Isolation and Characterization of Human Repetin, a Member of the Fused Gene Family of the Epidermal Differentiation Complex

Marcel Huber,* Georges Siegenthaler,† Nicolae Mirancea,‡ Ingo Marenholz,§ Dean Nizetic,¶ Dirk Breitkreutz,‡ Dietmar Mischke,§ and Daniel Hohl*

*Department of Dermatology, University Hospital of Lausanne, Lausanne, Switzerland; †Department of Dermatology, University Hospital of Geneva, Geneva, Switzerland; ‡Division of Differentiation and Carcinogenesis, German Cancer Research Center, Heidelberg, Germany; §Institute for Immungenetics, Humboldt University, Berlin, Germany; ¶Institute of Cell and Molecular Science, Barts & The London, Queen Mary's School of Medicine, London, UK

The human repetin gene is a member of the ''fused'' gene family and localized in the epidermal differentiation complex on chromosome 1q21. The ''fused'' gene family comprises profilaggrin, trichohyalin, repetin, hornerin, the profilaggrin-related protein and a protein encoded by c1orf10. Functionally, these proteins are associated with keratin intermediate filaments and partially crosslinked to the cell envelope (CE). Here, we report the isolation and characterization of the human repetin gene and of its protein product. The repetin protein of 784 amino acids contains EF (a structure resembling the E helix-calcium-binding loop-F helix domain of parvalbumin) hands of the S100 type and internal tandem repeats typical for CE precursor proteins, a combination which is characteristic for ''fused'' proteins. Repetin expression is scattered in the normal epidermis but strong in the acrosyringium, the inner hair root sheat and in the filiform papilli of the tongue. Ultrastructurally, repetin is a component of cytoplasmic non-membrane ''keratohyalin'' F-granules in the stratum granulosum of normal epidermis, similar to profilaggrin. Finally, we show that EF hands are functional and reversibly bind Ca^{2+}. Our results indicate that repetin is indeed a member of the fused gene family similar to the prototypical members profilaggrin and trichohyalin.

Keywords: keratinocyte/terminal differentiation/lq21/fused gene/repetin
J Invest Dermatol 124:998–1007, 2005

The skin protects our body against diverse environmental hazards. The frontline defense against microorganisms, physical stress, ultraviolet radiation, chemical irritation, or water loss is constituted by the epidermal barrier (Eckert *et al*, 1997; Presland and Dale, 2000; Fuchs and Raghavan, 2002). This barrier is established during epidermal terminal differentiation, a complex biochemical process changing keratinocytes from the proliferating phase in the basal layer to the mitotically inactive cells producing the cornified cell envelope (CE), the outermost protecting structure of our body. Thereby, keratinocytes migrate through the four epidermal layers, constantly modulating their proteomic profiles to the specific needs of the respective cell layer. During final maturation of the keratinocyte, the formation and subsequent dissolution of keratohyalin granules, the simultaneous alignment of keratin intermediate filament bundles and the formation of the CE are major structural changes. Their components constitue the bulk of proteins synthesized during late epidermal differentiation. The functional consequences of disturbed epidermal differentiation leading to diverse inherited disorders of keratinization are well demonstrated by the various genetic defects of epidermally expressed keratins, connexins, calcium pumps, enzymes, or protease inhibitors (Arin *et al*, 2002). The crucial role of the late stages of differentiation for epidermal homeostasis is demonstrated by lamellar ichthyosis resulting from keratinocyte transglutaminase deficiency (Huber *et al*, 1995).

In recent years, a number of genes specifying structural proteins expressed late during epidermal differentiation, e.g. proteins forming keratohyalin granules and CE precursor proteins, have been identified and found to be clustered on chromosome 1q21 (Backendorf and Hohl, 1992; Volz *et al*, 1993; Marenholz *et al*, 1996, 2001). Therefore, this region has been named the epidermal differentiation complex (EDC) (Mischke *et al*, 1996). The proteins encoded by the EDC genes can be, based on the primary sequence, combined into three groups.

The members of the first group are precursor proteins of the CE. The CE is a structure of 15 nm width consisting primarily of loricrin, involucrin, and small-proline rich proteins (SPRR), and to a lesser extent of proteinase inhibitors, keratins, desmosomal components, and keratohyalin proteins (Steinert and Marekov, 1995; Robinson *et al*, 1997; Steinert, 2000) that are crosslinked by the action of keratinocyte transglutaminase at the cell periphery (for a recent review, see Grenard *et al*, 2001). Simultaneously, ceramides are extruded from lamellar bodies into the intercellular space and become covalently attached to an involucrin scaffold on the outer surface of the CE (Swartzendruber *et al*, 1988; Marekov and Steinert, 1998). This leads to the formation of a bicomposite protein–lipid structure that progressively replaces the plasma membrane. Loricrin, involucrin, SPRR, the more recently identified xp5 or late

Abbreviations: CE, cell envelope; SPRR, small-proline rich proteins

envelope proteins (LEP) (Zhao and Elder, 1997; Marshall *et al*, 2001), and the distantly related NICE-1 (Marenholz *et al*, 2001) are encoded by genes with similar structures. The proteins encoded have homologies in the terminal protein domains and contain a variable number of internal tandem repeats specific for each protein, and are major precursors for the building of the CE. The human EDC contains one gene for loricrin and involucrin, 11 SPRR genes (two SPRR1, seven SPRR2, one SPRR3, and one SPRR4), 16 xp5/LEP genes (Marshall *et al*, 2001), and 1 NICE-1 gene (Marenholz *et al*, 2001). It is therefore thought that this gene family emerged from a common ancestor (Backendorf and Hohl, 1992).

