Studier, W. E. (1973) *J. Mol. Biol. 79*, 227-236.
Tabor, S., & Richardson, C. C. (1981) *Proc. Natl. Acad. Sci. U.S.A. 78*, 205-209.
Tabor, S., & Richardson, C. C. (1987) *J. Biol. Chem. 262*, 15330-15333.
Tabor, S., Huber, H. E., & Richardson, C. C. (1987) *J. Biol. Chem. 262*, 16212-16223.
Weiner, J. H., Bertsch, L. L., & Kornberg, A. (1975) *J. Biol. Chem. 250*, 1972-1980.
Weinstein, I. B., Jeffrey, A. M., Jennette, K. W., Blobstein, S. H., Harvey, R. G., Harris, C., Autrup, H., Kasai, H., & Dakanishi, K. (1976) *Science 193*, 592-595.
Weiss, B., Live, T. R., & Richardson, C. C. (1968) *J. Biol. Chem. 243*, 4530-4542.

Complete Nucleotide Sequence of the Gene for Human Heparin Cofactor II and Mapping to Chromosomal Band 22q11[†,‡]

Ruth Herzog,[§] Steffi Lutz,[§] Nikolaus Blin,[§] Jayne C. Marasa,[∥] Morey A. Blinder,[∥] and Douglas M. Tollefsen*[,∥]

Institut für Humangenetik, Universität des Saarlandes, 6650 Homburg/Saar, BRD, and Division of Hematology-Oncology, Department of Internal Medicine, Washington University, St. Louis, Missouri 63110

Received July 26, 1990; Revised Manuscript Received October 15, 1990

ABSTRACT: Heparin cofactor II (HCII) is a 66-kDa plasma glycoprotein that inhibits thrombin rapidly in the presence of dermatan sulfate or heparin. Clones comprising the entire HCII gene were isolated from a human leukocyte genomic library in EMBL-3 λ phage. The sequence of the gene was determined on both strands of DNA (15 849 bp) and included 1749 bp of 5′-flanking sequence, five exons, four introns, and 476 bp of DNA 3′ to the polyadenylation site. Ten complete and one partial Alu repeats were identified in the introns and 5′-flanking region. The HCII gene was regionally mapped on chromosome 22 using rodent–human somatic cell hybrids, carrying only parts of human chromosome 22, and the chronic myelogenous leukemia cell line K562. With the cDNA probe HCII7.2, containing the entire coding region of the gene, the HCII gene was shown to be amplified 10–20-fold in K562 cells by Southern analysis and in situ hybridization. From these data, we concluded that the HCII gene is localized on the chromosomal band 22q11 proximal to the breakpoint cluster region (BCR). Analysis by pulsed-field gel electrophoresis indicated that the amplified HCII gene in K562 cells maps at least 2 Mbp proximal to BCR-1. Furthermore, the HCII7.2 cDNA probe detected two frequent restriction fragment length polymorphisms with the restriction enzymes *Bam*HI and *Hin*dIII.

Heparin cofactor II (HCII)[1] is a plasma glycoprotein that inhibits thrombin rapidly in the presence of dermatan sulfate or heparin (Tollefsen, 1989). HCII cDNA clones have been isolated, sequenced, and expressed in *Escherichia coli* (Inhorn & Tollefsen, 1986; Blinder et al., 1988; Blinder & Tollefsen, 1990; Derechin et al., 1990) and COS cells (Ragg, 1986; Ragg et al., 1990). A dysfunctional variant of HCII ($HCII_{Oslo}$) with decreased affinity for dermatan sulfate was characterized by amplification and sequence analysis of a portion of the HCII gene (Blinder et al., 1989). Restriction fragment length polymorphisms (RFLPs) with *Bam*HI and *Msp*I have been identified in the HCII gene by Southern blot analysis (Blinder et al., 1988; Turner et al., 1990). In the present investigation, we determined the complete nucleotide sequence of the HCII gene, including portions of the 5′- and 3′-flanking sequences, observed a new *Hin*dIII polymorphism, and identified sites in the gene at which the *Bam*HI and *Hin*dIII polymorphisms may occur.

The HCII gene has been assigned to human chromosome 22 by hybridization of a cDNA subclone to DNA from flow-sorted chromosomes (Blinder et al., 1988). Chromosome 22, although minute in size, is involved in inherited as well as in acquired diseases. About a dozen diseases including seven different malignancies have been characterized at the cytogenetic or phenotypic level or both (Kaplan et al., 1987; Kaplan & Emanuel, 1989). Analysis of the molecular pathology of chromosome 22 is greatly facilitated by the use of regionally assigned DNA probes that recognize RFLPs. Therefore, we regionally mapped the polymorphic HCII gene using somatic cell hybrids containing only parts of chromosome 22 and the chronic myelogenous leukemia (CML)-derived cell line K562. In addition, a more detailed analysis was achieved by using pulsed-field gel electrophoresis for a long-range map surrounding the HCII locus.

Materials and Methods

Probes. HCII7.2 and HCII7.3 are 1.6- and 0.6-kb *Eco*RI fragments subcloned into pGEMblue, corresponding to 5′ and 3′ portions of the HCII cDNA, respectively (Blinder et al.,

[†] Supported by DFG Grant Bl 166/3-3 to N.B. and by NIH Grant HL-14147 (Specialized Center of Research in Thrombosis) to D.M.T.
[‡] The nucleic acid sequence in this paper has been submitted to GenBank under Accession Number J05309.
* Address correspondence to this author at the Division of Hematology–Oncology, Box 8125, Washington University Medical School, 660 South Euclid Ave., St. Louis, MO 63110.
[§] Universität des Saarlandes.
[∥] Washington University.

[1] Abbreviations: HCII, heparin cofactor II; RFLP, restriction fragment length polymorphism; serpin, serine proteinase inhibitor; PCR, polymerase chain reaction; CML, chronic myelogenous leukemia; BCR, breakpoint cluster region; Mbp, megabase pair(s).

0006-2960/91/0430-1350$02.50/0

1988). IGLV (pVL3.3) is a 3.3-kb *Bam*HI fragment of a genomic clone (V4A) containing a gene of the human immunoglobulin λ light-chain gene cluster subcloned into pUC12 (Julier et al., 1985). D22S10 (22c1-18) is an anonymous 2-kb human genomic *Sau*3A fragment subcloned into pUC12 (Hofker et al., 1985). BCR is a 1.9-kb *Hin*dIII/*Bgl*II fragment of the breakpoint cluster region (Groffen et al., 1984). IGLV, D22S10, and BCR have been localized to chromosomal band 22q11 (Kaplan & Emanuel, 1989). phom37H is a human anonymous single-copy 1.4-kb *Hin*dIII fragment subcloned into pTZ19 that was mapped previously to 22q12-qter (Metzdorf et al., 1990). The inserts were labeled with [α-^{32}P]dATP or [α-^{32}P]dCTP by the random primer method (Feinberg & Vogelstein, 1984).

