Ann. Rev. Biochem. 1988. 57:593–625

MOLECULAR AND CELLULAR BIOLOGY OF INTERMEDIATE FILAMENTS[1,2]

Peter M. Steinert[3] *and Dennis R. Roop*[4]

National Cancer Institute, National Institutes of Health, Bethesda, Maryland 20892

CONTENTS

[1]The US Government has the right to retain a nonexclusive royalty-free license in and to any copyright covering this paper.

[2]Abbreviations used: IF, Intermediate-sized filament(s); IFAP, Intermediate filament–associated protein; GFAP, Glial fibrillary acidic protein; kbp, kilo base pairs; kd, kilodaltons; STEM, scanning transmission electron microscopy; TPA, 12-*O*-tetradecanoyl-phorbol 13-acetate.

[3]Dermatology Branch

[4]Laboratory of Cellular Carcinogenesis and Tumor Promotion

INTRODUCTION: SCOPE OF THIS REVIEW

Within the last 10 years, intermediate filaments (IF) has become a topic of considerable interest to cell and molecular biologists. It now appears that IF and IF-like proteins are prominent components of the cytoskeleton and nuclear envelope (karyoskeleton) of most eukaryotic cell types (1–9). A large body of data indicates that IF proteins constitute an extremely heterogeneous multigene family, of for example, 30 or more protein chains per mammalian species. In this regard, they are far more complex than other principal cytoskeletal structural proteins, which form microtubules (tubulin) or microfilaments (actin). While we do not yet fully understand the role of IF in organisms, this complexity means that IF are a functionally diverse component of cells. Yet IF proteins are all built according to a common tripartite plan. They possess a central α-helical 'rod' domain of conserved secondary structure that forms the basis of their related structures and morphologies. They possess amino-terminal and carboxyl-terminal 'end' domains of widely varying size and chemical character. It is now widely suspected that the properties of these end domains in large part define the functional diversity of the IF family. Biochemists and biophysicists are interested in solving the structure of this complex and unique class of fibrous proteins. Molecular biologists are interested in how the various genes of IF proteins are regulated during development and differentiation. Cell biologists are interested in the relationship between differential IF gene expression in specialized cells and the dynamic organization and functions of IF in cells.

Our current understanding of IF follows an hierarchical progression of complexity beginning with the structure of the IF protein chain; the formation of α-helical coiled-coils; the assembly of these into IF; the regulation of their gene expression; the distribution of IF in cells; and their interactions with other IF-associated proteins (IFAP) as well as the cell periphery and the nucleus. By and large, our understanding of IF declines progressively with these increases in organizational complexity. In this review, we summarize key data in the order of this progression, by presenting known facts and emphasizing areas of controversy and uncertainty, with the hope that this will stimulate a wider research effort into the understanding of the function of IF.

DISTRIBUTION OF IF AND IF-LIKE PROTEINS

Early work defined five distinct subclasses of IF in various differentiated vertebrate cells (1, 2) (Table 1), but recent studies based on analyses of sequence data indicate that these subclasses may more properly be classified into the distinct sequence types I–IV (Refs. 4, 5, 8–11) (Table 1): some 15

different acidic keratins of epithelia of 44–60 kd are type I; about 15 different neutral-basic keratins of epithelia (50–70 kd) are type II; the single proteins vimentin (54 kd) of "mesenchymal" cells and many cultured cells, desmin (53 kd) of myogenic cells, and the glial fibrillary acidic protein of glial cells and astrocytes (51 kd) are type III; the three neurofilament proteins of the nervous system NF-L (60–70 kd), NF-M (105–110 kd), NF-H (135–150 kd), and also possibly a protein of 57 kd located in peripheral neurons (12) are type IV. Each of these four sequence types forms cytoplasmic filamentous structures 10–15 nm wide classically known as intermediate-sized filaments (IF). More recently, the IF family has grown to include the lamins, which form the nuclear lamina complex or karyoskeleton on the inner surface of the nuclear membrane of eukaryotic cells (13–15). The lamins do not form 10–15-nm wide cytoplasmic structures, but rather form a quasi-tetragonal meshlike lattice composed of fibers about 10 nm wide in interphase cells in vivo (14) and paracrystalline arrays in vitro (14, 16–18) that are built slightly differently from cytoplasmic IF. The lamins consist of at least four proteins of 60–75 kd

Table 1 Distribution of different IF and IF-like proteins[a]

Origin	Common name	Sequence type	Estimated number of chains	Size (kd)	Ref.
All epithelia	acidic keratins[b]	I	~15	40–60	
All epithelia	neutral-basic[b] keratins	II	~15	50–70	
Various "mesenchymal" cells	vimentin	III	1	53	1–11, 25
Myogenic cells	desmin	III	1	52	
Glial cells and astrocytes	glial fibrillary acidic protein	III	1	51	
Most neurones[c]	neurofilaments	IV	≥4 vertebrate	57–150	
			≥2 invertebrate	60–200	
Nuclear lamina of all eukaryotes	lamins	V	≥4 vertebrate	60–70	14–24
			≥1 invertebrate	60–70	
Plant cells	[d]		—		26
Yeast and other simple eukaryotes	[d]		—		27, 28
Flagellae	tektins	—	3	47–55	29–32

[a] The IF contents of only a few major cell types are listed. Many cells, especially tumor cells, are known to coexpress vimentin as well as their more specific IF type (6,7). Further, nucleated cells also express type V lamins. Intermediate filament–associated proteins (IFAP) will be discussed below.

[b] The term "cytokeratin" has been widely used to describe the IF of epithelia. We believe this term is a misnomer since the IF of epithelia are not known to occur outside the cytoplasmic cytoskeleton.

[c] Axons of many *Arthropoda* do not contain neurofilaments (33).

[d] Protein detected by use of the "universal" IF antibody (34), the epitope of which is presumably a canonical feature of the rod domain.

in mammalian and lower vertebrate cells (13, 17, 19) or fewer in invertebrates (20, 21). Analyses of available sequences establish that the lamins form a distinct type V class of IF-like proteins (22–24).

Less conclusive data suggest that IF or IF-like proteins are even more widespread. All classes and sequence types are present in primitive vertebrates. At least some are present in many differentiated cell types of invertebrates and simple metazoa (6, 7, 25). Further, IF may be present in cultured plant cells (26), various single-cell eukaryotes including yeast (27) and slime molds (28), and as tektins (47–55 kd), a component of the axonemes of eukaryotic flagellae (29–32). Further detailed studies including isolation and sequencing of such proteins will be required to confirm these possibilities. One report (36) that clathrin light chains are IF-like cannot be sustained on the basis of limited sequence homology to only a short tract of sequence common to known IF proteins (the L12 linker region; see below). Nevertheless, we fully anticipate that the IF family will grow considerably in the future.

STRUCTURE OF IF

The IF Protein Chain

Partial or complete amino acid sequence data are available for a large number of IF protein chains of many vertebrate species (see Refs. 10, 11 for recent lists). The single feature that distinguishes an IF protein chain is its central α-helical rod domain of 310–315 (types I–IV) or 356 (type V) residues that has been remarkably conserved in its size, secondary structure, and to a greater or lesser extent, its sequence. IF protein chains differ primarily in the sizes and sequences of their amino- and carboxyl-terminal end domains.

