The EMBO Journal vol.7 no.8 pp.2343-2350, 1988

Human lysosomal acid phosphatase: cloning, expression and chromosomal assignment

Regina Pohlmann, Christiane Krentler, Bernhard Schmidt, Wolfgang Schröder[1], Gerhard Lorkowski[2], Jan Culley[2], Günther Mersmann[2], Carola Geier, Abdul Waheed, Stephen Gottschalk, Karl-Heinz Grzeschik[3], Andrej Hasilik[2] and Kurt von Figura

Biochemie II, Universität Göttingen, Gosslerstr. 12d, D-3400 Göttingen, [2]Physiologisch-Chemisches Institut, Universität Münster, Waldeyerstr. 15, D-4400 Münster, [1]Institut für Biochemie, Pharmaforschungszentrum der Bayer AG, Postfach 101701, D-5600 Wuppertal, and [3]Institut für Humangenetik, Universität Münster, Vesaliusweg 12-14, D-4400 Münster, FRG

Communicated by B.Dobberstein

A 2112-bp cDNA clone (λCT29) encoding the entire sequence of the human lysosomal acid phosphatase (EC 3.1.3.2) was isolated from a λgt11 human placenta cDNA library. The cDNA hybridized with a 2.3-kb mRNA from human liver and HL-60 promyelocytes. The gene for lysosomal acid phosphatase was localized to human chromosome 11. The cDNA includes a 12-bp 5′ non-coding region, an open reading frame of 1269 bp and an 831-bp 3′ non-coding region with a putative polyadenylation signal 25 bp upstream of a 3′ poly(A) tract. The deduced amino acid sequence reveals a putative signal sequence of 30 amino acids followed by a sequence of 393 amino acids that contains eight potential glycosylation sites and a hydrophobic region, which could function as a transmembrane domain. A 60% homology between the known 23 N-terminal amino acid residues of human prostatic acid phosphatase and the N-terminal sequence of lysosomal acid phosphatase suggests an evolutionary link between these two phosphatases. Insertion of the cDNA into the expression vector pSVL yielded a construct that encoded enzymatically active acid phosphatase in transfected monkey COS cells.

Key words: lysosomal acid hydrolyase/human chromosome 11

Introduction

Acid phosphatase (EC 3.1.3.2) is represented by a number of enzymes that can be differentiated according to structural, catalytic and immunological properties, tissue distribution and subcellular location (Yam *et al.*, 1980; Waheed *et al.*, 1985). The lysosomal and prostatic acid phosphatases belong to the group of acid phosphatases that are sensitive to inhibition by L-tartrate. Both consist of identical subunits with an apparent M_r of 48 000–52 000, but are immunologically distinct (Saini and van Etten, 1978; Luchter-Wasyl and Ostrowski, 1974; Gieselmann *et al.*, 1984; Waheed *et al.*, 1985). The lysosomal acid phosphatase (LAP) is found in all tissues and has played a pivotal role in the discovery of lysosomes (de Duve, 1983; Farquhar and Palade, 1981). It is synthesized as a precursor with an apparent M_r of 69 000 and processed to a family of M_r 43 000–57 000 polypeptides (Lemansky *et al.*, 1985; Waheed and van Etten, 1985). Like other lysosomal enzymes, LAP contains mannose-6-phosphate residues that serve as a recognition marker for targeting to lysosomes. LAP, however, may become segregated into lysosomes also in the absence of the mannose-6-phosphate recognition marker. This is indicated by its normal activity in I-cell fibroblasts, which are deficient in the enzyme generating the mannose-6-phosphate recognition marker (Gieselmann *et al.*, 1984). Other peculiarities of LAP are its occurrence in non-lysosomal organelles and partial association with the lysosomal membrane. The membrane-associated forms, which account for ~40% of LAP activity, differ in structure and metabolism from the soluble forms (Lemansky *et al.*, 1985; S.Gottschalk, unpublished data). In order to identify eventually the signals that determine the transport and compartmentalization of LAP we have attempted to define its primary structure. In the present study we report on the molecular cloning and expression of a cDNA encoding human LAP chromosomal assignment of its gene and homology to prostatic acid phosphatase.

Results

Purification and protein sequencing of LAP

The LAP purified from human placenta (Gieselmann *et al.*, 1984) was microsequenced. The sequence of the N-terminus (21 amino acids) and of six tryptic peptides (70 amino acids in total) are shown in Table I.

Isolation and characterization of cDNA clones for LAP

By screening 1.2×10^7 plaques of a human hepatoma cDNA library in the λgt11 expression vector with antibodies against LAP four putative clones were isolated. The DNA of one of these clones, λAP2, hybridized to the oligonucleotide probes 140 and 141 that were constructed using partial amino acid sequence data of LAP peptides 86b and 91b (Table II). When λAP2 was used to lysogenize *Escherichia coli* strain BNN103, a fusion protein was detectable in the cell lysate that was ~40 kd, larger in size than β-galactosidase (Figure 1, lane 2) and which reacted with antibodies against LAP (Figure 1, lane 4). AP2 (insert of λAP2) was subcloned into M13mp18. It contains three internal *Pst*I sites yielding fragments of ~75, 300, 600 and 950 bp (Figure 2A, lane 3). The 300-bp fragment hybridized to oligonucleotide 140 (Figure 2B) and the 600-bp insert to oligonucleotide 141 (Figure 2C).

AP2 was sequenced by the dideoxy chain termination method using the strategy shown in Figure 3. AP2 comprises 1908 bp (corresponding to nucleotides 193–2100 in Figure 4) and encodes for 359 amino acids. The deduced sequence shows perfect colinearity to the LAP peptides 64, 58b, 70,

Table I. Protein sequence analysis of LAP

Peptide	Sequence	cDNA nucleotides
N-Terminus	R S L R F V T L L Y R H G D R S P V K T Y	91– 153
85b	K T Y P K D P Y Q E E E W P Q/E G	145– 192
58b	Y H G F L N T S Y H	259– 288
86b	F N P N I S W Q P I P V	388– 424
70	D T L F C E Q T H G L R	622– 657
91b	F L F G I Y Q Q A E K	724– 756
64	D W Q Q E C Q L A	1093–1119

Amino acids are identified by the single-letter code. The numbering of nucleotides corresponds to the numbers in Figure 4.