Isolation of HCII Genomic Clones. A human leukocyte genomic library in EMBL-3 λ phage (Clontech Laboratories) was propagated in *E. coli* LE392 and plated at a density of 50000 phage per plate (150 mm) on 10 plates. Duplicate nitrocellulose (Schleicher & Schuell) lifts were hybridized to ^{32}P-labeled HCII7.2 and HCII7.3 (Blinder et al., 1988) to select clones containing the 5′ and 3′ ends of the HCII gene. Positive phage were plaque-purified. Phage λ DNA was purified by the liquid-lysis method of Garber et al. (1983). Inserts were isolated by digestion of the purified EMBL-3 λ DNA with *Sal*I, followed by agarose gel electrophoresis. The restriction fragments were electroeluted from the gel by using a Schleicher & Schuell Elutrap and ligated into the *Sal*I site of Bluescript KS phagemid (Stratagene). Large-scale preparations of double-stranded plasmid DNA were purified according to Davis et al. (1986) using poly(ethylene glycol).

DNA Sequence Analysis. Plasmid DNA was subjected to alkaline denaturation for 5 min at 85 °C according to Toneguzzo et al. (1988). Sequence analysis was performed by a modification (Hattori & Sakaki, 1986) of the dideoxynucleotide method of Sanger et al. (1977), in which modified T7 DNA polymerase (Sequenase, United States Biochemical Corp.), deoxyadenosine [α-^{35}S]thiotriphosphate, and 6% denaturing polyacrylamide gels were employed. Primers were synthesized in the Protein Chemistry Facility of Washington University to obtain overlapping sequences on both strands of DNA. The sequences of difficult areas of high G/C content or other secondary structure were determined by inclusion of 7-deaza-2′-deoxyguanosine 5′-triphosphate or dITP (Barnes et al., 1983; Gough & Murray, 1983; Mizusawa et al., 1986) in the sequencing reactions. DNA sequences were analyzed on a VAX/VMS computer using software obtained from the University of Wisconsin Genetics Computer Group (Devereux et al., 1984).

Polymerase Chain Reaction. Genomic DNA was amplified by the polymerase chain reaction (PCR) as described in a previous report (Blinder et al., 1989). The PCR products were purified by agarose gel electrophoresis and sequenced directly or cloned into Bluescript KS prior to sequencing.

Somatic Cell Hybrids and K562 Cells. The regional localization of the HCII gene on chromosome 22 was performed with the mouse–human cell hybrid PgMe25Nu, which contains only chromosome 22 as the human complement (de Klein et al., 1982), and two Chinese hamster–human cell hybrids, 1/22 AM-27 (Geurts van Kessel et al., 1980) and 11/22 A3EW2-6A (Geurts van Kessel et al., 1985), containing translocation chromosomes carrying only parts of human chromosome 22. K562 is a CML-derived cell line with amplification of the BCR-*abl* fusion gene and other sequences of 22q11 on a marker chromosome (Heisterkamp et al., 1983; Selden et al., 1983) (gift of P. Ambros, Wien).

Southern Blot Analysis. The DNA of the somatic cell hybrids and placenta was isolated according to Müllenbach et al. (1989). Human DNA and DNA of K562 cells were isolated by preparing agarose blocks (Van Ommen & Verkerk, 1986). White blood cells were obtained by isotonic lysis of whole blood, and K562 cells were collected by using standard procedures. Prior to digestion, agarose blocks containing 4–5 μg of genomic DNA were equilibrated in 0.5 mL of the appropriate restriction buffer for 2 h at room temperature. Upon removal of the restriction buffer, the blocks were melted at 65 °C for 5 min and then kept at 37 °C for 10 min. Thereafter, 25–50 units of the restriction enzyme was added, and the samples were incubated for 6–8 h. The liquid blocks were dry-loaded onto 0.6–0.8% agarose gels and allowed to solidify at 4 °C for 10 min before electrophoresis. After transfer to a Zetaprobe membrane (Bio-Rad) using 0.5 M NaOH/1 M NaCl, the filters were neutralized in 0.5 M Tris-HCl/3 M NaCl, pH 7.0, and air-dried. Hybridization was carried out at 65 °C overnight in the following solution: 4× SET (0.6 M NaCl, 8 mM EDTA, and 0.12 M Tris-HCl, pH 7.4), 10% dextran sulfate, 1% SDS, 0.1% sodium pyrophosphate, and 0.1 mg/mL sheared denatured herring sperm DNA. Filters were washed to a final stringency of 0.1% SDS/0.1× SSC (15 mM NaCl/1.5 mM sodium citrate, pH 7.0) at 65 °C.

Pulsed-Field Gel Electrophoresis. Partial and complete digests of K562 DNA in agarose blocks were performed in 50 μL of restriction buffer for 6–8 h. DNA was separated in the hexagonal Pulsaphor apparatus (Pharmacia) as described previously (Metzdorf et al., 1990).

In Situ Hybridization. Metaphase chromosome spreads were prepared by using standard procedures. In situ hybridization was performed according to Mattei et al. (1985) using ^{3}H-labeled DNA with minor modifications. Following hybridization and extensive washing, the slides were dipped in photographic emulsion (Kodak NTB 2) and exposed for 14 days at 4 °C. Chromosomes were GTG-banded (Seabright, 1971).

RESULTS

Isolation of HCII Genomic Clones. Two overlapping clones comprising the complete HCII gene were isolated from a human leukocyte genomic library in EMBL-3 λ phage. Approximately 500000 plaques were screened with cDNA probes HCII7.2 and HCII7.3 (Figure 1A); 21 plaques were positive with both probes, and 3 were positive with HCII7.2 only. *Sal*I digestion of inserts from clones that were positive only with HCII7.2 yielded 9.2- and 5.6-kb fragments, whereas inserts from clones that were positive with both probes yielded fragments of 7.4, 4.2, and 2.5 kb. Only the 2.5-kb fragment hybridized to HCII7.3, indicating that this fragment contained the 3′ end of the gene. Each of the *Sal*I fragments was ligated into Bluescript KS phagemid and amplified for further restriction mapping and sequence analysis. Restriction mapping of the five cloned fragments with *Dra*II, *Bam*HI, *Eco*RI, *Pst*I, and *Sma*I indicated that the 9.2- and 4.2-kb fragments overlapped, as did the 7.4- and 5.6-kb fragments, resulting in the arrangement shown in Figure 1B.

