Proc. Natl. Acad. Sci. USA
Vol. 84, pp. 4403–4407, July 1987
Biochemistry

Human α_2-HS-glycoprotein: The A and B chains with a connecting sequence are encoded by a single mRNA transcript*

CHONG-CHOU LEE, BARBARA H. BOWMAN, AND FUNMEI YANG

Department of Cellular and Structural Biology, The University of Texas Health Science Center at San Antonio, San Antonio, TX 78284

Communicated by Eloise R. Giblett, March 16, 1987 (received for review January 12, 1987)

ABSTRACT **The α_2-HS-glycoprotein (AHSG) is a plasma protein reported to play roles in bone mineralization and in the immune response. It is composed of two subunits, the A and B chains. Recombinant plasmids containing human cDNA AHSG have been isolated by screening an adult human liver library with a mixed oligonucleotide probe. The cDNA clones containing AHSG inserts span approximately 1.5 kilobase pairs and include the entire AHSG coding sequence, demonstrating that the A and B chains are encoded by a single mRNA transcript. The cDNA sequence predicts an 18-amino-acid signal peptide, followed by the A-chain sequence of AHSG. A heretofore unseen connecting sequence of 40 amino acids was deduced between the A- and B-chain sequences. The connecting sequence demonstrates the unique amino acid doublets and collagen triplets found in the A and B chains; it is not homologous with other reported amino acid sequences. The connecting sequence may be cleaved in a posttranslational step by limited proteolysis before mature AHSG is released into the circulation or may vary in its presence because of alternative processing. The AHSG cDNA was utilized for mapping the *AHSG* gene to the 3q21→qter region of human chromosome 3. The availability of the AHSG cDNA clone will facilitate the analysis of its genetic control and gene expression during development and bone formation.**

The human α_2-HS-glycoprotein (AHSG) is a plasma protein that consists of two subunits, an A chain of 282 amino acids and a B chain of 27 amino acids (1, 2). The A chain and B chain are linked through a disulfide bond. The molecule has interesting and unique features; there are 29 amino acid doublets in circulating AHSG with Ala-Ala or Ala-Ala-Ala occurring six times and Pro-Pro occurring six times (2). Collagen-like sequences found in the complement component C1q were also detected in AHSG—i.e., Gly-Xaa-Pro (three times) and Gly-Pro-Xaa (two times) (2).

The *AHSG* gene is in a linkage group including genes for transferrin and plasma cholinesterase (3). The assignment of the transferrin gene to human chromosome 3 placed *AHSG* and the ceruloplasmin gene there also (4). Cox and Franke (5) localized the *AHSG* gene on human chromosome 3 by immunological studies with somatic cell hybrids.

In the matrix of adult and fetal bone, AHSG is concentrated up to 300-fold compared to other plasma glycoproteins (6, 7). The concentration of AHSG progressively falls in bone throughout childhood to adult life (8); fetal bone was reported to contain at least 10 times more AHSG than does adult bone (9), while neonatal bone contains 7 times more AHSG than does adult bone (10).

In addition to its presence in adult and fetal bone, the AHSG has been associated with several other biological functions. The protein has been observed to demonstrate opsonic properties (11), to bind DNA and promote endocytosis by macrophages (12), and to participate in chemotaxis leading to the recruitment of peripheral monocytes for osteoclast development (13).

AHSG is one of the few negative acute-phase reactants in human plasma; its level during an inflammatory response decreases significantly (14). Its concentration is also decreased in certain malignancies (15) and Paget's disease (16). High concentrations of AHSG are correlated with the presence of low density, immature, or newly formed bone (17, 18).

