Cloning, sequencing, and chromosomal localization of human term placental alkaline phosphatase cDNA

(DNA sequence analysis/synthetic oliogodeoxynucleotide probe/amino acid sequence identity/isozymes/serine hydrolase)

WING KAM*, ERIC CLAUSER^{†§}, YOUNG S. KIM*, YUET WAI KAN[‡], AND WILLIAM J. RUTTER[†]

*Gastrointestinal Research Laboratory, Veterans Administration Medical Center, San Francisco, CA 94121; †Hormone Research Laboratory, University of California, San Francisco, CA 94143; and ‡Howard Hughes Medical Institute, University of California, San Francisco, CA 94143

Contributed by William J. Rutter, August 15, 1985

ABSTRACT A human term (third trimester) placental alkaline phosphatase (PLAP; EC 3.1.3.1) cDNA was isolated from a human placental $\lambda gt11$ cDNA library. The expression library was screened by using rabbit antibodies against PLAP and oligonucleotide probes. DNA sequence analysis of a positive clone with an insert of 2.7 kilobase pairs allowed us to predict the complete amino acid sequence of PLAP (530 residues), which coincided with the reported 42 N-terminal amino acid sequence of PLAP except at position 3. Contrary to the previous supposition that there was no amino acid sequence homology between PLAP and Escherichia coli alkaline phosphatase (471 residues), we found 30% overall homology, with regions of strong homology including the putative active site and the metal-binding sites. The 44-residue C-terminal extension of PLAP has a stretch of 17 hydrophobic amino acids, which presumably anchors the protein to the plasma membrane, a change perhaps necessary for the transition from a bacterial periplasmic enzyme to a mammalian membraneassociated enzyme. We have also localized PLAP-related DNA sequences mainly on chromosome 2 and to a lesser degree on chromosome 17. It seems likely therefore that the PLAP gene resides on chromosome 2 and other member(s) of the alkaline phosphatase family may exist (on this chromosome and) on chromosome 17.

Alkaline phosphatase [orthophosphoric-monoester phosphohydrolase (alkaline optimum), EC 3.1.3.1] is a metalloenzyme that hydrolyzes phosphate esters with a peculiarly high pH optimum (pH 10–10.5). Its presence in nearly all living organisms, with the exception of some higher plants, argues for its physiologic importance, but its exact physiological role is not known (1). Human alkaline phosphatase may consist of five tissue-specific forms or isozymes: term placental (PLAP; Regan isozyme); intestinal (IAP), tissue unspecific (found in liver, bone, and kidney), PLAP-like (Nagao isozyme), and IAP-like (Kasahara isozyme) (2). Of these, PLAP is of interest because of its genetic polymorphism, reported in 1961 (3), and its use as a tumor marker owing to its ectopic expression in some human cancer tissue and cell lines (2, 4, 5).

PLAP is a homodimer, membrane-associated glycoprotein enzyme with a 58- to 70-kDa subunit (2, 6, 7). It has immunological cross-reactivity to IAP and PLAP-like but not to tissue-unspecific alkaline phosphatase, and it differs from IAP in heat stability, enzyme inhibition, and monoclonal antibody reactivity (2). Surprisingly little is known about the structure of human alkaline phosphatases, except for a recent determination of the N-terminal 42 amino acids of PLAP (7). A better understanding of the structural differences in the various forms of this human enzyme and their expression may improve their clinical diagnostic values. Here we report



FIG. 1. Schematic representation of human PLAP cDNA clone $\lambda AP27$. Untranslated sequences are represented by a thick line. Coding sequences are represented by a rectangle: the black portion represents the putative signal peptide; the hatched region encodes the protein.

the cloning, complete nucleotide sequencing, and chromosomal localization of related sequences of a PLAP cDNA. Perhaps because the prokaryotic alkaline phosphatase is soluble and the PLAP is membrane bound, it has been assumed that the molecules were quite different. However, the predicted amino acid sequence of PLAP shows homology with the recently reported sequence of *Escherichia coli* alkaline phosphatase (8, 9), especially in the regions of the putative active site and the metal-binding sites.

MATERIALS AND METHODS

Screening of Human Placental cDNA Expression Library with Antibody and Mixed Oligonucleotide Probes. An oligo(dT)-primed human placental \gt11 cDNA library, kindly provided by Ebina et al. (10), was screened with a polyclonal antibody to PLAP (Dako, Santa Barbara, CA), using an immunoscreening kit, CLIK (Clontech, Palo Alto, CA) (11). Further screening was performed by hybridizing positive clones from the antibody screening with two sets of ³²P-labeled, 26-base oligonucleotide probes (12) (5'-TTCC-AGAAGTCGGGGTTCTCCTCC3' and 5'-TTCCAGAA- ${}_{G}^{A}TC_{T}^{C}GG_{G}^{A}TT_{T}^{C}TC_{T}^{C}TC_{T}^{C}TC - 3')$, which were synthesized with a Beckman System 1 automated DNA synthesizer (13) on the basis of the amino acid sequences of the N terminus (residues 5 to 13, see Fig. 1). These probes were labeled at the 5' end with $[\gamma^{-32}P]ATP$ (7000 Ci/mmol; 1 Ci = 37 GBq) and T4 polynucleotide kinase by the standard method (ref. 14, pp. 122-123)

