Use of Articles in the Pachyonychia Congenita Bibliography

The articles in the PC Bibliography may be restricted by copyright laws. These have been made available to you by PC Project for the exclusive use in teaching, scholarship or research regarding Pachyonychia Congenita.

To the best of our understanding, in supplying this material to you we have followed the guidelines of Sec 107 regarding fair use of copyright materials. That section reads as follows:

Sec. 107. - Limitations on exclusive rights: Fair use
Notwithstanding the provisions of sections 106 and 106A, the fair use of a copyrighted work, including such use by reproduction in copies or phonorecords or by any other means specified by that section, for purposes such as criticism, comment, news reporting, teaching (including multiple copies for classroom use), scholarship, or research, is not an infringement of copyright. In determining whether the use made of a work in any particular case is a fair use the factors to be considered shall include - (1) the purpose and character of the use, including whether such use is of a commercial nature or is for nonprofit educational purposes; (2) the nature of the copyrighted work; (3) the amount and substantiality of the portion used in relation to the copyrighted work as a whole; and (4) the effect of the use upon the potential market for or value of the copyrighted work. The fact that a work is unpublished shall not itself bar a finding of fair use if such finding is made upon consideration of all the above factors.

We hope that making available the relevant information on Pachyonychia Congenita will be a means of furthering research to find effective therapies and a cure for PC.
Chapter 7

Keratins
E. Birgitte Lane
CRC Cell Structure Research Group, Department of Anatomy and Physiology, Medical Sciences Institute, University of Dundee, Dundee DD1 4HN, United Kingdom

INTRODUCTION

Keratins are everywhere. They are major components of household dust, and fragments of epidermal keratins are an almost unavoidable contamination in gel electrophoresis experiments. Before becoming thus displaced and disseminated, these proteins reach levels of 80% or more of the total cell protein in terminally differentiated keratinocytes of the skin, which is itself the largest organ of the human body. They are therefore likely to constitute a significant contaminating component in connective tissue sampling of the skin. Keratins are the major filamentous proteins of epithelial cells, and occur at their greatest density and complexity in the keratinocytes of the epidermis. They are intracytoplasmic, but at the cell periphery they link into the desmosome junctions which form cell-to-cell and cell-to-substrate attachment plaques. Thus, the keratin—desmosome interaction creates a three-dimensional network through the epithelium which must be of significant importance in the structural properties of the whole tissue. This chapter, therefore, briefly outlines what is understood about the keratin intermediate filaments in relation to their structure and tissue expression patterns and what, if anything, we can deduce from such data about their function.

KERATIN NOMENCLATURE

Intracellular keratin filaments have been recognized back to the earliest days of microscopy. The characteristic filament (tonofilament) bundles which they form within epidermal keratinocytes take up histochemical dyes strongly and are easily visible by light microscopy in the middle layers of the epidermis, where the bundles were referred to as tonofibrils. These bundles weave through the cell cytoplasm and loop into desmosomes [1] at the cell periphery, giving the cells a spiky appearance in tissue sections which resulted in the name stratum spinosum being given to this region of the epidermis. Electron microscopy subsequently revealed these fibrils to be made up of individual tonofilaments, which are now known to be composed of keratins [2], members of the large family of intermediate filament proteins.

The term keratin was originally used to refer to the material of the stratum corneum, the insoluble waterproof layer of dead cellular material, which is composed of terminally dif-

THE INTERMEDIATE FILAMENT GENE FAMILY

Keratin filament proteins are gene products of the much larger family of the intermediate filament proteins, which are expressed in one form or another in nearly every cell type of the vertebrate body. The intermediate filament gene family is now known to consist of at least 40 members, expressed in various human tissues (Fig. 1). Three-quarters of all these are keratin genes. The latest product of this gene family to be characterized is keratin 20 [6], and it seems probable that a small number of minor intermediate filament proteins still remain undiscovered.

The keratins were identified and characterized using biochemical analysis from two-dimensional gel electrophoresis: Although it is easy to extract intermediate filament proteins on the basis of their insolubility, a large number of related proteins coexist within a similar range of sizes and charges, making identification difficult with one-dimensional separation alone. The use of specific monoclonal antibodies has steadily gained in application over the last 10 years as an alternative way of identifying intermediate filament proteins, both by immunoblotting as well as by in situ immunohistochemistry techniques. The ultimate characterization of their
identity depends upon the sequencing of messenger RNA or
the genes themselves, and thence the deduction of their pri-
mary amino acid sequences and prediction of their secondary
structures. Very few intermediate filament proteins still re-
main unsequenced today.

