Genomic Characterization of the Human Type I Cuticular Hair Keratin hHa2 and Identification of an Adjacent Novel Type I Hair Keratin Gene hHa5

Michael A. Rogers,* Hermelita Winter,* Lutz Langbein,† Thomas Krieg,‡ and Jürgen Schweizer*

Research Program *2 and †1, German Cancer Research Center, Heidelberg, and ‡Department of Dermatology, University of Cologne, Cologne, Germany

Hair keratins, a subset of the keratin multigene family expressed in hard keratinizing structures, previously have been thought to comprise four members of each subfamily, designated Ha1-4 (type I) and Hb1-4 (type II), which are differentially expressed in the cuticle and cortex of the hair follicle. This report describes the genomic cloning and sequencing of the human type I cuticular hair keratin hHa2, as well as the identification of a previously unknown human type I hair keratin gene. The 12.5-kilobase pair genomic clone ghkI2.12, obtained by hybridization of a human genomic deoxyribonucleic acid library with a 3′-complementary deoxyribonucleic acid probe of hHa2, as well as the partially overlapping 14.4-kilobase pair genomic clone ghkI2.17, isolated using a 5′-fragment of clone ghkI2.12, allowed the characterization of the entire hHa2 gene. The gene displays the same exon/intron structure as two previously characterized type I mouse and sheep hair/wool keratin genes with strict positional conservation of the six introns in the region coding for the central α-helix. At the 5′-extremity of clone ghkI2.17, i.e., approximately 8.0 kilobase pairs upstream of the hHa2 gene and oriented in the same transcriptional direction, lies the gene for a hitherto unknown human type I hair keratin. Clone ghkI2.17 contains partial sequence information for this gene beginning with intron 5 and extending to the end of the gene. Screening of a human scalp complementary deoxyribonucleic acid library with a 3′-fragment of the gene yielded a full length complementary deoxyribonucleic acid clone of the new hair keratin, which in continuation of the current nomenclature for hair keratins was termed hHa5. Remarkably, the hHa5 gene, which contains an additional 7th intron in its 3′-noncoding region, is expressed mainly in supramatricial cells and lowermost cortical cells of the hair bulb and thus constitutes a very early component of hair morphogenesis. Our results confirm the type specific clustering of keratin genes and indicate that the human type I hair keratin subfamily contains more members than previously assumed. ***Key words: keratin genes/structure/organization/expression. J Invest Dermatol 107:633–638, 1996***

The keratin multigene family comprises more than 30 individual but structurally related members that, by virtue of their sites of expression, are traditionally divided into epithelial-type keratins (soft α-keratins) and hair-type keratins (hard α-keratins). Two-dimensional gel electrophoresis studies of hair follicle protein extracts from various species, including man, have consistently led to the identification of eight major hair keratins (Heid *et al*, 1986; Lynch *et al*, 1986). Four of them, migrating in a narrow molecular weight range of 59–63 kilodaltons (kDa), belong to the basic to neutral type II subfamily and are designated Hb1-4. The remaining four keratins with molecular masses between 44–48 kDa represent the acidic type I subfamily Ha1-4. Together with a minor hair keratin pair, Hax/Hbx, the complement of hair keratins is therefore assumed to comprise 10 members (Heid *et al*, 1986; Lynch *et al*, 1986). Until now, the concept of eight major hair keratins seemed to be confirmed by recombinant DNA techniques which have led to the structural elucidation of four type II sheep wool keratins (Powell *et al*, 1992; Powell and Beltrame, 1994) as well as four type I mouse hair keratins (Bertolino *et al*, 1988, 1990; Winter *et al*, 1994). Expression studies of the murine type I hair keratins revealed that three structurally highly related keratins, mHa1, mHa3, and mHa4, are sequentially expressed in the hair cortex (Winter *et al*, 1994), as are three of the four type II wool keratins (Powell *et al*, 1992). The synthesis of the fourth type I member, mHa2, which is structurally unrelated to the three cortex keratins, is limited to the cuticle of the hair (Winter *et al*, 1994). For a long time, sequence elucidations of human hair keratins have lagged considerably behind those of the animal keratins. Recently, however, our laboratory succeeded in determining the primary structure not only of the type I hair keratins hHa1 (Fink *et al*, 1995), hHa2 (Rogers *et al*, 1995a), and hHa4 (unpublished results), as well as of two isoforms of hHa3 (Yu *et al*, 1993; Rogers *et al*, 1995b), but also of the type II hair keratins hHb1 (Rogers *et al*, 1995a) and hHb2-4 (Winter H, Rogers MA, Schweizer J: unpublished observations). In all cases, the assignment of the individual human hair keratins

Manuscript received February 12, 1996; revised May 22, 1996; accepted for publication June 7, 1996.

Reprint requests to: Michael A. Rogers, German Cancer Research Center, Research Program 2, 5th Floor, Room 526B, Im Neuenheimer Feld 280, 69120 Heidelberg, Germany.

Abbreviation: LEF1, lymphoid enhancer factor 1.

0022-202X/96/$10.50

was based on their striking sequence homology with the respective mouse and sheep keratins.