Subcloning of λ gt11 Inserts and Preparation of Cloned cDNA. Insert cDNA from *Eco*RI digestion of the positive clones was ligated to the *Eco*RI-digested pMT21 vector (15). Plasmid DNA was prepared from the transformed *E. coli* HB101 cells and was purified by gradient centrifugation in cesium chloride (ref. 14, pp. 88–94).

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. \$1734 solely to indicate this fact.

Abbreviations: PLAP, placental alkaline phosphatase; IAP, intestinal alkaline phosphatase; kb, kilobase(s).

[§]Present address: Institut National de la Santé et de la Recherche Médicale U36, 17 Rue du Fer a Moulin, Paris 75005, France.

TGCACTCCAGGCTGCCCGGGTTTGC	-17 The field leu leu leu leu leu leu leu ATG CTG CTG CTG CTG CTG CTG CTG CTG	49	330 ser arg ala tyr arg ala leu thr glu thr ile met phe asp asp AGC AGG GCT TAC COG GCA CTG ACT GAG ACG ATC ATG TTC GAC GAC 1	1084
A gly leu arg leu gln leu se GCC CTC ACC CTA CAC CTC TC	1 ar leu gly ile ile leu val glu glu CC CTG CGC ATC ATC CTA GTT GAG GAG	94	340 ala ile glu arg ala gly gln leu thr ser glu glu asp thr leu GCC ATT GAG AGG GCG GCC CAG CTC ACC AGC GAG GAG GAG AGG CTG 1	1129
10 glu asn pro asp phe trp as GAG AAC CCC GAC TTC TCC AA	20 sn arg glu ala ala glu ala leu gly AC CGC GAG GCA GCC GAG GCC CTG GGT	139	360 ser leu val thr ala asp his ser his val phe ser phe gly gly AGC CTC GTC ACT GCC GAC CAC TCC CAC GTC TTC TCC TTC GGA GGC 1	1174
ala ala lys lys leu gln pr GCC GCC AAG AAG CTG CAG CC	30 To ala gln thr ala ala lys asn leu TT GCA CAG ACA GCC GCC AAG AAC CTC	184	370 380 tyr pro leu arg gly ser ser phe ile gly leu ala ala gly lys TAC CCC CTG CGA GOG AGC TCC TTC ATC GOG CTG GCC GCT GOC AAG 1	1219
40 ile ile phe leu gly asp gl ATC ATC TTC CTG GGC GAT GC	50 Y met gly val ser thr val thr ala X ATG COG GTG TCT ACG GTG ACA CCT	229	390 ala arg asp arg lys ala tyr thr val leu leu tyr gly asn gly GCC COG GAC AGG AAG GCC TAC ACG GTC CTC CTA TAC GGA AAC GGT 1	1264
ala arg ile leu lys gly gl GCC AGG ATC CTA AAA GGG CA	60 In lys lys asp lys leu gly pro glu IG AAG AAG GAC AAA CTG GGG CCT GAG	274	400 410 pro gly tyr val leu lys asp gly ala arg pro asp val thr glu CCA GGC TAT GTG GTC AAG GAC GGC GGC CGG GGT GTT ACC GAG 1	1309
ile pro leu ala met asp ar ATA CCC CTG GCC ATG GAC CC	rg phe pro tyr val ala leu ser lys C TTC CCA TAT GTG GCT CTC TCC AAG	319	420 ser glu ser gly ser pro glu tyr arg gln gln ser ala val pro AGC GAG AGC GGG AGC CCC GAG TAT COC CAG CAG TCA GCA GTG CCC 1	1354
B ACA TAC AAT GTA GAC AAA CA	s val pro asp ser gly ala thr ala AT GTG CCA CAC AGT GCA GCC ACA GCC	364	430 leu asp glu glu thr his ala gly glu asp val ala val phe ala CTG GAC GAA GAG ACC CAC GCA GOC GAG GAC GTG GCG GTG TTC GCG 1	1399
thr ala tyr leu cys gly va ACG GCC TAC CTG TGC GGG GI	11 lys gly ash phe gln thr ile gly C AAG GOC AAC TTC CAG ACC ATT GOC	409	arg gly pro gln ala his leu val his gly val gln glu gln thr COC GOC CCC CAG CCG CAC CTG GTT CAC CGC GTG CAG GAG CAG ACC 1	444
leu ser ala ala ala arg ph TTG AGT GCA GCC GCC GCC TT	the asn gln cys asn thr thr arg gly T AAC CAG TGC AAC ACG ACA CGC GGC	454	470 phe ile ala his val met ala phe ala ala cys leu glu pro tyr TTC ATA GCG CAC GTC ATG GCC TTC GCC GCC TGC CTG GAG CCC TAC 1.	489