Characteristic features of the rod domain are illustrated in Figure 1. It consists of four α-helical tracts that possess a quasi-heptad sequence of the form $(a\text{-}b\text{-}c\text{-}d\text{-}e\text{-}f\text{-}g)_n$ where more than 75% of the *a* and *d* positions are occupied by apolar residues, whereas positions *b, c, e, f,* and *g* are often polar or charged residues. This heptad sequence favors the formation of a coiled-coil. The four α-helical tracts are termed 1A, 1B, 2A, and 2B and are interrupted by three 'linkers': L1 joins 1A and 1B to form segment 1; L2 joins 2A and 2B to form segments 2, and segments 1 and 2 are joined by linker L12. The exact sizes and sequences of these α-helical tracts and linkers define the types I–V mentioned above. Protein chains of the same sequence type typically show 70–95% homology, but chains of different types show 30% or less homology (10, 11). The α-helical tracts in types I–IV chains are of invariant size; L1 is of variable length and sequence and is non-α-helical; L12 is likely to form a short β structure; L2 is α-helical but does not conform to a heptad sequence. Tract 1B and segment 2 possess highly significant

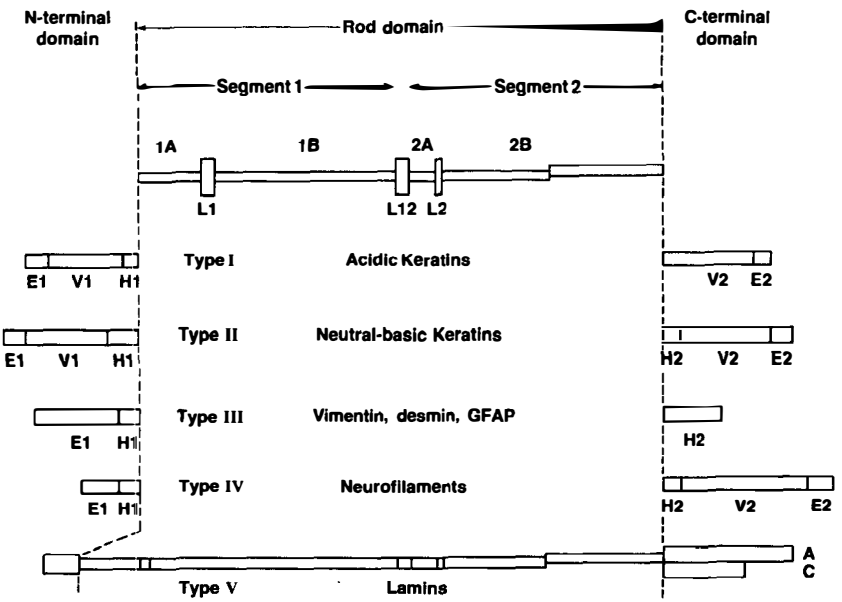


Figure 1 IF chain subdomainal organization. All chains contain a central rod domain flanked by end domains. In types I–IV IF chains, the α-helical tracts are of invariant size: 1A is 35 residues long, 1B is 101, 2A is 19, and 2B is 121. In type V lamins, 1B is six heptads longer due to an insertion of six heptads. This corresponds precisely to the intron position at position 42 conserved in types I–III IF genes (24, 36; see Figure 3). In types I–IV IF linker L1 is 8–14 residues long, L12 is 16 or 17, and L2 is always 8. In type V lamins, the linkers are α-helical (see text). The end domains may also be subdivided into subdomains based on homologous (H), variable (V), or end (E) sequences. See Refs. 10, 11 for more details. Redrawn from Ref. 4.

periodicities of 9.55 and 9.85 residues, respectively, in the distributions of charged residues, which are thought to be critical features in determining higher orders of IF structure (see below). The segments 1 and 2 are each about 22 nm long and when joined by L12 give an entire rod domain length of about 47 nm. The rod domains of type V lamins, however, are built slightly differently (11, 24). Segment 1 is exactly 42 amino acids longer due to the insertion of six heptads in the 1B tract (24, 35), so that this segment is 28 nm long, giving a total rod domain length of about 53 nm. Sequences comparable to the linkers are present, but L1 complies with a heptad sequence and L12 and L2 are α-helical. Thus in lamins, the rod domain is α-helical throughout its entire length, and in this regard is more similar to tropomyosin (24). Also, segments 1 and 2 are more highly charged than they are in type I–IV chains and contain a significantly different periodicity in the distributions of charged residues. Nevertheless, all IF chain types, including the lamins, contain certain other common 'canonical' features: they all contain highly conserved

sequences of 2–3 heptads near the beginning of tract 1A, and of 4–5 heptads at the end of segment 2; and they all possess a precisely conserved break in the regularity of the heptad substructure near the middle of tract 2B (Figure 1). We propose that in order for a protein chain to be designated as a member of the IF family in the future, it should possess a rod domain of secondary structure and properties closely similar to those summarized here.

Analyses (4, 5, 11) of end-domain sequences have revealed patterns that permit classification into subdomains, the organization of which appear to be highly characteristic of each rod domain sequence type. This classification scheme is based on sequences of (*a*) high homology (H subdomains), (*b*) special variability or content of exact or inexact peptide repeats (V subdomains), and (*c*) highly charged termini (E1 and E2 subdomains). These organization schemes are summarized schematically in Figure 1. Type II keratins possess H1 and H2 subdomains on either side of the rod domains of 36 and 20 amino acids, respectively, that have been conserved. Type I keratins, however, possess a short and variable H1 subdomain, but do not possess a defined H2 subdomain (11). Distal to these subdomains, both keratin types expressed in stratified squamous epithelia and terminally differentiated epidermal tissues possess V1 and V2 subdomains that (*a*) are enriched in glycines and/or serines of the sequence form aliphatic-(glycine/serine)$_n$ and (*b*) that are frequently configured as exact or inexact tandem or displaced peptide repeats that vary in size from zero residues in the V2 subdomains of the smallest known keratin (37) to more than 100 residues in the largest known keratins (11, 38). The keratins of simple epithelia and lower vertebrates possess differing repeats rich in serines and threonines (39). The E1 and E2 subdomains of keratin chains are all basic and have been conserved in the same keratin of different species but are highly variable between different keratins such that they provide unique epitopes that have proven useful for the production of chain-specific antibodies (40). Interestingly, the E1 and V1 subdomains are usually larger, more basic, and more glycine/serine-rich than the V2 and E2 subdomains. The keratins of epidermal derivatives (hair, nail, etc) do not possess well-defined V subdomains, but rather have extended cysteine-rich E subdomains instead (41, 42). The amino-terminal domains of type III IF chains are highly basic and hypervariable in sequence, thus forming a well-defined E1 subdomain, except for a short homologous H1 subdomain of 20 amino acids adjacent to the beginning of their rod domains (11). The carboxyl-terminal domains of all type III chains are related and thus form an H2 subdomain of about 55 residues. The more limited sequence data for the type IV neurofilament chains do not yet define a precise subdomain organization, but they do reveal E1 and H1 subdomains at the amino-termini of these chains. Their carboxyl termini possess a short H1 adjacent to the end of the rod domain, and a long V2 subdomain, which in

NF-L and NF-M is enriched in lysine and glutamate residues (11, 43–49). In addition, the human NF-M chain possesses six tandem copies of a 13-residue repeat in the V2 subdomain, and also contains a region of heptad sequences capable of forming a coiled-coil on its E2 end (48). The two human lamin A and C chains for which sequence information is available so far (22, 23), presumably derived by differential splicing from a single gene since they have an identical 30-amino-acid amino-terminal end, share the first 179 residues of their carboxyl-terminal end, while the larger A chain contains an additional 98 residues. These end domains also possess aliphatic-(glycine/serine)$_n$ repeats similar to those of the keratins (23), which may explain why certain anti-keratin antibodies cross-react with lamins and vice versa (16, 17). More lamin chain sequences, including that of the lamin B chain, will be necessary to explore their reported heterogeneity as defined by peptide mapping and amino acid composition (16, 17, 19).