At present, complete gene structures of hair keratins are limited to two type II and two type I cortex keratins from different species (Wilson *et al*, 1988; Kaytes *et al*, 1991; Powell *et al*, 1992; Jones *et al*, 1996; Korge *et al*, 1996). The exon–intron organization of the two type II hair keratin genes corresponds to that found in type II epithelial keratin genes, i.e., nine exons interrupted by eight positionally conserved introns (Steinert *et al*, 1985; Powell *et al*, 1992; Jones *et al*, 1996; Korge *et al*, 1996). In contrast, the two type I hair keratin genes differ from type I epithelial keratin genes by the lack of a 7th intron in the region encoding the carboxy terminus (Steinert *et al*, 1985; Wilson *et al*, 1988; Kaytes *et al*, 1991). In this paper, we describe the structure and genomic organization of the human type I cuticular keratin gene hHa2. In addition, we show that another unusual and hitherto unknown functional type I hair keratin gene lies 5′ of and in close proximation to the hHa2 gene.

MATERIALS AND METHODS

Isolation of Genomic Clones For the isolation of genomic clones for the human hair keratin hHa2 gene, a human genomic library (DNA of circulating lymphocytes, partially digested with *Hae*III and cloned into the *Bam*HI site of the lambda DashII Vector; Stratagene, La Jolla, CA) was screened with a ^{32}P-labeled 250-bp *Sph*I–*Sty*I 3′-fragment of the previously published human hHa2 complementary deoxyribonucleic acid (cDNA) (Rogers *et al*, 1995a). Out of eight genomic clones isolated and purified by two rounds of rescreening, one clone, termed ghkI2.12, which also hybridized to a 144-bp *Eco*RI–*Nco*I 5′-fragment of the murine mHa2 cDNA (Winter *et al*, 1994) was further characterized. Subsequently, a 1.2-kilobase pair (kbp) *Not*I–*Pst*I fragment, derived from the 5′-end of clone ghkI2.12, was used in a second screening of the genomic library to isolate clone ghkI2.17, which extended the 5′-flanking region of the hHa2 gene and, in addition, contained 3′-sequences of a new hair keratin gene.

Restriction Enzyme Mapping Lambda DNA of genomic clones ghkI2.12 and ghkI2.17 was digested with *Hin*dIII and *Hin*dIII–*Not*I. The orientation of the resulting *Hin*dIII and *Hin*dIII–*Not*I subfragments of ghkI2.12 and ghkI2.17 was determined by means of Southern blot analysis with 3′- and 5′-specific probes of the hHa2 and mHa2 cDNAs, as well as sequence comparison of the isolated fragments with known type I hair keratins and finally by Southern hybridization of *Xho*I–*Xba*I digested ghkI2.12 using the *Hin*dIII and *Hin*dIII–*Not*I fragments of ghkI2.12 as probes.

Sequence Analysis The relevant restriction fragments of the two lambda clones were subcloned into Bluescript II KS+ and sequenced from both ends, initially using M13 and T3 primers and then 17mer oligonucleotides as walking primers. Sequencing was carried out according to the chain termination method of Sanger *et al* (1977) using both T7 Sequencing Kit (Pharmacia, Piscataway, NJ) and Sequenase Sequencing Kit (United States Biochemicals, Cleveland, OH). DNA assembly and data base comparisons were performed using the Heidelberg UNIX Sequence Analysis Resource (HUSAR, German Cancer Research Center).

Human Scalp cDNA Library The preparation of a cDNA library with polyA$^+$ ribonucleic acid (RNA) from normal human scalp, cloned directionally into the lambda ZapII vector (Stratagene Cloning Kit) has been described previously (Rogers *et al*, 1995a). The library was screened using a 2.5-kb *Pst*I–*Sal*I fragment of the genomic clone ghkI2.17, containing 3′-noncoding sequences as well as sequences coding for part of the carboxy terminus of the new type I human hair keratin. A 1.1-kb partial cDNA clone was isolated and sequenced. From this clone, a 300-bp *Pst*I–*Xho*I 3′-fragment was derived and used for a rescreening of the library, which yielded the full length clone phkI-5.

***In Situ* Hybridization** Surgically removed samples from human scalp were snap-frozen in isopentane cooled with liquid nitrogen. Cryostate sections (nominally 5 μm) were mounted on glass slides treated with 3-aminopropyl triethoxysilane (Sigma, St. Louis, MO), air-dried, and stored at −80°C. A 300-bp *Pst*I–*Xho*I fragment of cDNA clone phkI-5, comprising 3′-noncoding sequences and part of sequences coding for the carboxy terminus was subcloned into Bluescript II KS+ and used for the generation of both ^{35}[S]cytidine triphosphate ($_{\alpha}$CTP)-labeled antisense and sense probes. *In situ* hybridization was carried out essentially as described previously (Langbein *et al*, 1994).

A confocal laser scanning microscope (LSM 410 UV, Carl Zeiss, Oberkuchen/Jena, Germany) was used for recording *in situ* hybridizations.