asn glu val 11e ser val me AAC GAG GTC ATC TCC GTG AT	at asn arg ala lys lys ala gly lys G AAT COG CCC AAG AAA CCA COG AAG	499	thr ala cys asp leu ala pro pro ala gly thr thr asp ala ala ACC GCC TGC GAC CTG GCG CCC GCC GCC ACC ACC GAC GCC GCG 11	534
ser val gly val val thr th TCA GTG GGA GTG GTA ACC AC	ar thr arg val gln his ala ser pro	544 C	• his pro gly arg ser val val pro ala leu leu pro leu leu ala • CAC CCG CCC CCC CTC CTC CTC CTC CTC CTC CT	579
ala gly thr tyr ala his th GCC GGC ACC TAC GCC CAC AC	170 nr val asn arg asn trp tyr ser asp 3G GTG AAC CGC AAC TOG TAC TCG GAC	589	gly thr leu leu leu leu glu thr ala thr ala pro OP <u>GGC ACC CTC CTC CTC CTC CTC</u> GAG ACC GCC ACT GCT CCC TGA GTGTCCCCGT 10	1627
ala asp val pro ala ser al GCC GAC GTG CCT GCC TCG GC	180 a arg gln glu gly cys gln asp ile CC CGC CAG GAG GGG TGC CAG GAC ATC	634	CCCTOBODETCCTGCTTCCCCATCCCGGAGTTCTCCTGCTCCCCACCTCCTGTGTGTCCTG CCTOGCCTCCAGCCCGAGTCGTCATCCCCGGAGTCCCTATACAGAGGTCCTGCCATGGAA 1'	.687 1 747
190 ala thr gln leu ile ser as GCT ACG CAG CTC ATC TCC AA	n met asp ile asp val ile leu gly C ATC GAC ATT GAT GTG ATC CTA GGT	679	CCTTCCCCTCCCCGTCCCCTCTCCCCACCCATCACCCCAAACCCTGCCCCTTCGCT 11 CCTCTCCCACTCCCTACCCCAACCCCACCACCCACGCACTGCACCCTGCCCTGCCA 1:	.807 1867
gly gly arg lys tyr met ph GGA GGC CGA AAG TAC ATG TT	e arg met gly thr pro asp pro glu T CGC ATG GGA ACC CCA CAC CCT GAG	724	CCCCAGGAAAGGAGGGGGCTCAGGCCATCCACCCACCCAC	.927 1987
tyr pro asp asp tyr ser gl TAC CCA GAT GAC TAC AGC CA	n gly gly thr arg leu asp gly lys A GGT GGG ACC AGG CTG GAC GGG AAG	769	CAGTECTTGAATCACCTGTGGGACTTGAGGACTCGGGATCTTCAGGACGCCTGGAGAAGG 21 GTGGTTTCCTGCCACCCTGCTGGCCAAGGACGCTCCTGGGGTGGGGATCACCAGGGGGAT 2	2047 2107
asn leu val gln glu trp le AAT CTC GTC CAC GAA TOG CT	u gly glu arg gln gly ala arg tyr C GOC GAA CGC CAG GGT GCC CGG TAC	814	TTTGACACAGCCTTCGGCCGCCCCCACTAAGTTAATTCCACACCCCCGTGTACCCCCCACGG 2: GGGCCCTCTGCCCCCAAAAGGCTTGCCCCAAATCTCAACCTTCTCAGACGTTCCATAC 2:	2227
* 250 val trp asn arg thr glu le GTG TOG AAC COC ACT GAG CI	u met gln ala ser leu asp pro ser °C ATG CAG GCT TCC CTG GAC CCG TCT	859	CCCCACATGCCAATTTCAGCACCCAACTGAGATCCGAGGAGCTCCTOGGAAGCCCTOGGT 2 GCAGGACACTGGTCGAGAGCCAAAGGTCCCTCCCCAGACATCTGGACACTGGGCATAGAT 2	287 2347
27 val thr his leu met gly le GTG ACC CAT CTC ATG GGT CT	0 nu phe glu pro gly asp met lys tyr C TTT GAG CCT CGA GAC ATG AAA TAC	904	TTCTCAAGAAGGAAGACTCCCCTCCCCCAGGCCCTCTGCTCTCCTGGGAGACAAAGC 24 AATAATAAAAGGAAGTGTTTGTAATCCCAGCACTTTGGGAGGCCGAGGTGGGCGGATCAC 24	!407 }467
280 glu ile his arg asp ser th GAG ATC CAC CGA GAC TCC AC	290 ar leu asp pro ser leu met glu met A CTG GAC CCC TCC CTG ATG GAG ATG	949	GAGGTCAGGAGATGGAGACCATCCTGGCTAACACGGTGAAACCCCTTATCTATGCGCCGG 28 TAGTCCCAGGTACCCAGGAGGCTGAAGCAGGATAATCGCTTGAACCCCGGGGGGGG	:527 2587
thr glu ala ala leu arg le ACA GAG GCT GCC CTG CGC CT	300 Du leu ser arg his pro arg gly phe TG CTG AGC AGA CAC CCC CGC GGC TTC	994	TOCAGTGAGCCGAGGTCATGCCACTGCACTGCAGCCTGGGGGGAGAGCGAGATTCTGC 24 CTCAAAAAATAAACAAATAAATAAATAAATAAATAAATA	2647 2698
310 phe leu phe val glu gly gl TTC CTC TTC GTC GAG GGT GG	320 y arg ile asp his gly his his glu T CGC ATC GAC CAT GGT CAT CAT GAA	1039		