Table II. Partial sequences of tryptic acid phosphatase peptides used for construction of mixed oligonucleotide probes (probes 139 and 140 for peptide 86b were synthesized in two portions)

Peptide	Amino acid sequence	Probe	Oligonucleotide
86b	Trp Gln Pro Ile Pro	139	3′ACC GT T GG T A G GG 5′ C C T
		140	3′ACC GT T GG T A G GG 5′ C G T
91b	Ile Tyr Gln Gln Ala	141	3′TAT G A T T AT GT GT CG 5′ C G C C

86b and 91b, which comprise a total of 54 amino acid residues. AP2 contains a 831-bp-long 3′ untranslated region with a potential polyadenylation signal (AATGAA) 25 bp upstream of a poly(A) tract. AP2 lacks the information for the N-terminus of LAP and represents therefore a partial cDNA.

Further screening was done with the 308-bp *Eco*RI–*Pst*I fragment from the 5′ end of AP2. Screening of the hepatoma-, placenta-, HepG2- and fibroblast cDNA libraries yielded 1, 42, 2 and 2 plaque-purified clones. Of these clones 10 had additional nucleotides at the 5′ end, as analysed by restriction mapping and sequencing (not shown). The most complete information was contained in the clone λCT29 isolated from the placenta cDNA library. The insert of λCT29 encodes 2112 bp (for sequencing strategy see Figure 3 and for nucleotide sequence Figure 4). The sequencing of the 5′ end of CT29, which is rich in G and C, proved to be difficult. The G at position 30 could only be detected by sequencing with T7 DNA polymerase using dITP or by chemical sequencing. Repeated sequencing of this region in different clones with reverse transcriptase consistently missed the G at position 30. This G was further confirmed by sequencing a genomic clone for LAP (unpublished). CT29 encodes a 12-bp 5′ untranslated region followed by a 1269-bp coding region that is initiated with a GTGATGG sequence. This sequence with a purine in position −3 and a G in position +4 conforms with the consensus sequence for an initiator methionine in eukaryotic mRNAs (Kozak, 1986). The GCAGCY (Y is C or T) sequence, which was found in five lysosomal enzymes overlapping the ATG start codon or following within four nucleotides (Quinn *et al.*, 1987), is not present in LAP. The 3′ non-coding region of CT29 is identical to that of AP2, except for a slightly longer poly(A) tract.

Deduced amino acid sequence of LAP

The 423 amino acid sequence (Figure 4) shows colinearity with the N-terminus of LAP and the LAP peptides shown in Table I. The sequence starts with a putative signal peptide of 30 amino acids at the N-terminus. The signal peptide has the following structural features characteristic of signal peptides (von Heijne, 1986): (i) a central hydrophobic domain (Arg-Ala-Ala-Leu-Leu-Gln-Leu-Leu-Leu-Gly-Val-Asn-Leu-Val-Val) 5′ flanked by a positively charged residue; (ii) a signal peptidase recognition sequence (Thr-Arg-Ala) at the putative cleavage site; and (iii) secondary structure breaking Pro residues in positions −4 and −5. The 393 amino acids following the putative signal peptide predict a size of 43 000 daltons and begin with a sequence that corresponds to the N-terminus of placental LAP. The predicted size is in good agreement with the size of the deglycosylated precursor of membrane associated LAP, which is 44 000 daltons in human fibroblasts (unpublished). The sequence of LAP contains eight potential *N*-glycosylation sites. Hydrophobic amino acids (Leu, Val and Ile) represent 20% of the total amino acids. The hydrophobicity profile (Figure 5) reveals a hydrophobic sequence of 27 amino acids (residues 379–405) close to the C-terminus of the LAP which is of sufficient length to span the membrane.

Homology

The predicted amino acid sequence was compared for homology with the known sequence of nine human, one mouse and two rat lysosomal enzymes, and two different human alkaline phosphatases (see Materials and methods). Using a block size of 30 residues a maximum of 33% homology was found for one LAP block with α-galactosidase A. For all other sequences the block homologies were ≤30%. A striking homology was found between LAP and human prostatic acid phosphatase, for which the sequence of the 23 N-terminal residues is known (Taga *et al.*, 1983). The homology between the N-termini of LAP and prostatic acid phosphatase is 60% (Figure 6).

mRNA-analysis and chromosomal mapping

When poly(A) RNA from human liver and HL-60 promyelocytes was hybridized with AP2, RNA species with a length of 2.3 kb were detected [shown for liver poly(A) RNA in Figure 7]. AP2 hybridized with a 21-kb fragment of *Eco*RI-digested human DNA and with a 5-kb fragment of *Eco*RI-digested mouse DNA (Figure 8). In a panel of

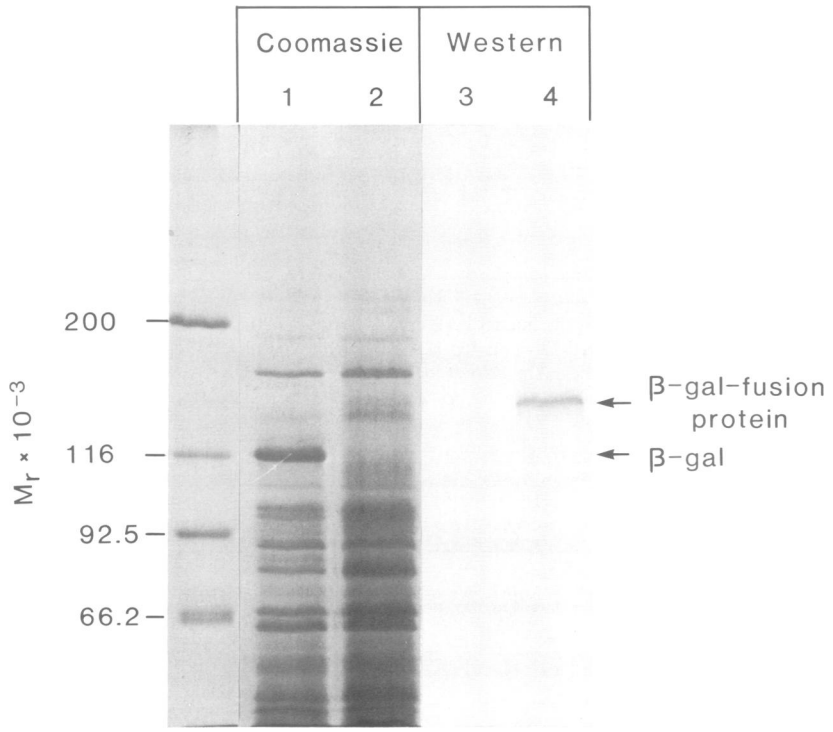

Fig. 1. Immunodetection of a β-galactosidase–human acid phosphatase fusion protein. Freeze-thaw lysates of BNN 103 λgt11 (**lanes 1** and **3**) and BNN 103 λAP2 (**lanes 2** and **4**) were subjected to SDS–PAGE. The gel was stained with Coomassie blue (**lanes 1** and **2**) or subjected to Western blot analysis using rabbit anti-human phosphatase Ig as first antibody (**lanes 3** and **4**). The migration of *E.coli* β-galactosidase (M_r 116 000) and the λAP2-encoded fusion protein (M_r 160 000) are indicated.