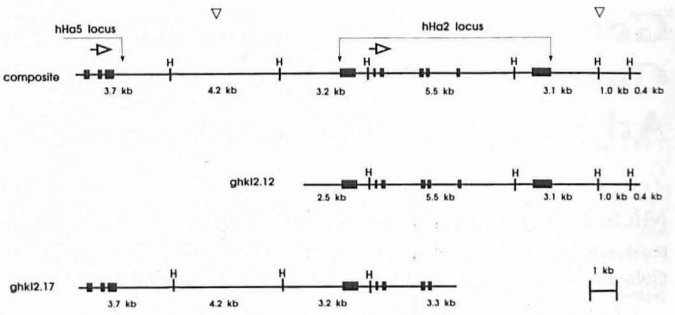

Figure 1. Composite gene domains with partial restriction map (*H*, *Hin*dIII) and indication of the loci of the human type I hair keratin gene hHa2 and of the new type I hair keratin gene hHa5. The 21-kb gene domain was assembled from the two genomic clones ghkI2.12 and ghkI2.17 that were used for the elucidation of the structural organization of the two hair keratin genes. Exons are shown as ■; ▷ indicate the direction of transcription of the genes. The numbers below the genomic clones and the composite gene domain indicate distances in kbp between *Hin*dIII and *Hin*dIII–*Not*I fragments. ▽ above the composite gene domain demarcate the size of the hHa2 gene sequence. The sequence is available from the EMBL data library, accession number X90761. For further details, see *Results*.

The instrument allows epi-illumination for reflection microscopy to detect the hybridization signals and transmitted light microscopy for hematoxyline staining both at a wavelength of 543 nm using a He-Ne-laser. Both images were combined by an overlay in false colors (transmission images in green, reflection images in red). Photographs were taken with a special high resolution image order system (Focus Graphics, Farchant, Germany). Video prints were made with a sublimation printer (Mitsubishi, Tokyo, Japan).

RESULTS

Isolation, Restriction Mapping, and Characterization of a hHa2 Genomic Clone Successive screening of a human genomic library with a 250-bp *Sph*I–*Sty*I probe that comprised 3′ noncoding sequences of the hHa2 cDNA (Rogers *et al*, 1995a), as well as with a 144-bp *Eco*RI–*Nco*I 5′-probe derived from the murine Ha2 cDNA (Winter *et al*, 1994) encompassing sequences coding for the amino terminus and the beginning of the α-helix of the hair keratin, yielded genomic clone ghkI2.12, which was further characterized. Digestion of ghkI2.12 with both *Hin*dIII (no restriction site in the vector) and *Not*I (unique restriction sites in both linkers) resulted in three *Hin*dIII fragments of 5.5, 3.1, and 1.0 kb as well as two *Hin*dIII–*Not*I fragments of 2.5 and 0.4 kb, respectively. Southern blot hybridization of these fragments with the specific 3′-probe of the hHa2 cDNA localized the 3′-end of the hHa2 gene to the 3.1. *Hin*dIII fragment, whereas the 2.5-kb *Hin*dIII–*Not*I fragment was detected by the 5′-probe of the mHa2 cDNA. Hybridization with the total hHa2 cDNA clone localized the hHa2 exons to both the 2.5-kb *Hin*dIII–*Not*I fragment and the 5.5- and 3.1-kb *Hin*dIII fragments. The correct assignment of the remaining 1.0-kb *Hin*dIII fragment was achieved by digestion of ghkI2.12 with *Xba*I (restriction sites in both linkers) and *Xho*I (no restriction site in the vector), followed by Southern blot hybridization of the resulting fragments using each of the individual *Hin*dIII and *Hin*dIII–*Not*I fragments as hybridization probes. This analysis yielded a 2.8-kb *Xba*I fragment that reacted with the 3.1- and 1.0-kb *Hin*dIII fragments as well as with the 0.4-kb *Hin*dIII–*Not*I fragment, thus leading to the overall order of the various ghkI2.12 fragments indicated in **Fig 1**.

Structural Organization of the hHa2 Gene Sequencing of ghkI1.12 from the 2.5-kb *Hin*dIII–*Not*I fragment through the 3.1-kb *Hin*dIII fragment confirmed that the complete locus of hHa2 is present in this area. The gene is composed of seven exons and six introns. (The hHa2 gene and flanking sequences, spanning 14117 nucleotides, is available from the EMBL data library, accession number X90761.) Intron positions were determined by comparison

of the genomic sequence with the hHa2 full length cDNA sequence (Rogers *et al*, 1995a). All of the exon–intron boundaries exhibit the minimal consensus sequence for donor and acceptor splice sites (Smith *et al*, 1989) with dinucleotides *gt* and *ag* at the respective 5′- and 3′-ends of the introns. A putative ATA box (CATAAA), identical to that present in the promoter of the type II wool keratin gene K2.9 (Powell *et al*, 1992) and related to that occurring in the murine Ha1 gene (Kaytes *et al*, 1991), and to a type I wool keratin gene (Wilson *et al*, 1988), is located 211 bp upstream of the ATG initiation of translation codon (nucleotides 4678–4680). Upstream of this ATA box is the sequence AAACCAAA (nucleotides 4004–4011) that has been identified in the promoters of some of epidermal keratins genes (Blessing *et al*, 1987). More importantly, the 5′-flanking sequence of the hHa2 gene contains the lymphoid enhancer factor 1 (LEF1) DNA binding site, CTTTGAA (nucleotides 4308–4314), which seems to be characteristic for hair follicle associated genes (Zhou *et al*, 1995). It should be mentioned that the hHa2 gene is subject to three natural polymorphisms that concern nucleotides c 9043, gt 9064/65, and t 9080. The two polymorphisms in the region coding for the helix termination motif lead to threonine–methionine and serine–arginine substitutions, respectively (for details, see Winter *et al*, 1996).