Despite these recognizable divisions into subdomains, the secondary and higher-order structures of these end domains remain poorly understood. In the case of epidermal keratin IF, solid state NMR experiments have shown that the glycine-rich end-domain sequences have little or no structural order (50), perhaps because they may form Omega-loop-like folds (51). We anticipate that clarification of the functions of IF will require detailed knowledge of the structure of the end domains of their constituent chains.

Oligomers of IF Chains

The first hierarchical step in IF assembly involves the formation of a two-chain coiled-coil molecule (Figure 2*b*). This is stabilized largely by hydrophobic interactions of the apolar residues of the *a* and *d* positions of the heptad repeat, which are thus buried along the axis of the coiled-coil. Several lines of data indicate that this molecule forms by aligning two compatible chains in parallel and in exact axial register. Firstly, this arrangement is the only way to maximize coiled-coil formation between the precisely conserved α-helical tracts of the two chains (52–54). Alignment of the two chains is probably also governed by ionic interactions of residues in the other heptad positions. Analyses of the distributions of ionic charges along the chains reveals that the maximum number of favorable interactions (that is, salt bridges) form only when the two chains are aligned in parallel and within ± 7 residues (52–55). Chemical amino acid sequencing of two-chain α-helix-enriched species obtained by limited proteolysis of keratin IF reveal that the two chains must be in close axial register (56, 57). Type III (58, 59), certain type IV (60, 61), and type V (19, 62) chains possess one or more cysteine residues in identical positions along their chains that can be oxidatively cross-linked to disulfide bonds in vitro, which could occur only if the two

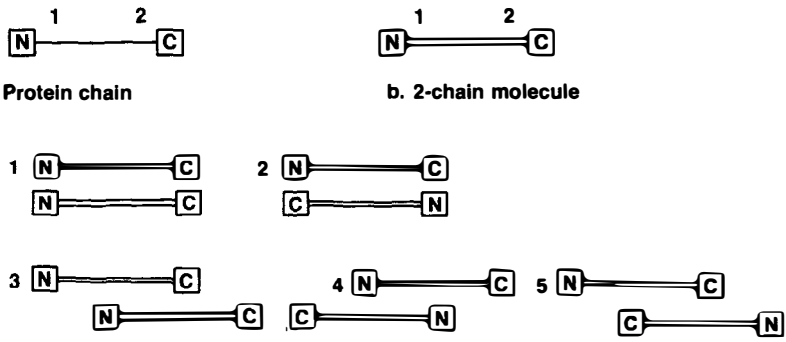


Figure 2 Oligomers of IF chains. (*a*) Universal IF protein chain. (*b*) Two-chain coiled-coil molecule formed by two compatible in-register parallel chains. (*c*) Proposed models for the four-chain complex formed by in-register (1, 2) or half-staggered (3–5), parallel (1, 3), or antiparallel (2–5) molecules. (*d*) Possible eight-chain complex formed by two antiparallel four-chain complexes. Note that this structure can be disrupted in two ways to generate the two currently favored models (2, 4) for the four-chain complex. (*e*) Partial drawing of a brick-wall-like generic lattice for an entire IF containing 32 chains/47 nm. In this case, the end domains are not included because of uncertainties of polarity. Nevertheless, both parallel and antiparallel alignments would form a similar generic lattice. The axial bandings at 22 and 47 nm observed by electron microscopy and X-ray diffraction can be easily generated as shown. It is also possible the lattice is tilted slightly to generate left- or right-handed pitch for the major helix of the IF (91). These models assume precise registration of the entire rod-domain segments. However, ionic interaction studies (52–54, 100) cast doubt on the involvement of segment 1A in stabilizing adjacent molecules, but it is also possible the canonical sequences of segment 1A interact with the canonical sequences at the end of segment 2B (25, 91).

chains are within one residue of axial registration. Finally, elegant electron microscope data (14) of the lamin molecule reveal an elongated rod 52 nm long with twin knobs on one end, corresponding to the large carboxyl-terminal end domains, which could only form if the two chains were parallel and in approximate register.

Keratin IF are obligate heteroplymers, requiring at least one type I and one type II chain for IF assembly in vivo and in vitro (63–65). It has been suggested, although not yet rigorously proven, that this requirement is imposed at the level of the two-chain coiled-coil molecule (56, 57); that is, the molecule is a heterodimer. However, since stable type I and type II homodimers can form in vitro and thus presumably in vivo as well (65–67), these issues of the chain composition of the molecule and the level at which the obligatory type I-type II composition is imposed remain unresolved. In vitro reassembly experiments have demonstrated that virtually any type I keratin can copolymerize with any other type II chain (63–65). In contrast, type III chains form homopolymer IF in vivo and in vitro, but they may also form facultative heteropolymers with other type III chains (58, 68–70). Of the mammalian type IV chains, NF-L, but not NF-M or NF-H, form homopolymer IF in vitro, but the latter chains copolymerize with NF-L to form heteropolymer IF (71–74). Similarly, the smaller invertebrate neurofilament chains form homopolymer IF and copolymerize with the larger chains (75, 76). To date, no type V lamin chain has been shown to self-assemble, but mixtures of purified A and C chains (14) or all of the chains (16–18), form ordered paracrystalline arrays in vitro.

However, it is not yet known why the various chain types prefer to form homopolymer or heteropolymer IF. Since each rod domain sequence type has a characteristic set of end-terminal subdomains, it seems possible that certain of the end domain sequences may specify assembly preferences. Interestingly, the smallest known IF chain does not possess a carboxyl-terminal end at all (except for a short extension of its rod domain) (Ref. 37), implying that such putative assembly regulatory sequences lie mostly in the amino-terminus. Along this same theme, IF particles can be assembled in vitro from desmin chains lacking carboxyl-terminal but not amino-terminal sequences (77). Likewise, the integrity of the amino-terminal end of vimentin is required for in vitro assembly (78).

The next hierarchical step in IF assembly involves the association of a pair of coiled-coil molecules to form a four-chain complex. This concept was first introduced for wool keratin IF based on chemical cross-linking of α-helical-enriched peptides derived on proteolytic digestions of solubilized keratin proteins (56, 59, 79–81), and has now been shown to apply to other keratin (57, 82), vimentin (83), and desmin (84) IF. Further, in solutions of low ionic

strength, vimentin from native (85) and reassembled (83) IF forms a tetramer as do keratins in solutions of 4-M urea or 2-M guanidine-hydrochloride (67). These data firmly establish that earlier concepts of a three-chain building block for IF are incorrect, presumably due to errors in molecular weight calculations (86–89).