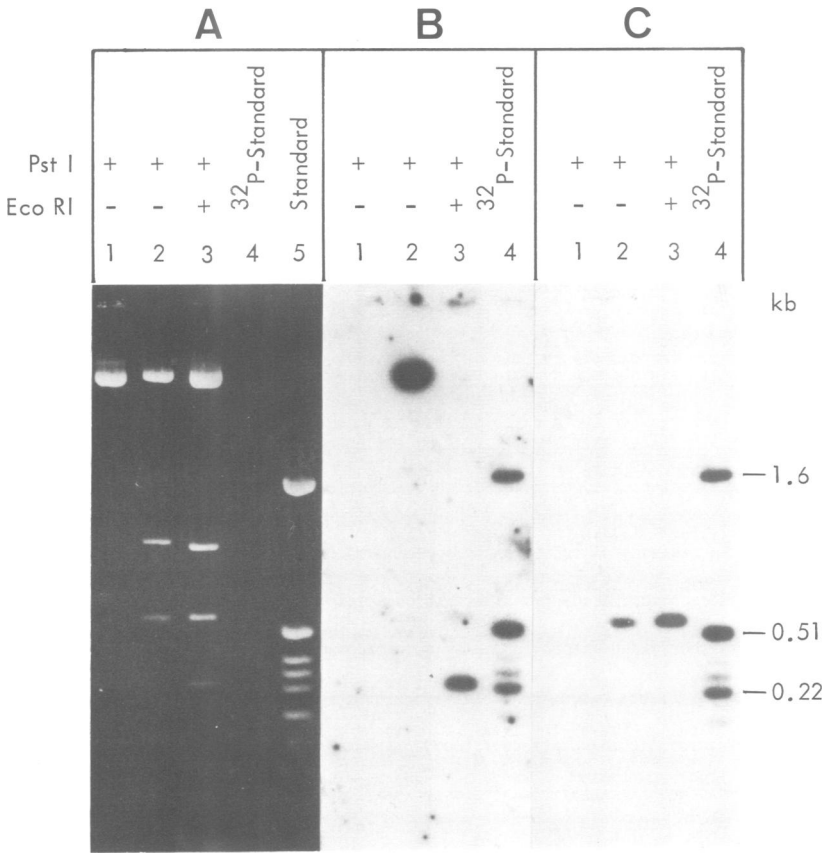

Fig. 2. Hybridization of AP2 with oligonucleotide probes. AP2 subcloned into M13mp18 was digested with *Pst*I and *Eco*RI and electrophoresed in a 0.7% agarose gel. The gel was stained with ethidium bromide (**A**) or hybridized wtih oligonucleotide 140 (**B**) and oligonucleotide 141 (**C**). **Lane 1**, *Pst*I digest of M13mp18; **lane 2**, *Pst*I digest of M13mp18-AP2; **lane 3**, *Eco*RI plus *Pst*I digest of M13mp18-AP2; **lane 4**; ^{32}P-labelled, *Hin*fI digest of pBR322; **lane 5**, *Hin*fI digest of pBR322.

