. 63–82, North-Holland, Amsterdam

Delvin, E. & Granetto, R. (1970) Biochim. Biophys. Acta **220**, 93–100

Erickson, A. H. & Blobel, G. (1983) Biochemistry **22**, 5201–5205

Erickson, A. H., Conner, G. E. & Blobel, G. (1984) in Molecular Basis of Lysosomal Storage Disorders (Barranger, J. A. & Brady, R. O., eds.), pp. 235–249, Academic Press, New York

Glaser, J. H. & Sly, W. S. (1973) J. Lab. Clin. Med. **82**, 969–977

Goldberg, D. E. & Kornfeld, S. (1981) J. Biol. Chem. **256**, 13060–13067

Grunstein, M. & Hogness, D. (1975) Proc. Natl. Acad. Sci. U.S.A. **72**, 3961–3965

Guise, K. S., Korneluk, R. G., Wayne, J., Lamhonwah, A.-M., Quan, F., Palmer, R., Ganschow, R. E., Sly, W. S. & Gravel, R. A. (1985) Gene **34**, 105–110

Hasilik, A. & Neufeld, E. F. (1980) J. Biol. Chem. **255**, 4946–4950

Himeno, M., Ohhara, H., Arakawa, Y. & Kato, K. (1975) J. Biochem. (Tokyo) **77**, 427–438

Jefferson, R. A., Burgess, S. M. & Hirsh, D. (1986) Proc. Natl. Acad. Sci. U.S.A. **83**, 8447–8451

Levy, G. A., McAllan, A. & Mersh, C. A. (1958) Biochem. J. **69**, 22–27

Lipman, D. J. & Pearson, W. R. (1985) Science **227**, 1435–1441

McPhaul, M. & Berg, P. (1986) Proc. Natl. Acad. Sci. U.S.A. **83**, 8863–8867

McPhaul, M. & Berg, P. (1987) Mol. Cell. Biol. **7**, 1841–1847

Medda, S. & Swank, R. T. (1985) J. Biol. Chem. **260**, 15802–15808

Messing, J. (1983) Methods Enzymol. **101**, 20–78

Natowicz, M., Baenziger, J. U. & Sly, W. S. (1982) J. Biol. Chem. **251**, 4412–4420

Nishimura, Y., Rosenfeld, M., Krebich, G., Gubler, V., Sabatini, D., Adesnik, M. & Andy, R. (1986) Proc. Natl. Acad. Sci. U.S.A. **83**, 7292–7296

Okayama, H. & Berg, P. (1982) Mol. Cell. Biol. **2**, 161–168

Oshima, A., Kyle, J. W., Miller, R. D., Hoffman, J. W., Powell, P. P., Grubb, J. M., Sly, W. S., Tropak, M., Guise, K. S. & Gravel, R. A. (1987) Proc. Natl. Acad. Sci. U.S.A. **84**, 685–689

Palmer, R., Gallagher, P. M., Boyko, W. L. & Ganschow, R. E. (1983) Proc. Natl. Acad. Sci. U.S.A. **80**, 7596–7600

Pfister, K., Watson, G., Chapman, V. & Paigen, K. (1984) J. Biol. Chem. **259**, 5816–5820

Rosenfeld, M. G., Kreibich, G., Popov, D., Kato, K. & Sabatini, D. D. (1982) J. Cell Biol. **93**, 135–143

Sprague, J., Condra, J. M., Arneiter, M. & Lazzarini, R. A. (1983) J. Virol. **45**, 773–781

Strauss, E. C., Kobori, J. A., Siu, G. & Hood, L. E. (1986) Anal. Biochem. **154**, 353–360

Watson, G., Felder, M., Rabinow, L., Moore, K., Laberca, C., Tietze, C., Molen, G. V., Bracey, L., Brabant, M., Cai, J. & Paigen, K. (1985) Gene **36**, 15–25

Received 14 May 1987/2 September 1987; accepted 4 November 1987