Mutations in the H1 and 1A Domains in the Keratin 1 Gene in Epidermolytic Hyperkeratosis

Jun-Mo Yang, Constantin C. Chipev, John J. DiGiovanna,* Sherri J. Bale, Lyuben N. Marekov, Peter M. Steinert, and John G. Compton

Skin Biology Branch, National Institute of Arthritis and Musculoskeletal and Skin Diseases; and *Dermatology Branch, National Cancer Institute, National Institutes of Health, Bethesda, Maryland, U.S.A.

In the autosomal dominant disorder epidermolytic hyperkeratosis, the structural integrity of the keratin intermediate filaments is altered in the suprabasal layers of the epidermis. We and others have used genetic linkage studies and mutation analysis to establish that single amino acid substitutions in either the keratin 1 or keratin 10 chains can cause epidermolytic hyperkeratosis. However, a larger database of mutations is required to better understand the relationship between specific mutations in these keratin chains and their effect on keratin filament structure. A larger database will also provide a catalog that may be useful for genetic counseling purposes. In this paper, we report the identification of three new mutations of the keratin 1 chain of epidermolytic hyperkeratosis probands in highly conserved residues in the H1 or beginning of the 1A rod domain segments. These correspond to regions involved in molecular overlaps between neighboring molecules in keratin filaments. Using an in vitro assay, synthetic peptides bearing these substitutions show diminished capacity to disassemble preformed filaments in vitro in comparison to the wild type peptides. Moreover, analyses of all mutations in epidermolytic hyperkeratosis known to date demonstrate remarkable clustering in the molecular overlap region. We conclude that non-conservative substitutions in the overlap region are likely to interfere with normal keratin filament structure and function, leading to pathology. Key words: genodermatoses/intermediate filaments/epidermolysis bullosa simplex/epidermis. J Invest Dermatol 102: 17–23, 1994

The quantitatively major differentiation products of epidermal cells are the cytoskeletal keratin intermediate filaments (KIF). Normal human epidermis expresses two major type I/type II pairs of KIF chains: type II keratin 5 (K5) and type II keratin 14 (K14) in the proliferative-potential basal cells, and type II K1 (and K2) and type I K10 in the suprabasal terminally differentiating epidermal cells. Together, these chains form the cytoplasmic KIF complement of these cells that in large measure is responsible for the physical properties of the entire epidermis [1,2]. Recent exciting molecular genetic and molecular biology experiments have demonstrated that mutations in the KIF chains expressed in the epidermis are etiologically responsible for at least three distinct types of autosomal dominant skin diseases, epidermolysis bullosa simplex (EBS), epidermolytic hyperkeratosis (EH), and palmoplantar hyperkeratosis (reviewed in [3-7]).

In EH, the integrity of the suprabasal layers of the epidermis is seriously disrupted, resulting in an altered differentiation program and highly abnormal, thickened stratum corneum [8,9]. Several types of data have now led to the understanding that the major underlying cause of this genetic disorder is mutation in either the K1 or K10 chain. Cells of affected EH epidermis contain abnormal clumps of the KIF in the cytoplasm often found around the nucleus and near the cell periphery at desmosomal junctions [8-10]. Using this as an important clue to the disease etiology, our laboratory tested keratins, among other candidate genes, for linkage to EH in one large family, and found complete co-segregation of the disorder with the type II keratin gene locus on chromosome 12q [11]. Because K1 is the major type II keratin expressed in suprabasal epidermis, we reasoned that a defect in this gene may underlie the pathology. Indeed, sequence analyses of DNA obtained from all affected members of the family revealed a single point mutation in residue position 17 of the H1 region of the keratin 1 gene that changed a CTT (leucine) to a CCT (proline), but not in any unaffected members of the family [12]. Simultaneously, two other research groups reached the conclusion that defects in the K1 or K10 gene could cause EH. One of these employed transgenic technology using a chimeric gene construct bearing the likely promoter region of the K10 gene and part of the K10 coding sequences coupled to a deletion form of the K14 coding sequences, to direct expression of the dominant-negative phenotype to the suprabasal layers of the epidermis. Newborn transgenic mice displayed an abnormal epidermal phenotype and changes in the suprabasal layers reminiscent of the clinical features of EH in humans [13]; thus it was concluded that K10 is a likely candidate for the disease etiology. Subsequent sequence analyses of DNA from affected patients of two occurrences of EH revealed the same arginine to histidine mutation at residue position 10 (R10H) of the 1A rod domain segment in the K10 protein [14]. A concurrent study by another group reported point mutations in either the K1 or K10 chain in three EH patients [15]: arginine to histidine at position 10 (R10H) of the 1A rod domain segment of K10, leucine to serine at position 15 (L15S) of the 1A rod domain segment of K10, and glutamic acid to glutamine at position 118 (E118Q) of the 2B rod domain segment of K1.