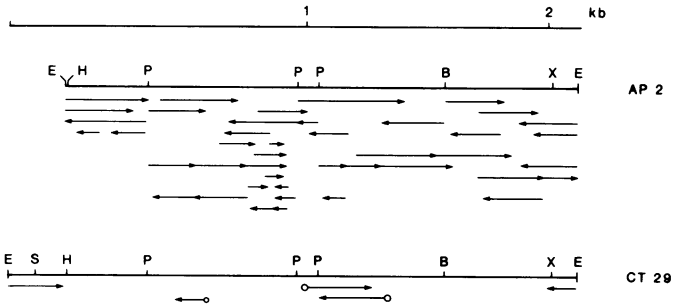

Fig. 3. Partial restriction map and sequencing strategy of the human LAP cDNA clones AP2 and CT29. The main restriction sites used in subcloning and DNA sequencing are shown (B = *Bam*HI; E = *Eco*RI; H = *Hin*cII; P = *Pst*I; S = *Sma*I; X = *Xba*I). Further sites that had been used are *Acc*I, *Ava*II, *Alu*I, *Bst*NI and *Hin*fI (not indicated). The arrows indicate the extent of DNA sequencing and the direction of the strand sequenced. The arrows starting from open circles indicate sequencing primed with synthetic oligonucleotides.

```
-10        1          10         20         30         40         50         60         70         80
ATTACAACGGTG ATG GCG GGC AAG CGG TCC GGC TGG AGC CGG GCG GCT CTC CTC CAG CTC CTT CTC GGC GTG AAC CTG GTG GTG ATG CCG CCC
             Met Ala Gly Lys Arg Ser Gly Trp Ser Arg Ala Ala Leu Leu Gln Leu Leu Leu Gly Val Asn Leu Val Val Met Pro Pro  27

        90         100        110        120        130        140        150        160        170
ACC CGG GCC CGG AGT CTG CGC TTC GTT ACC TTG CTG TAC CGC CAT GGA GAC CGT TCA CCA GTG AAG ACA TAT CCC AAG GAC CCC TAT CAG
Thr Arg Ala Arg Ser Leu Arg Phe Val Thr Leu Leu Tyr Arg His Gly Asp Arg Ser Pro Val Lys Thr Tyr Pro Lys Asp Pro Tyr Gln  57

       180        190        200        210        220        230        240        250        260
GAA GAA GAA TGG CCC CAG GGG TTT GGT CAG TTA ACC AAG GAG GGG ATG CTA CAG CAC TGG GAA CTG GGC CAG GCC CTG CGG CAG CGC TAT
Glu Glu Glu Trp Pro Gln Gly Phe Gly Gln Leu Thr Lys Glu Gly Met Leu Gln His Trp Glu Leu Gly Gln Ala Leu Arg Gln Arg Tyr  87

       270      *  280        290        300        310        320        330        340        350
CAC GGC TTC CTA AAC ACC TCT TAT CAC CGG CAA GAG GTT TAT GTG CGA AGC ACA GAC TTT GAC CGG ACT CTC ATG AGT GCT GAG GCC AAC
His Gly Phe Leu Asn Thr Ser Tyr His Arg Gln Glu Val Tyr Val Arg Ser Thr Asp Phe Asp Arg Thr Leu Met Ser Ala Glu Ala Asn 117

       360        370        380        390      * 400        410        420        430        440
CTG GCT GGA CTC TTC CCT CCC AAC GGG ATG CAG CGC TTC AAC CCG AAC ATC TCG TGG CAG CCT ATT CCT GTG CAC ACT GTG CCC ATC ACT
Leu Ala Gly Leu Phe Pro Pro Asn Gly Met Gln Arg Phe Asn Pro Asn Ile Ser Trp Gln Pro Ile Pro Val His Thr Val Pro Ile Thr 147
                                                          *                                                  *
       450        460        470        480        490        500        510        520        530
GAG GAC AGG CTG CTG AAG TTC CCG TTG GGC CCA TGT CCC CGT TAT GAG CAG CTG CAG AAC GAG ACC CGG CAG ACA CCA GAG TAT CAG AAT
Glu Asp Arg Leu Leu Lys Phe Pro Leu Gly Pro Cys Pro Arg Tyr Glu Gln Leu Gln Asn Glu Thr Arg Gln Thr Pro Glu Tyr Gln Asn 177

       540        550        560        570 *      580        590        600        610        620
GAG AGT TCT CGG AAT GCA CAA TTT CTG GAC ATG GTG GCC AAC GAG ACA GGG CTT ACA GAC CTG ACA CTG GAG ACC GTC TGG AAT GTC TAT
Glu Ser Ser Arg Asn Ala Gln Phe Leu Asp Met Val Ala Asn Glu Thr Gly Leu Thr Asp Leu Thr Leu Glu Thr Val Trp Asn Val Tyr 207

       630        640        650        660        670        680        690        700        710
GAC ACA CTC TTC TGT GAG CAA ACG CAC GGG CTG CGC CTG CCG CCC TGG GCC TCA CCC CAA ACC ATG CAG CGT CTC AGC CGG CTA AAG GAC
Asp Thr Leu Phe Cys Glu Gln Thr His Gly Leu Arg Leu Pro Pro Trp Ala Ser Pro Gln Thr Met Gln Arg Leu Ser Arg Leu Lys Asp 237

       720        730        740        750        760        770        780        790        800*
TTC AGC TTC CGC TTC CTC TTC GGA ATC TAC CAG CAG GCG GAG AAG GCC CGG CTT CAG GGG GGA GTC CTG CTG GCT CAG ATA AGG AAG AAC
Phe Ser Phe Arg Phe Leu Phe Gly Ile Tyr Gln Gln Ala Glu Lys Ala Arg Leu Gln Gly Gly Val Leu Leu Ala Gln Ile Arg Lys Asn 267

       810        820        830        840        850        860        870        880        890
CTG ACC CTA ATG GCG ACC ACC TCC CAG CTC CCC AAG CTG CTG GTT TAC TCT GCG CAC GAC ACT ACC CTG GTT GCC CTG CAA ATG GCA CTG
Leu Thr Leu Met Ala Thr Thr Ser Gln Leu Pro Lys Leu Leu Val Tyr Ser Ala His Asp Thr Thr Leu Val Ala Leu Gln Met Ala Leu 297

       900        910        920        930        940        950        960      *  970        980
GAT GTC TAC AAT GGT GAA CAA GCC CCC TAC GCC TCC TGC CAC ATA TTT GAA CTG TAC CAG GAA GAT TCT GGG AAT TTC TCA GTG GAG ATG
Asp Val Tyr Asn Gly Glu Gln Ala Pro Tyr Ala Ser Cys His Ile Phe Glu Leu Tyr Gln Glu Asp Ser Gly Asn Phe Ser Val Glu Met 327

      990  *     1000       1010       1020       1030       1040       1050       1060       1070
TAC TTT CGG AAC GAG AGT GAC AAG GCC CCC TGG CCG CTC AGC CTG CCT GGC TGC CCT CAC CGC TGC CCA CTG CAG GAC TTC CTT CGC CTC
Tyr Phe Arg Asn Glu Ser Asp Lys Ala Pro Trp Pro Leu Ser Leu Pro Gly Cys Pro His Arg Cys Pro Leu Gln Asp Phe Leu Arg Leu 357

      1080       1090       1100       1110       1120       1130       1140       1150       1160
ACA GAG CCC GTC GTG CCC AAG GAT TGG CAG CAG GAG TGC CAG CTG GCA AGC GGT CCT GCA GAC ACA GAG GTG ATT GTG GCC TTG GCT GTA
Thr Glu Pro Val Val Pro Lys Asp Trp Gln Gln Glu Cys Gln Leu Ala Ser Gly Pro Ala Asp Thr Glu Val Ile Val Ala Leu Ala Val 387

      1170       1180       1190       1200       1210       1220       1230       1240       1250
TGT GGC TCC ATC CTC TTC CTC CTC ATA GTG CTG CTC CTC ACC GTC CTC TTC CGG ATG CAG GCC CAG CCT CCT GGC TAC CGC CAC GTC GCA
Cys Gly Ser Ile Leu Phe Leu Leu Ile Val Leu Leu Leu Thr Val Leu Phe Arg Met Gln Ala Gln Pro Pro Gly Tyr Arg His Val Ala 417

      1260       1270       1280       1290       1300       1310       1320       1330       1340       1350       1360
GAT GGG GAG GAC CAC GCC TGA CAACCACTCAGCCCCCTTCCCTCCACCTCCTAGGGGAGGTGGGCTGGGCCCTCGCTCCTGACTGTTGCTGCTCCCCAGCCCATGGACAGGAG
Asp Gly Glu Asp His Ala end                                                                                          423

  1370       1380       1390       1400       1410       1420       1430       1440       1450       1460       1470       1480
ATCCTGGGTTGGGCCTCCCTCTGATGACCCCAGCCAGATGAGCGAGTGGGGCTCAGCGTGGCCCATGGTGCCTGTCACTCAGCATTCCCATGCCTGATGTTTACCAAGTGCTGTGTTGGA

  1490       1500       1510       1520       1530       1540       1550       1560       1570       1580       1590       1600
CACTGGCTTTCTCCAAACAGGATTTGCCTCCTCCACGCTCCCTACACACCTGAGATGTAAACTGGCAGTCAGTGTTCACTCAGGACCTAGGATTAGAAAATGGCAGAGTTGGTGCTGGAT

  1610       1620       1630       1640       1650       1660       1670       1680       1690       1700       1710       1720
CCACCTTGCACTTCTATCAAGCCCTGTTCTTTTTCCTCCAGCCTGAAGTCTTCGGCAAATAGCTCAGAGGGACACGGTCTTGCCTCTCAGTGCTTATTTTAGTGGGAAAAACAGCTAATA

  1730       1740       1750       1760       1770       1780       1790       1800       1810       1820       1830       1840
CCAGGGGTACAAACATTGGCTCCCAAGGAACTGGATCACCCAACAGCCAGCCAGCCACATTTCCCTGTGTCTGGCTAGAGCCACCATTAGACTCAGACAGAATGCTTCAGGAATCGTTGT

  1850       1860       1870       1880       1890       1900       1910       1920       1930       1940       1950       1960
CACCCCTTCAACTGGAGCAGGACGGAAGGTTGTCTGTACTTGGGAGGGAGTGGGGAGTGGTGGGAAGGGAGTCGCTTGTCACACGAATCAGGAAACTGCTCTCCCTCAGCTGGGCTGGGG

  1970       1980       1990       2000       2010       2020       2030       2040       2050       2060       2070      2080
TCTCCAGGGACCTGAGCTACATGCAGGTTGTGAGCTGGAAAAGAAAGTTCTAGACTGTGGCCCAGAATAGGGCTGGGGCAGCTCCCAGAGAATGAATAGTGCTGTTTCCTATTGGACTGA

  2090       2100
TAAAAAAAAAAAAAA
```