Because the four-chain complex seems to be the smallest stable polymer that can exist in solution, solving its structure will provide important clues on the arrangement of molecules within the IF. Calculations of ionic interaction scores between neighboring rod domains reveal that the largest number of favorable interactions occur between 1B and 2 segments when the two molecules are aligned in register or half-staggered and either parallel or antiparallel (Figure 2*c*), although higher scores were obtained in the antiparallel models (10, 11, 52–54). Several studies have attempted to determine the most likely structure(s) for the four-chain complex. By use of electron microscopic techniques, IF preparations containing four-chain complexes, as determined by physicochemical data, appeared to be about 48–50 nm long (81, 83, 91) (compatible with models 1 or 2 of Figure 2*c*) or about 70 nm (92) long (models 3–5). An antibody elicited against segment 2 of the desmin rod formed 50-nm long dumbbell-shaped structures (91), denoting antibody binding to both ends, which is compatible only with model 2. However, a significant number of particles had bound antiboby to only one end (91), which is compatible with models 1 and 2. Proteolytic digestion experiments on wool and epidermal keratin IF have recovered four-chain particles containing sequences from the 1B tract of the rod domain only (56, 57, 80), which is consistent with models 1 or 4. Physicochemical (14) and ionic interaction scoring (18, 24) data on type V lamins indicate that these assemble into 52-nm long four-chain particles in vitro and presumably in vivo as well (models 1 or 2). Under higher salt conditions, 70-nm long vimentin particles could be visualized by electron microscopy (83), but these were interpreted to represent the next hierarchical step, the formation of an eight-chain polymer from a half-staggered association of two antiparallel 50-nm long four-chain complexes (Figure 2*d*). Thus current data, while still inconclusive, seems to favor the antiparallel arrangements of models 2 and 4 of Figure 2*c*. The coexistence of both of these can be envisaged simply at the eight-chain level as in Figure 2*d* by supposing that different experimental conditions favor the disruption of the same structure along alternate boundaries (90). Nevertheless, further biochemical and biophysical experiments are needed. This issue is important because if IF are built from equal numbers of oppositely directed molecules, they will be nonpolar structures and are thereby fundamentally different from other polar cytoskeletal structural proteins such as microfilaments and microtubules (4, 8, 90, 91). On the other hand, if IF contain

parallel molecules, they will have a definite polarity, which may have considerable functional significance.

Possible Models for the Structure of Types I–IV IF: Intimations of Uniformity

The precise ways by which the four-chain (or eight-chain) complexes assemble into higher-order structures remain to be settled. A large body of evidence indicates, however, that types I–IV IF are built according to a common plan, or limited variety of plans, based on the highly conserved secondary structures of the rod domains and predicted secondary structures of two- and four-chain molecules generated as described above. Several lines of data, including scanning transmission electron microscopy (STEM), X-ray diffraction, ionic interaction scoring, and electron microscopy, have recently converged on rather similar models.

Digital analyses of STEM images of unstained IF have provided important quantitative information on the numbers of molecules in types I–III IF. It was shown that these IF are polymorphic, consisting of at least three different mass forms (93–98). A major form in native vimentin IF and all reassembled IF so far studied consists of 32 protein chains, that is, eight four-chain complexes per 47 nm, as well as minor forms containing 5–6 and 11–12 four-chain complexes per 47 nm. The STEM data established that within the same polymorph, IF contain the same numbers of protein chains per unit length, which argues strongly that the structures of types I–III IF are very similar. Moreover, these data provide quantitative constraints on possible models of IF structure.

Analyses of the meridional reflections generated by X-ray diffraction of the well-ordered keratin IF of the porcupine quill have shown that these IF contain an axial periodicity of 47 nm, with seven or more probably eight quasi-equivalent units on a basic helix of 22 nm (53, 54, 99, 100). Interestingly, these periods correspond precisely to the lengths of the entire rod domain and segments 1 and 2, respectively, determined from analysis of sequence information. Surface lattice models have been constructed from these X-ray diffraction data in conjunction with ionic interaction scoring of neighboring molecules. The model that affords the best compliance with both sets of data places a four-chain complex on each quasi-equivalent lattice point so that each is antiparallel with respect to its neighbor (5, 53, 54, 100). However, in order to comply with the mass data, it was necessary to invoke a dislocation along the axis of the IF. This was accomplished by kinking the adjacent molecules at the region of one or more of the several interruptions along the coiled-coils, such as at the L12 linker, which is thought to form a

short β sheet conformation and may specify registration of adjacent molecules (5, 53, 54, 100).

Electron microscopy of shadowed IF and subfilamentous forms of IF has provided support for aspects of this model. Axial bandings of periods of 22 nm and 46–48 nm have been demonstrated in each of the types I–IV IF (93, 96, 101–106). While high-resolution X-ray diffraction patterns have not been possible for most IF, similar or 'generic' surface lattice models that comply with available morphological and biochemical data have been constructed that involve arrays of half-staggered antiparallel four-chain complexes to form brick-wall-like models (83, 91, 101) (Figure 2*e*).

Several other electron microscope studies have described a variety of subfilamentous forms during in vitro assembly or disassembly of IF by modifications to the composition, pH, and ionic strength of the solution (98, 107–112). These have suggested that IF structure may be organized by different hierarchical levels of subfilaments: the four-chain complexes form 3-nm wide protofilaments, two of which entwine to form a 4.5-nm protofibril, and then 2–6 protofibrils, most commonly 4, assemble to form the polymorphic variants of intact IF (109, 111). This view of the organization of IF is an important elaboration of the generic surface lattice because it addresses polymorphism in a straightforward way (5) and requires an understanding of end-to-end associations of the four-chain complexes that have not yet been experimentally addressed. Furthermore, this model predicts that different sets of interactions, between rod-domain segments or end domains or both, stabilize IF at each level of hierarchy.

Taken together, the available data clearly suggests that the coiled-coil segments of IF chains form the basic 'core' of IF and are largely responsible for the common morpholgies of types I–IV IF. Analyses of both equatorial X-ray diffraction data (99, 113, 114) and radial density profiles of STEM images (96, 105) suggest this core is 8–9 nm in diameter.