Manuscript received August 20, 1993; accepted for publication October 4, 1993.

Reprint requests to: Dr. Peter M. Steinert, National Institutes of Health, Building 6, Room 425, Bethesda, Maryland 20892.

Abbreviations: EBS, epidermolysis bullosa simplex; EH, epidermolytic hyperkeratosis; KIF, keratin intermediate filaments; PASA, polymerase chain reaction amplification of specific alleles.

0022-202X/94/$06.00 Copyright © 1994 by The Society for Investigative Dermatology, Inc.
cause EH in another family [16]. Finally, linkage analyses in three other families have been reported in which the disease gene was mapped to the 12q region that contains the type II genes [17], thereby again implicating the K1 gene in disease etiology.

Thus five different mutations that cause EH have been identified to date in either the K1 or K10 chains. There are a number of important reasons it is necessary to generate a larger body of data on mutations that cause disease. First, for clinical diagnosis, it would be useful to determine whether there is a correlation between disease presentation and the type or location of mutation in either keratin chain. If indeed there are “hotspots” for mutations in the keratin chains as suggested [5,6], this would in turn facilitate identification of further mutations in the future simply because investigators will know where to look. Second, the creation of such a “catalog” of mutations will be of benefit for the future development of rational therapeutic approaches for the disease. Similarly, the availability of this information will be useful in genetic counseling of patients whose planned children are at risk of inheriting the disease-causing mutation. Additionally, an important spinoff may be a better understanding of KIF structure and function, which in turn may aid in diagnosis and treatment.

To establish a catalog, we are in the process of analyzing the DNA from a large number of families with EH. In this paper, we report three new mutations in the K1 chain resulting in amino acid substitutions in different probands, and provide evidence that such substitutions are likely to severely affect KIF structure and function in vitro, and by extrapolation, in vivo as well.

MATERIALS AND METHODS

Source of DNA DNA was isolated as described [18] from freshly drawn blood or transformed lymphocyte cell lines obtained from both affected and unaffected family members, as well as from additional unaffected unrelated persons.

Direct PCR Amplification of Keratin 1 Gene Exons Oligonucleotide primers for amplification of K1 exons in five fragments, and primers for direct sequencing of these fragments are given below, based on the published human K1 gene sequence [12,19,20]. Some primers were as used in a previous study [12] and contained 5’ sequences encoding a restriction enzyme site (lowercase letters) in addition to K1 specific sequences (uppercase letters); plus sign means sense strand and minus sign means antisense strand. The primers are located in the TATA box and the initiation codon (19), 5’-TGCTTTGGGGTAGAGGAGTGTGATTCCCTCCTCT. Primer GII(+), located in the V1 subdomain, 5’-TTTGGTTGCGTTGGTGGAATTT. Primer J(+), also located in V1, 5’-TGGTTTGGGGTGGAATTT. Primer 4n(+), located in intron 1, 5’-TGGTTTGGGGTGGAATTT.

DNA Fragment I (Including Exon 1): Primer 1(+) located between the TATA box and the initiation codon [19], 5’-TGGTTTGGGGTAGAGGAGTGTGATTCCCTCCTCT. Primer GI(+), located in the V1 subdomain, 5’-TTTGGTTGCGTTGGTGGAATTT. Primer J(+), also located in V1, 5’-TGGTTTGGGGTGGAATTT. Primer 4n(+), located in intron 1, 5’-TGGTTTGGGGTGGAATTT.