FIG. 2. Nucleotide and predicted amino acid sequences of the term human PLAP. Nucleotides are numbered at the right; amino acids are numbered throughout. The amino acid sequence used to design the synthetic DNA probe is underlined with a broken line. Sites of potential asparagine-linked glycosylation are indicated by asterisks. Boxes indicate the putative signal peptide (box A, as indicated on the left), phosphate-binding serine active site (box B), and transmembrane region (box C). The poly(A) addition sequences (AATAAA) are underlined.

Restriction Endonuclease Mapping. Restriction mapping of the cDNA insert of $\lambda gt11$ was performed by single or double endonuclease digestions of the recombinant phage, according to the suppliers' recommendations (Bethesda Research Laboratories, New England Biolabs, and Boehringer-Mannheim).

DNA Sequence Analysis. The entire insert and the restriction fragments were subcloned in phage M13mp18 and M13mp19. Single-stranded phage DNA was prepared as described (16). The entire DNA sequence of one of these inserts, λ AP27, was determined by sequencing of both strands by the dideoxynucleotide chain termination technique (17) using deoxyadenosine 5'-[α -[³⁵S]thio]triphosphate (650 Ci/mmol; Amersham) or [α -³²P]dCTP (3000 Ci/mmol; Amersham), and the 17-mer universal primer and, in selected areas, specific oligonucleotide primers.

RNA Blot Hybridization Analysis. Poly(A)⁺ RNA from human placenta, human PLC/PRF/5 hepatoma cells, human liver, and mouse testis were electrophoresed on a 1% agarose/2.2 M formaldehyde gel, blotted to nitrocellulose filter according to Thomas (18), and hybridized with nicktranslated ³²P-labeled 1.9-kilobase (kb) Kpn I-digested fragment from λ AP27 (under *high* stringency conditions) (ref. 14, pp. 387–389). Autoradiography was performed overnight and for 10 days at -70°C.

Genomic Mapping and Chromosome Localization of the PLAP. Various restriction endonuclease-treated human

-17	MLLLLLGLRLQLSLGIILVEEENPDFWNREAA EALGAAK	PLAP
-22'	MKQSTIALALLPLLFTPVTKAR TPEMPVLENRAAQGDITAPGGAR	E. coli AP
25	K L Q P A Q T A A K N L I I F L G D G M G V S T V T A A R I L K G Q K K D K L G	PLAP
24	R L T G D Q T A A L R D S L S D K P A K N I I L L I G D G M G D S E I T A A R N Y A E G A G G F F K	E. coli AP
65	PEIPLAM-DRFPYVALSKTYNVDKHVPDSGATATAYLCGVKGNFQTIGLS	PLAP
74'	GIDALPLTGQYTHYALNKKTGKPDYVTDSAASATAWSTGVK-TYNG-ALG	E. coli AP
114	A A A R F N Q C N T T R G N E V I S V M N R A K K A G K S V G V V T T T R V Q H A S P A G T Y A H T	PLAP
122'	V D I H E K D H P T I L E M A K A A G L A T G N V S T A E L Q D A T P A A L V A H V	E. coli AP
164	V N R N W Y S D A D V P A S A R Q E G C Q D I A T Q L I S N M D I D V I L G G G R K Y M F R	PLAP
164'	T S R K C Y G P S A T S Q K C P G N A L E K G G K G S I T E Q L L N A R A D V T L G G G A K - T F A	E. coli AP
210	M G T P D PEY P D D Y S Q G G T R L D G K N L V Q E W L G E R Q G A R Y V W N R T	PLAP
213'	E T A T A G E W Q G K T L R E Q A Q A R G Y Q L V S D A A S L N S V T E A N Q Q K P L L G L F A D G	E. coli AP
252	E L M Q A SLDPS VTH L M G L F EPG D M K Y E I H R D S T L D P SL M E M T E A A L R L L S R	PLAP
263'	N M P V R W L G P K A T Y H G N I D K P A V T C T P N P Q R N D S V P T L A Q M T D K A I E L L S K	E. coli AP
302	H P R G F F L F V E G G R I D H G H H E S R A Y R A L T E T I M F D D A I E R A G Q L T - SE E D T	PLAP
313'	N E K G F F L Q V E G A S I D K Q D H A A N P C G Q I G E T V D L D E A V Q R A L E F A K K E G N T	E. coli AP
351	L S L V T A D H S H V F S F G G Y P L R G S S F I G L A A G K A R D R K A Y T V L L Y G N G P G Y V	PLAP
363'	L V I V T A D H A H A S Q I V A P D T K A P G L T Q A L N T K D G A V M V M S Y G N	E. coli AP
401	L K D G A R P D V T E S E S G S P E Y R Q Q S A V P L D E E THAGE D V A V FARG P Q A H L V H	PLAP
405'	S E E D S Q E H T G S Q L R I A Y G P H A A N V V	E. coli AF
451 431'	GVQEQTFIAHVMAFAACLE PLAP GLTDQTDLFYTMKAALGLK E. coli AP	