MATERIALS AND METHODS

Synthesis of Oligonucleotides of Mixed Sequence. Oligodeoxynucleotide probes were constructed on the basis of amino acid sequences of human AHSG B chain reported by Schmid and co-workers (1, 2). Residues 22–27 of the B chain, Ile-Arg-His-Phe-Lys-Val were used to construct 128 mixed 17-mer oligonucleotide probes: 5′-ACYTTRAARTGNCHH-AT-3′. (Single-letter abbreviations are as follows: R is A or G; Y is C or T; H is T or G; N is A, G, C, or T.) The probes were synthesized by P-L Biochemicals. Purified oligonucleotides were labeled at the 5′ end with [γ-^{32}P]ATP and polynucleotide kinase (19).

Isolation of AHSG cDNA with Oligonucleotides of Mixed Sequence. A human liver cDNA library (20) was screened with the oligonucleotide probes. The hybridization conditions and plasmid DNA purification and isolation were carried out as described (4).

DNA Sequence Determination. Both strands of the AHSG cDNAs were sequenced by the dideoxy chain-termination method (21). Template DNAs for dideoxy sequencing were prepared by inserting the restriction enzyme fragments of the AHSG cDNA inserts into phage M13 cloning vectors. The complete AHSG cDNA sequences were assembled from sequences of overlapping restriction enzyme fragments.

Primer Extension. For primer extension analysis, a 27-mer synthetic oligonucleotide was used as primer, and poly(A)$^+$ RNA isolated from human liver and the cultured human hepatoma cell line Hep 3B2 (22) was used as template. The 5′-end-^{32}P-labeled oligonucleotide (0.6 ng, 8×10^8 cpm/μg) was hybridized to 10 μg of poly(A)$^+$ RNA in 20 μl of 1× reverse transcriptase buffer (50 mM Tris·HCl, pH 8.3/75 mM KCl/10 mM dithiothreitol/3 mM $MgCl_2$) at 60°C for 1 hr. Primer extension with Moloney murine leukemia virus reverse transcriptase (Bethesda Research Laboratories) was subsequently carried out in 50 μl of 1× reverse transcriptase buffer containing actinomycin D at 50 μg/ml, 0.5 mM dNTPs, bovine serum albumin at 100 μg/ml, and 8000 units of Moloney murine leukemia virus reverse transcriptase per ml

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.

Abbreviation: AHSG, α_2-HS-glycoprotein.
*A brief report of this work has been presented (31).