DNA Fragment II (Including Exon 2): Primer 5(+)(intron 1, EcoRI), 5’-ggggaattcGAGCGAGTGAGTGAGTTTTGAGGAGTGTGATTCCCTCCTCT. Primer 13(+) located in exon 2 (BamHI), 5’-ggggaattcGAATTCATGGAGTGTGATTCCCTCCTCT. Primer 6(-) located in exon 2 (EcoRI), 5’-ggggaattcGAATTCATGGAGTGTGATTCCCTCCTCT. Primer 4n(-), located in intron 1, 5’-TGGTTTGGGGTGGAATTT.

DNA Fragment III (Including Exons 3, 4, 5, and 6): Primer 6(+)(intron 2 (EcoRI), 5’-ggggaattcGAGCGAGTGAGTGAGTTTTGAGGAGTGTGATTCCCTCCTCT. Primer E4sp(+)(intron 3, 5’-CCCTCAGCAGTTAGTATAAGG). Primer E5sp(-) in exon 4, 5’-GGTTTGGTGCTTGGTTCA. Primer 10(+) in exon 4 (BamHI), 5’-ggggaattcAACTGCAGCTGAGCTTGCTTATGTTAGTATAAGG. Primer 13(+) in exon 5, 5’-TGGTTTGGGGTGGAATTT. Primer V1(-) in exon 6, 5’-ggggaattcGAATTCATGGAGTGTGATTCCCTCCTCT.


Conditions of PCR About 400 ng of DNA was used per 100 µl reaction with 1.5 mM MgCl₂, 200 µM dNTPs, a 0.1 µM concentration of each of the primers, and 2.5 units of Taq polymerase (AmpliTaQ, Perkin Elmer Cetus). Amplifications were done on a DNA Thermal Cycler (Perkin-Elmer) for 35 cycles (1, denaturation for 60 seconds at 94°C; 2, a 90-second transition to 55°C; 3, annealing for 30 seconds at 55°C; 4, transition for 30 seconds to 72°C; and 5, extension for 90 seconds at 72°C). Amplifications were preceded by a hot start: 5–10 min incubation of the DNA and primer mix at 95–100°C, and equilibration to 80–85°C, at which time the premixed dNTPs, PCR buffer, MgCl₂, and Taq polymerase were added and cycling started. To increase the yield of PCR product for direct sequencing, amplification was performed on a 1:100 to 1:1000 dilution of the first amplification reaction. The primers used in the first PCR reactions were as follows: fragment I, primer 1(+)/4n(-) or GII(+)/4n(-); fragment II, primer 5(+)/6(-); fragment III, primer 6(+)/10(+); fragment IV, primer 4411(+)/2470(+); and fragment V, primer 13(+)/15(+). The amplified DNA fragments were resolved on 1.5% agarose gels and purified by the GeneClean II procedure (Bio 101) or by the Magic Prep DNA purification system (Promega).

Sequencing Procedures Direct sequencing of the PCR-amplified DNA was done in reactions using the single-strand binding protein (US Biochemicals) to prevent strand reassociation [20]. A 1:8 dilution of the labelling mix and abram 300–500 ng of gel-purified DNA fragment were used for each direct sequencing reaction with the Sequenase II kit (US Biochemicals). Standard 6% denaturing polyacrylamide gels were run. The primers used in direct sequencing reactions were as follows: exon 1, primer 1(+)/4n(-); exon 2, primer 5(+)/6(-); exon 3, primer 6(+); exon 4, primer E4sp(+); exon 5, primer E5sp(-); exon 6, primer 10(+); exon 7, primer 108(+)/12(+); and exons 8 and 9, primer 13(+)/15(+). These primers permitted analysis of most of the sequence of K1, except for the variable V1 and V2 end domain regions.