Intermediate filaments fall into six classes (type I to type
VI) on the basis of amino acid sequence data. All these pro-
teins are intracellular. Most are cytoplasmic like the keratins,
and form long sinuous filaments of close to 10 nm in di-
ameter. All intermediate filaments are characteristically very
stable structures, as demonstrated by the traditional purifi-
cation route of washing the cells in weak detergent and col-
clecting the insoluble material. The tissue distribution of
expression of individual, or groups of, intermediate filament
proteins is quite specific and predictable, suggesting that these
stable cytoskeletal proteins play an important role in tissue
differentiation. As to exactly what this role is, either in de-
termining cell behavior in the early stages of commitment to
differentiation, or in defining or restricting cell behavior in
the later stages of differentiated tissue function, we can only
extrapolate hypotheses from circumstantial evidence ob-
tained by careful observation.

Of the six types of intermediate filament proteins, the ker-
atins make up the type I and type II classes. They are the
characteristic intermediate filament proteins of epithelia, and
as cellular proteins they exist only as type I--type II pairs,
not singly. Type III filament proteins are expressed predomi-
nantly in mesenchymal cells, and include vimentin (in very
many cell types, including fibroblasts, endothelial, and he-
matopoietic cells), desmin (in muscle cells), glial fibrillary
acidic protein (GFAP) (in astroglia), and peripherin (in parts
of the nervous system). Type IV intermediate filament pro-
teins are the neurofilament proteins NF-L, NF-M, NF-H and
α-internexin, expressed in neurons. The type V proteins are
the nuclear lamins, which are intranuclear. Lamin proteins
are probably the oldest of the intermediate filament classes
in evolutionary terms [7]. They appear to have a broader
tissue expression range than the cytoplasmic intermediate
filaments and are expressed in many different cell types: B-
type lamins (lamins B₁, B₂) are constitutive, but A-type lamins
(lamins A, C) are absent from many embryonic and some
adult tissues. The recently described filament protein nestin
is the only type VI protein, and is the largest intermediate
filament protein known [8].

The central question in intermediate filament biology is,
why need there be such heterogeneity? Are we simply ob-
serving evolutionary redundancy? Cell biologists have been
persistently unable to demonstrate, in a testable experimental
context, a defined role for intermediate filaments in cells and
tissues which could explain the need for such heterogeneity
or multiplicity of closely related proteins. This is probably
because the essential function of intermediate filament pro-
teins lies in the fully differentiated tissue, which we cannot
adequately reproduce in a controlled tissue culture situation.
Nevertheless, the tissue specificity of intermediate filament
expression is a most useful characteristic of differentiated
cells, and the generation of monoclonal antibodies to inter-
mediate filaments has had a big impact on cell typing in the
laboratory and in clinical pathology [9].

KERATIN HETEROGENEITY

Keratins are now classified individually by number, after
Moll et al. [10], identified by their migration in two-
dimensional gel electrophoresis on the basis of charge and
molecular weight. Keratin 1 is the largest and most basic,
and keratin 19 the smallest and one of the more acidic (ker-
atin 20 is a late addition). Prior to this, keratins had been
described by molecular weight only, which led to much con-
fusion since one-dimensional gel electrophoresis alone has
insufficient resolving power to discriminate between all isotypes. The keratins of type I (‘soft’ keratins 9–20, plus five trichocyte keratins found in hair, nail, Hassall’s corpuscles in the thymus, and in filiform papillae on the dorsal surface of the tongue) and type II (‘soft’ keratins 1–8, plus five more trichocyte keratins) appear to be structurally more complex than other intermediate filament classes, in the sense that they will only polymerize into filaments if both type I and type II proteins are present. This obligate heteropolymeric character contrasts with other intermediate filament proteins that all have at least some capacity for homopolymerization. The evolution of heteropolymerization in keratins may be associated with greater physical resilience of the filament network so formed, but insufficient data are available to be sure of this. It has been shown by a number of groups that single vertebrate keratins of either type are unstable, and are rapidly degraded in the absence of a copolymerizing complementary keratin of the other type [11–13], and it is possible that this could be one mechanism used to balance the quantities of two coexpressed keratins in vivo. There is also some evidence for a feedback mechanism which results in changes in the level of mRNA of one keratin in response to changes in expression levels of its copolymerizing partner [14]. Many type I keratins can apparently polymerize with many type II keratins in vitro, but in situ the keratins are predominantly coexpressed in defined pairs, with each pair characteristic of a particular epithelial phenotype. The paradigm of paired keratin expression was first described by Sun and co-workers [15], and this (like Moll et al.’s numerical classification) represented a significant advance in interpretation of the complexity of keratin expression.