DNA Sequence Analysis. The sequence of 15849 contiguous base pairs of DNA comprising the HCII gene is shown in Figure 2. The structure was determined by sequencing the 9.2-, 7.4-, and 2.5-kb *Sal*I fragments using a series of synthetic oligonucleotide primers that hybridized to the DNA at intervals of 200–300 nucleotides to yield overlapping sequences on each strand of DNA. Both strands of DNA were sequenced in their entirety for the region spanned by the two arrows shown in Figure 2. Both of the internal *Sal*I sites in the λ

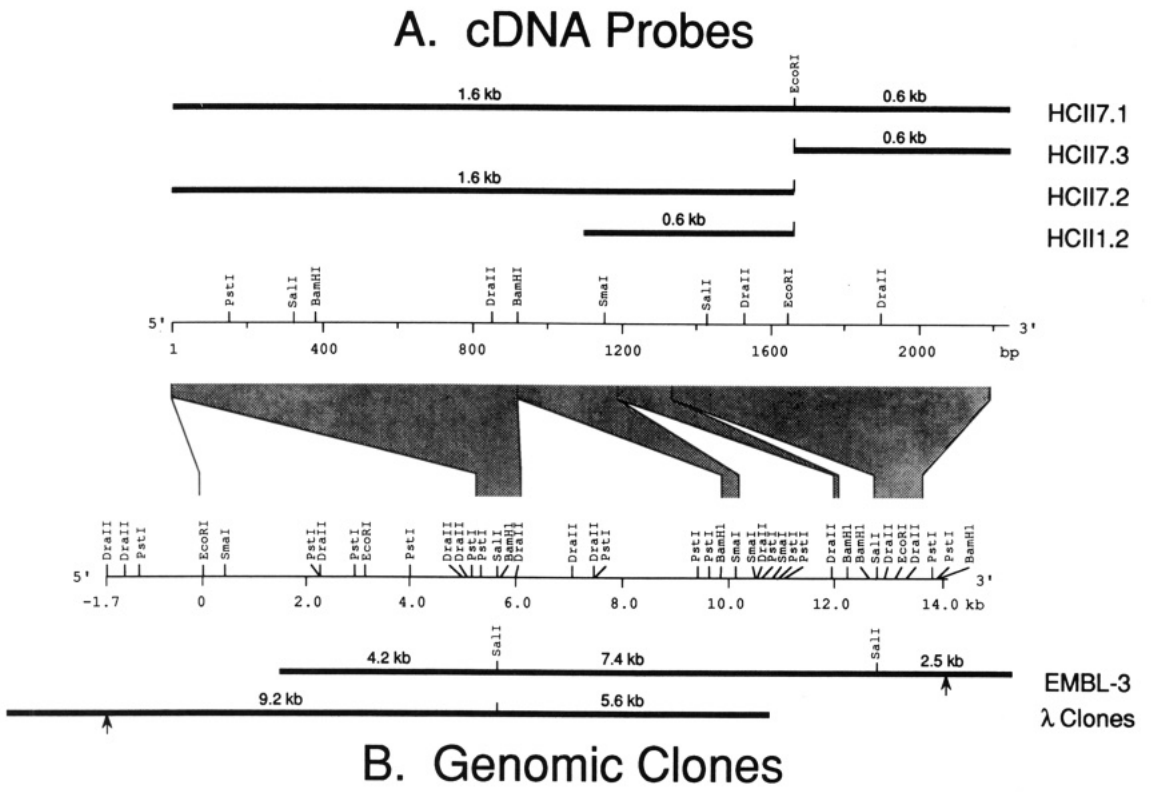


FIGURE 1: HCII cDNA probes and genomic clones. (A) Partial restriction map of human HCII cDNA clones. HCII7.1 was isolated from a partial *Eco*RI digest of a human liver λgt11 clone (Blinder et al., 1988). HCII7.2 and HCII7.3 were prepared from HCII7.1 by cleavage of the internal *Eco*RI site. HCII1.2 was isolated separately from the same library (Inhorn & Tollefsen, 1986). Nucleotides are numbered according to Blinder et al. (1988). (B) Partial restriction map of two clones isolated from a human genomic library in EMBL-3 λ phage by hybridization to cDNA probes HCII7.2 and HCII7.3. The arrows indicate the boundaries of the DNA sequence shown in Figure 2. Nucleotides are numbered according to Figure 2. The shaded areas represent exon regions. The fragment sizes of the cDNA probes and the *Sal*I genomic fragments are indicated by the numbers above each solid bar.

clones occurred in exons (see below) and were located by inspection of the cDNA sequence (Blinder et al., 1988). To rule out the presence of an unsuspected intron at one or both of the *Sal*I sites, genomic DNA was amplified by PCR with primers spanning these sites (data not shown). The sequences of the PCR products were consistent with the structure shown in Figure 2.

Nucleotide +1 represents the first nucleotide of the longest reported cDNA (Ragg & Preibisch, 1988). Alignment of the genomic and cDNA sequences indicated that the gene consists of five exons and four introns. All of the intron–exon boundaries obey the AG/GT rule and are identical with those reported for "human leuserpin 2" (i.e., HCII) by Ragg and Priebisch (1988). The largest intron (5151 bp) occurs between the first and second exons in the 5′-noncoding region of the gene, whereas the other three introns (3770, 1757, and 728 bp in length) interrupt the coding sequence of the gene. Computer analysis revealed seven Alu repetitive sequences on the coding strand and 3.5 Alu repeats on the complementary strand of DNA. The repetitive sequences occurred in the 5′-flanking region and the first three introns of the HCII gene (Figure 2). Table I indicates differences between the genomic sequence in Figure 2 and the previously reported cDNA (Blinder et al., 1988; Ragg, 1986) and partial genomic (Ragg & Preibisch, 1988) sequences; these differences may represent polymorphisms in the HCII gene.