Screening Assays for K1 Mutations We developed assays that did not require DNA sequencing to facilitate the detection of each identified mutation in patients and to screen for the occurrence of that mutation in the normal population. The V11G substitution of family 1027 results in the loss of a Mae III restriction enzyme site in DNA amplified from exon 1 using primers GII(+)(and 4n(−)). Completed PCR reactions (50 µl) were adjusted as recommended by Boehringer-Mannheim Biochemicals with NaCl, MgCl₂, and Tris-HCl and digested with 2 units of Mae III for 2–4 h, and analyzed on 3% agarose gels (NuSieve 3:1, FMC Corp.). The N85 substitution in family 1002 results in the creation of a new Dde I site. The altered Dde I cleavage pattern was examined as above, except that 2 units of enzyme were added directly to the completed PCR reaction. For the S13P substitution in family 3138, a PCR amplification of specific alleles (PASA) assay was developed [21]. In this assay, PCR amplification is obtained specifically from the wildtype allele using primers that differ only at the 3' most nucleotide (see primers listed above). The primers SPw(−), specific for the wildtype base T, and SpSmu(+) specific for the mutation C, were each used with GII(+) as the second primer. Amplifications were performed with 0.1 mM dNTP, 1.5 mM MgCl₂, 120 ng DNA, and 0.1 µM primers in 25 µl for 30 cycles of 60-second denaturation at 95°C, 60-second annealing at 55°C, and 60-second extension at 72°C.

Linkage Analysis Linkage analysis in family 1027 was performed using the computer program LINKAGE, subprogram MLINK on a 386 personal computer [12]. EH was modeled as a fully penetrant disorder with allele frequency of 0.001. Analysis was robust over a range of disease and marker allele frequencies as all individuals included in the analysis were genotyped.

Synthetic Peptides The following peptides, based on published sequences of human K10 [22] and mouse and human K1 [23] chains, were synthesized and purified by high-pressure liquid chromatography, and their compositions confirmed and concentrations ascertained by amino acid analysis: 1) wildtype K1 H1, 1–36, PVCSPPQGQETQINQLQPLNVEIDPEIQVKVSKRE; 2) mutant K1 H1, 1–36, PVCSPPQGQETQINQLQPLNVEIDPEIQVKVSKRE; 3) wildtype K1 1A, 1A, 1–16, PVCSPPQGQETQINQLQPLNVEIDPEIQVKVSKRE; 4) wildtype K1 1A, 1A, 1–18, PVCSPPQGQETQINQLQPLNVEIDPEIQVKVSKRE; 5) wildtype K1 1A, 1A, 1–18, 35, VRFLEQONQKVLTQKWK; and 6) wildtype K1 1A, 1A, 18–35, VRFLEQONQKVLTQKWK.

KIF Disassembly Assay In Vitro Human KIF containing the K1/K5 K10/K14 chains were prepared from freshly excised foreskin epidermis as described previously [24], and equilibrated into a buffer of 5 mM Tris-HCl.
(pH 7.6) containing 1 mM ethylenediaminetetraacetic acid and 1 mM di-thiothreitol. Peptides were dissolved in the buffer at 3 ng/ml immediately before use, and added in 3 μl aliquots to 0.25 ml of KIF so as to effect a twofold molar excess of peptide over KIF heterodimer (110 kDa). At molar ratios exceeding about 2:1, peptides 1–3 began to aggregate the KIF instead, apparently due to an ionic zipper effect [25]. Changes in turbidity were monitored by measuring light scattering at 310 nm in a Beckman DU-65 spectrophotometer at 37°C for 60 min [12, 24, 25]. We have determined empirically by electron microscopy [11, 25] that for KIF 2–5 μm in length, turbidity due to light scattering is approximately proportional to the square of KIF length; thus minor changes in KIF length will result in significantly reduced turbidity.

RESULTS

During the course of the analysis of several EH families, we have discovered three hitherto undescribed mutations in the K1 gene that are the likely cause of the disease.

Clinical Description of Patients The diagnosis of EH was confirmed by histopathology from skin biopsy specimens obtained from at least one affected person of each family. Patients in family 1027 had primarily palmar and plantar involvement and truncal sparing. There was thick hyperkeratosis of the palms and soles, with a relatively smooth surface and sharp border delineated by a red halo. In some patients, there was restriction of full extension of the palms and digits. Blistering from trauma or irritation (such as restrictive clothing) was reported, although it was a mild component of the disease and tended to decrease with age. Children had some generalized involvement, particularly of the flexures and scaling at the corner of the lips, all of which improved over time. Patients 1002 and 3138 had severe palmar-plantar hyperkeratosis with contractions of the palms and digits. One patient had surgical amputations of two digits secondary to infection. Both patients were erythrodermic and had generalized hyperkeratosis, most severe over the joints, both flexor and extensor surfaces. Inter-articular areas and most of the trunk were covered with fine white scale. The analysis and confirmation of these mutations are set forth below by family.