The major keratins show a characteristic tissue distribution (Fig. 2), even to the resolution of subpopulations of cells within a tissue showing specific phenotypes. In some way this must reflect physiological differences between these cell populations. Some keratin pairs, which can be described as primary or essential keratins, are constitutively expressed as the baseline keratin phenotype in appropriate tissues. These primary keratins are keratins 5 and 14 and keratins 8 and 18. The secondary keratins are differentiation-specific, and their expression can be modulated by a variety of external factors, such as retinoids and growth factors.

Keratin distribution patterns confirm the existence of two distinct categories of epithelial cells, simple epithelial cells and basal or stratifying epithelial cells [16]. Simple epithelial cells retain contact with the basal lamina and also retain a free apical surface, and simple epithelial cells express type I keratin 18 and type II keratin 8, the primary simple epithelial keratins. Simple epithelial keratins 8 + 18 are the first keratins to be expressed in embryogenesis (Fig. 3), and persist alone in a small number of adult epithelia (e.g., hepatocytes). In contrast to these are basal cells, from complex glandular epithelia to stratifying squamous epithelia, which all have a keratinocytelike phenotype and are characterized by basal lamina contact in the absence of a free apical surface. All of these basal cells synthesize type I keratin 14 and type II keratin 5, the primary keratins of stratifying epithelia. Keratin filaments in this type of cell are shown in tissue culture in Figure 4. Basal cells appear to synthesize these keratins only while they are in contact with a basal lamina; when such cells begin to lose this basal lamina contact, as in the course of epidermal differentiation, a change occurs in keratin synthesis. Production of keratins 5 and 14 slows or ceases [17], and secondary or differentiation-specific keratins are synthesized. The basal keratins, however, are not immediately turned over, and keratins 5 and 14 persist at detectable levels at least as far up as the stratum granulosum [16,18]. Keratins are very long-lived and stable proteins, and even simple epithelial keratins 8 and 18 have been calculated to have a half-life of around 100 h [19].

When a cell begins to leave the basal cell tissue compartment, there are several options for essentially the same cell type, depending on its body location. Thus, keratinocyte cells could express keratins 1 + 10 in cornifying differentiation, as in skin and cornifying regions in the mouth, or keratins 4 + 13 in noncornifying stratification, such as occurs in mucosal epithelia of the eyes, ears, and nose, and in oral, urogenital, and anal epithelia [20]. Corneal epithelial cells express keratins 3 + 12 [21]. Areas of rapidly turning over epithelium express keratins 6 + 16 (psoriatic lesions, cultured keratinocytes, hair follicles) [22], as secondary keratins which can be simultaneously present with another keratin phenotype [23]. The identification of the precise location of these cells within the tissues has come from the use of monoclonal antibodies to individual keratins. A number of minor keratins are also expressed in these and other tissues: Keratin 19 is present in a wide range of simple and stratified epithelia [24], keratins 15 and/or 17 can be found in basal cells [25], keratin 20 is expressed by certain simple epithelia in the gastrointestinal tract [6], keratin 2 occurs in certain cornifying epithelia, keratin 7 occurs in a subset of glandular epithelia [26], and keratin 9 is expressed in regions of pressure or abrasion, such as palm and footsole skin [27]. The total pattern of keratin expression now becomes rather intricate and complex, and it is not unreasonable to suppose that there are probably as many distinct keratin expression phenotypes among epithelial tissues as there are physically distinct epithelia, and that these two parameters may be related.