FIG. 3. Comparison of homologous regions of human PLAP and *E. coli* alkaline phosphatase (AP). Unprimed and primed residue numbers are shown on left for PLAP and *E. coli* alkaline phosphatase, respectively. Identical residues are boxed. A \bullet between the two amino acids represents conservative replacements; a - represents insertion to produce alignment. The sequences are shown in single-letter code, which corresponds to the three-letter code as follows: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; Y, Tyr.

DNAs were electrophoresed in a 1% agarose gel. Southern transfer to a nitrocellulose filter (19) was followed by hybridization under high stringency conditions with the nick-translated ³²P-labeled 1.9-kb *Kpn* I restriction fragment derived from λ AP27 (ref. 14, pp. 387–389). Autoradiography was performed overnight and for 10 days. The chromosomal location of the gene was analyzed by direct hybridization of the same probe as used for the Southern blot to the DNA of chromosomes resolved by dual laser chromosome sorting as previously described (20).

RESULTS AND DISCUSSION

Isolation of Human PLAP cDNA and Comparison with E. coli Alkaline Phosphatase. The human placental λ gt11 cDNA library was screened at high cell density with the polyclonal antibody to PLAP. Out of 150,000 recombinant phages, 68 clones reacting with the antibody were isolated, suggesting that PLAP mRNA is at most 0.27% of the mRNA from human placenta. After rescreening at lower cell density, EcoRI digestion of DNA from 29 of the cells exhibiting the strongest signals during antibody screening yielded 23 cDNA inserts ranging from 1.9 to 2.7 kb.

To confirm the identity of these positives (since the antibody used was not affinity purified) and to identify potentially full-length PLAP cDNA clones, we used ³²Plabeled oligonucleotide probes corresponding to the Nterminal peptide sequence to screen the positives obtained by immunoscreening. Nine out of 23 clones hybridized to both sets of probes and all had inserts of approximately 2.7 kb. The restriction map of λ AP27 is shown in Fig. 1. Cleavage sites for Xba I, HindIII, and Sal I were not found in the insert. Both strands of $\lambda AP27$ (2698 base pairs) were sequenced (Fig. 2). The longest open reading frame, which starts with a methionine codon preceded by 25 nucleotides of 5' untranslated sequence and ends with a TGA codon, encodes a protein of 530 amino acids. This coding sequence indicates a $M_{\rm r}$ of 57,504, which is in good agreement with the 60,000 $M_{\rm r}$ estimate for the unglycosylated polypeptide from in vitro translation (6). The predicted amino acid composition and tryptic peptide maps also agree with those already determined (1, 7, 21). Although the initiating methionine codon does not have the typical initiation consensus sequences around it $[CC_G^ACCAUGG (22)]$, we assume it is the initiation site for translation because it is followed by a predicted signal peptide just prior to the known N-terminal amino acids. The predicted M_r of 1872 of this signal peptide agrees with the previously estimated value of 2000 determined in expression experiments (6). This leucine-rich signal peptide is similar to the signal peptides of mouse epidermal growth factor precursor and human pancreatic polypeptide hormone precursor (23, 24). Except for the Pro/Leu difference at position 3, representing a single-base difference, the predicted sequence precisely matches the published N-terminal sequence of 42 amino acids (7). Although the possibility of errors in the amino acid sequence determination cannot be ruled out, the discrepancy could be due to a substitution $(C \rightarrow T)$ that occurred as a result of reverse transcriptase errors during cDNA synthesis or of the well-studied genetic polymorphism at the PLAP locus (2). The first four N-terminal amino acids are hydrophobic; it is common for hydrophobic amino acids to be found next to the signal peptide (2).