Extension of the 5′-Flanking Region of the hHa2 Gene The region 5′ to the ATG initiation of translation codon of the hHa2 gene located on genomic clone ghkI2.12 comprises only 1250 bp. In order to obtain more sequence information for the 5′-flanking region of the gene, a 1.2-kb *Not*I–*Pst*I fragment derived from the 2.5-kb *Hin*dIII–*Not*I fragment of ghkI2.12 (see **Fig 1**) was used as a probe for a second screening of the genomic library. One of the positive clones, termed ghkI2.17, was digested with *Hin*dIII and *Hin*dIII–*Not*I, yielding two *Hin*dIII–*Not*I fragments of 4.2 kb and 3.2 kb as well as two 3.7- and 3.3-kb *Hin*dIII–*Not*I fragments **(Fig 1)**. Southern blot hybridization with a probe containing solely α-helix coding sequences of the hHa2 cDNA showed labeling of the 3.2 *Hin*dIII fragment but also of the two *Hin*dIII–*Not*I fragments at the extremities of ghkI2.17, thus indicating the presence of a second type I keratin gene locus in the 5′-region of clone ghkI2.17. In order to determine the area of overlap between ghkI2.12 and ghkI2.17, the *Hin*dIII and *Hin*dIII–*Not*I fragments of ghkI2.17 were subcloned and partially sequenced. Subsequent sequence comparisons showed that the 3.3-kb *Hin*dIII–*Not*I fragment of ghkI2.17 was part of the 5.5-kb *Hin*dIII fragment of ghkI2.12 and that the 3.2-kb *Hin*dIII ghkI2.17 fragment was contained in the 2.5-kb *Hin*dIII–*Not*I fragment of ghkI2.12 (see **Fig 1**). In addition, this analysis allowed the determination of the overall order of the fragments on ghkI2.17 and thus allowed depiction of the organization of the entire 21.1-kb gene domain covered by the two genomic clones (**Fig 1**, *uppermost scheme*). Finally, clone ghkI2.17 was used to extend the sequence information of the 5′-flanking region of the hHa2 gene to about 4.7 kb (see EMBL data library, accession number X90761).

A New Hair Keratin Gene Is Located Upstream of the hHa2 Locus The preceding hybridization analysis has provided evidence for the location of a type I keratin gene upstream of the hHa2 gene on the distal 3.7-kb *Hin*dIII–*Not*I fragment of ghkI2.17. This was confirmed by sequencing the fragment from its 5′-end and subsequent amino acid translation of the elucidated sequence. As shown in **Fig 2**, the open reading frame starts with an amino acid segment that is almost identical in both sequence and length with that of exon 6 of hHa2 (see EMBL data library, accession number X90761). It thus corresponds to the end of the α-helical 2B subdomain. At the gene level the corresponding coding sequence is flanked by intron sequences that would positionally conform to introns 5 and 6 of the hHa2 gene (see EMBL data library, accession number X90761). The high content of carboxy terminal cysteines (8 residues) and prolines (10 residues) clearly points to a classification of the protein as a hair keratin (Winter *et al*, 1994). In order to further characterize this keratin, we isolated a 2.5-kb *Pst*I–*Sal*I fragment that comprised both carboxy terminus coding sequences

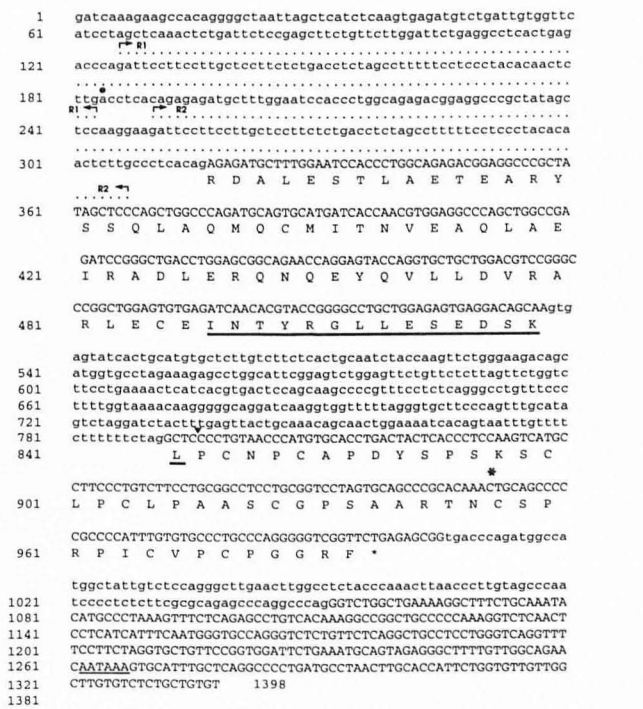

Figure 2. Partial nucleotide sequence of a new type I human hair keratin gene, hHa5, located on genomic clone ghkI2.17 (see Fig 1) Nucleotides in exons and amino acid translations are shown in capital letters. ▼ marks the end of the α-helix whose termination motif is underlined, as is the polyadenylation signal. ⌐, designated R1 and R2, respectively, demarcate two near direct repeats of 119 nucleotides with one base variation (the adenine marked by a ● in R1 is replaced by cytosine in R2). ∗ denotes a *Pst*I restriction site used for the generation of a specific 3′-probe of the gene. The sequence is available from the EMBL data library; accession number X90762.