Despite considerable advances toward understanding IF structure, a number of questions remain. 1. As discussed above, are IF polar or nonpolar structures, or under certain circumstances, both? 2. What role do the end domains have in specifying IF assembly in vitro and in vivo? 3. What is the handedness of the 22-nm repeat that specifies how four-chain complexes (or protofilaments) entwine to form the IF? This is thought, but not yet proven, to be left-handed (103), which would require protofibrils to entwine in a right-handed mode (97, 98, 111). 4. How do the surface lattice models close to form an intact IF? Both annular and close-packed models have been advanced (53, 54, 99, 100). X-ray diffraction (63, 99, 102) and certain electron microscopy (93, 96, 101–106) data demonstrate that the α-helixes are aligned almost parallel to the long axis of the IF, which means that the pitch angle of the basic helix of the IF must be low. Radial density profiles of

STEM images of various native and reassembled IF have not revealed the existence of a significant hollow space in the center of IF (96, 105), which thus favors a close-packed arrangement and is consistent with the concept of entwining of protofilaments and protofibrils. 5. What is the nature of the interactions that stabilize the four-chain complexes both laterally and longitudinally? This issue is important, for example, in understanding the known phosphorylation-dependent disassembly-reassembly of vimentin IF and the nuclear lamina complex during mitosis (see below). 6. When visualized by negative staining in vitro or in situ in whole cells or tissues, IF appear as long flexuous structures. Thus the packing of neighboring structural units in IF must permit a considerable degree of flexibility that has yet to be explained. Does this flexibility have its basis in the existence of several discontinuities in the regularity of the rod domain? Preliminary solid state NMR experiments using ^{13}C- and ^{2}H-leucine-labeled keratin IF indicate that while the peptide backbone of the coiled-coils is relatively rigid, surprisingly, the leucine side-chains buried along the axis of the coiled-coil are very mobile (50). This means that the coiled-coil does not form a rigid rod and is consistent with the observed flexibility of IF. Thus solid and solution state NMR technology may offer a new approach to understanding IF structure. In addition, the methods recently developed to computationally straighten images of IF have provided digitized diffraction patterns of quality comparable to the X-ray diffraction images of ordered keratin IF (104), and thus show promise in solving remaining questions of IF structure. The widespread availability of cDNA and genomic clones to all classes of IF chains should now permit construction of deletion mutants by which the precise functions of the various portions of the chains can be explored.

Structure of Type V Lamins: Variations on a Common Theme

Lamins appear to be arranged as an orthogonal lattice network in *Xenopus* of about 52 nm spacing (14), which corresponds to the length of the rod domains of human lamin A and C chains. Unlike types I–IV IF, however, lamins form a series of polymorphic structures in vitro, ranging from wide paracrystalline particles in solutions of low ionic strength (14, 16–18) to thinner 10-nm wide particles at higher ionic strength (14) with a pronounced periodicity of 25–27 nm. Electron microscope, physicochemical, and ionic charge scoring studies all show that lamins form two- and four-chain molecules similar to those of types I–IV IF. A simple model has been presented that can explain all of these polymorphic structures in terms of arrays of in-register or half-staggered antiparallel molecules, the numbers of which vary due to conformational changes in the large carboxyl-terminal end domains as the ionic strength varies (18). The model places adjacent lamin molecules so that the consensus sequence of the 1A tract of one molecule is juxtaposed with the consensus

sequence of the 2B tract of its neighbor. Thus axial and lateral aggregation may be specified by ionic and apolar interactions of adjacent molecules.

Peripheral Locations of Functional End Domains

Several lines of data indicate that large portions of the end domains of IF chains project out from the 8–9-nm diameter IF core. The 22-nm and 46–48-nm axial repeats visualized by electron microscopy and X-ray diffraction are due to periodic perturbations on the IF surface, quite possibly the end domains. Analysis of radial density profiles of STEM images of many IF indicates that their true diameters are 12–16 nm, depending on the chain composition, that is, the size of the end domains (4, 94–97, 104, 105). Since the average density of the outer radial 2–3 nm is lower than the central core, this region presumably acquires or excludes stain poorly so that it is rarely visualized by most electron microscopic techniques (105). Proteolytic digestion experiments on both keratin IF (38, 115) and neurofilaments (106, 116–119) can release substantial amounts of end domain sequences without apparent disruption of the morphological integrity of the IF core. In addition, high-resolution electron microscopy of neurofilaments in vitro or in situ has revealed rodlike extensions projecting from the IF core (106, 117, 120) that correlate with the large carboxyl-terminal end-domain sequences of their chains.

Accordingly, it is now widely accepted that the IF core of conserved structure provides a scaffold from which the hypervariable end domains of the constituent chains protrude from where they specify the functions of IF. Consequently, the elaborate mechanisms involved in the regulation of IF gene expression are directed toward the production of a pattern of exposed end domains of properties most suited to the functions of the IF required by the cell. Future work directed toward an understanding of the functions of IF in cells will require information on the structures of the protruding end domains and how they interact with other molecules within the cytoplasm and at the nuclear and cell surface.

Postsynthetic Modications of IF Chains

All IF protein chains are known to be subject to modifications for reasons that are not yet generally understood but that presumably regulate IF function and their dynamic organization in cells (see Ref. 6 for recent review). These include acetylation of the amino-terminal residue, cross-linking by transglutaminases and other unusual cross-links (121), phosphorylation, proteolysis by Ca^{2+}-activated or other proteases, and perhaps glycosylation (122). Most known modifications occur on end-domain sequences of the IF chains where the responsible enzymes would be expected to have greater access. For example, in the same proteolytic digestion experiments mentioned above (38,

106, 115–119), the bulk of the phosphate could be removed in end-domain sequences. The available neurofilament sequences have demonstrated the presence of many tandem peptide repeats in their V2 subdomains, which contain likely sites of phosphorylation at serines (48, 123). In the case of the NF-M and NF-H chains, the number of repeats correlates well with the degree of phosphorylation of the chains (48, 123).

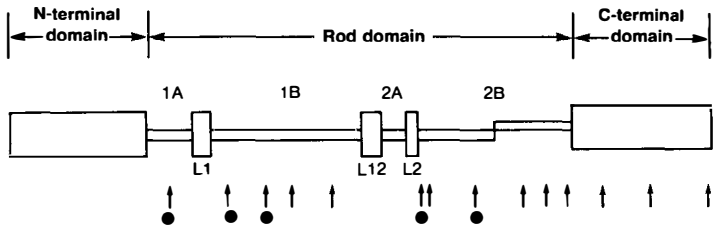
Of special interest is the probable role of phosphorylation in the reversible dissociation of the nuclear lamina complex during mitosis. The lamins become soluble and lose their association with chromatin during prophase. In CHO cells, for example, this coincides precisely with the acquisition on the lamin chains of 2–3 mol/mol of phosphate (15, 124–126). The lamins lose this phosphate and reassemble around the chromosomes as the nuclear envelop re-forms during telophase. However, the nature of the protein kinases and phosphatases are not yet known (127). Solution chemistry experiments suggest that the soluble form of the hyperphosphorylated lamins is a tetramer (14). The insertion of several additional highly charged phosphate groups on the end (or rod) domains would be sufficient to dissociate the lamina structure into its constituent four-chain complexes by electrostatic repulsion (18). Interestingly, the reversible rearrangement of the vimentin network in some cultured cells during various cellular activities including mitosis has also been attributed to a wave of phosphorylation (128–130). These data provide important molecular insights into how cells might regulate the structural organization of IF and the IF-like lamina complex, and thus provide an exciting system for further study of the regulation and function of IF in cells.