BamHI and HindIII Polymorphisms. As reported previously (Blinder et al., 1988), hybridization of the partial cDNA probe HCII1.2 to genomic DNA detected a *Bam*HI RFLP. In the present study, the *Bam*HI RFLP was confirmed with HCII7.2, which yielded two polymorphic bands of 2.8 and 2.4 kb and three invariant bands of 9.5, 4.6, and 1.3 kb (data not shown). The frequencies obtained with a total of 33 individuals were 0.56 for allele 1 (2.8 kb) and 0.44 for allele 2 (2.4 kb). In addition, the probe HCII7.2 comprising the entire coding sequence of the HCII gene detected a *Hin*dIII RFLP that was not apparent with the probe HCII1.2. Southern blots of *Hin*dIII-digested genomic DNA from 20 unrelated individuals probed with HCII7.2 yielded 3 polymorphic fragments; allele 1 consisted of a 20-kb fragment, whereas allele 2 comprised 2 fragments of 2.5 and 17.5 kb. The frequencies of alleles 1 and 2 were 0.55 and 0.45, respectively. Mendelian inheritance of the *Hin*dIII polymorphism is depicted in Figure 3.

A partial restriction map of the sequence of the HCII gene is shown in Figure 4. The regions of the gene to which the cDNA probe HCII7.2 hybridizes are also indicated. It is unlikely that restriction fragments derived from exon 1 are

Table I: Differences in Previously Reported Sequences

	nucleotide			
position	gene (this study)	cDNA (Blinder et al., 1988)	cDNA (Ragg, 1986)	gene (Ragg & Preibisch, 1988)
−1040	C			T
−992	C			*a*
−800	A			AT[b]
−181	G			A
36	A	C		A
5210	C			*a*
5718	T (Gly)[c]	C (Gly)	T (Gly)	
5924	A (Lys)	G (Arg)	A (Lys)	
12915	T (His)	C (His)	C (His)	T (His)

[a] Nucleotide deletion. [b] Nucleotide insertion. [c] Codon in which nucleotide occurs.

detected by HCII7.2, since the probe contains only 9 bp of DNA capable of hybridizing to this region. The *Bam*HI fragments predicted from the gene sequence correspond reasonably well to the *Bam*HI (allele 2) fragments observed on Southern blots probed with HCII7.2. Point mutations (asterisk) abolishing the *Bam*HI site in intron 4 would result in the pattern observed for *Bam*HI allele 1 and thereby explain the RFLP.

The single *Hin*dIII fragment predicted from the gene sequence corresponds to that observed for *Hin*dIII allele 1 by Southern blot analysis. A single base pair substitution at nucleotide position 6259 (G → T) would create a *Hin*dIII site in intron 2, resulting in the 2.5-kb fragment of *Hin*dIII allele 2 (Figure 4). The restriction map reported by Ragg and Priebisch (1988) included a *Hin*dIII site near the 5′ end of intron 2, suggesting that a mutation at this location has occurred.

Chromosome Mapping. For regional assignment of the HCII gene, Southern blots of *Hin*dIII-digested DNA isolated from the somatic cell hybrids were hybridized to the radiolabeled probe HCII7.2. Both hamster and mouse genomic DNA hybridized weakly to the probe, yielding two fragments each (15 and 3.5 kb for hamster; 15 and 2.9 kb for mouse). The HCII7.2 probe hybridized strongly to a 20-kb *Hin*dIII fragment from the hamster–human hybrid AM-27, which contains the proximal part of human chromosome 22, but did not hybridize to the hamster–human hybrid A3EW2-6A, containing the distal part of human chromosome 22 (Figure 5). These results indicate that the HCII gene is located on 22p-q11.2. In a control experiment, the same Southern blot was hybridized to an anonymous chromosome 22 specific probe (phom37H) known to map to 22q12-qter (Metzdorf et al., 1990). phom37H hybridized to A3EW2-6A, but not AM-27, thereby mapping phom37H even more distal on chromosome 22 (q13.1-qter).

Since the short arm of chromosome 22 is essentially made up of heterochromatin, we speculated that the HCII gene is located within 22q11. Therefore, we decided to sublocalize the HCII gene with the CML-derived cell line K562. K562 cells have amplified the BCR-*abl* fusion gene and sequences of 22q11 on a marker chromosome (Heisterkamp et al., 1983; Selden et al., 1983). Hybridization of the HCII7.2 probe to *Hin*dIII-digested DNA from K562 cells gave a 10–20-fold greater signal as compared to normal placenta DNA (Figure 5), indicating that the HCII gene maps within the amplified region in 22q11. Hybridization to the control probe phom37H showed the same signal intensities of bands in placenta and K562 DNA. The results are summarized in Figure 6.

In addition, in situ hybridization to metaphase chromosomes from K562 cells was performed. In 32 mitoses evaluated, 519 grains were associated with metaphase chromosomes; 29.1% (n = 151) of all grains were found on the normal chromosome 22, and a strong accumulation (256 grains; 49.1%) on an acrocentric marker chromosome was noted. Distribution of the grains on chromosome 22 was 7.9% (n = 12) on 22pter-p11.1, 78.1% (n = 118) on 22cen-q12, and 13.9% (n = 21) on 22q13-qter. This result confirmed our localization with the somatic cell hybrids and K562 cells.