The amino acid sequences of PLAP and E. coli alkaline phosphatase bear a remarkable homology throughout the molecule, including the signal peptide. PLAP has 30% identity with the 471 residues of E. coli alkaline phosphatase (Fig. 3). The PLAP residues corresponding to residues in the proposed α -helix and the A to J strands within the β -sheet of E. coli alkaline phosphatase are within conserved areas (25, 26). We therefore presume that the general structures of the bacterial alkaline phosphatase and PLAP are similar. All but one of the residues in the putative catalytic active site and metal-binding sites are identical. Thus, serine residue 92 (PLAP) and residue 102' (E. coli alkaline phosphatase) (the ' designates E. coli alkaline phosphatase residues) which bind to a phosphate group (27) and the other important amino acids believed to be involved in the metal-binding sites (25, 26) [Arg (residue 166, 166'), His (residues 320, 331'; 358, 370'; 360, 372'; and 432, 412'), Asp (residues 42, 51'; 357, 369')] are all identical and within highly conserved areas. The exception is residue Thr-155', which is replaced by Ser-155. The residues immediately adjacent to serine (residue 92) correspond to the characteristic sequence Asp/Glu-Ser-Ala/Gly found in most of the serine hydrolases with the exception of subtilisin (1). More significantly, both E. coli alkaline phosphatase and PLAP have eight carboxyl anions and two amino cations in the putative active site region (assuming the histidines are neutral). The resulting net charge of -6 equals the combined charge of three divalent metal ions (25).

Two longer insertions of 16 residues (396-411), and 13 residues (419-431) shift one of the histidines at a putative metal-binding site (residues, 432, 412') of PLAP further toward the C terminus relative to the equivalent position in *E. coli* alkaline phosphatase. Since the corresponding residues (412-418, 405'-411') between these two insertions are not involved in the proposed α -helix- β -sheet structure or in the metal-binding or active sites in the *E. coli* alkaline phosphatase, they may form a loop-out structure and the overall structure of PLAP could resemble that of the *E. coli* enzyme. There are two N-linked glycosylation sites (residues 122, 249), in agreement with previous evidence that PLAP is only slightly glycosylated (28); both occur at insertion/deletion sites. The positions for cysteine, however, are not conserved in the PLAP.

PLAP is 59 residues longer than *E. coli* alkaline phosphatase, in part because of a net gain from insertion/deletion in the body of the protein but also because of a 44-residue extension (487-530) at the C terminus with a stretch of 17 hydrophobic amino acids (491-507). These C-terminal amino acids presumably anchor the protein to the plasma membrane (29), and they explain this fundamental difference between the bacterial periplasmic enzyme and the mammalian mem-



FIG. 4. (Left) Size of PLAP mRNA by gel analysis of poly(A)⁺ RNA from human and mouse. Samples are 4 μ g each of total poly(A)⁺ RNA from human placenta (lane a), human PLC/PRF/5 hepatoma cell (Alexander cell) (lane b), human liver (lane c), 1 μ g of mouse testis total poly(A)⁺ RNA (kindly provided by R. Cann) (lane e), and DNA markers (HindIII-digested λ phage DNA and HincIIdigested ϕ X174 phage DNA) (lane d). (Right) Southern gel analysis of undigested and various restriction-enzyme-digested human genomic DNAs. Eighteen micrograms was loaded in each lane of a 1% agarose gel. Lanes: a, undigested human DNA; b, size markers of HindIII-digested λ DNA and HincII-digested ϕ X174 DNA; c-i, human DNA digested by Kpn I (c), Xba I (d), Sst I (e), Sal I (f), Pst I (g), HindIII (h), and EcoRI (i).

brane-associated enzyme. This assignment of a C-terminal anchor agrees with previous experimental evidence suggesting that IAP and PLAP are inserted at their C termini (2).

Until recently, it was generally assumed that membranespanning domains are α -helical and that at least 20 amino acids are required to span the lipophilic core of the membrane (30). The putative transmembrane domain in the β subunit of the T-cell receptor, like PLAP, also contains an uncharged segment of only 17 amino acids (31, 32). In fact, recent reports using deletion mutations (33) and oligonucleotidedirected mutagenesis (34) generating deletions in the membrane anchor suggest that as few as 12 to 14 amino acids may allow attachment to the membrane.

The strong homology of PLAP with the *E. coli* enzyme through most of the molecule and the sharp divergence at the C terminus containing the anchor suggests that the mammalian enzyme may have been formed by transposition of an exon containing the C-terminal domain onto the 3' terminus of the primitive bacterial alkaline phosphatase type gene. Thus, the structure of the PLAP gene is of interest. An intron might be found in the region between the two domains at approximately residue 470.