and 3′-noncoding sequences (see **Fig 2**). In northern blots with RNA from both human scalp and hairless epidermis, this fragment detected a single messenger RNA (mRNA) of about 1.8 kb only in RNA from human scalp (results not shown). Subsequently, this fragment was used for the screening of a cDNA library constructed with polyA$^+$ RNA from human scalp which ultimately yielded the full length keratin clone phkI-5 shown in **Fig 3**. The encoded keratin comprises 425 amino acid residues and has a calculated molecular mass of 47.6 kDa. The expression of its mRNA in the human hair follicle was investigated by radioactive *in situ* hybridization using a specific 3′-probe of phkI-5. As shown in the follicular sections of **Fig 4*a*** and ***b***, transcripts of the keratin occur first in supramatricial cells of the hair bulb slightly below the critical zone of Auber (Auber, 1952). Expression of the mRNA ceases rather abruptly in the lowermost cortex region of the hair shaft, three to four cell layers above the apex of the dermal papillae. No hybridization signals were, however, observed in matrix cells and in trichocytes bordering on the dermal papillae **(Fig 4*b*)**. The cuticle of the hair shaft as well as the adjacent inner and outer root sheaths were also free of label **(Fig 4*a*)** as was the interfollicular scalp epidermis (results not shown).

DISCUSSION

Until now, two type I hair keratins, the murine Ha1 and a 47.6 kDa wool keratin, have been characterized at the genomic level (Wilson *et al*, 1988; Kaytes *et al*, 1991). mHa1 is the largest member of the murine type I hair keratin subfamily, and according to the sequential features of its carboxy terminus, the 47.6 kDa wool keratin represents the ortholog of murine hair keratin mHa3 (Winter *et al*, 1994). In this paper, we present the genomic sequence of a human

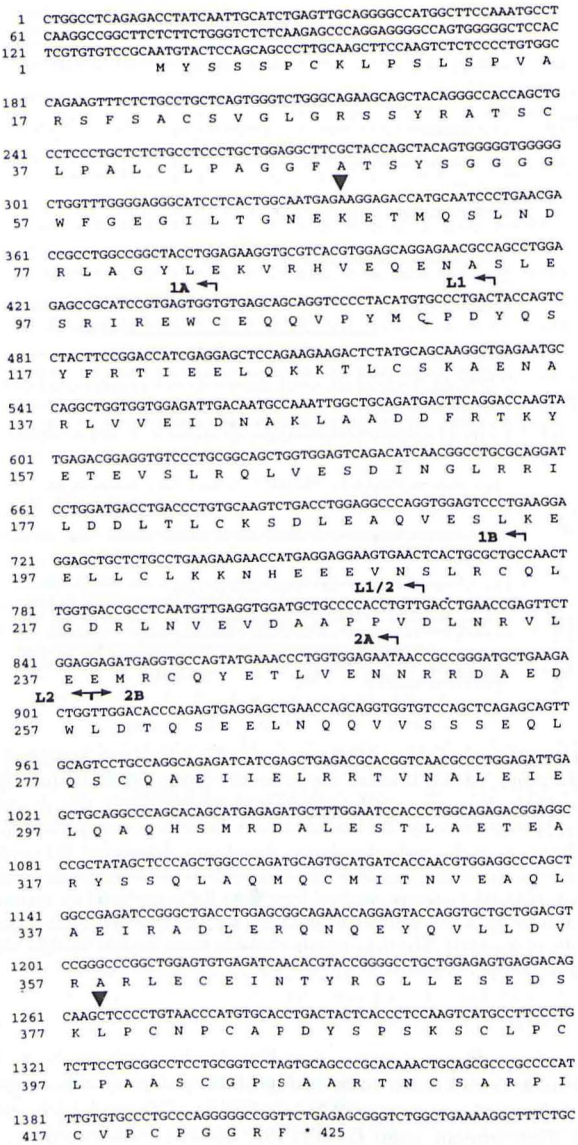

Figure 3. Nucleotide sequence of cDNA clone phkI-5 and derived amino acid sequence of the encoded keratin hHa5. ▼ demarcate the α-helical domain whose subdomains are indicated by ¬¬. The polyadenylation signal is *underlined*. The sequence is available from the EMBL data library; accession number X90763