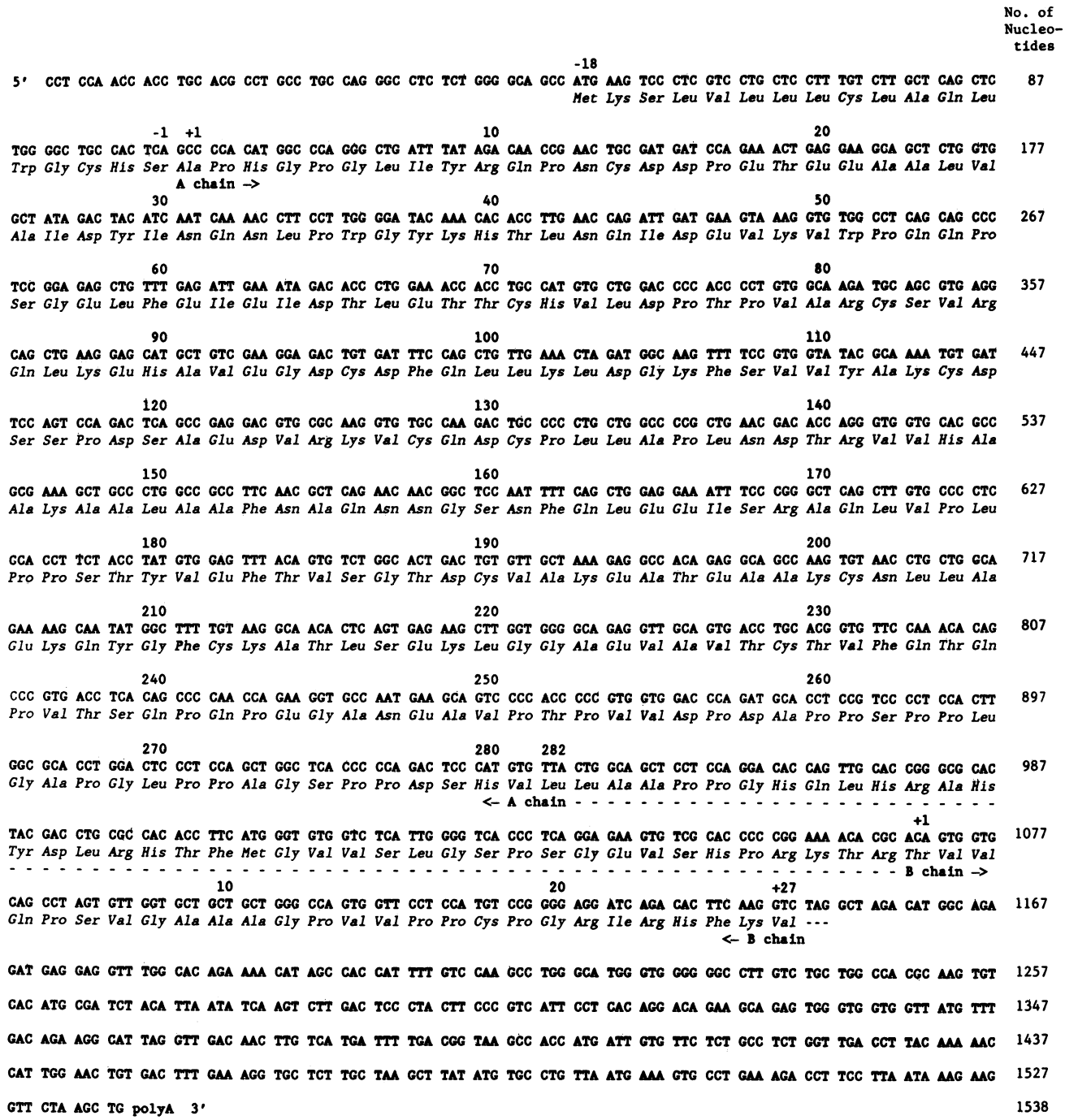

FIG. 1. Complete nucleotide sequence of AHSG cDNA and the deduced amino acid sequence. The putative signal peptide appears before the amino-terminal alanine, residue 1. Dotted lines denote the DNA sequence encoding a connecting amino acid fragment that does not appear in circulating AHSG.

at 37°C for 1 hr as recommended by the vendor. The resulting cDNA was purified by precipitation with ethanol and was sized next to a DNA sequence ladder on an 8 M urea/8% polyacrylamide gel.

Gene Mapping. Chromosome mapping of *AHSG* was performed with human–mouse somatic-cell hybrids (23). DNAs from the cell hybrids were isolated and digested with *Hin*dIII and were analyzed by Southern filter hybridization as described (24). The 1.5-kilobase (kb) cDNA insert was isolated by *Pst* I digestion, agarose gel electrophoresis, and electroelution of the human fragments. The human cDNA insert was ^{32}P-labeled by oligonucleotide priming (25) and was hybridized to the cell-hybrid DNA as described (24).

Southern Blot Analysis of DNA Polymorphisms. High molecular weight DNAs were isolated from human lymphocyte nuclei (26). Filter hybridization of lymphocyte DNA with human AHSG cDNA was carried out at 42°C as described (4).