Fig. 4. Nucleotide sequence and deduced amino acid sequence of human LAP cDNA clones AP2 and CT29. The nucleotides are numbered in the 5′–3′ direction, starting with the first base following the linker used for cDNA cloning. The arrow indicates the N-terminus of the LAP. The signal peptide (amino acids 1–30) and the hydrophobic domain (amino acids 379–405) are boxed and the eight possible N-glycosylation sites are indicated by stars. The polyadenylation signal in the 3′ untranslated region is underlined. Bold underlines indicate amino acid sequences obtained by microsequencing (Table I). AP2 encompasses nucleotides 193–2110 of CT29.

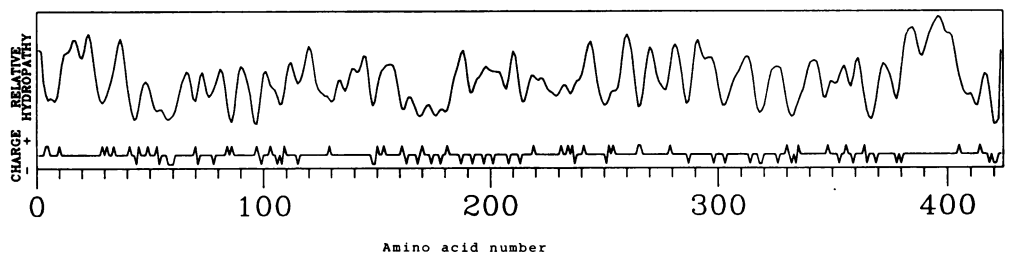

Fig. 5. Hydropathy plot of human LAP. The hydropathy indices of seven consecutive amino acids were calculated.

LAP	^{1}Arg	Ser	Leu	Arg	Phe	Val	Thr	Leu	Leu	^{10}Tyr	Arg	His	Gly	Asp	Arg	Ser	Pro	Val	Lys	^{20}Thr	Tyr	Pro	Lys
PAP	^{1}Lys	Glu	...	Lys	...	...	...	...	Val	^{10}Phe	...	...	...	...	...	...	...	Ile	Asp	^{20}...	Phe	...	Thr

Fig. 6. N-terminal amino acid sequences of human LAP and prostatic acid phosphatase (PAP). The N-terminal PAP sequence was published by Taga *et al.* (1983). Identical amino acids are represented by three dots (...) and non-homologous amino acids are indicated by the three-letter code.

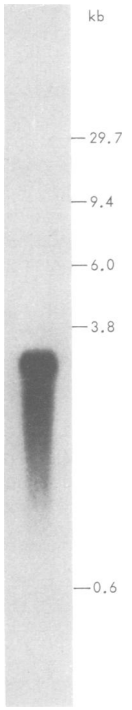

Fig. 7. Northern blot analysis of human liver RNA. Ten micrograms of human poly(A)$^+$ selected liver mRNA was electrophoresed in a 1% agarose gel and analysed by Northern blotting. ^{32}P-Labelled AP2 was used as a probe. Hybridization was performed as described in Materials and methods. The size of the acid phosphatase mRNA is ~2.3 kb.

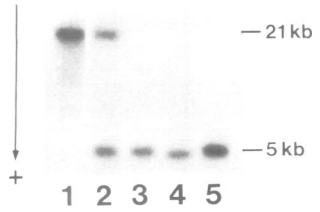

Fig. 8. Hybridization of AP2 with *Eco*RI-digested human and mouse DNA. Fragments of *Eco*RI-digested genomic DNA on a Southern blot were detected with nick-translated AP2. **Lane 1** contains human DNA (GM3104), **lanes 2–4** contain DNA from different human mouse somatic cell hybrids, **lane 5** contains mouse DNA (RAG).

human–mouse cell hybrids the human 21-kb fragment segregated with chromosome 11 (Table III).

Expression of LAP in COS cells

CT29 cDNA was inserted into the SV40 based mammalian expression vector pSVL, which allows transient expression in COS cells. Two days after transfection of COS cells with pSVL or pSVL-CT29 the activity of LAP was determined (Table IV). The LAP activity in cells transfected with pSVL-CT29 was 5- to 8-fold higher than in non-transfected or pSVL-transfected COS cells. About 40% of the LAP-activity was associated with membranes and required Triton X-100 for solubilization. This result demonstrates that CT29 encodes a functional LAP.

Discussion

A cDNA, CT29, was isolated that encoded human LAP. Authenticity of CT29 for LAP was shown by the colinearity of the nucleotide-deduced amino acid sequence with that determined by direct chemical sequencing of human placental LAP. Insertion of CT29 into the SV40-based mammalian expression vector pSVL and expression of the construct in monkey COS cells showed that CT29 contains the necessary nucleotide sequence for expression of an enzymatically active LAP in heterologous cells. In a panel of mouse–human somatic cell hybrids the homologous sequences in human genomic DNA segregated with human chromosome 11. This suggests that the phosphatase encoded by CT29 is identical with the human acid phosphatase ACP2, which has been localized to chromosome 11 using electrophoretic characterization of acid phosphatase activity in somatic cell hybrids (Bruns and Gerald, 1974; Busby *et al.*, 1976; Shows *et al.*, 1976; Jones and Kao, 1978; Kao *et al.*, 1978).

A homology of 60% was observed for the N-termini of LAP and human prostatic acid phosphatase. The available sequence information for prostatic acid phosphatase is restricted to the N-terminal 23 residues (Taga *et al.*, 1983). The homology may indicate that LAP and prostatic acid phosphatase are part of a gene family for tartrate inhibitable acid phosphatases.