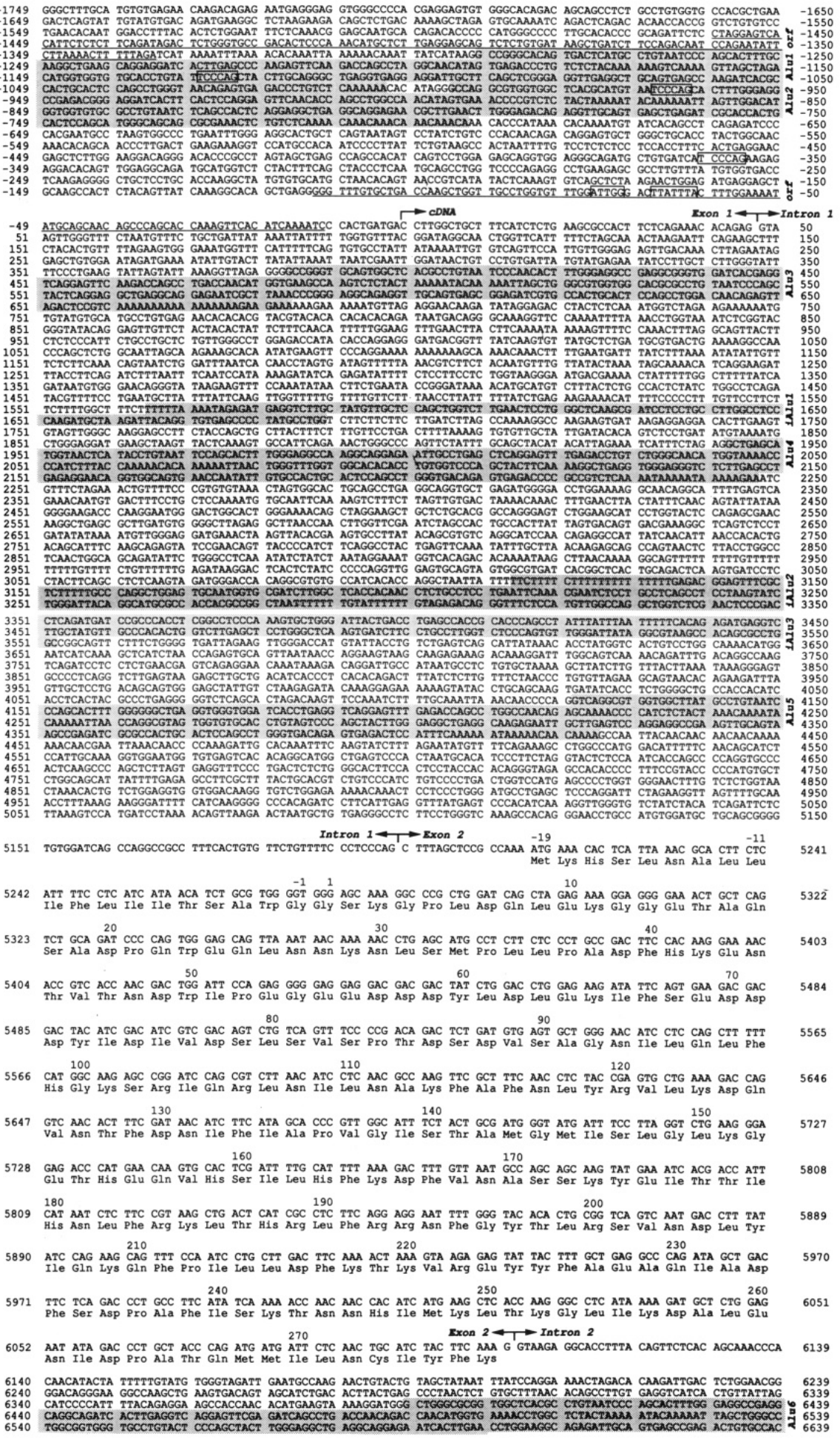
Neighborhood analysis of the HCII gene with probes known to be amplified in K562 cells as well as IGLV, BCR, and D22S10 (Budarf et al., 1989) was performed by using pulsed-field gel electrophoresis. A series of complete and partial digests of K562 DNA with the rare cutting enzymes *Nru*I, *Not*I, *Mlu*I, and *Spl*I produced fragments in the range of 1.5–7.0 megabasepairs (Mbp). Various overlapping fragments with BCR and D22S10 indicate that the HCII gene maps at least 2 Mbp proximal to BCR and is localized in the vicinity of D22S10 (data not shown).

Discussion

We have isolated clones comprising the entire HCII gene from a human leukocyte genomic library in EMBL-3 λ phage. Sequence analysis of two overlapping clones confirmed the organization of the HCII gene proposed by Ragg and Preibisch (1988). As these investigators pointed out, the positions of the introns in the HCII gene are highly conserved in the genes for α_1-antitrypsin and angiotensinogen; the introns are not conserved in genes for some of the other homologous members of the serpin family, including antithrombin III. The HCII gene contains 10 complete Alu repetitive sequences and 1 partial Alu repeat (Schmid & Jelinek, 1982). In addition, 7 open reading frames ≥300 nucleotides in length were identified in the 5′-flanking region and in introns 1, 2, and 4. Portions of three of the open reading frames located at nucleotide positions −1459 to −1334, −112 to −11, and 7379–7501 appear to be nonrandom coding sequences as defined by Fickett (1982); however, none of the open reading frames codes for a known protein or peptide.

The transcription initiation site of the HCII gene has not been determined with certainty and may, in fact, be heterogeneous. S1 nuclease protection experiments revealed 6 potential initiation sites spanning ~100 nucleotides (Ragg & Preibisch, 1988). The TATA-like sequence TTATTTA occurs at positions −67 to −61 downstream from the first transcription initiation site, and an inverted CCAAT sequence (ATTGG) occurs at positions −75 to −71. Whether these elements comprise the promoter of the HCII gene remains to be determined. Further inspection of the 5′-flanking sequence demonstrates three inverted acute-phase squences (TCCCAG) (Adrian et al., 1986; Fowlkes et al., 1984) at −1128 to −1123, −967 to −962, and −360 to −355 that may mediate an increase in HCII biosynthesis during the acute-phase reaction (Sandset & Andersson, 1989).

The HCII gene sequence was localized to chromosomal band 22q11, a region in the human genome of remarkable association with human pathology. Chromosomal band 22q11 represents only 0.25% of the haploid autosomal length. Despite its small cytologic size, it bears significance on a genetic scale as indicated by its involvement in a number of clinically relevant human diseases. These include the t(9;22) of CML, the t(8;22) of Burkitt lymphoma variant, the t(11;22)(q23;q11) of Ewing's sarcoma and neuroepithelioma (Emanuel et al., 1986), the cat eye syndrome, and DiGeorge syndrome (Carey et al., 1990; Kaplan et al., 1987; Kaplan & Emanuel, 1989). We have demonstrated that the HCII gene is coamplified in the CML-derived cell line K562, which contains an amplified BCR-*abl* fusion gene and surrounding sequences such as IGLV and D22S10 (Budarf et al., 1989). Amplification of these sequences is not a characteristic feature of CML and, to our knowledge, is restricted to K562 cells; thus, we would not expect overexpression of HCII to occur in CML. By long-range mapping, we have placed the amplified HCII gene in the vicinity of D22S10 and BCR in K562 cells. Since the physical distance of BCR and D22S10 is not precisely known (Rouleau et al., 1989), from our data we cannot discriminate whether the HCII gene is localized proximal or distal to D22S10. In addition, we have shown that the HCII gene is highly polymorphic with two restriction enzymes, *Hin*dIII and *Bam*HI, thus making it useful for linkage analysis within 22q11. A third possible RFLP with *Msp*I was recently reported (Turner et al., 1990). The HCII gene should aid in


cDNA
Exon 1
Intron 1
Intron 1
Exon 2
Exon 2
Intron 2
Alu2
Alu1
orf
orf
Alu3
Alu4
ΔAlu1
ΔAlu2
ΔAlu3
Alu5
Alu6