RESULTS AND DISCUSSION

Five cDNA clones that contained sequences corresponding to the synthetic oligonucleotide probes were isolated from the human cDNA library. Upon rescreening with these clones, one clone, PHS-39, was identified that contained a full-length 1538-base-pair (bp) insert including the entire AHSG coding sequence. It consisted of a 5′ untranslated region of 48 bp, a 54-bp sequence encoding a putative 18-amino-acid leader sequence, a 1047-bp sequence encoding the amino acid

sequence, and 389 bp in the 3′ untranslated region (Fig. 1). The amino acid sequence deduced from the cDNA was identical to that determined from protein-sequence analyses of AHSG reported earlier (1, 2), with one exception. Residue 36 was predicted to be tryptophan by the cDNA sequence, instead of lysine. This difference could be the result of an inherited variation in the AHSG protein, although the substitution would require more than one nucleotide change.

The most important point revealed from the deduced sequence of AHSG cDNA is the presence of a heretofore unknown sequence of 40 amino acids in the region between the amino acid sequences reported for the A and B chains. Apparently the connecting 40-amino-acid fragment is cleaved by limited proteolysis in a posttranslational step that produces A and B chains joined by a disulfide bond. Fig. 2 is a model of how the mature AHSG protein could have been formed by limited proteolysis. The sequence of the connecting fragment originates behind the A-chain leucine in the 282 position and begins with another leucine residue. There are distinct similarities of the connecting fragment and the A and B chains; as summarized in Table 1, three additional amino acid doublets are within its sequence, an Ala-Ala pair in its second and third residue, a Pro-Pro pair in residue 4 and 5, and a Val-Val pair in residues 23 and 24. In addition, the collagen-like sequence, Gly-Xaa-Pro, is also present in the connecting fragment (residues 27–29). The similarity of the connecting fragment to the A and B chains raises the question of its biological significance and the role of alternative processing occurring during development.

A computer-aided search of the protein sequences available through the SEARCH data base (27) revealed no proteins having sequences homologous to mature AHSG or to the 40-amino-acid connecting fragment between the A and B chains. The closest identity of amino acid sequences was noted between the connecting fragment and the Epstein–Barr virus BFRF1 protein (28); 12 of the 40 amino acid residues are identical with those of the BFRF1 sequence. The similarity is of interest, since it has been reported that AHSG binds to a receptor on lymphocytes transformed by Epstein–Barr virus but does not bind to normal lymphocytes (29). The transformed lymphocyte receptor was reported to have the same molecular weight and isoelectric point as the Epstein–Barr virus-determined nuclear antigen.

To determine whether the 1538-bp AHSG cDNA insert represents the full-length human AHSG cDNA, primer extension analysis of human liver and hepatoma cell (Hep 3B2) RNA was conducted with a synthetic 27-mer oligonucleotide complementary to nucleotides 46–72 of clone PHS-39 (Fig. 1). When either human liver RNA or human Hep 3B2 RNA was used as template, electrophoresis of the resulting cDNA products on a denaturing polyacrylamide gel displays two major fragments of 108 and 105 nucleotides. When liver RNA was used as a template, the 105-nucleotide fragment was produced in greater quantity than the 108-nucleotide fragment. However, when the Hep 3B2 RNA was used as template, more of the 108-nucleotide fragment was synthesized. This may reflect different preferential initiation sites for AHSG RNA synthesis in the liver versus in the Hep 3B2 cells. Fig. 3 is the primer extension analysis of AHSG transcripts. The size of the cDNA was compared to a DNA sequence ladder. Thus, the cDNA insert in clone PHS-39 appears to begin 33 or 36 nucleotides 3′ to the mRNA cap sites. Therefore, the full-length human AHSG mRNA without the poly(A) tail is predicted to be either 1571 or 1574 nucleotides.

Table 1. Similarities of the A and B chains and the connecting fragment of AHSG

Peptide	Residues	Amino acid doublets or triplets, no.	Collagen repeating sequences,* no.
A chain	282	25 (1)†	4
Connecting peptide	40	3	1
B chain	27	4	1

*Gly-Pro-Xaa and Gly-Xaa-Pro.
†Indicates the Leu-Leu doublet from the junction of the A chain and connecting fragment of AHSG.