The 1027 Family This is a three-generation family with six affected individuals (Fig 1a). Linkage analyses with the polymorphic markers developed in the V2 glycine loop sequences of K1 and K10 described previously [11, 26, 27] documented an obligatory recombination between the disease phenotype and the K10 locus at 17q (denoted by asterisk in Fig 1a). This excluded the possibility that a defect in K10 accounted for the disease in this family. The KRT1 (K1) gene, however, was fully informative and showed complete cosegregation with a maximum lod score of 2.1 at θ = 0. Direct sequencing of all exons of the K1 gene (except the V1 and V2 regions of exons 1 and 9, respectively) in an affected individual amplified by PCR revealed heterogeneity at a single base position in exon 1 GGC in one K1 allele and GTT in the other allele. This leads to a substitution of valine for glycine in residue position 11 of the H1 domain (V11G) (Fig 1b). DNA amplified from exon 1 of an unaffected individual did not show the same change. We then devised a test for the presence of this point mutation throughout the pedigree and in the normal control population. The T to G substitution destroys a restriction enzyme recognition site for Mde III (GTNAC) in the disease allele. The PCR product of exon 1 is 497 bp long, which for normal alleles is cleaved by Mde III into fragments of 297, 107, and 93 bp (Fig 1c). However, the loss of the site due to the mutation gives a composite fragment of 390 bp (297 + 93) from the disease allele, as well as the fragments expected from the normal allele. Thus individuals heterozygous for this mutation will display four bands upon Mde III digestion (Fig 1d). Using the appearance of the composite 390-bp band as the assay, we documented that all affected members of this family possessed the mutant allele, but only the normal allele was present in the eight unaffected members of this family (right set of lanes) (Fig 1d).

**Figure 1.** A valine to glycine mutation in family 1027. A) Pedigree of three generations. Below each person is indicated the genotype with respect to the polymorphic markers used for either the K10 (KRT10) or K1 (KRT1) loci. For example, person I-1 has size alleles 1 and 3 at the K10 locus, and the 1 and 2 size alleles at the KRT1 locus. These polymorphisms are due to variable numbers of glycine loops in the V2 subdomain regions of the K1 and K10 genes [26, 27]. By following the co-segregation of alleles through the family, it can be seen that in person II-4 (asterisk), there is an obligate recombination between the disease locus and the K10 locus; she inherited the K10-3 allele from her affected father, whereas her affected sib inherited the KRT10-1 allele and her unaffected sib inherited the KRT10-3 allele. On the other hand, there is complete cosegregation with the KRT1 (K1) locus, with all affected individuals inheriting the KRT1-2 allele from I-1, whereas none of the affecteds did. Individuals III-2 and III-6 were not tested (ND) as they were under 1 year of age, although they clearly had the EH phenotype. B) Portion of a sequencing gel of an affected (left set of lanes) and unaffected (right set of lanes) family member, which shows an additional band in the G position, representing a nucleotide substitution in one allele resulting in a valine to glycine change. C) This mutation causes the loss of a Mde III restriction enzyme site in affected persons. In the PCR product used for this work, affected persons possess a novel composite 390 bp band, which is used here to demonstrate that all unaffected persons of the family and unrelated unaffected persons do not carry this substitution. Markers are 123-bp ladder (Gibco-BRL).
Additional Nucleotide Sequence Variations in the K1 Gene

During the course of sequencing numerous normal K1 alleles in this study, we found three hitherto undescribed single nucleotide sequence variations. None of these segregate with EH suggesting they represent functionally equivalent alleles in the normal human population. Two silent codon variants were in the 2B rod domain segment, one at lysine 66 (codons AAA or AAG, estimated allele frequencies of 0.95 and 0.05, respectively), and the other at arginine 91 (CGT or CGC, estimated allele frequencies of 0.5 each). The third nucleotide variation is in the second position of codon 11 of the E2 end domain segment (AAG lysine or AGG arginine, estimated allele frequency of 0.75 and 0.25, respectively).

Materials and Methods)

Only the affected individual gave a PCR product with both the mutant and wildtype-specific primers, whereas both unaffected parents were negative with the primer bearing the mutation (Fig 3a). The PCR product for the mutant allele was not detected in 47 unrelated unaffected individuals using the mutated primer.