The second group comprises 14 membres of the S100 family whose genes flank the EDC (Heizmann, 2002). S100 proteins are calcium-binding proteins because of the presence of two EF hands. They are regulatory proteins primarily involved in different steps of the calcium signal transduction pathway regulating cell shape, cell cycle progression, and differentiation (Eckert *et al*, 2004). They may play a role in the pathogenesis of epidermal diseases such as psoriasis, skin cancer, and skin inflammation (Eckert *et al*, 2004). Some S100 proteins (e.g. S100A10 and S100A11) have been isolated from purified CE and thus are crosslinked by transglutaminases (Robinson *et al*, 1997).

The third group combines EF hands and internal tandem repeats, the reason why these proteins are called "fused" members of CE precursor proteins. This group comprises profilaggrin, trichohyalin, repetin (Krieg *et al*, 1997), hornerin (Makino *et al*, 2001), and the protein encoded by c1orf10 (Xu *et al*, 2000). Profilaggrin, trichohyalin, and possibly other protein products of the fused gene group are components of cytoplasmic non-membrane "keratohyalin" granules in the stratum granulosum of normal epidermis, hair follicles, and mucosal keratinizing epithelia (Dale *et al*, 1994). Functionally, they are associated to keratin intermediate filaments and partially crosslinked to the CE. Profilaggrin is processed to functional filaggrin units in the terminal phase of epidermal differentiation (Resing *et al*, 1995). Further degradation of the filaggrin monomers to amino acids is thought to have an important role in the water retention capabilities of the skin. The "fused" gene, repetin, was recently cloned from mouse epidermis (Krieg *et al*, 1997). In this paper, we report the isolation and characterization of the human repetin gene and its protein product.

Results

Isolation of the gene and chromosomal assignment The mouse repetin gene was localized to chromosome 3F, a region that is syntenic to human chromosome 1q21 (Krieg *et al*, 1997). To clone the human homologue, we screened 40,000 clones from a chromosome 1-specific cosmid library (Nizetic *et al*, 1994) with a mouse repetin probe (Genbank X99251 nt 7198–7441). One positive cosmid, 31H23, was further investigated by Southern blot analysis using the same mouse repetin probe localizing the human repetin gene to a 7 kbp fragment flanked by *Hind*III and *Eco*RI sites (Fig 1*A*). Sequence analysis of this fragment yielded 6679

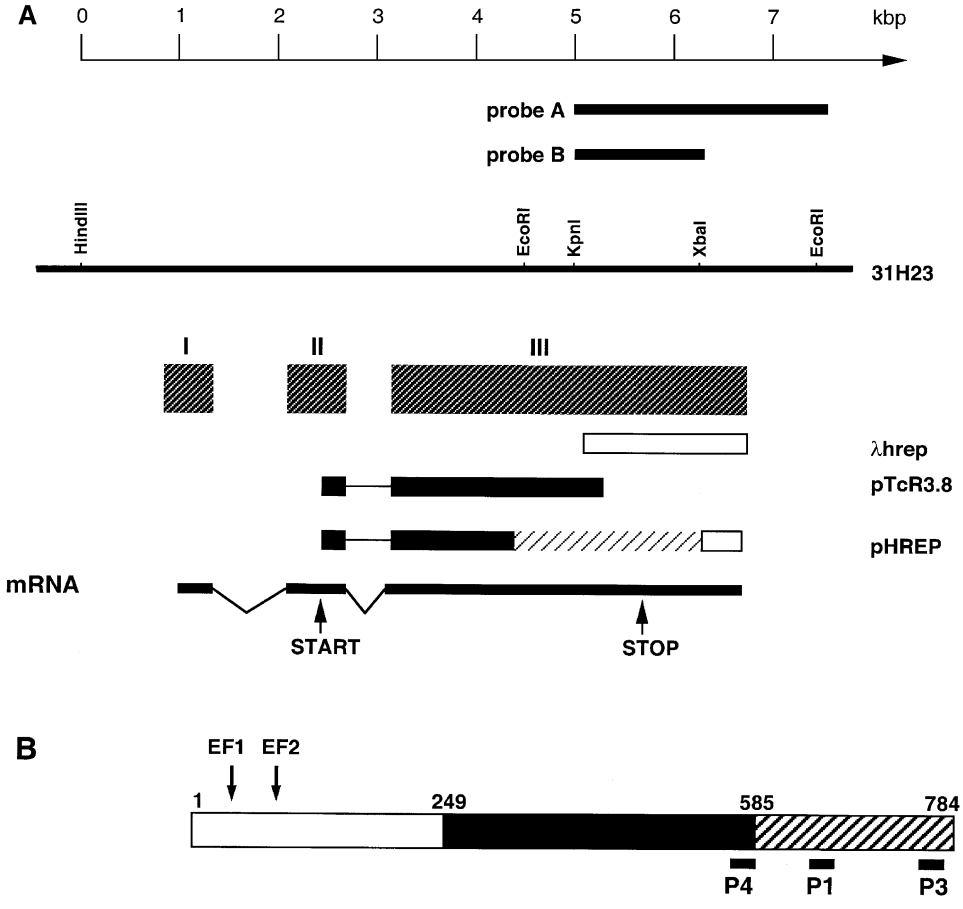

Figure 1
Genomic organization of the human repetin gene. (*A*) The genomic structure of cosmid 31H23 containing the repetin gene with exons I–III and the different cDNA clones are shown. Probes A and B designate DNA fragments used for cDNA cloning and hybridization experiments. (*B*) Depicts the repetin domain structure with the N-terminal part (amino acids 1–248) containing EF hands, the central repetitive domain (amino acids 249–584), and the C-terminal part (amino acids 585–784). P1, P3, and P4 indicate the positions of three peptides that were used for production of affinity-purified polyclonal antibodies.