There are 1083 nucleotides of 3' untranslated sequence following the coding sequence. Many other membrane protein cDNAs have 1-3 kb of 3' untranslated sequence (10). The 3' sequence contains six polyadenylylation signals (AATA-AA). Multiple polyadenylylation sites have been described in other gene systems (35, 36), but it is unusual to find five (two overlap) within a short stretch of 38 nucleotides (2654-2691) at the 3' end.

RNA Blot Hybridization Analysis. By this analysis, probably only one species of PLAP mRNA, 2.8 kb, was detected (Fig. 4 *Left*). Since the mRNA is only 0.1 kb larger than our cDNA, the poly(A)⁺ stretch at the 3' end of the sequence may be part of the poly(A)⁺ tail. The analysis also indicates the absence or low abundance of PLAP mRNA or other closely related mRNAs in human liver, Alexander cells, and mouse testis. Thus, liver alkaline phosphatase probably does not have close sequence homology with PLAP, as is reflected by the lack of immunological cross-reactivity between these isozymes (2).

Genomic Mapping and Chromosome Localization. At least three loci have been postulated for the various forms of



FIG. 5. Autoradiogram of dot-blot filters of sorted chromosomes.

human alkaline phosphatase on the human chromosome(s) (37). The chromosomal location of the PLAP gene was analyzed by direct hybridization to the DNA of chromosomes resolved by dual laser chromosome sorting (20). A strong signal is found on chromosome 2 (Fig. 5), which, interestingly, is also the location of the acid phosphatase gene (38). However, with longer exposure a weaker signal also appeared on chromosome 17, and we speculate that this may be the location of a related gene(s). Although we cannot assign the chromosomal location of the PLAP gene with certainty because the number of closely related genes and the degree of their cross-hybridization are not known, it is likely that the PLAP gene is located on chromosome 2 and that one or more of the other closely related genes (IAP or PLAP-like genes) is located on chromosome 17 and perhaps on chromosome 2 also. Southern blot analysis of human DNA treated with several hexanucleotide specific restriction endonucleases revealed several bands hybridizing to a probe containing the entire coding sequence of the cDNA (Fig. 4 Right). The interpretation of the data is also limited by our lack of knowledge of the number of cross-reacting genes. Clarification of both sets of data must await isolation and sequence information from the other alkaline phosphatase genes.

PLAP has no known causal relationship with any human disease, but presumably is essential for normal fetal development, especially in the last trimester, when the level of the enzyme is high. Recently, a monoclonal antibody specific for IAP has been used for prenatal diagnosis of cystic fibrosis (39). The availability of PLAP cDNA and knowledge of its chromosomal location will allow further study of this gene family and of its possible relationship with human diseases.

We thank Drs. James Byrd, James Gum, Charles Craik, James Ou, Yousuke Ebina, Jeffrey Edman, Mark Selby, and Jeannette Bonifas for helpful advice; Dr. Daniel Cashman and Ms. Jennifer Barnett for oligonucleotide synthesis; Dr. Roger Lebo and Ms. Mei-Chi Cheung for chromosomal mapping; Drs. Robert Stroud and Janet Finer-Moore for assistance in homology matching; and Ms. Jennifer Gampell and Ms. Leslie Spector for preparation of the manuscript. W.K. is a recipient of a Veterans Administration Research Associateship. E.C. is a recipient of a National Institutes of Health International Fogarty Fellowship, and a Fondation de la Recherche Medicale (Paris) Award. Y.S.K. is a Medical Investigator of the Veterans Administration. Y.W.K. is an Investigator of the Howard Hughes Medical Institute. This work was supported by the Veterans Administration Medical Research Service (W.K. and Y.S.K.) and the National Institutes of Health (Grants GM 28520 and AM 21344 to W.J.R.).

- 1. McComb, R. B., Bowers, G. N., Jr., & Posen, S. (1979) Alkaline Phosphatase (Plenum, New York).
- 2. Stigbrand, T. & Fishman, W. H., eds. (1984) Progress in

Clinical and Biological Research (Liss, New York), Vol. 166. Boyer, S. H. (1961) Science 132, 1002–1004.

 Fishman, W. H., Inglis, N. I., Stolbach, L. L. & Kraut, M. J. (1968) Cancer Res. 29, 150-154.

3.