Additional Nucleotide Sequence Variations in the K1 Gene

The 1002 Family

This family consisted of a single affected individual and her two normal parents (Fig 2a). We amplified exons 1, 2, 5, and 7 of the K1 gene, which encode virtually the entire rod domain including the H1, 1A, L12, and 2B segments. By direct sequencing, we found a single nucleotide substitution in one allele in residue position 8 of the 1A rod domain segment (AAC to AGC, asparagine to serine) (N8S) in the affected child II-1 (Fig 2b). DNA from both unaffected parents did not show this mutation (only parent I-1 is shown). This nucleotide substitution creates a Dde I restriction enzyme site (CTNAG), which was then used as a test for the presence of the mutation in PCR-amplified DNA. In this analysis, the normal alleles yield Dde I fragments of 347 and 150 bp, but the affected allele yields fragments of 347, 103 (Fig 2c), and 47 bp. Because neither of the unaffected parents displays the new Dde I fragments, the A to G substitution in the affected child represents a new mutation. Similarly, 50 other normal unrelated individuals also lacked the fragments of 103 and 47 bp.

The 3138 Family

This family also consisted of a single affected child and two normal parents (Fig 3a). Direct sequence analyses of DNA amplified from exons 1, 2, 5, and 7 from the affected individual II-1 identified a single nucleotide substitution in one allele of K1 at position 13 of the 1A rod domain segment (TCC to CCC, serine to proline) (S13P) (Fig 3b). Neither of the parents showed this change. This was confirmed as a new mutation as both K1 alleles from the parents showed only the T at this position (only parent I-1 is shown). The nucleotide substitution removes an Mnl I restriction enzyme recognition site, but because there are several Mnl I sites in the PCR product of exon 1 utilized in this work, the pattern was too complex for use. Therefore, a PASA assay was used to selectively amplify the wildtype allele or the allele carrying the mutation (see Materials and Methods). Only the affected individual gave a PCR product with both the mutant and wildtype-specific primers, whereas both unaffected parents were negative with the primer bearing the mutation (Fig 3a). The PCR product for the mutant allele was not detected in 47 unrelated unaffected individuals using the mutated primer.
K10 to subfilamentous particles containing only one to four molecules within 90 min [12,25]. However, the peptide bearing the V11G amino acid substitution seen in family 1027 was much less effective; electron microscopy revealed (data not shown) that the slight decrease in turbidity (Table I, line 2) was due to reductions in KIF length from >2 to 0.5–1 μm. Even after 6 h, the KIF were still largely intact. The full-length wildtype 1A peptide also promoted complete disassembly of preformed KIF to subfilamentous particles, generally only one molecule in size [25]. This property is retained within the sequence of the first 18 residues rather than the second 18 (Table I, compare line 4 with lines 3 and 5) as peptide 5 containing residues 18–35 of the K1 chain had only a minor effect on KIF length. Notably, peptides bearing the N8S substitution of family 1002 or the S13P substitution of family 3138 (Table I, lines 6, 7) only shortened KIF, and did not promote massive disassembly, even after 6 h.

These data show that the three substituted peptides do not function like the wildtype peptides presumably because their structures have been altered. Because of their altered structures they are unable to effectively compete with the native sequences in the KIF in the same way as the wildtype peptides. By extension, we infer that the full-length mutant K1 chains are unable to function normally in the KIF in vivo because of altered structure, promoting the pathology of EH.

**DISCUSSION**

This laboratory has initiated a major project to identify the likely disease-causing mutation in a collection of EH cases for the purpose of building a “catalog” with two major purposes in mind. First, if it turns out that the mutations are clustered in specific sequence regions of the keratin chains, such data would make further characterization of mutations for this disease far easier and systematic because workers will then know where to focus attention in the future. Thus prospects for genetic counseling and prenatal diagnosis are likely to be enhanced. Second, the occurrence of the mutations in the keratin chains may provide important clues to the structure of KIF and their dynamic functions in cells. An important spinoff of such models may be improved therapeutic interventions, management, or even cure of the disease.