IF GENE ORGANIZATION AND COMPLEXITY

IF Gene Structure

The study of IF genes has been a subject of intense research in recent years in order to answer fundamental questions relating to their complexity, evolution, and differential gene expression. Figure 3 illustrates the intron/exon structures of several representatives of types I–IV IF genes: type I genes for mouse keratin 10 (59 kd) (131), bovine 54 kd (132), human keratins 14 (133) and 17 (134), and bovine keratin 19 (37); type II genes for human keratins 1 (135), 3 (136), and 6a/b (137); type III genes for hamster vimentin (138) and desmin (139) and mouse GFAP (140); and the type IV genes for mouse (46) and human (49) NF-L and human NF-M (48). Other genes have also been sequenced whose intron/exon structures comply with the types shown in Figure 3, including a type I gene from *Xenopus laevis* (141), the mouse keratin 1 (type II) gene (H. Nakazawa, D. Roop, unpublished), the type III human (142) and chicken (143) vimentin genes, and mouse NF-M (144).

As we had described previously (4, 131, 135, 145), types I–III IF genes



Mouse 59 kD (I)			14	42		94		4		46		119		113
Bovine 54 kD (I)			14	42		94		4		46		119		84
Human K14 52 kD (I)			14	42		94		4		46		120	15	
Human K17 46 kD (I)			14	42		94		4		46		120	15	
Bovine K19 44 kD (I)			14	42		94		4		46				
Human K1 67 kD (II)	17			42	62	94		7		46		119	12	
Human K3 63 kD (II)	17			42	62	94		4		46		119	12	
Human K6b 60 kD (II)	17			42	62	94		4		46		119	11	
Hamster Vimentin 53 kD (III)				42	62	94		4		46		120	14	42
Hamster Desmin 52 kD (III)				42	62	94		4		46		120	14	41
Mouse GFAP 50 kD (III)				42	62	94		4		46		120	14	42
Mouse NF-L 68 kD (IV)											79111			95
Human NF-L 68 kD (IV)											79111			95
Human NF-M 105 kD (IV)											79111			

Figure 3 Location of introns in selected types I–IV IF genes. The arrows indicate the positions where introns splice the coding regions of a universal model of an IF chain. The numbers refer to the residue position along the rod-domain tract or the carboxyl-terminal end. The arrows with closed circles mark those introns that occur at the first residue of the heptad repeat. See text for references. The nomenclature used for the human chains and their mouse equivalents is from Refs. 159, 160. Although the gene structure for type V lamins is not yet known, human A and C chains contain an additional six heptads (42 residues) in their 1B tract, which are inserted at the conserved intron position 42 in types I–III IF genes (24, 36). The wool keratin type II 'orphan' exon (42) spans residues 46–119 of the 2B tract. Figure redrawn from Refs. 4, 5, 145.

show striking similarities in their organization, including certain type-specific intron locations in their rod domains, several common introns, and certain introns that splice the beginning of heptads. Contrary to our earlier predictions, however, type IV neurofilament genes are organized completely differently (Figure 3).

Most genes so far characterized appear to exist as a single copy per haploid genome, with the exception of human keratin 6 (137). A pseudogene has been reported for human keratin 14 (146) and the mouse endo A (keratin 8) gene (147). It should be pointed out that in most studies the identification of a single-copy gene was based on the finding of a single band on a genomic

Southern blot. Experiments using more quantitative methods including sequencing will be necessary to exclude the possibility of multiple identical copies or copies with only a minor degree of polymorphism (148).

Models for IF Gene Evolution

Based on the remarkable conservation of the protein chain structures of types I–IV IF and of the genes for types I–III IF, we initially suggested that IF genes arose from a common ancestor (4, 131, 135, 145). We proposed that a primordial gene was assembled from smaller units encoding multiple heptad repeats (perhaps 28 residues or 84 bp) that were separated by intervening introns. The amplification of simple DNA sequences has previously been suggested to play an important role in the evolution of other multigene families including the collagens (149), immunoglobulins (150), and albumin: α-fetoproteins (151). The conservation of certain features of the rod domains of types I–IV IF suggested that during or after assembly of the primordial gene, certain events occurred prior to duplication of the genes for the sequence types. These included the formation of the discontinuity of the heptad polarity near the middle of the 2B tract and the insertion of three linkers in the rod domain (Figure 1). These four features have been highly conserved in position, but not necessarily in sequence, and presumably provide some important evolutionary advantage in IF formation, stability, or function. According to this view, duplication events followed by the loss of some introns and divergence of exons would have generated the types I–IV IF progenitor genes, and subsequently, their various members.

However, recent data on the type IV neurofilament genes require modifications to this hypothesis as their introns do not coincide with those for types I–III genes (Figure 3). Lewis & Cowan (46) suggested that the unexpected intron/exon structure of the mouse NF-L gene may have occurred by the transposition of an mRNA intermediate, by mechanisms previously proposed to explain the origin of processed pseudogenes (152–154). This model would require the insertion of at least two introns into the transposed progenitor neurofilament gene prior to duplication, since the human NF-L (49), NF-M (48), and NF-H (155; J.-P. Julien, personal communication) genes contain introns in the same positions in the 2B tract as does the mouse NF-L gene (46). More recently, an alternative hypothesis was presented that suggested that neurofilament genes diverged from a progenitor IF gene prior to the types I–III IF genes (48, 49), at least 700 Myr ago. That is, the primordial gene initially diverged into two progenitors, one for the type IV gene lineage, the other for types I–III gene lineage. This view is consistent with recent surveys of IF proteins in lower eukaryotes (8, 25). In this scenario, it is possible the type IV neurofilament progenitor lost numerous introns and

acquired others after divergence. Alternatively, in the primordial IF gene, the repeating heptad elements may not have been separated by introns, but were inserted into the separate type I–III and type IV progenitors after divergence.

The question arises as to what is or was the IF primordial gene. The type V lamins may provide an important clue. Although information on their gene structure is not yet available, their protein chain sequences reveal numerous interesting but subtle differences from the types I–IV cytoplasmic IF chains, including insertion of precisely six heptads in the 1B tract at an intron position conserved in types I–III IF chains, notable differences in the linker sequences, and variations in the periodicities in the distributions of charged residues (Figure 1). Lamins emerged very early during eukaryote cell evolution, perhaps 1000 Myr ago, presumably in order to provide a structural complex or matrix ('karyoskeleton') with which to coordinate chromatin structure during the then novel process of mitosis (13, 15). Thus modern lamins may represent the primordial or early progenitor form of the cytoplasmic IF.

The likely evolutionary mechanisms that have generated the diverse IF genes will probably be debated for some time. Nevertheless, it seems likely that they appeared as metazoan evolution progressed in response to the requirements of specific differentiated types of cells. As discussed earlier in this review, these specialized functions are most probably achieved through the hypervariable end domains of the different IF chains. To date, there are few obvious hints from gene structure as to how the end domain sequences were recruited onto the various type-specific rod domains. Interestingly, all type I–III IF genes, except the smallest type I bovine keratin 19 gene (37), possess an intron near the end of the rod domain (Figure 3). The keratin 19 gene is normally expressed only in simple epithelia, and it was suggested that it may represent the ancestral IF gene prior to recruitment of the carboxyl-terminal end domains (37). As noted earlier, the V1 and V2 subdomains of keratins and V2 subdomains of neurofilament chains possess conserved and nonconserved peptide repeats. The human NF-M chain, for example, contains six repeats of a 13-residue peptide (48) and the keratins contain exact or inexact repeats of the form aliphatic-(glycine/serine)$_n$ (38, 145). We have previously noted that many of the variations in the sequences of this repeat could have occurred by single base changes in the most commonly used codons (145). As more complex differentiated cell-types arose, it is likely that the end domains lengthened and diverged by frequent amplification events and subsequent divergence (33, 136, 145). Furthermore, coexpressed type I and II keratins possess similar V1 and V2 subdomains (4, 38, 145). However, the mechanism(s) involved in recruiting similar V1 and V2 subdomain sequences remains an unresolved evolutionary enigma, especially since the coexpressed partners seem to be widely scattered on the genome.