- Kam, W. K., Bresalier, R. S. & Kim, Y. S. (1984) in Progress in Clinical and Biological Research, eds. Stigbrand, T. & Fishman, W. H. (Liss, New York), Vol. 166, pp. 207-222.
- 6. Ito, F. & Chou, J. Y. (1983) Biochem. Biophys. Res. Commun. 111, 611-618.
- Ezra, E., Blacher, R. & Udenfriend, S. (1983) Biochem. Biophys. Res. Commun. 116, 1076-1083.
- 8. Kikuchi, Y., Yoda, K., Yamasaki, M. & Tamura, G. (1981) Nucleic Acids Res. 9, 5671-5678.
- Bradshaw, R. A., Cancedda, F., Ericsson, L. H., Neumann, P. A., Piccoli, S. P., Schlesinger, M. J., Shriefer, K. & Walsh, K. A. (1981) Proc. Natl. Acad. Sci. USA 78, 3473-3477.
- Ebina, Y., Ellis, L., Jarnagin, K., Edery, M., Graf, L., Clauser, E., Ou, J.-H., Masiarz, F., Kan, Y. W., Goldfine, I. D., Roth, R. A. & Rutter, W. J. (1985) *Cell* 40, 747-758.
 Huynh, T. V., Young, R. A. & Davis, R. W. (1985) in *DNA*
- Huynh, T. V., Young, R. A. & Davis, R. W. (1985) in DNA Cloning Volume 1: A Practical Approach, ed. Glover, D. M., (IRL, Oxford), pp. 49-78.
- Wallace, R. B., Schold, M., Johnson, M. J., Dembek, P. & Itakura, K. (1981) Nucleic Acids Res. 9, 3647-3656.
- 13. Crea, R. & Horn, T. (1980) Nucleic Acids Res. 8, 2331-2348.
- Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) Molecular Cloning: A Laboratory Manual (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY).
- Arias, C. F., Carlos, F., Lopez, S., Bell, J. R. & Strauss, J. H. (1984) J. Virol. 50, 657–661.
- 16. Messing, J. (1983) Methods Enzymol. 101, 20-78.
- 17. Sanger, F., Nicklen, S. & Coulson, A. R. (1977) Proc. Natl. Acad. Sci. USA 74, 5463-5467.
- 18. Thomas, P. S. (1983) Methods Enzymol. 100, 255-266.
- 19. Southern, E. M. (1975) J. Mol. Biol. 98, 503-517.
- 20. Lebo, R., Gorin, F., Fletterick, R. J., Kao, F. T., Cheung,
- M. C., Bruce, B. I. & Kan, Y. W. (1984) Science 225, 57–59.
 21. Badger, K. S. & Sussman, H. H. (1976) Proc. Natl. Acad. Sci. USA 73, 2201–2205.
- 22. Kozak, M. (1984) Nucleic Acids Res. 12, 857–872.
- Scott, J., Urdea, M., Quiroga, M., Sanchez-Pescador, R., Fong, N., Selby, M., Rutter, W. J. & Bell, G. I. (1983) Science 221, 236-240.
- Schwartz, T. W., Hansen, H. F., Hakanson, R., Sundler, F. & Tager, H. S. (1984) Proc. Natl. Acad. Sci. USA 81, 708-712.
- Wyckoff, H. W., Handschumacher, M., Murthy, K. & Sowadski, J. M. (1983) Adv. Enzymol. Relat. Areas Mol. Biol. 55, 453-480.
- 26. Coleman, J. E. (1984) Ann. N.Y. Acad. Sci. 429, 26-48.
- Whitaker, K. B., Byfield, P. G. H. & Moss, D. W. (1976) Clin. Chim. Acta 71, 285-291.
- 28. Millan, J. L. & Stigbrand, T. (1983) Biochem. Biophys. Res. Commun. 111, 611-618.
- Sabatini, D. D., Kreibich, G., Morimoto, T. & Adesnik, M. (1982) J. Cell Biol. 92, 1-22.
- 30. Tanford, C. (1978) Science 200, 1012-1018.
- Hedrick, S., Nielsen, E. A., Kavalen, J., Cohen, D. I. & Davis, M. (1984) Nature (London) 308, 153-158.
- Yanagi, Y., Yoshikai, Y., Legett, K., Clark, S. P., Aleksander, I. & Mak, T. (1984) Nature (London) 308, 146-149.
- Davis, N. G., Boeke, J. D. & Model, P. (1985) J. Mol. Biol. 181, 111-121.
- 34. Adams, G. A. & Rose, J. K. (1985) Cell 41, 1007-1015.
- 35. Tosi, M., Young, R. A., Hagenbuchle, O. & Schibler, W. (1981) Nucleic Acids Res. 9, 2313–2323.
- Reynolds, G. A., Basu, S. K., Osborne, T. F., Chin, D. J., Gil, G., Brown, M. S. & Goldstein, J. L. (1984) Cell 38, 275-285.
- 37. Harris, H. (1984) Annu. Rev. Genet. 17, 279-314.
- 38. McKusick, V. A. (1985) Clin. Genet. 27, 207-239.
- 39. Brock, D. J. H., Barron, L. & Bedgood, D. (1984) Prenatal Diag. 4, 421-426.