**KERATIN STRUCTURE**

The α-Helix

Intermediate filaments all share the same general pattern of molecular structure, although their precise amino acid sequences vary considerably and their encoding DNA sequences vary even more. The earliest X-ray crystallography data on intermediate filaments was obtained from wool keratins and suggested that the main components in these structures were α-helical proteins. This was reinforced by secondary structure predictions from the first amino acid sequences that became available for intermediate filaments [28], initially for type III proteins and shortly after for a type I keratin (keratin 14). Although they are predominantly α-helical, their structure is quite different from that known for the collagen family.

Every protein in the intermediate filament family contains a rodlike central α-helical domain, which is the structural subunit of the ultimate polymeric filament (Fig. 5). A single polypeptide chain is unstable and dimerizes rapidly to form a coiled-coil two-chain molecule, a type I–type II heterodimer in the case of keratins [29,30]. This instability and dimerization is due to the presence of a characteristic strip of hydrophobic residues, running down along the α-helical cylinder with a slight twist in its course. Formation of a coiled-coil therefore sequesters the hydrophobic faces of two interacting polypeptides, with their hydrophobic side chains interdigitating like the teeth of a zipper, away from the aqueous environment of the cytoplasm. This tendency to form a coiled-coil [4] is a characteristic that can be simply and confidently predicted from the amino acid sequence by looking for the
Figure 2. Summary of keratin classification, based on Moll et al. [10], giving the relative molecular weights observed on gel electrophoresis for each polypeptide. The major coexpressed keratin pairs are shown, based on Sun et al. [78], and their characteristic tissue distribution is indicated as deduced from many biochemical and immunohistochemical studies. Shaded keratins are the primary (essential) keratins for simple and stratified epithelia.

presence of suitably hydrophobic residues (usually leucine, valine, isoleucine, or methionine, or bulky residues such as phenylalanine or tryptophan) occupying every first and fourth position (the $a$ and $d$ positions) of each seven residues over a stretch of sequence. The $\alpha$-helical conformation is, in fact, the most common secondary structure seen in proteins [31].

What makes the intermediate filament $\alpha$-helical structure unusual is its length: The "rod" domain of the intermediate filament proteins extends for a fairly consistent minimum of 310 amino acids, which are framed by consensus sequences at either end. This domain contains four interruptions of the helix, again conserved in their size and position, three of which consist of stretches with a low probability of $\alpha$-helical conformation, and are referred to as linkers (L1, L12, and L2), dividing the rod domain into helices 1A, 1B, 2A, and 2B (nomenclature of subdomains following Steinert [32]). The fourth interruption is a skip residue or a probable helix phase reversal (the "stutter") in the middle of helix 2B. This consensus structure of intermediate filament polypeptides is illustrated in Figure 5. From a single polypeptide chain, a filament of around 8–10 nm in diameter is built up like rope by repeated coiled-coiling, forming dimers then tetramers, to protofilaments of 2–3 nm and then 4.5 nm, and then to the intermediate filament with some 32 polypeptide chains in four protofilaments within the diameter of one 10 nm filament.

Knowledge of the length of these domains has been used to interpret banding of negative staining seen on paracrystals of type III intermediate filament proteins, to deduce the probable orientation and packing of the $\alpha$-helix in filament assemblage [33]. The data suggest alternate antiparallel strands, with a substantial overlap of the amino-terminal half of the rod and a very short overlap at the carboxy-terminal end of the rod (see Fig. 5), between adjacent strands. Computer predictions suggest that these interactions between the rod domains are primarily due to charged patches on the surface of the helix, but it has not been possible to predict the preferred alignment this way since there are many possibilities. Although paracrystal data must be treated cautiously, since they are basically artifactual, models derived from such analyses do fit very well with other lines of evidence. For example, the known points of highest sequence conservation (and therefore structural importance?) coincide with the overlaps, particularly at the end of helix 2B (the "helix termination peptide"). Binding studies [34] and transfection experiments [35,36] have defined certain regions as essential for filament assembly, which substantially coincide with paracrystal-predicted overlaps [33]. Recently, we have also found that immunogenic "hotspots," which one would expect still to be exposed after polymerization, are indeed located in the predicted nonoverlapping helical span (unpublished observations from this laboratory).