IF Gene Linkage and Chromosomal Localization

To date there are a number of reports concerning the organization of IF genes. Current data indicates that genes of the same sequence type are often linked into clusters. There are no substantiated reports of close linkage of genes of different sequence types. Three human type I genes including those for keratin 17 are located on a 20-kbp fragment (134). Preliminary data from overlapping human cosmid fragments suggest that there may be as many as five type I genes including keratins 14 and 17 within 60 kbp (156). A cosmid clone containing the human keratin 1 (type II) gene appears to be linked to another unspecified type II gene (156). Clusters of three type I and two type II sheep wool keratin genes have been located on separate cosmid clones (42). Interestingly, in this study, an 'orphan' exon encoding the latter half of the 2B rod domain (see Figure 3) was found between the two type II wool genes. In other studies, two type I *Xenopus* genes are separated by 10 kbp (141), four bovine type II genes are located within 11 kbp (157), and the mouse NF-L and NF-M genes are located within 20 kbp (144). By use of rodent-human somatic cell hybrids, the single-copy human vimentin and desmin genes are located on chromosomes 10 and 2, respectively (158). Chromosomal localization studies by restriction fragment length polymorphism analyses have demonstrated that the mouse keratin 1 and 6 (type II) genes are located on chromosome 15, and the keratin 10 (type I) gene is located on chromosome 11 (J. Compton, D. Roop, unpublished). Similarly, by use of somatic cell hybrids, the human keratin 1 and 10 genes have been localized to chromosomes 12 and 17, respectively (S. Lessin, P. Steinert, unpublished). Characterization of the entire complement of the more than 30 IF genes and their distributions poses a formidable challenge but will be required to fully understand their mechanisms of gene evolution and expression.

IF GENE EXPRESSION

While there are relatively few studies concerning the regulation of expression of types III–V IF (but see Refs. 4–8, 13 for review), much more is known about the complex keratin family. The keratins consist of 25–30 protein chains (159, 160) divided approximately equally into two distinct subfamilies of acidic type I and neutral-basic type II chains. Certain general principles of their expression have been established (159, 165), of which the most striking is that at least one member of each subfamily is always coexpressed in any given epithelial tissue. This means that keratin IF in vivo are heteropolymers, as demonstrated in in vitro assays (63–67). The pair(s) of keratins coexpressed is highly characteristic of the epithelium. Few data are available on the structural/functional significance of this phenomenon, but several interesting reports have recently shed light on aspects of keratin IF gene expression that will have relevance to all IF types.

Developmentally Regulated IF Gene Expression

To date there is no evidence that any type of IF chain in mammals is expressed exclusively during embryonic or fetal development (166–168). However, the mouse keratin 6 gene is an interesting special case. This gene is usually expressed in adult epidermal tissues following induction of hyperplasia (169, 170). Recently, it has been shown that keratin 6 is expressed at day 15 in the periderm (the precursor of the epidermis) of mouse embryos and persists until the periderm is lost at day 18 after a fully stratified epidermis has formed (169). Also, this gene is constitutively expressed in newborn mice homozygous for the pf and Er mutations that exhibit hyperplasia and defects in epidermal differentiation (171).

In *Xenopus,* however, there is considerable evidence for the existence of at least three different sets of keratin genes: those expressed in eggs; those in developing embryos; and those in adult tissues only (172–175). For example, Sargent & Dawid entered the IF field quite by accident when they initiated a study to identify genes that were expressed in early frog embryogenesis but not oogenesis (176). They isolated several type I keratin clones from a subtracted cDNA library containing mRNA species that were transcribed in gastrulae following fertilization (174, 175). By use of RNA blotting and in situ hybridization experiments, they have demonstrated that these genes are expressed in embryos as early as 7 hr after fertilization and throughout development to tadpoles, but are no longer detectable after metamorphosis (141, 177). These data indicate that several keratin genes are expressed very early in a cell-type-specific manner. This system therefore offers considerable promise in elucidating the inductive mechanisms involved in keratin (and other IF) gene expression during development.

Other studies in *Xenopus* have shown that different sets of type V lamin chains are expressed during oogenesis, spermatogenesis, and development of differentiated somatic cells (178–180). In this case, the synthesis of these cell-type specific lamins is probably related to changes in nuclear organization (13).

Transcriptional Versus Translational Regulation

The majority of data concerning expression of IF genes suggests that they are regulated transcriptionally and/or posttranscriptionally (that is, mRNA stability) rather than at the level of translation. However, there are two reports indicating that regulation may occur at the translational level. One concerns the presence of nontranslated mRNA encoding the type II keratin 1 in certain transplantable squamous cell carcinomas of mouse skin and forestomach (181). This is not a general regulatory mechanism for squamous cell carcinomas, since RNA blots (182) and in situ hybridization (183) fail to detect appreciable amounts of keratin 1 mRNA. A second report suggests that the

human type II keratin 6 gene is constitutively expressed in normal epidermis but is not translated (184). Translation is presumed to occur in response to a hyperproliferative stimulus such as wounding. This mechanism of regulation does not occur in mouse epidermis. mRNA for the mouse keratin 6 gene is not detectable in normal epidermis, as judged by in situ hybridization and RNA blot analyses, although it can be detected in epidermis within 7 hr after topical application of the hyperplasiogen 12-*O*-tetradecanoyl-phorbol 13-acetate (TPA) (170). In related experiments, mRNA for the mouse type I keratin 16 gene, usually coexpressed with the type II keratin 6, is not detectable in normal mouse epidermis by in situ hybridization, but is detectable shortly after exposure to hyperproliferative stimuli such as treatment with TPA (185). It is unclear at this time whether the mechanism(s) regulating the expression of keratin 6 (and 16) are different between the two species or whether there are technical reasons to explain the different observations.

Identification of Sequences Regulating IF Gene Expression

Little information is available concerning factors that regulate the expression of IF genes, although several effectors have been implicated, including environment (186–192), vitamin A (193–199), estrogens (200–202), and tumor promotors such as TPA (142, 182, 203–206). In no case has the mechanism of action of these putative effectors been determined. Nevertheless, because IF genes are obviously tightly regulated with regard to cell type and state of differentiation, sequences must exist that are important in the regulation of their expression.

Several attempts have been made to identify potential regulatory sequences of IF genes by comparisons of 5'-flanking sequences (37, 134, 137, 141–143, 158). To date, these comparisons have revealed in a variety of IF genes homologies to the simian virus 40 core transcription enhancer (134, 137, 158) and GC box elements that bind the specific cellular transcription factor Sp1 (37, 158). In addition, a putative epidermal gene-specific consensus sequence of the form AARCCAA (where R = a purine) has been identified (157), but the sequence has not yet been proven to be a specific enhancer (157).