type I hair keratin, hHa2. Whereas the mouse and sheep keratin genes are expressed in the cortex of the hair shaft, the expression of the Ha2 gene is restricted to the cuticle of the hair (Winter H, Rogers MA, Langbein L, Schweizer J, 1994, and unpublished observations). Regarding their genomic organization, the three type I hair keratin genes are identical in that they are divided into seven exons by six introns. In all three genes, the six introns are located in the region coding for the α-helix and they interrupt this central gene domain at exactly the same phase of the triplet codons including the lysine codons split by introns 2 and 6. Previously, a high degree of length and sequence conservation has been noted for the small intron 4 of the mHa1 and 47.6 kDa wool keratin gene, as well as of two other, partially characterized, murine type I hair keratin genes, but not for type I epithelial keratin genes. It was therefore suggested that intron 4 sequences might be involved in the regulation of type I hair keratin expression (Kaytes *et al*, 1991). Because, however, a comparison of the hHa2 intron 4 with those of the mouse and sheep hair keratin genes does not confirm such a length and sequence conservation, (results not shown), a general relevance of intron 4 for the regulation of type I hair keratin gene expression can be excluded. A functionally better defined regulatory sequence, 5′-CTTTGAA-3′, is present in the promotor region of the hHa2 gene. This sequence was originally described in the promotors of five hair follicle associated genes and referred to as HK1 motif (Rogers and Powell, 1993). Only recently, it was found that the HK1 motif corresponds to the DNA binding site of LEF1 (core consensus sequence, CTTTGA/TA/T/A), which is actively synthesized in both the developing and mature hair follicle (Zhou *et al*, 1995). Two minimal promoters of hair follicle associated genes containing the LEF1 binding motif were stimulated by LEF1 in chloramphenicol-acetyl-transferase reporter gene assays of epidermal keratinocytes. Moreover, transgenic experiments with altered LEF1 expression revealed striking abnormalities in both hair patterning and morphogenesis (van Genderen *et al*, 1994; Zhou *et al*, 1995). In the meantime the LEF1 binding site has been discerned in 13 of 13 published promoters of hair associated genes. In general, the first such motif is positioned about 160–250 bp upstream from the TATA box (Zhou *et al*, 1995). The site seems most optimal for factor binding if the core motif is preceded by a C and contains A's instead of T's at its 3′ end (Giese *et al*, 1992; Giese and Grosschedl, 1993). All these criteria are perfectly fulfilled by the LEF1 motif in the proximal hHa2 promoter. Besides the first LEF1 binding site upstream of the TATA box, five more LEF1 site related sequence motifs have been localized in a 3688-bp 5′-region of the mHa1 gene (Zhou *et al*, 1995). We were not able to detect additional LEF1 consensus binding sequences in the 4700-bp-long 5′-flanking region of the hHa2 gene. This may indicate that, in regulatory terms, the first motif upstream of the TATA box may be the most important one.

The extension of the hHa2 promotor region led to the detection of another, partial keratin gene sequence about 8 kbp upstream of the hHa2 gene. The open reading frame of the partial gene yielded amino acid sequences for the terminal part of the α-helix and the carboxy terminus that displayed the typical cysteine and proline enrichment of hair keratins (Winter *et al*, 1994). Sequence comparisons of the carboxy terminus with that of hHa1, hHa2, hHa3-I, hHa3-II, and hHa4 revealed no identity with these keratins, thus indicating the occurrence of a hitherto unknown hair keratin gene. The expression of the new gene in the hair follicle was demonstrated by the isolation of a corresponding cDNA clone from a scalp cDNA library and by *in situ* hybridization studies. In numerical continuation of the current nomenclature for hair keratins, we propose to name the new hair keratin hHa5. Preliminary data from our laboratory indicate, however, that hair keratins hHa1–hHa5 do not yet constitute the full complement of functional type I hair keratins in man.

A comparison of the partial hHa5 gene sequence with the corresponding cDNA clone reveals the presence of an intron in the 3′-noncoding region of the gene (see **Figs 2** and **3**), which, together with the six conserved α-helical introns, brings the overall number of introns of the hHa5 gene to seven. In this respect, the hHa5 gene deviates from the type I genes of hHa2, mHa1, and the 47.6 kDa wool keratins but resembles type I cytokeratin genes, which consistently possess a 7th intron at varying positions of the region coding for the carboxy terminus (Steinert *et al*, 1985). Interestingly, however, a shift of the 7th intron into the 3′-noncoding region has recently also been reported for the human type I cytokeratin K9 gene (Reis *et al*, 1994).

Another remarkable feature involving intron sequences of the hHa5 gene concerns the occurrence of two near direct 119-bp repeats that, except for one base, are completely conserved **(Fig 2)**. The first repeat and half of the second repeat are located in intron 5, whereas the remaining half of the second repeat encodes the beginning of exon 6 **(Fig 2)**. This configuration implies that the

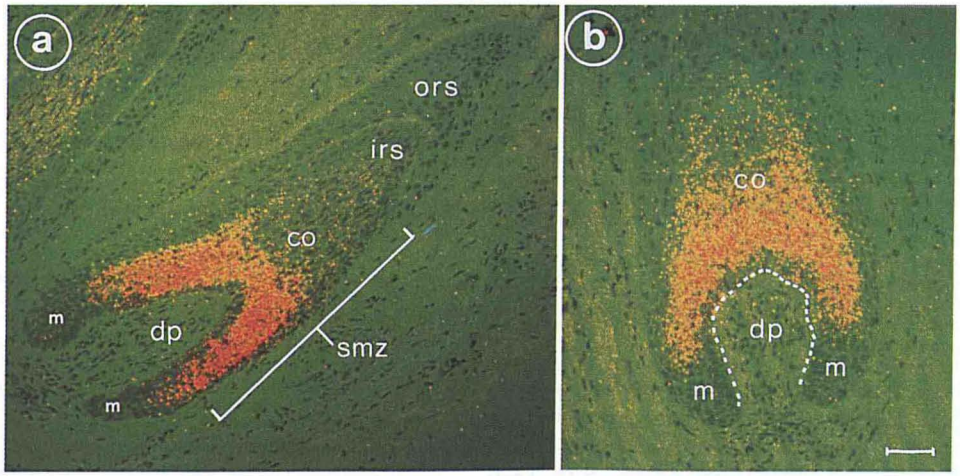

Figure 4. Radioactive *in situ* hybridization with a specific 3′-fragment of clone phkI-5 to slightly oblique sections of two different human anagen hair follicles *(a,b)*. ORS, outer root sheath; IRS, inner root sheath; CO, cortex; DP, dermal papillae; M, matrix; SMZ, supramatricial zone. *Scale bar*, 100 μm.