Head and Tail Domains

The function of the nonhelical parts of the filament proteins is not fully understood, but more information is becoming available all the time. Much can be predicted, albeit with caution, about the secondary structure of the terminal domains. Some workers have suggested the presence of "om-
Figure 3. Simple epithelial or embryonic keratins K8 and K18 are the first to be expressed in the embryo. K8 and K18 persist into adult tissues in all simple epithelial cells, and even as the sole keratins in some tissues such as hepatocytes. They are shown here by immuno-fluorescence using an antibody which reacts with keratin 18 (L665), in all cells of the preimplantation mouse blastocyst. (Unpublished material from collaborative experiments with Martin Johnson.)

ega loops" [37] or "glycine loops" [38] centered on the Phe or Tyr residues between Gly-Gly-Gly-Ser repeats in the tail domains of epidermal keratins. The linker regions within the α-helical rod domain are highly conserved in their size and location, and also show partial sequence conservation. Linkers may function as hinge regions to give the filament flexibility, since they are close to predicted boundaries or gaps between the staggered overlaps of the helical domains. A size restriction may be necessary to avoid steric hindrance of packing between the helical domains. The nonhelical domains at the amino- and carboxy-termini, or the head and tail domains respectively, may protrude outside the body of the polymerized filament, since light proteolysis of intact epidermal filaments yields fragments with the amino acid constitution of terminal domains [39]. The abundance of glycines in the head and tail domains of epidermal keratins 1 and 10 also argues for high flexibility in these domains. By immunoelectron microscopy, there is little doubt that at least part of the terminal domains of neurofilaments protrude: Regular extensions at right angles to the filament are observed [40], and terminal-specific antibodies can be seen to bind to such extended protrusions [41].

Although it has been shown for type III intermediate filament proteins that the head domains are essential for filament polymerization [42], it is also to be expected that the head and tail domains contain the key to the tissue specificity of intermediate filament expression. The amino acid sequence of the α-helical rod domain is recognizably conserved between proteins within the same class of intermediate filaments, irrespective of where the protein is expressed, but the head and tail domains vary substantially. This is particularly so within the keratins. Furthermore, when terminal domains are compared between members of a coexpressed type I-type II keratin pair, a degree of relatedness in amino acid composition is seen, although the rod domains of these two keratins of different types will be quite distinct.

One interpretation of all these observations would be that the head and tail domains may have a specific role in interaction with other cytoplasmic components which are important for differentiation. However, recent evidence suggests that tail domains, as well as the head domains, play an important part in correct filament polymerization. Forced expression of keratin pairs from which either head or tail domains had been deleted from one or both of the polymerizing pair of keratins suggested that tail domains were necessary for efficient filament network formation, and that at least one of the two interacting keratins needed both end domains intact in order to produce normal-looking filaments [13]. These experiments suggested the existence of a homotypic end-to-end interaction between keratin head and tail domains. The location of putative binding site(s) in the head/tail domains is not known, but such binding sites (which may be repeated, and may involve the glycine loops [38]: see Fig. 5) could be the key to the tissue-specific types of sequence observed here. Sequence variation could result in different qualities of the final filament network, by altering for example the rigidity or flexibility of the filaments, or the ease with which the filaments could be dissociated or remodeled. For example, the large number of such repeat loops, as potential head-to-tail binding sites, on keratins 1 and 10 may result in a very strong bonding within the filaments, as opposed to the fewer repeats in simple epithelial keratins which are also easier to dissociate in vitro. The unusual presence of cysteine residues (rare in "soft" keratins) in the head and tail domains of trichocyte keratins [43,44], on the other hand, opens the possibility for disulphide cross-linking to make a more rigid network in hair and nail cells.