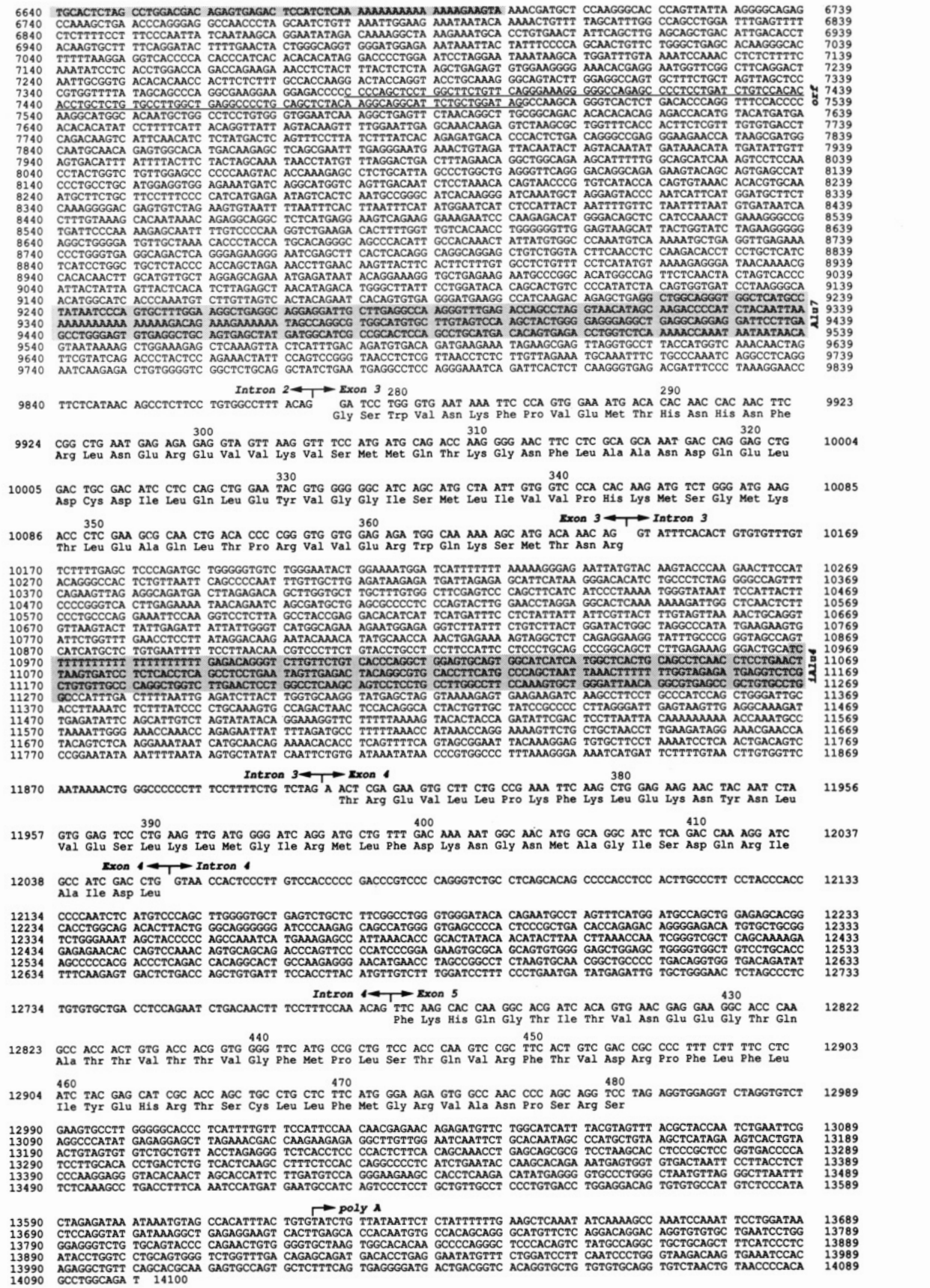


FIGURE 2: Structure of the HCII gene. The genomic clones shown in Figure 1 were sequenced by the dideoxynucleotide method on both strands. Nucleotide +1 corresponds to the 5′ end of the longest cDNA clone isolated by Ragg and Preibisch (1988) and does not indicate the transcription initiation site. The upstream sequence is numbered from −1 to −1749. The intron–exon boundaries and the polyadenylation site are indicated by arrows. The first exon includes only the 5′-noncoding sequence. The translated amino acid sequence of the remaining exons is shown with the amino acids numbered above their respective codons. The shaded boxes indicate Alu repeats (*Alu*) on the coding strand and inverted Alu repeats (*iAlu*) on the complementary strand. *iAlu1* is a partial repeat. The three nonrandom open reading frames (*orf*) are underlined. The open boxes indicate the positions of inverted acute-phase sequences (TCCCAG), the TATA-like sequence (TTAATTTA), and the inverted CCAAT sequence (ATTGG).

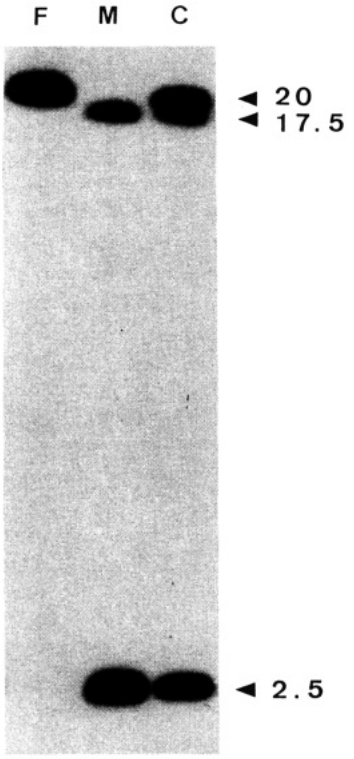


FIGURE 3: Mendelian inheritance of the *Hin*dIII RFLP. Genomic DNA was digested with *Hin*dIII, electrophoresed on a 0.6% agarose gel, transfered to a Zetaprobe membrane, and hybridized with the radiolabeled probe HCII7.2. Fragment sizes are indicated in kilobases. F, father; M, mother; C, child.