Chromosomal mapping of the *AHSG* gene to human chromosome 3 was confirmed by somatic-cell analysis using radiolabeled AHSG cDNA. The *AHSG* gene was mapped by Southern blot analysis of human–mouse hybrids after *Hin*dIII

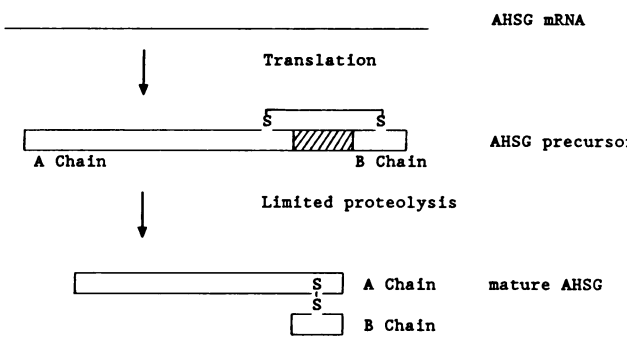

FIG. 2. Model for maturation of human AHSG. The connecting amino acid fragment (hatched bar) is not seen in mature AHSG; therefore, it is proposed that the connecting fragment is removed in a posttranslational step by limited proteolysis. The subunits formed are held together through a disulfide bond formed between the A and B chains.

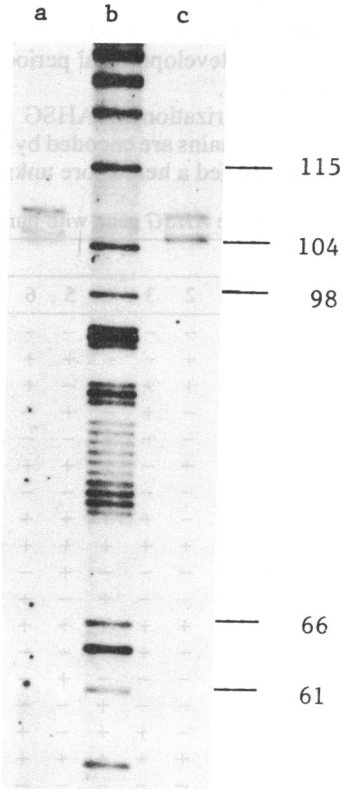

FIG. 3. Primer extension analysis of AHSG transcripts in human liver (lane c) and human hepatoma (Hep 3B2) cells (lane a). Ten micrograms of poly(A)$^+$ RNA was used in each reaction. The resulting cDNA was sized next to a DNA sequence ladder (lane b), and the sizes of some of the markers (in nucleotides) are indicated.

digestion. Digestion of human genomic DNA yielded fragments of 6.6, 1.5, and 1.3 kbp that hybridized with AHSG cDNA. The presence of fragments hybridizing with AHSG cDNA was determined for 19 characterized hybrids (Table 2). The *AHSG* gene segregated exclusively with chromosome 3; all other chromosomes showed discordancy. One hybrid retaining a region of chromosome 3 as a translocation, XTR-22, localized the *AHSG* locus to region 3q21→qter. This is of particular interest because two other plasma proteins important in the binding and transport of heavy metals, ceruloplasmin (30) and transferrin (4), have been mapped to the same region of human chromosome 3. A DNA polymorphism consisting of 3.3- and 2.8-kb bands was observed after hybridization of the AHSG cDNA probes with human lymphocyte DNA digested with *Pvu* II. Fig. 4 shows the *AHSG* polymorphism in a panel of DNAs from 11 unrelated individuals.