The open reading frame of CT29 encodes 423 amino acids, of which the first 30 amino acids constitute a putative signal peptide. The N-terminus of placental LAP follows directly the putative signal peptide, suggesting that translocation of LAP is followed by removal of the signal peptide and that the proteolytic maturation, which is characteristic

Table III. Segregation of AP2 in a human–mouse somatic cell hybrid panel

Hybrid Clone	Human chromosome																							Chromosome marker	AP2
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	X	Y	
B82 MS2 1a-14-1		+	+	+			+	+			+		+		+	+	+	+	+		+	+	+	+		+
RAG ANLY 1	+		+	+	+	+						+		*				+			+					
RAG PI 7-2	+		+		+	+	+	+	+		*	+		+	*			+				+	*		Xp$^+$	+
RAG PI 5-15	+		+	+	+	+	+	+			+	+		+	*	+	+	+	+	+	+	+	*		Xp$^+$	+
RAG GM194 7	+		*	+	+			+		+		*							+		+		*		Xq$^+$	
RAG G0 4			+	+			+	+					+	+	+			+				+	*		Xp$^-$	
A9 SU 1-2		+	*			+							+	+				+			+				3p$^-$	
A9 MS89 9c-7			+	+			+				+				+	+		+		*	+	+	*		Xq$^+$	+
RAG GM194 6-13			*	+			+					+	+	+	*								*		3q$^-$,7q$^-$	
RAG MH 8-7	+		+	+			*	+				+	+	+	*			+					+		15q$^-$	
RAG GM194 5-5			*	+	+	+	+			+		+	+	+		+				+	+	+	+		3q$^-$	
RAG GM610 4-5-1	*		+	+	+	*						+	*	+	+			*		+		+	+		6q$^-$,1q$^-$	

+, Indicates that there is evidence for the presence of this chromosome both from cytogenetics and enzyme analysis. *, Indicates that a fragment of this chromosome is retained. Empty spaces indicate that no evidence for the presence of this chromosome has been obtained so far.

Table IV. LAP activity in transfected COS-cells

Transfected DNA (μg)	LAP (mU/mg protein)
–	8.6
pSVL (7)	9.4
pSVL-CT29 (7)	49.6
pSVL-CT29 (15)	78.6

of lysosomal enzymes (for review see Hasilik and von Figura, 1984), does not involve removal of N-terminal sequences of LAP. The sequence predicted for the translocated precursor of LAP (lacking the signal peptide) contains eight potential N-glycosylation sites. The extent, by which the N-glycosylation sites of LAP are utilized, may vary. In pulse-labelled fibroblasts soluble and membrane-associated precursors of LAP are observed, in which N-linked oligosaccharides account for masses between 14 000 and 40 000 daltons (unpublished).

Close to the C-terminus of the predicted sequence a hydrophobic stretch of 27 amino acids (residues 379–405) is found that could function as a membrane-spanning domain. In cells like fibroblasts ~40% of LAP activity is tightly associated with the lysosomal membrane and requires detergents for solubilization (Lemansky *et al.*, 1985; S. Gottschalk, unpublished data). The relation between membrane-associated and soluble LAP is unknown. The precursor as well as the mature forms of soluble and membrane-associated LAP differ in size, while no antigenic differences between soluble and membrane-associated LAP are detectable with antibodies against soluble LAP (unpublished). It is not clear whether the soluble and membrane-associated forms are products of a single mRNA or of different RNAs originating from one or several LAP genes. In the accompanying study (Waheed *et al.*, 1988) we describe the synthesis, processing and transport of CT29-encoded LAP in stably transfected baby hamster kidney cells.

Materials and methods

cDNA library screening

λgt11 cDNA expression libraries made from human hepatoma (De Wet *et al.*, 1984) (provided by Dr J.O.Brien, UCSD Medical School, San Diego, CA), Hep G 2 cells (provided by Dr M.Mueckler, Washington University Medical School, St Louis, MO) and human placenta (Clontech, Palo Alto, CA) and an Okayama-Berg cDNA library made from human SV40-transformed fibroblasts (provided by Dr R.Gravel, Hospital for Sick Children, Toronto) were used for screening. Antibody screening with affinity-purified polyclonal rabbit antibody against human placental acid phosphatase (Gieselmann *et al.*, 1984; Lemansky *et al.*, 1985) was done as described (De Wet *et al.*, 1984). Plaque and colony hybridization was done according to Maniatis *et al.* (1982).

Western blot

For detection of fusion proteins the procedure of De Wet *et al.* (1984) was used.

Oligonucleotide probes

Mixed oligonucleotide probes (see Table II) were synthesized with an Applied Biosystem solid phase synthesizer. These probes were kindly provided by Dr B.v.Wilcken-Bergmann (Universität Köln).

Oligonucleotide hybridization

Restriction-endonuclease-digested DNA was separated by electrophoresis in a 0.7% agarose gel, stained with ethidium bromide, dried, denatured and subjected to hybridization as recommended by Kidd (1984).

DNA probes

cDNA probes were labelled with [α-^{32}P]dATP (111 TBq/mmol, Amersham) utilizing random primers (Feinberg and Vogelstein, 1984). Oligonucleotides were labelled at their 5′ ends using [^{32}P]ATP (185 TBq/mmol, Amersham) and T4-polynucleotide kinase (New England Biolabs).

Isolation and electrophoresis of RNA

RNA was prepared from human liver and HL-60 promyelocytes by the guanidinium isothiocyanate method of Chirgwin *et al.* (1979). Poly(A)$^+$ RNA was selected by oligo(dT)–cellulose chromatography according to Aviv and Leder (1972). RNA samples were electrophoresed on 1% agarose gels after denaturation in 6.6 M glyoxal according to Thomas (1983). Transfer to Hybond-N membranes, prehybridization and hybridization were carried out according to the supplier's manual (Amersham). Filters were hybridized in 6 × SSC, 5 × Denhardt's, 0.1% SDS, 100 μg/ml salmon sperm DNA, 50% formamide at 42°C for 16 h and washed three times with 2 × SSC, 0.1% SDS and once with 0.2 × SSC, 0.1% SDS at 68°C for 30 min (Maniatis *et al.*, 1982).

DNA sequencing

Restriction fragments were subcloned into M13mp18 or mp19 using standard methods (Messing *et al.*, 1981) and sequencing by the dideoxy chain termination method of Sanger *et al.* (1977) with [α-^{35}S]dATP (22 TBq/mmol, Amersham) according to Biggin *et al.* (1983). The T7 DNA polymerase sequencing system was obtained from Renner, Dannstadt. Sequencing was carried out according to the supplier's manual. Chemical sequencing was performed using the method of Maxam and Gilbert (1980).