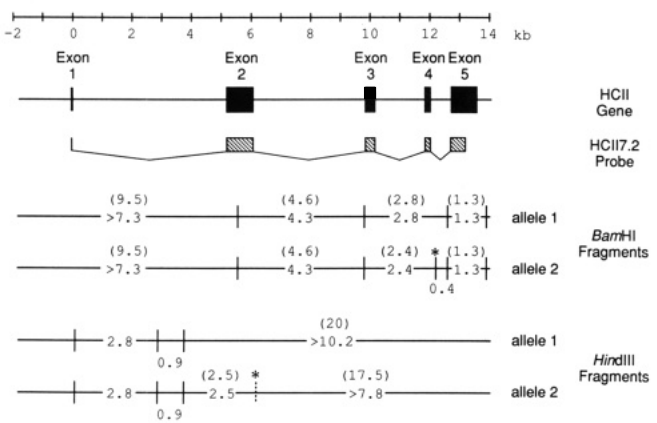


FIGURE 4: Proposed restriction site polymorphisms in the HCII gene. The solid boxes indicate the locations of exons in the HCII gene derived from the data in Figure 2. The shaded boxes indicate the portions of the exons spanned by the probe HCII7.2. *Bam*HI and *Hin*dIII sites predicted from the genomic sequence are indicated by solid vertical lines. The dashed vertical line indicates the *Hin*dIII site that would result from a G → T substitution at position 6259. The proposed polymorphic sites are indicated by asterisks. The numbers in parentheses indicate the apparent lengths of the restriction fragments observed on Southern blots. The numbers not in parentheses indicate the fragment sizes as determined from the DNA sequence.

constructing a long-range map of the chromosomal band 22q11, thus allowing a more precise localization of the various breakpoints within 22q11 (Budarf et al., 1989; McDermid et al., 1989).

A role for HCII in human physiology has not been established. Although patients have been reported in whom inherited HCII deficiency was associated with thrombosis (Sié et al., 1985; Tran et al., 1985), the incidence of HCII deficiency in the general population appears to be similar to that in patients with thromboembolic disorders (Andersson et al., 1986). Experiments in vitro indicate that the thrombin–HCII reaction is stimulated by dermatan sulfate proteoglycans synthesized by fibroblasts or vascular smooth muscle cells (McGuire & Tollefsen, 1987). When the vascular endothelium is disrupted, HCII may interact with dermatan sulfate in the subendothelium to control the mitogenic (Bar-Shavit et al., 1986; Glenn et al., 1980) and chemotactic (Bar-Shavit et al., 1983) activities of thrombin, as well as its effects on blood coagulation. Furthermore, cleavage of HCII by neutrophil proteases releases peptides with potent chemotactic activity for neutrophils and monocytes, suggesting that HCII may regulate the acute inflammatory response (Hoffman et al., 1989).

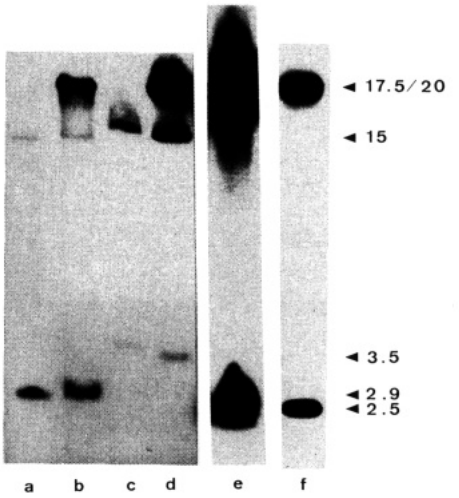


FIGURE 5: Sublocalization of the HCII gene on chromosome 22. DNA from mouse liver (a), PgMe25Nu (b), A3EW2-6A (c), AM-27 (d), K562 cells (e), and human placenta (f) was digested with *Hin*dIII, electrophoresed on a 0.8% agarose gel, and hybridized to the radiolabeled HCII7.2 probe. Sublocalization was performed with the aid of the rodent–human somatic cell hybrids in lanes b–d and K562 cells. Amplification (10–20-fold) of the HCII gene in K562 cells is shown by comparison to human placenta. Fragment sizes are given in kilobases.

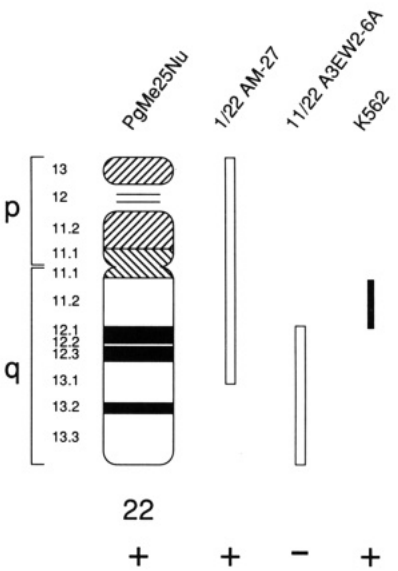


FIGURE 6: Portions of chromosome 22 retained in the somatic cell hybrids and amplified in K562 cells. The presence (+) or absence (−) of the human HCII hybridizing or coamplified fragments is indicated at the bottom of the figure.

ACKNOWLEDGMENTS

We are grateful to V. Klein for excellent technical assistance, to E. Göttert for help with the in situ hybridization, and

to K. D. Zang for supporting the chromosome 22 studies.

Registry No. DNA (human heparin cofactor II gene), 131010-07-6; heparin cofactor II (human protein moiety reduced), 120718-35-6; heparin cofactor II (human protein moiety mature form reduced), 102055-84-5; heparin cofactor II, 81604-65-1.

REFERENCES

Adrian, G. S., Korinek, B. W., Bowman, B. H., & Yang, F. (1986) *Gene 49*, 167–175.

Andersson, T. R., Larsen, M. L., Handeland, G. F., & Abildgaard, U. (1986) *Scand. J. Haematol. 36*, 96–102.

Barnes, W. M., Bevan, M., & Son, P. H. (1983) *Methods Enzymol. 101*, 98–122.

Bar-Shavit, R., Kahn, A., Wilner, G. D., & Fenton, J. W., II (1983) *Science 220*, 728–731.

Bar-Shavit, R., Kahn, A. J., Mann, K. G., & Wilner, G. D. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 976–980.

Blinder, M. A., & Tollefsen, D. M. (1990) *J. Biol. Chem. 265*, 286–291.

Blinder, M. A., Marasa, J. C., Reynolds, C. H., Deaven, L. L., & Tollefsen, D. M. (1988) *Biochemistry 27*, 752–759.

Blinder, M. A., Andersson, T. R., Abildgaard, U., & Tollefsen, D. M. (1989) *J. Biol. Chem. 264*, 5128–5133.

Budarf, M., Sellinger, B., Griffin, C., & Emanuel, B. S. (1989) *Am. J. Hum. Genet. 45*, 128–139.