3′-splice site involved in the elimination of intron 5 as well as its flanking regions are also present in the first repeat. The relevance of this repeat in the correct splicing of the hHa5 gene is not known, but the accidental use of the 3′-splice site in the first repeat would lead to a hHa5 keratin devoid of the terminal part of its α-helix and thus unable to form functional filaments with its type II partner. Whether such an event occurs and whether such a splice variant could be causally related to the occurrence of spontaneous, non-heritable hair disorders remains to be seen.

The particular characteristics of the hHa5 gene also carry over to the hHa5 protein, which displays unique features when compared with the known human type I hair keratins. Thus, the cortex keratins hHa1, hHa3-I, hHa3-II, and hHa4 possess amino termini that are closely conserved in both length and sequence, as well as carboxy termini that exhibit extended sequence homologies adjacent to their α-helices. In contrast, the head and tail portions of hHa5 virtually lack sequence homology with those of the cortex keratins. In this respect, hHa5 resembles the cuticular hair keratin hHa2, which also exhibits complete amino- and carboxy-terminal sequence deviation from the cortex keratins (Rogers *et al*, 1995a). Remarkably, hHa5, and hHa2 share another feature, namely, their early expression in the hair follicle. The first transcripts of hHa5 can be detected just above the matrix cells of the hair bulb (see **Fig 4**). Similarly, the onset of mRNA expression of both mHa2 and hHa2 begins at the height of the critical zone of Auber, i.e., the widest point of the hair bulb (Winter *et al*, 1994; Rogers MA, Langbein L, Winter H, Schweizer J: unpublished observations). Therefore, both expression patterns differ substantially from those of the structurally related cortex keratins whose lowermost expressed type I member, Ha1, and type II member, K2.9, can first be detected three to four cell layers above the apex of the dermal papillae (Kaytes *et al*, 1991; Powell *et al*, 1992; Bowden *et al*, 1994; Rogers MA, Langbein L, Winter H, Schweizer J: unpublished observations).

Interestingly, there is evidence from some sources that the genes of functionally and/or sequentially related keratins tend to be grouped closer to each other than distantly related keratin genes (Bader *et al*, 1988; Rosenberg *et al*, 1988; Filion *et al*, 1994, Powell and Beltrame, 1994; Troyanovsky and Leube, 1994; Yoon *et al*, 1994). Apparently, the hHa2 and hHa5 genes, separated from each other by only 8 kb, represent a subcluster of early expressed hair keratin genes that may, however, comprise additional as yet unknown hair or hair related keratin genes.

We thank Christian Wolff and Silke Praetzel for excellent technical assistance and Margit MacLeod for typing the manuscript. This work was supported by the Deutsche Forschungsgemeinschaft, Grant Schw 539/1-1.

REFERENCES

Auber L: The anatomy of follicles producing wool-fibres with special reference to keratinization. *Trans R Soc Edinburgh* 62:191–254, 1952

Bader BL, Jahn L, Franke WW: Low level expression of cytokeratins 8, 18 and 19 in vascular smooth muscle cells of human umbilical cord and in cultured cells derived therefrom, with an analysis of the chromosomal locus containing the cytokeratin 19 gene. *Eur J Cell Biol* 47:300–319, 1988

Bertolino AP, Checkla DM, Heitner S, Freedberg IM, Yu DW: Differential expression of type I hair keratins. *J Invest Dermatol* 94:297–303, 1990

Bertolino AP, Checkla DM, Notterman R, Sklaver I, Schiff TA, Freedberg IM, DiDona GJ: Cloning and characterization of a mouse type I hair keratin cDNA. *J Invest Dermatol* 91:541–546, 1988

Blessing M, Zentgraf H, Jorcano JL: Differentially expressed bovine cytokeratin genes. Analysis of gene linkage and evolutionary conservation of 5′-upstream sequences. *EMBO J* 6:567–575, 1987

Bowden PE, Hainley S, Parker G, Hodgkins MB: Sequence and expression of human hair keratin genes. *J Dermatol Sci* 7:s152–s163, 1994

Filion M, Sarafian V, Lussier M, Belanger C, LaPointe L, Royal A: Arrangement of a cluster of three mouse type I keratin genes expressed sequentially during esophageal-type epithelial cell differentiation. *Genomics* 24:303–310, 1994

Fink P, Rogers MA, Korge B, Winter H, Schweizer J: A cDNA encoding the human type I hair keratin hHa1. *Biochim Biophys Acta* 1264:12–14, 1995

Giese K, Cox J, Grosschedl R: The HMG domain of lymphoid enhancer factor 1 bends DNA and facilitates assembly of functional nucleoprotein structures. *Cell* 69:185–195, 1992

Giese K, Grosschedl R: LEF-1 contains an activation domain that stimulates transcription only in a specific context of factor-binding sites. *EMBO J* 12:4667–4676, 1993