The role of AHSG in bone has not become clear despite many valuable observations of its presence, temporal concentrations, and influence on cell differentiation in bone-forming tissue (6–13; 15–18). Dickson and Bagga (8) suggested that AHSG protein was more likely to be related to bone formation and turnover than to be involved in stabilization of the mineralized matrix of bone. High concentrations of AHSG are correlated with the presence of low-density, newly formed bone; furthermore, the AHSG content of fetal bone is more than 10 times higher than in adult bone (9). Availability of the AHSG clone will provide the means to analyze the role of AHSG in bone formation during early development. In preliminary experiments, *in situ* hybridization of radiolabeled human AHSG cDNA and a panel of human and rat tissues has revealed that AHSG mRNA transcripts are found not only in human and rat liver and human hepatoma cell line Hep 3B2 but also in rat embryonal cartilage (unpublished results). Further studies are needed to determine whether or not AHSG protein is produced by bone-forming cells and to analyze the developmental periods in which the *AHSG* gene is expressed.

In summary, characterization of AHSG cDNA has revealed that the A and B chains are encoded by a single mRNA transcript and has identified a heretofore unknown sequence of 40 amino acids connecting the A and B chains. The characteristics of the connecting fragment are strikingly similar to the A and B chains. The *AHSG* gene has been mapped to human chromosome 3q21→qter, and a DNA polymorphism of the *AHSG* gene has been detected.

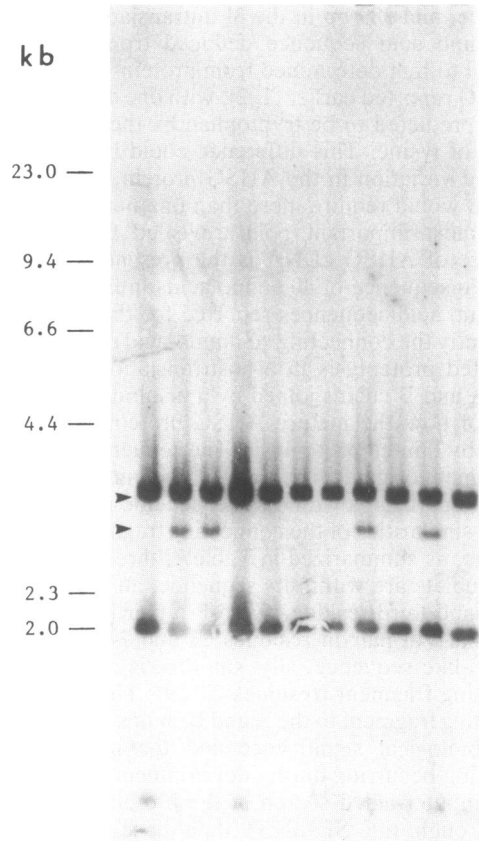

FIG. 4. *AHSG* DNA polymorphism detected with *Pvu* II shown in a panel of lymphocyte DNA digests from 11 unrelated individuals. DNA polymorphic fragments, 3.3 and 2.8 kb, are indicated by arrowheads.

Table 2. Segregation of the *AHSG* gene with human chromosome in human–mouse somatic cell hybrids