Much of the information regarding structure (above) and most of the information about tissue distribution of keratins (below) and other intermediate filaments has depended on the remarkable insolubility of these proteins. Washing cells in culture with soapy water (more accurately, 1% nonionic detergent) removes most things apart from the major components of the cytoskeleton; an additional high-salt wash (1.5 M potassium chloride) will remove residues of the actin and tubulin systems, leaving almost exclusively intermediate filaments as the insoluble residue. Complete dissociation of the type I-type II keratin complexes in the detergent-insoluble residue requires up to 9 M urea for epidermal keratins 5 and 14, plus reducing agents for the trichocyte keratins. Intermediate filament proteins appear to be remarkably inert, and are not significantly affected by the standard drug disruption techniques used on actin and tubulin cytoplasmic fibrous systems. Microinjection of certain keratin-specific antibodies is probably still the cleanest way to get rid of keratin filaments in an epithelial cell [45–48], and it is striking that tissue-cultured cells appear to be undisturbed by this loss. Presumably the insolubility and molecular stability is relevant to intermediate filament function in vivo, and all these observations support the probability that the importance of these proteins lies in their role in a fully differentiated threedimensional tissue, a role which is not relevant in a monolayer of cultured cells.
Figure 4. Keratinocytes express the greatest range of keratins at their highest organizational complexity. Keratin filaments are visualized here in tissue culture, by immunofluorescence of a squamous cell carcinoma-derived line, TR146 cells [79]. The position of the desmosomes is indicated, although they are unstained, by the symmetrical focusing of keratin filament bundles in neighboring cells onto points around the cell periphery.

![Diagram of keratin structure]

Figure 5. Diagram to show the structure of the keratin heterodimer (top), composed of one type I (cross-hatched boxes) and one type II (open boxes) polypeptide in a parallel and in-register alignment along the α-helical rod domain. Below is a representation of the possible packing of four such dimers (two tetramers) in the beginning of polymerization to a 10 nm filament, with antiparallel strands close to a half-stagger alignment. The head and tail domains may interact to reinforce this polymerisation. The nonhelical terminal domains are drawn to indicate the presence of glycine loop structures as proposed by Steinert et al. [38], although the actual conformation of these, and of the “linker” nonhelical regions, is still a matter for speculation.
EVOLUTION OF INTERMEDIATE FILAMENTS

If the sequence of evolution of the members of the intermediate filament gene family were to be known, then it should be possible to predict the relative functional importance of different classes of intermediate filaments, and thus to deduce the sort of pressures that may have driven this evolution, and thence what the primary functions of these proteins are. Evidence has been obtained for the existence of intermediate filament-like proteins in organisms as far apart as insects and annelids, and now even within the plant kingdom [49,50], although no sequence information is yet available for plant intermediate filament proteins.

Attempts to sequence such proteins from distant animal phyla have, however, yielded success as far away as the Molusca and Annelida. Sequencing of cytoplasmic intermediate filaments from invertebrates [7] revealed a striking resemblance of these proteins to the type V nuclear lamin proteins of vertebrates, in the conservation of certain key features considered to be unique and characteristic of lamins among vertebrate intermediate filament proteins. These include the presence of an additional run of amino acids within helix 1 which is missing from all vertebrate cytoplasmic intermediate filaments sequenced, as well as certain features in the tail domain in the region associated with targeting lamins to the nucleus. Sequencing of an intermediate filament gene from Helix aspersa [51] and comparison with a Xenopus laevis frog lamin gene [52] revealed that this similarity extended to the conservation of intron positions, making it unlikely that the similarity of the protein sequence simply represented convergent evolution. It therefore seems likely that intermediate filament genes evolved as laminlike sequences, and/or that the conservation of lamin sequences is particularly critical, suggesting that intranuclear lamin intermediate filament proteins play an essential role which evolved very early in the evolution of multicellular organisms. This role could be associated with the maintenance of intact nuclear structure between cell divisions. Compartmentalization of DNA replication away from the cytoplasm, so that its timing can be regulated, is probably a prerequisite for the development of cellular specialization and differentiation at the first stages of evolution of a multicellular organism; the lamins may be reinforcing this compartment structure.

There is very little evidence regarding the sequence of evolution of the multiple classes of intermediate filament genes [54]. The high degree of conservation between mammalian species makes it clear that even the complex keratin groups evolved before mammalian speciation. In all vertebrate species where intermediate filaments have been looked at at the level of primary amino acid sequences, individual specific proteins are clearly recognizable, and primary simple epithelial keratins and types III, IV, and V proteins can be identified. Examination of the gene structure has revealed further information: A gross discrepancy in the position of introns between the neurofilament genes and those of other intermediate filaments suggests that neurofilament genes have been evolving separately from other intermediate filament genes for a significant time in evolutionary terms [53].