Heid HW, Werner E, Franke WW: The complement of native alpha-keratin polypeptides of hair-forming cells: a subset of eight polypeptides that differ from epithelial cytokeratins. *Differentiation* 32:101–119, 1986

Jones DO, Zimojic D, Popescu N, Bowden PE: Sequencing and chromosomal localization of two human hair-specific type II keratin genes. *J Invest Dermatol* 106:907, 1996

Kaytes PS, McNab AR, Rea TJ, Groppi V, Kawabe TT, Buhl AE, Bertolino AP, Hatzenbuhler NT, Vogeli G: Hair-specific keratins: characterization and expression of a mouse type I keratin gene. *J Invest Dermatol* 97:835–842, 1991

Korge BP, Richards G, Pünter C, Mauch C, Junginger H, Lin J-P, Rogers M, Winter H, Hamm H, Schweizer J, Krieg T, Traupe H: Monilethrix links to the keratin type II cluster at 12q13 and cloning of a possible candidate gene. *J Invest Dermatol* 106:843, 1996

Langbein L, Heid HW, Moll I, Franke WW: Molecular characterization of the body site-specific human epidermal cytokeratin 9: cDNA cloning, amino acid sequence, and tissue specificity of gene expression. *Differentiation* 55:57–71, 1994

Lynch MH, O'Guin WM, Hardy C, Mak L, Sun T-T: Acidic and basic hair/nail ("hard") keratins: their colocalization in upper cortical and cuticle cells of the

human hair follicle and their relationship to "soft" keratins. *J Cell Biol* 103:2593–2606, 1986

Powell B, Crocker L, Rogers G: Hair follicle differentiation: expression, structure and evolutionary conservation of the hair type II keratin intermediate filament gene family. *Development* 114:417–433, 1992

Powell BC, Beltrame JS: Characterization of a hair (wool) keratin intermediate filament gene domain. *J Invest Dermatol* 102:171–177, 1994

Reis A, Hennies HC, Langbein L, Digweed M, Mischke D, Drechsler M, Schrock E, Royer-Pokora B, Franke WW, Sperling K, Kuster W: Keratin 9 gene mutations in epidermolytic palmoplantar keratoderma (eppk). *Nat Genet* 6:174–179, 1994

Rogers GE, Powell BC: Organization and expression of hair follicle genes. *J Invest Dermatol* 101:50s–55s, 1993

Rogers MA, Nischt R, Korge B, Krieg T, Fink TM, Lichter P, Winter H, Schweizer J: Sequence data and chromosomal localization of human type I and type II hair keratin genes. *Exp Cell Res* 220:357–362, 1995a

Rogers MA, Schweizer J, Krieg T, Winter H: A novel human type I hair keratin gene: evidence for two keratin hHa3 isoforms. *Mol Biol Rep* 20:155–161, 1995b

Rosenberg M, RayChaudhury A, Shows TB, Le Beau MM, Fuchs E: A group of type I keratin genes on human chromosome 17 characterization and expression. *Mol Cell Biol* 8:722–736, 1988

Sanger F, Nicklen S, Coulson AR: DNA sequencing with chain terminating inhibitors. *Proc Natl Acad Sci USA* 75:5463–5467, 1977

Smith CWJ, Patton JG, Nadal-Ginard B: Alternative splicing in the control of gene expression. *Annu Rev Genet* 23:527–577, 1989

Steinert PM, Steven AC, Roop DR: The molecular biology of intermediate filaments. *Cell* 42:411–420, 1985

Troyanovsky SM, Leube RE: Activation of the silent human cytokeratin 17 pseudogene-promoter region by cryptic enhancer elements of the cytokeratin 17 gene. *Eur J Biochem* 225:61–69, 1994

van Genderen C, Okamura RM, Farinas I, Quo R-G, Parslow TG, Bruhn L, Grosschedl R: Development of several organs that require inductive epithelial–mesenchymal interactions is impaired in LEF-1–deficient mice. *Genes Dev* 8:2691–2703, 1994

Wilson BW, Edwards KJ, Sleigh MJ, Byrne CR, Ward KA: Complete sequence of a type-1 microfibrillar wool keratin gene. *Gene* 73:21–31, 1988

Winter H, Rogers MA, Mahr B, Cremer M, Krieg T, Navsaria HA, Leigh IM, Schweizer J: The region coding for the helix termination motif and the adjacent intron 6 of the human type I hair keratin gene hHa2 contains three natural, closely spaced polymorphic sites. *J Invest Dermatol* 106:544–548, 1996

Winter H, Siry P, Tobiasch E, Schweizer J: Sequence and expression of murine type I hair keratins mHa2 and mHa3. *Exp Cell Res* 212:190–200, 1994

Yoon S-J, LeBlanc-Straceski J, Ward D, Krauter K, Kucherlapati R: Organization of the human keratin type II gene cluster at 12q13. *Genomics* 24:502–508, 1994

Yu J, Yu D-W, Checkla DM, Freedberg IM, Bertolino AP: Human hair keratins. *J Invest Dermatol* 101:56s–59s, 1993

Zhou P, Byrne C, Jacobs J, Fuchs E: Lymphoid enhancer factor 1 directs hair follicle patterning and epithelial cell fate. *Genes Dev* 9:700–713, 1995