Hybrid	*AHSG*	Chromosome																							Translocation
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	X	
ICL-15	−	−	−	−	−	−	−	−	−	−	−	−	+	−	−	−	−	+	−	−	+	+	−	−	
TSL-2	−	−	+	−	−	+	+	−	−	−	+	−	+	−	−	−	−	−	+	−	+	+	−	+	17/3
ATR-13	+	+	+	+	+	−	+	+	+	−	+	−	+	−	+	+	−	+	+	−	−	−	−	−	5/X
JSR-17S	+	+	−	+	−	+	−	−	+	+	+	+	+	+	+	+	+	+	+	−	+	+	+	−	$7q^-$
WIL-5	−	−	−	−	+	−	−	−	−	−	−	−	−	−	−	−	−	+	+	−	−	+	−	+	
WIL-6	−	−	+	−	+	+	+	+	+	−	+	+	−	−	+	−	−	+	−	+	+	+	−	+	
WIL-2	−	−	−	−	−	−	−	−	+	−	−	−	+	−	−	−	−	+	−	−	−	+	−	+	
WIL-8X	+	−	−	+	+	+	+	+	−	−	+	+	+	−	+	−	−	+	+	+	+	+	−	+	
JWR-26C	+	−	+	+	+	+	+	+	−	+	+	+	+	−	+	+	+	+	+	−	+	+	−	+	$1p^-$
NSL-9	−	−	−	−	−	+	−	−	+	−	+	−	+	+	+	+	+	+	−	−	+	+	+	−	17/9
JWR-22H	−	−	−	−	+	−	+	−	−	−	+	+	−	−	+	−	−	+	+	−	+	+	−	−	2/1
REW-7	+	+	+	+	+	+	+	+	+	−	+	+	+	+	+	+	−	+	+	+	+	+	+	+	
WIL-14	+	−	−	+	−	−	−	−	+	−	+	−	−	−	+	−	−	+	−	−	−	−	−	+	
WIL-13	−	−	−	−	−	+	−	−	−	−	−	−	−	−	−	−	−	+	+	−	−	+	−	−	
REW-11	−	−	−	−	+	−	−	−	−	−	−	+	−	+	−	−	+	−	−	−	+	+	−	+	
XER-11	+	+	−	+	+	−	+	+	+	−	+	−	+	−	−	+	+	+	+	+	+	+	+	−	11/X X/11
XER-7	+	+	+	+	+	+	+	+	+	+	+	−	+	+	+	+	−	−	+	+	−	−	−	+	11/X
XTR-3BSAGB	−	−	−	−	−	−	−	−	−	−	−	−	+	−	−	−	−	−	−	−	+	−	−	−	3/X*
XTR-22	+	−	+	−	+	+	+	−	+	−	+	+	−	−	−	−	−	−	+	+	+	+	+	−	X/3*

+ and −, Presence and absence, respectively, of an intact human chromosome.

*Translocation chromosomes from the human parent segregating in the XTR series of hybrid cells: 3/X, (3pter → 3q21::Xq28 → Xqter) and X/3, (Xpter → Xq28::3q21 → 3qter). Neither XTR-3BSAGB nor XTR-22 contains an intact chromosome 3; however, the presence of the *AHSG* gene correlates with the presence of the 3q21→3qter segment in the X/3 translocation.

We thank Dr. S. H. Orkin (Harvard Medical School) for the human liver cDNA library; Dr. Susan Naylor for advice and help in chromosome localization; Drs. Frank Weaker, Damon Herbert, and John Chirgwin for valuable suggestions; Victor Luna, Rod Cupples, and James Buchanan for their technical help; and Betty Russell for manuscript preparation. This work was supported in part by grants from the National Institutes of Health (GM33298, HD16584) and the American Cancer Society (NP-470).