Another avenue of investigation which might be useful is the determination of the relative chromosomal location of genes for intermediate filaments, at least in the case of those genes still evolving in vertebrates. Lamin gene locations are not known, and genes for type III and IV proteins are scattered and yield no information. Keratin genes, however, are significantly clustered [55–59], suggesting gene duplication events may have opened the way for diversification of keratins. In man, there are clusters of type I epidermal keratins on the long arm (at 17q12-21) and the short arm (17p11-12) of chromosome 17, and at least one cluster of type II keratins which is on chromosome 12, at 12q11-13. Thus, although type I–type II pairs of keratins appear to be specifically coexpressed and coexpressed, the genes for the two paired proteins are distant in their location. Any genomic regulation must thus operate via a trans-acting intermediate, or alternatively the pairs of coexpressed gene products may be coregulated posttranscriptionally or posttranslationally. Similar clusters of keratin genes occur on the homologous chromosomes of mouse [60], where it is interesting to see that they are known to lie close to major clusters of homeobox genes. It is not known whether or not both of the type I keratin clusters are functional, but some nonfunctional pseudogenes for epidermal keratins have been identified, as well as a very large number of processed pseudogenes for the simple epithelial keratins 8 and 18 which are expressed in early embryos. Interestingly, these primary simple epithelial keratins, keratins 8 and 18, do not follow the same chromosomal distribution pattern. These two genes were only localized recently, because the processed (intronless) pseudogenes masked the location of the true genes in conventional DNA hybridization searches. This problem was overcome by using the polymerase chain reaction technique to amplify sequences of the two genes which specifically contained introns [61,62]. Running such reactions with human-specific primers on mouse-human somatic cell hybrids revealed that both keratin 8 and keratin 18 genes were located on chromosome 12, the chromosome known to carry the type II gene cluster; the type I keratin 18 was thus separated from the clusters of epithelial type I keratins. We do not know the relative proximity of the genes for keratins 8 and 18 on chromosome 12; we have not been able to rule out the possibility that they could even be coregulated in some way. This distribution of the keratin genes would be consistent with the evolution of both type I and type II genes from a keratin 8-like precursor. The chromosomal localization data can be taken as circumstantial evidence that the primary embryonic or simple epithelial keratins 8 and 18 evolved before those of stratified epithelia like epidermis.

The identity of epidermal keratins is sometimes difficult to pin down between frog and man, and even between mouse and man there are variations in the secondary keratins of stratified epithelia [63]. This suggests that these proteins involved in the quality of the external surface have been evolving later—a readily acceptable concept, since changing habitat provides scope for speciation, and changing appearance can have rapid effects on mate selection, which is an early point of action of natural selection. One would therefore expect the secondary, differentiation-specific keratins to be the latest evolving intermediate filaments, right at the front line of evolutionary change. More data on the sequences of hair and feather keratin genes may show if this is the case.

THE QUESTION OF INTERMEDIATE FILAMENT FUNCTION

Intermediate filaments are considered as one of the three filamentous components of the cytoskeleton. The dynamic...
functions of the other two systems, the actin- and tubulin-based systems, contrast sharply with the polymer stability and lack of cytoplasmic monomer pool observed for the intermediate filaments. All the observations of expression pattern in vivo indicate a structural role for these proteins, and one that is essential in tissue differentiation. From the structural integration of muscle cells to the maintenance of the attenuated axonal cytoplasm in a functional state in the neuron, a case for a supporting cytoskeleton can be seen, with distinctly different requirements in each tissue. The case is probably strongest of all for the keratins, since the structural role of epithelia as boundary tissues is unequivocal, and the association of keratin filaments with desmosomes does give a three-dimensional integrity across the whole tissue which could help it to resist deformation. This three-dimensional network also has the built-in potential to transduce mechanical signals over long distances. Across the wide variety of different situations in which an epithelial sheet has to perform a structural supporting role while contributing to optimal tissue shape compatible with function, the physical requirements of the component cells will vary substantially. Physical requirements such as plasticity or rigidity may or may not be accompanied by other functional demands, such as production of specific secretory products, as in glands. Biochemical properties of intermediate filament protein complexes observed in vitro are likely to reflect biological differences in situ; a more easily disruptable cytoskeleton (keratins 8 and 18) might be more appropriate in secretory cells, where there is a lot of cytoplasmic vesicle traffic, and a less easily disruptable cytoskeleton such as keratins 5 and 14 might be better fitted to the physically stressed epidermis [16]. As discussed above, there is probably sufficient variability in the amino acid sequence of the head and tail domains to give the degree of variation in filament properties which is observed from one isotype to another. Our rudimentary but expanding understanding of the function of intermediate filament subdomains now suggests that the fine-tuning of the quality of the filaments formed can be mediated by the nature of the head and tail domains, which may contain the clue to tissue-specific expression. It is probable, in other words, that the intermediate filaments come closer to what one would expect of a cytoskeleton than any other "cytoskeletal" system.