Amino acid sequencing

Acid phosphatase (1 mg) purified from human placenta (Gieselmann *et al.*, 1984) was reduced and carboxymethylated. Tryptic digestion of the purified acid phosphatase was carried out at a trypsin to protein ratio of 1/100. Peptides were separated by reverse-HPLC using an LKB gradient system and a Baker RP18 wide-pore column (4.6 × 250mm). The column was eluted with a solvent gradient from 0 to 70% acetonitrile in 0.1% trifluoroacetic acid. Single fractions were rechromatographed with the same acetonitrile gradient in 0.1% ammoniumacetate. The sequence analysis of the peptides was performed on an Applied Biosystems gasphase sequencer (Hewick *et al.*, 1981) with off-line analysis of PTH amino acids (Beyreuther *et al.*, 1983).

Transfection

Transfection was performed using 6×10^5 COS cells/60-mm dish. DNA (7–15 μg) was added according to the methods of Gorman *et al.* (1982) and Graham and van der Eb (1973). L-Tartrate inhibitable phosphatase activity in the cell homogenate, saponin and Triton X-100 extracts of cells was measured 48 h after transfection using *p*-nitrophenylphosphate as substrate (Lemansky *et al.*, 1985).

Computer analysis

The hydrophobicity profile of amino acid sequences was generated by a program obtained from and described by Novotny and Auffray (1984) using the hydrophobicity parameters given by Rose and Roy (1980). Routinely three smoothing cycles and a window size of 7 were used. The homology between the amino acid sequence coded by CT29 and those published for other hydrolases was tested. A program using the algorithm of Wilbur and Lipman (1983) was applied for analysis of homology between LAP and two human placental alkaline phosphatases (Millán, 1986; Kam *et al.*, 1985) and the following lysosomal enzymes: cathepsin D (Faust *et al.*, 1985), cathepsin B (Chan *et al.*, 1986; Fong *et al.*, 1986), α-fucosidase (Fukushima *et al.*, 1985), α-galactosidase A (Bishop *et al.*, 1986; Tsuji *et al.*, 1987), β-hexosaminidase α- and β-chain (Myerowitz *et al.*, 1985; O'Dowd *et al.*, 1985), β-glucuronidase (Guise *et al.*, 1985), glucocerebrosidase (Sorge *et al.*, 1985; Tsuji *et al.*, 1986), α-glucosidase (Martiniuk *et al.*, 1986), all of human origin; of rat lysosomal β-glucuronidase (Nishimura *et al.*, 1986), rat cathepsin H (Takio *et al.*, 1983) and of mouse major excreted glycoprotein, which is believed to be the analogue for cathepsin L (Troen *et al.*, 1987).

Chromosomal localization

The panel of 27 mouse–human cell hybrids generated by fusion of mouse RAG or A9 cells and human fibroblasts originating from various translocation carriers, has been characterized by isozyme and cytogenetic analyses for the presence of human chromosomes and chromosome fragments (Balazs *et al.*, 1984; Camerino *et al.*, 1984; Zabel *et al.*, 1984). Hybrid cell DNAs were digested with *Eco*RI, separated by agarose (0.8%) electrophoresis and transferred to Biodyne A membranes (Pall Filtrationstechnik, D-6072 Dreieich). Hybridization of nick-translated AP2 was as in Maniatis (1982).

Acknowledgements

We thank Drs J.O'Brien, J.De Wet, N.Dewji (UCSD, San Diego) and A.Unterbeck (Universität Köln) for introducing us into the methodology, Dr B.v.Wilcken-Bergmann (Universität Köln) for the synthesis of oligonucleotides, and Dr C.Peters (ZMBH-Heidelberg) for chemical sequencing. This work was supported by the Deutsche Forschungsgemeinschaft and the Fonds der Chemischen Industrie.