Carey, A., Roach, S., Williamson, R., Dumanski, J., Nordenskjold, M., Rouleau, G. A., Blin, N., Jalbert, P., & Scambler, P. (1990) *Genomics 7*, 299–306.

Davis, L. G., Dibner, M. D., & Battey, J. F. (1986) *Basic Methods in Molecular Biology*, pp 99–101, Elsevier, New York.

de Klein, A., Geurts van Kessel, A., Grosveld, G., Bartram, C. R., Hagemeijer, A., Bootsma, D., Spurr, N. K., Heisterkamp, N., Groffen, J., & Stephenson, J. R. (1982) *Nature 300*, 765–767.

Derechin, V. M., Blinder, M. A., & Tollefsen, D. M. (1990) *J. Biol. Chem. 265*, 5623–5628.

Devereux, J., Haeberli, P., & Smithies, O. (1984) *Nucleic Acids Res. 12*, 387–395.

Emanuel, B. S., Nowell, P. C., McKeon, C., Croce, C. M., & Israel, M. A. (1986) *Cancer Genet. Cytogenet. 19*, 81–92.

Feinberg, A. P., & Vogelstein, B. (1984) *Anal. Biochem. 137*, 266–267.

Fickett, J. W. (1982) *Nucleic Acids Res. 10*, 5303–5318.

Fowlkes, D. M., Mullis, N. T., Comeau, C. M., & Crabtree, G. R. (1984) *Proc. Natl. Acad. Sci. U.S.A. 81*, 2313–2316.

Garber, R. L., Kuroiwa, A., & Gehring, W. J. (1983) *EMBO J. 2*, 2027–2036.

Geurts van Kessel, A. H., Westerveld, A., de Groot, P. G., Meera Khan, P., & Hagemeijer, A. (1980) *Cytogenet. Cell Genet. 28*, 169–172.

Geurts van Kessel, A., Turc-Carel, C., de Klein, A., Grosveld, G., Lenoir, G., & Bootsma, D. (1985) *Mol. Cell. Biol. 5*, 427–429.

Glenn, K. C., Carney, D. H., Fenton, J. W., II, & Cunningham, D. D. (1980) *J. Biol. Chem. 255*, 6609–6616.

Gough, J. A., & Murray, N. E. (1983) *J. Mol. Biol. 166*, 1–19.

Groffen, J., Stephenson, J. R., Heisterkamp, N., de Klein, A., Bartram, C. R., & Grosveld, G. (1984) *Cell 36*, 93–99.

Hattori, M., & Sakaki, Y. (1986) *Anal. Biochem. 152*, 232–238.

Heisterkamp, N., Stephenson, J. R., Groffen, J., Hansen, P. F., de Klein, A., Bartram, C. R., & Grosveld, G. (1983) *Nature 306*, 239–242.

Hoffman, M., Pratt, C. W., Brown, R. L., & Church, F. C. (1989) *Blood 73*, 1682–1685.

Hofker, M. H., Breuning, M. H., Bakker, E., van Ommen, G. J., & Pearson, P. L. (1985) *Nucleic Acids Res. 13*, 7167.

Inhorn, R. C., & Tollefsen, D. M. (1986) *Biochem. Biophys. Res. Commun. 137*, 431–436.

Julier, C., Lathrop, M., Lalouel, J. M., Reghis, A., Szajnert, M. F., & Kaplan, J. C. (1985) *Cytogenet. Cell Genet. 40*, 664.

Kaplan, J. C., Aurias, A., Julier, C., Prieur, M., & Szajnert, M. F. (1987) *J. Med. Genet. 24*, 65–78.

Kaplan, J. C., & Emanuel, B. S. (1989) *Cytogenet. Cell Genet. 51*, 372–383.

Mattei, M. G., Philip, N., Passage, E., Moisan, J. P., Mandel, J. L., & Mattei, J. F. (1985) *Hum. Genet. 69*, 268–271.

McDermid, H. E., Budarf, M. L., & Emanuel, B. S. (1989) *Genomics 5*, 1–8.

McGuire, E. A., & Tollefsen, D. M. (1987) *J. Biol. Chem. 262*, 169–175.

Metzdorf, R., Herzog, R., Göttert, E., & Blin, N. (1990) *Hum. Hered. 40*, 145–152.

Mizusawa, S., Nishimura, S., & Seela, F. (1986) *Nucleic Acids Res. 14*, 1319–1324.

Müllenbach, R., Lagoda, P. J., & Welter, C. (1989) *Trends Genet. 5*, 391.

Ragg, H. (1986) *Nucleic Acids Res. 14*, 1073–1088.

Ragg, H., & Preibisch, G. (1988) *J. Biol. Chem. 263*, 12129–12134.

Ragg, H., Ulshofer, T., & Gerewitz, J. (1990) *J. Biol. Chem. 265*, 5211–5218.

Rouleau, G. A., Haines, J. L., Bazanowski, A., Colella-Crowley, A., Trofatter, J. A., Wexler, N. S., Conneally, P. M., & Gusella, J. F. (1989) *Genomics 4*, 1–6.

Sandset, P. M., & Andersson, T. R. (1989) *J. Intern. Med. 225*, 311–316.

Sanger, F., Nicklen, S., & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 5463–5467.

Schmid, C. W., & Jelinek, W. R. (1982) *Science 216*, 1065–1070.

Seabright, M. (1971) *Lancet 2*, 971–972.

Selden, J. R., Emanuel, B. S., Wang, E., Cannizzaro, L., Palumbo, A., Erikson, J., Nowell, P. C., Rovera, G., & Croce, C. M. (1983) *Proc. Natl. Acad. Sci. U.S.A. 80*, 7289–7292.

Siẽ, P., Dupouy, D., Pichon, J., & Boneu, B. (1985) *Lancet 2*, 414–416.

Tollefsen, D. M. (1989) in *Heparin: Chemical and Biological Properties, Clinical Applications* (Lane, D. A., & Lindahl, U., Eds.) pp 257–273, Edward Arnold, London.

Toneguzzo, F., Glynn, S., Levi, E., Mjolsness, S., & Hayday, A. (1988) *BioTechniques 6*, 460–469.

Tran, T. H., Marbet, G. A., & Duckert, F. (1985) *Lancet 2*, 413–414.

Turner, J., Grundy, C. B., Kakkar, V. V., & Cooper, D. N. (1990) *Nucleic Acids Res. 18*, 1664.

Van Ommen, G. J. B., & Verkerk, J. M. H. (1986) in *Human Genetic Diseases* (Davies, K. E., Ed.) pp 113–133, IRL Press, Washington.