INHERITED DISORDERS OF KERATINS

The identification of a heritable disorder is usually very informative in understanding the true function of a protein. The fact that there have been, until now, no known heritable disorders of keratins can be interpreted in two ways. It could indicate that keratin heterogeneity merely reflects redundancy if defects are passing undetected, or conversely, it could indicate their supreme importance if defects are proving lethal early in development, and not becoming apparent for that reason. A third, and increasingly plausible, reason could be that the system has evolved a buffering capacity, in the form of the heteropolymerization requirement, which can tolerate minor protein defects and still function: Our experiments have shown that as long as one of the two interacting keratins is intact, long filaments can form in tissue culture cells even if the second keratin carries substantial deletions [13]. Evidence for the existence of heritable keratin defects is, however, now emerging, which may at last demonstrate an essential role of keratins in epidermis.

The inherited skin disorders in the epidermolysis bullosa group, in which skin blistering occurs upon minimal trauma, are divided clinically into three types: functional and dystrophic, both of which arise from connective tissue abnormalities which result in subepidermal blistering, and the simplex forms in which there is intraepidermal cell disruption (see Chapter 17 for more details). Epidermolysis bullosa simplex (EBS) of the Dowling-Meara form (also known as EBS-heteropiferms) is characterized at the electron microscopic level by the occurrence of cytoplasmic, nonmembrane-bound electron-dense bodies [64], which have recently demonstrated univocally to be composed of basal cell keratins (keratins 5 and 14) [65]. These bodies, or tonofilament clumps, are found in blistered skin [64,66], but have now also been demonstrated in nonblistered skin of affected individuals [65], so they are not caused by the blistering but may in fact predispose toward it. Immunohistochemical analysis of keratin expression in this group of diseases has given varying results depending on the antibody used, some showing no changes [67,68] and some showing altered antibody reactions [69]. Nevertheless, other sources of evidence support the hypothesis from the electron microscopy results that epidermolysis bullosa simplex may arise from mutations in the genes for epidermal keratins 14 or 5. Tonofilament clumping has been seen following introduction of a truncated keratin 14 gene into transgenic mice, where blisters reminiscent of human Dowling-Meara EBS were also produced [36], similar to clumping effects seen in tissue culture cells transfected with the truncated keratin 14 gene [35]. The keratin clumps are reminiscent of keratin filament disruption induced by microinjected antibodies [46] and cold shock [70,71], and those produced spontaneously during mitosis [72–74]. Direct attempts to identify the putative mutation in one or other of the basal cell keratins are now underway in at least three laboratories, and the first data are now becoming available. It appears that there are different keratin mutations for different clinical types of EBS, or possibly even within the same clinical subgroup. Bonifas et al. have identified a point mutation (leucine584 to proline) in helix 2 of keratin 14 in a family with Weber-Cockayne EBS (EBS-localisata) [75], while Coulombe et al. [76] have identified two isolated possible Dowling-Meara EBS patients with two different point mutations in helix 1A of the same keratin 14 (arginine125 to cysteine in one, and arginine125 to histidine in the second). Our laboratory, on the other hand, has analyzed a Dowling-Meara EBS family with a mutation in the helix termination peptide (helix 2B) of keratin 5 glutamic acid 475 to glycine [77]. Most mutations occur in conserved regions of the α-helical domains which would be predicted to be critical for filament structure. As the first example of pathological mutations in human keratin genes, the epidermolysis bullosa simplex diseases, with the lack of an effective keratin cytoskeleton rendering the basal cells susceptible to compression lysis, dramatically demonstrate the nature of the critical structural role for keratins in reinforcing the epidermis and adapting this epithelium to withstand one of the harshest of environments to which cells are ever exposed.

REFERENCES