References

Aviv,H. and Leder,P. (1972) *Proc. Natl. Acad. Sci. USA*, **69**, 1408–1412.
Balazs,J., Purello,M., Kurnit,D.M., Grzeschik,K.H. and Siniscalco,M. (1984) *Somatic Cell Molec. Genet.*, **10**, 385–397.
Beyreuther,K., Biesler,B,. Bowens,J., Diltrop,R., Neifer,K., Stüber,K., Zais,S., Ehring,R. and Zabel,P. (1983) *Modern Methods in Protein Chemistry*. Walter de Gruyter & Co., Berlin, pp. 303–325.
Biggin,M.C., Gibson,T.J. and Hong,G.F. (1983) *Proc. Natl. Acad. Sci. USA*, **80**, 3963–3965.
Bishop,D.F., Calhoun,D.H., Bernstein,H.S., Hantzopoulos,P., Quinn,M. and Desnick,R.J. (1986) *Proc. Natl. Acad. Sci. USA*, **83**, 4859–4863.
Bruns,G.A.P. and Gerald,P.S. (1974) *Science*, **184**, 480–481.
Busby,N., Courval,J. and Francke,U. (1976) *Cytogenet. Cell Genet.*, **16**, 105–107.
Camerino,G., Grzeschik,K.H., Jaye,M., de la Salle,H., Tolstoshev,P., Lecocq,J.P., Heilig,R. and Mandel,J.L. (1984) *Proc. Natl. Acad. Sci. USA*, **81**, 498–502.
Chan,S.J., San Segundo,B, McCormick,M.B. and Steiner,D.F. (1986) *Proc. Natl. Acad. Sci. USA*, **83**, 7721–7725.
Chirgwin,J.M., Przybyla,A.E., MacDonald,R.J. and Rutter,W.J. (1979) *Biochemistry*, **18**, 5294–5299.
de Duve,C. (1983) *Eur. J. Biochem.*, **137**, 391–397.
De Wet,J., Fukushima,H., Dewji,N., Wilcox,E., O'Brien,J. and Helinski,D.R. (1984) *DNA*, **3**, 437–447.
Farquhar,M.G. and Palade,G.E. (1981) *J. Cell Biol.*, **91**, 77–103.
Faust,I.L., Kornfeld,S. and Chirgwin,J.M. (1985) *Proc. Natl. Acad. Sci. USA*, **82**, 4910–4914.
Feinberg,A.P. and Vogelstein,B. (1984) *Anal. Biochem.*, **137**, 266–267.
Fong,D., Calhoun,D.H., Hsieh,W.T., Lee,B. and Wells,R.D. (1986) *Proc. Natl. Acad. Sci. USA*, **83**, 2909–2913.
Fukushima,H., De Wet,J.R. and O'Brien,J.S. (1985) *Proc. Natl. Acad. Sci. USA*, **82**, 1262–1265.
Gieselmann,V., Hasilik,A. and von Figura,K. (1984) *Hoppe-Seyler's Z. Physiol. Chem.*, **365**, 651–660.
Gorman,C.M., Moffat,L.F. and Howard,B. (1982) *Mol. Cell Biol.*, **2**, 1044–1051.
Graham,F.L. and van der Eb,A.J. (1973) *Virology*, **52**, 456–457.
Griffiths,G. and Simons,K. (1986) *Science*, **234**, 438–443.
Guise,K.S., Korneluk,R.G., Waye,J., Lamhonwah,A.M., Quan,F., Paler, R., Ganschow,R.E., Sly,W.S. and Gravel,R.S. (1985) *Gene*, **34**, 105–110.
Hasilik,A. and von Figura,K. (1984) In Dingle,J.T., Dean,R.T. and Sly,W. (eds), *Lysosomes in Biology and Pathology*. Elsevier, Amsterdam, pp. 3–16.
Hewick,R.M., Hunkapillar,M.W., Hood,L.E. and Dreyer,W. (1981) *J. Biol. Chem.*, **256**, 7990–7997.
Jones,C. and Kao,F.T. (1978) *Hum. Genet.*, **45**, 1–10.
Kam,W., Clauser,E., Kim,Y.S., Kan,Y.W. and Rutter,W.J. (1985) *Proc. Natl. Acad. Sci. USA*, **82**, 8715–8719.
Kao,F.T., Jones,C., Law,M.L. and Puck,T.T. (1978) *Cytogenet. Cell Genet.*, **22**, 474–477.
Kidd,V.J. (1984) *BRL-Focus*, **6** (3), 3–4.
Kozak,M. (1986) *Cell*, **44**, 283–292.
Lemansky,P., Gieselmann,V., Hasilik,A. and von Figura,K. (1985) *J. Biol. Chem.*, **260**, 9023–9030.
Luchter-Wasyl,E. and Ostrowski,W. (1974) *Biochim. Biophys. Acta*, **365**, 349–359.
Maniatis,T., Fritsch,E.F. and Sambrook,J. (eds) (1982) *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
Martiniuk,F., Mehler,M., Pellicer,A., Tzall,S., La Badie,G., Hobart,C., Ellenbogen,A. and Hirschhorn,R. (1986) *Proc. Natl. Acad. Sci. USA*, **83**, 9641–9644.
Maxam,A.M. and Gilbert,W. (1980) *Methods Enzymol.*, **65**, 499–560.
Messing,J., Crea,R. and Seeburg,P.H. (1981) *Nucleic Acids Res.*, **9**, 309–321.
Millán,J.L. (1986) *J. Biol. Chem.*, **261**, 3112–3115.
Myerowitz,R., Piekarz,R., Neufeld,E.F., Shows,T.B. and Suzuki,K. (1985) *Proc. Natl. Acad. Sci. USA*, **82**, 7830–7834.
Nishimura,Y., Rosenfeld,M.G., Kreibich,G., Gubler,U., Sabatini,D.D., Adesnik,M. and Andy,R. (1986) *Proc. Natl. Acad. Sci. USA*, **83**, 7292–7296.
Novotny,J. and Auffray,C. (1984) *Nucleic Acids Res.*, **12**, 243–255.
O'Dowd,B.F, Quan,F., Willerd,H.F., Lamhonwah,A.M., Korneluk,R.G., Lowden,J.A., Gravel,R.A. and Mahuran,D.J. (1985) *Proc. Natl. Acad. Sci. USA*, **82**, 1184–1188.
Quinn,M., Hantzopoulos,P., Fidanza,V. and Calhoun,D.H. (1987) *Gene*, **58**, 177–188.
Rose,G.D. and Roy,S. (1980) *Proc. Natl. Acad. Sci. USA*, **77**, 4643–4647.
Saini,M.S. and van Etten,R.L. (1978) *Arch. Biochem. Biophys.*, **191**, 613–624.
Sanger,F., Nicklen,S. and Coulson,A.R. (1977) *Proc. Natl. Acad. Sci. USA*, **74**, 5463–5467.
Shows,T.B., Brown,J.A. and Lalley,P.A. (1976) *Cytogenet. Cell Genet.*, **16**, 231–234.
Sorge,J., West,C., Westwood,B. and Beutler,E. (1985) *Proc. Natl. Acad. Sci. USA*, **82**, 7289–7293.
Taga,E.M., Moore,D.L. and van Etten,R.L. (1983) *The Prostate*, **4**, 141–150.
Takio,K., Towatari,T., Katnuama,N., Teller,D.C. and Titani,K. (1983) *Proc. Natl. Acad. Sci. USA*, **80**, 3666–3670.
Thomas,P.S. (1983) *Methods Enzymol.*, **100**, 255–266.
Troen,B.R., Gal,S. and Gottesman,M.M. (1987) *Biochem. J.*, **246**,

731–735.
Tsuji,S., Choudary,P.V., Martini,B.M., Winfield,S., Barranger,J.A. and Ginns,E.I. (1986) *J. Biol. Chem.*, **261**, 50–53.
Tsuji,S., Martin,B.M., Kaslow,D.C., Migeon,B.R., Choudary,P.V., Stubblefield,B.K., Mayor,J.A., Murray,G.J., Barranger,J.A. and Ginns,E.I. (1987) *Eur. J. Biochem.*, **165**, 275–280.
von Heijne,G. (1986) *Nucleic Acids Res.*, **14**, 4683–4690.
Waheed,A. and van Etten,R.L. (1985) *Arch. Biochem. Biophys.*, **243**, 274–283.
Waheed,A., van Etten,R.L., Gieselmann,V. and von Figura,K. (1985) *Biochem. Genet.*, **23**, 309–319.
Waheed,A., Gottschalk,S., Hille,A., Krentler,C., Pohlmann,R., Hauser,H., Geuze,H. and von Figura,K. (1988) *EMBO J.*, **7**, 2351–2358.
Wilbur,W.J. and Lipman,D.J. (1983) *Proc. Natl. Acad. Sci. USA*, **80**, 726–730.
Yam,L.T., Li,C.Y. and Lam,K.W. (1980) In Spring-Mills,E. and Hafez, E.S.E. (eds), *Male Accessory Sex Glands*. Elsevier/North Holland, Amsterdam, pp. 183–196.
Zabel,B.V., Nayler,S.L., Grzeschik,K.H. and Sakaguchi,A.Y. (1984) *Somatic Cell Molec. Genet.*, **10**, 105–108.

Received on January 25, 1988; revised on April 5, 1988