**Mutations in Keratins Causing EH Are Clustered** In this paper, we report three new mutations for EH that occur in the K1 chain. These occur only in affected patients and not in unaffected family members. We have described strategies for the analysis and sequencing of amplified DNA of patients, and tests to ensure that they are the likely causative mutation, not merely rare sequence polymorphisms within the normal human population. Each of the amino acid substitutions occurs in highly conserved sequence regions of the H1 and IA domains encoded by exon 1 of the K1 gene.

In addition we have reported one other mutation position in the K1 gene in the H1 domain of exon 1 sequences [12]. We know of two other mutations in the K1 chain ([16], W.H.I. McLean, personal communication). In other studies we have also identified the mutations in the K10 chain in a total of eight other families, representing five amino acid substitutions; that is, we have found mutations in either the K1 or K10 chain in 12 of 19 families. Together

**Table I.** Disassembly of Preformed KIF by Synthetic Peptidesa

<table>
<thead>
<tr>
<th>Peptide Number</th>
<th>Experiment</th>
<th>Light Scattering Remaining (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>KIF alone</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>Wildtype K1 H1</td>
<td>0 ± 2</td>
<td></td>
</tr>
<tr>
<td>Mutant K1 H1 V11G</td>
<td>77 ± 3</td>
<td></td>
</tr>
<tr>
<td>Wildtype full-length K1 1A</td>
<td>7 ± 1</td>
<td></td>
</tr>
<tr>
<td>Wildtype K1 1A residues 1-18</td>
<td>3 ± 1</td>
<td></td>
</tr>
<tr>
<td>Wildtype K1 1A residues 18-35</td>
<td>81 ± 1</td>
<td></td>
</tr>
<tr>
<td>Mutant K1 1A residues 1-18 NBS</td>
<td>74 ± 1</td>
<td></td>
</tr>
<tr>
<td>Mutant K1 1A residues 18-313P</td>
<td>77 ± 0</td>
<td></td>
</tr>
</tbody>
</table>

a The data are expressed as the percent of light scattering remaining after 60 min and are the average ± SD of two or three measurements on one batch of human K1/K10 KIF. We have used electron microscopy to assess these changes [12,24]. For example, 0% means the KIF have been totally disassembled to subfilamentous particles and no longer scatter light, whereas 77% means the KIF are still largely intact with minor shortening and/or unraveling. Wildtype peptides were designed from the published K1 sequence [19,23]. The three mutant peptides were designed based on the identified substitutions seen in the K1 chain in the three families studied here.
with existing published data [12-15], this means there are now 10 known different amino acid positions that are changed in the K1 or K10 chains in EH, from a total of 18 separate families reported (Table II). Compilation of these data (Table II) reveals a non-random distribution of mutations: two (11%) occur in H1, 14 (78%) (six positions) occur in the first half of the IA rod domain sequences, and two (11%) occur in sequences at the end of the 2B rod domain. Accordingly, we can speculate that most mutations in EH patients will be discovered following the amplification of only the first exon of the keratin chains, as described here for K1. Because 14 of 18 mutations occur in the first 15 residues of the IA rod domain, this region may constitute a “hot window” of increased mutational sensitivity in EH. Furthermore, mutations in the R10 position of K10 represent nearly half of all the cases of EH described so far. That is, this may represent a “hot spot” for mutations in EH [4-6]. Interestingly, this same position has been described as a hot spot for mutations in the K14 gene in EB [4-6,29]. However, we and others may have ascertained a biased set of Eh patients who were likely referred to research protocols and whose disease presentation may not be representative of the full spectrum of Eh patients. Nevertheless, the present insights should aid greatly the future identification of mutations in the keratin genes causing these genodermatoses.