1. Gejyo, F., Chang, J.-L., Burgi, W., Schmid, K., Offner, G. D. & Troxler, R. F. (1983) *J. Biol. Chem.* **258,** 4966–4971.
2. Yoshioda, Y., Gejyo, F., Marti, T., Rickli, E. E., Burgi, W., Offner, G. D., Troxler, R. F. & Schmid, K. (1986) *J. Biol. Chem.* **261,** 1665–1676.
3. Eiberg, H., Mohr, J. & Nielsen, L. S. (1983) *Cytogenet. Cell Genet.* **37,** 461 (abstr.).
4. Yang, F., Lum, J. B., McGill, J. R., Moore, C. M., Naylor, S. L., van Bragt, P. H., Baldwin, W. D. & Bowman, B. H. (1984) *Proc. Natl. Acad. Sci. USA* **81,** 2752–2756.
5. Cox, D. W. & Franke, U. (1985) *Hum. Genet.* **70,** 109–115.
6. Triffitt, J. T., Gebauer, U., Ashton, B. A., Owen, M. E. & Reynolds, J. J. (1976) *Nature (London)* **262,** 226–227.
7. Ashton, B. A., Hohling, H. J. & Triffitt, J. T. (1976) *Calcif. Tissue Res.* **22,** 27–33.
8. Dickson, I. R. & Bagga, M. K. (1985) *Connect. Tissue Res.* **14,** 77–85.
9. Wilson, J. M., Ashton, B. & Triffitt, J. T. (1977) *Calcif. Tissue Res.* **22,** 458–460.
10. Quelch, K. J., Cole, W. G. & Melick, R. A. (1984) *Calcif. Tissue Int.* **36,** 545–549.
11. van Oss, C. J., Gillman, C. F., Bronson, P. M. & Border, J. R. (1974) *Immunol. Commun.* **3,** 329–335.
12. Lewis, J. G. & André, C. M. (1980) *Immunology* **3,** 317–322.
13. Malone, J. D., Teitelbaum, S. L., Griffin, G. L., Senior, R. M. & Kahn, A. J. (1982) *J. Cell Biol.* **92,** 227–230.
14. Lebreton, J. P., Joisel, F., Raoult, J. P., Lannuzel, B., Rogez, J. P. & Humbert, G. (1979) *J. Clin. Invest.* **64,** 1118–1129.
15. Bradley, W. P., Blasco, A. P., Weiss, J. F., Alexander, J. C., Silverman, N. A. & Chretien, P. B. (1977) *Cancer* **40,** 2264–2272.
16. Ashton, B. A. & Smith, R. (1980) *Clin Sci.* **5,** 434–438.
17. Dickson, I. R., Poole, A. R. & Veis, A. (1976) *Nature (London)* **256,** 430–435.
18. Dickson, I. R. (1981) in *The Chemistry and Biology of Mineralized Connective Tissues*, ed. Weis, A. (Elsevier, New York), pp. 229–232.
19. Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY), pp. 122–127.
20. Prochownik, E. V., Markham, A. F. & Orkin, S. H. (1983) *J. Biol. Chem.* **258,** 8389–8394.
21. Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74,** 5463–5467.
22. Knowles, B. B., Howe, C. C. & Aden, D. P. (1980) *Science* **209,** 497–499.
23. Shows, T. B., Naylor, S. L. & Sakaguchi, A. Y. (1982) in *Advances in Human Genetics*, eds. Hirschhorn, K. & Harris, H. (Plenum, New York), Vol. 12, pp. 341–352.
24. Naylor, S. L., Sakaguchi, A. Y., Szoka, P., Hendy, G. N., Kronenberg, H. M., Rich, A. & Shows, T. B. (1983) *Somatic Cell Genet.* **9,** 609–616.
25. Feinberg, A. P. & Vogelstein, B. (1983) *Anal. Biochem.* **132,** 6–13.
26. Bell, G. I., Karem, J. H. & Rutter, W. J. (1981) *Proc. Natl. Acad. Sci. USA* **78,** 5759–5763.
27. Barker, W. C. & Dayhoff, M. O. (1982) *Proc. Natl. Acad. Sci. USA* **79,** 2836–2839.
28. Baer, R., Bankier, A. T., Biggin, M. D., Deininger, P. L., Farrell, P. J., Givson, T. J., Hatfull, G., Hudson, G. S., Satchwell, S. C., Sequin, C., Tuffnell, P. S. & Barrell, B. G. (1984) *Nature (London)* **310,** 207–211.
29. Lewis, J. G. & André, C. M. (1982) *FEBS Lett.* **143,** 332–336.
30. Yang, F., Naylor, S. L., Lum, J. B., Cutshaw, S., McCombs, J. L., Naberhaus, K. H., McGill, J. R., Adrian, G. S., Moore, C. M., Barnett, D. R. & Bowman, B. H. (1986) *Proc. Natl. Acad. Sci. USA* **83,** 3257–3261.
31. Lee, C.-C., Bowman, B. H. & Yang, F. (1986) *Am. J. Hum. Genet.* **39,** 617 (abstr.).