Most Mutations Occur in Critical Overlap Regions of the Keratin Chains

Recent studies have suggested advanced two-dimensional models for K1/K10 KIF structure [25] that are of relevance for mutations in EH. Cross-linking analyses showed there are three principal modes by which a pair of antiparallel molecules may be aligned in the KIF, and have provided estimates of the axial dimensions of these modes: A1a, in which the two molecules are staggered by about 16 nm so that their 1B rod domain segments are overlapped; A2a, in which the two molecules are staggered by about 28 nm so that their segments 2B are overlapped; and A1b, in which the two molecules are aligned in register (staggered by about 1 nm). Several cross-links established that the axial repeat length (45 nm) of molecules in a KIF is demonstrably less than the length of each molecule (46 nm). Thus these data permitted identification of a fourth mode of alignment, termed A2b, in which about the first 1 nm of the 2B rod domain segment of one molecule overlaps with the first 1 nm of the IA rod domain segment of a similarly directed molecule. This corresponds to a nominal overlap of about 10 amino acids of the 1A and 2B sequences. This means that the sequence regions H1, beginning of 1A, end of 2B and H2 all lie very close to and/or overlap each other within a KIF. Furthermore, we [24] have suggested that the H1 (and possibly H2) subdomains, immediately adjacent to the beginning and end of the rod domain, are involved in the alignment of nearest neighbor molecules in KIF, perhaps by ionic interactions of basic residues in H1 with the canonical acidic residues of the end of the 2B rod domain segment. In the A1a and A2a alignment modes, these sequence regions also overlap with the middle of the rod domain, near the L2 segment. These structural principles may explain why these five regions represent the most highly conserved sequences throughout the family of keratin and other IF chains [30-32].

Interestingly, all 12 substitutions known so far for EH occur in or near these conserved overlap sequence regions. Of the 12 substitutions, six lie within the nominal 10 residue 1A-2B A2b overlap window; three others lie just outside the window; and two occur in H1. No mutations have been observed yet in H2 or L2 sequences. However, little is yet known about how the molecules fold in three dimensions to form the intact KIF. It appears that inappropriate amino acid substitutions in either of the H1, 1A, or 2B sequences that participate in the A2b overlap are the most critical for KIF structure and function. Presumably key ionic, van der Walls and/or H-bonds are prevented (or introduced) that markedly affect molecular packing, stability, and thus KIF dynamics in the living epithelial cells [5,6]. Apparently, mutations lying in residue positions just outside of the overlap window interfere with KIF structure as much as those within the window, so that amino acids that should align or interact in three dimensions are unable to do so. Further structural studies are in progress to address these questions.

Other Mutations in EH

This laboratory has been unable to find mutations in the rod domain sequences of the K1 or K10 chains in seven of 19 (37%) probands of EH. Further work will be necessary to determine whether such mutations reside in the V1 or highly variable V2 end domain sequences, other keratin genes expressed in the epidermis such as K2e and K9, or even in other genes coexpressed in terminal differentiating epidermis.

Is There a Correlation Between Disease Presentation and Mutation in EH?

We note that the clinical presentations of the family members of four cases reported here and previously [12] involving the K1 chain all display severe palmar-plantar (palm and sole) hyperkeratosis. In contrast, truncal involvement is pronounced whereas severe palmar-plantar hyperkeratosis is absent in those families with K10 mutations (J.J. DiGiovanna and S.J. Bale, submitted). This observation raises the intriguing question as to whether it may be possible in the future to predict the nature and location of the causative mutation based upon the clinical presentation of the disease. Analyses of many additional cases of EH may establish whether this association is valid.

We are very appreciative to the families who participated in this study, and especially grateful for the assistance and encouragement of the Foundation for Ichthyosis and Related Skin Types (FIRST) that continues to provide. We thank George Poy for the timely and efficient synthesis of all of the oligonucleotides and peptides used in this work, and Inga Tokar and Sandra Santucci for nursing support. Byron Van Dyke kindly assisted with the genotyping of family 1027. We are especially grateful to Drs. Irwin McLean and Birgit Lane for sharing with us their unpublished data on mutations in an additional EH proband.

REFERENCES

NEWS RELEASE

In 1994, the Certifying Examination of the American Board of Dermatology will be held at the Holiday Inn O'Hare Airport in Rosemont, Illinois on October 9 and 10, 1994. The deadline for receipt of applications is May 1, 1994. The next Dermatopathology Special Qualification Examination will be held in Chicago, Illinois on October 20, 1995. The deadline for receipt of applications is July 1, 1995. The next Examination for Special Qualification in Dermatological Immunology/Diagnostic and Laboratory Immunology will be held in Rosemont, Illinois on October 20, 1995. The deadline for receipt of applications is April 1, 1995.

For further information on these examinations, please contact: Harry J. Hurley, M.D., Executive Director, American Board of Dermatology, Henry Ford Hospital, Detroit, MI 48202.