One problem confronting investigators attempting to identify sequences regulating the expression of IF genes expressed during differentiation has been the inability to induce high levels of expression of the genes in in vitro systems (207). The production of transgenic mice by microinjection of cloned DNA into fertilized mouse eggs provides an attractive alternative method of identifying *cis*-acting regulating elements responsible for tissue-specific gene expression (see Ref. 208 for review). This approach has recently been used to show that a 21.5-kbp fragment bearing the human NF-L gene contains the necessary information for correct expression of this gene in neuronal tissues (209). The transgenic mice expressed this gene appropriately at the RNA and

protein levels by all criteria measured, including assembly into neurofilaments in neuronal cell bodies and their processes (209). We have also recently produced transgenic mice by use of a 12.3-kbp fragment bearing the human type II keratin 1 gene (135). This fragment also contains sufficient information for tissue-specific expression of the gene at the protein level (D. Roop, P. Steinert, S. Chung, unpublished). Thus, experiments using transgenic mice appear to hold great promise in identifying specific regulatory elements for IF genes.

INTRACELLULAR ORGANIZATION OF IF: ASSOCIATIONS WITH THE PLASMA MEMBRANE AND NUCLEAR SURFACE

IF are among the most insoluble constituents of cells. Extraction of tissue or cultured cells by a variety of methods, including use of solutions of high salt with nonionic detergents, removes membranes and most other proteins including microtubules and microfilaments. What remains is a 'skeleton' consisting largely of the IF component and associated structures. Both electron microscopy and indirect immunofluorescence with anti-IF antibodies by light microscopy have been used to study the elaborate arrays of IF in such extracted cells. Generally, these arrays form complex interconnecting networks, emanating from a perinuclear ring, from which IF appear to connect to the nuclear surface and extend throughout the cytoplasm terminating at the plasma membrane (6, 7). Such associations between the plasma membrane and nuclear surface appear to provide a continuous link that may have important implications for the organization of the cytoplasm, cellular communication, and perhaps information transport into and out of the nucleus (210). Yet IF are not all static structures. Their dynamic organization, distribution, and interactions with other cellular components have been the subject of studies in many different systems, especially in cultured cells treated with drugs (211–213), or other chemicals (216, 217), injected with various anti-IF antibodies (214, 215, 218), grown under varied media conditions (219, 220), or observed during mitosis (220–223), cell spreading (223–225), etc. Many of these studies have noted that perturbations of the IF network result in the collapse of the IF onto or near the nuclear surface. Other studies show the IF or are reorganized from near the nuclear surface on recovery of the cells from treatment or after mitosis. These observations have raised the possibility of IF nuclear-organizing center(s) (226). Although the evidence for the association of IF between the nuclear and cell surface has been primarily morphological (6, 7, 16, 210, 227), recent studies have shed light on the molecular aspects of some of these associations.

Many cells possess a complex proteinaceous network or skeleton immediately subjacent to the cell membrane. Together with other proteins, ankyrin binds to the membrane as well as spectrin to which in turn the actin-containing microfilament network is attached (228–231). An association of IF with the spectrin component or related proteins has been suggested (232–234). However, it has more recently been demonstrated by in vitro binding assays on isolated human erythrocyte plasma membranes (inside-out vesicles) that both vimentin (235–238) and desmin (238) bind directly to the ankyrin and not spectrin components. Interestingly, this ocurs by way of the basic amino-terminal end domains of these IF chains (235–238), since use of peptide fragments devoid of these sequences did not bind ankyrin or to the vesicles. Further, it was shown that ankyrin binds to a small polymer (presumably the soluble tetramer) of vimentin, and this association prevented further assembly of vimentin into complete IF (236). Since the intact amino-terminal end domains of desmin (77) and vimentin (78) are required for IF assembly in vitro, these data suggest that ankyrin serves as a "capping" site for vimentin and desmin IF in vivo.

Several types of specialized cells possess junctions termed desmosomes that represent a complex region of cell-to-cell adhesion, whose structure, composition, and dynamics have been studied (210, 219, 220, 239–245). In the case of epithelia, keratin IF form a complex network coursing throughout the cytoplasm, which appears to originate from a perinuclear cage and abuts the desmosomal junction, thus appearing to integrate mechanically adjacent cells of the epithelium. On the intracellular surface, the desmosomal plaque consists of one or two desmoplakin proteins of 220–250 kd, to which bundles of keratin IF appear to attach. The dynamics of desmosome formation have been investigated using a model epidermal cell culture system in which the organization of the keratin IF cytoskeleton and desmosomes can be manipulated by varying the exogenous levels of Ca^{2+} (210, 219, 246–249). In such cells in low Ca^{2+}, preformed desmosomal components including the desmoplakins exist in the cytoplasm in close association with the keratin IF. Raising the media Ca^{2+} induces rapid redistribution of the components toward the cell membrane where desmosomes form (219, 244). This process appears to be coordinated with the microfilament network; that is, it may provide the motive force for the movement of the IF-desmoplakin complexes (246). Accordingly, the desmoplakins may also serve as IF-associated or "capping" proteins (244, 250, 251). However, the molecular mechanism of attachment remains to be established. Desmoplakins are known to be glycine-rich (252), and it is possible they may mediate attachment by interaction with the glycine-rich end domains of the keratin IF, and by analogy with vimentin and desmin binding to ankyrin, with the amino-terminal end domains present on even the smallest known keratin 19 chain. Hemidesmosomes connect a

basal epithelial cell to its underlying connective tissue, and thus are thought to be half of a desmosome. Although desmoplakins have also been implicated in keratin IF associations in hemidesmosomes (243), they appear not to be identical to those of desmosomes (251). It is thus possible the desmoplakins consist of a family of keratin IF-binding proteins, the members of which preferentially bond the keratin IF of varying protein chain composition expressed in different epithelial tissues, thus reflecting the specific structural and functional requirements of the epithelium. The combined use of structural, biochemical, and molecular biological approaches should clarify many of these tantalizing observations.

The avian erythrocyte nuclear membrane system has been utilized to study the observed morphological association of IF with the nuclear surface. By differential extraction studies, a specific in vitro binding of vimentin (253) or desmin (238) with the lamin B chain has been observed. The lamin B chain has been shown to associate with nuclear membrane fragments and thus may serve to coordinate the disassembly and reassembly of the nuclear lamina complex during mitosis (15). Vimentin or desmin IF-lamin B binding could be blocked by anti-lamin B antibodies, but not antilamin A or antiankyrin antibodies. Interestingly, the binding was localized to a 6.6-kd peptide from the carboxyl-terminal end of vimentin (253) and to a 31-residue region immediately following the rod domain, but not the terminal 24 residues of desmin (238). Furthermore, purified vimentin or desmin derivatives containing their carboxyl-terminal sequences could displace each other from binding the nuclear membranes. These data suggest that the lamin B chain may function as an IF attachment site.

Thus vimentin and desmin IF chains appear to bind to the nuclear lamina by way of their carboxyl-terminal ends and to the plasma membrane by way of their amino-terminal ends. These biochemical experiments provide direct evidence for the continuity of the IF network in cells by connecting the plasma membrane skeleton with the nuclear lamina or karyoskeleton. Moreover, these observations have important implications in vectorial assembly of the IF network in cells, as well as a functional/structural polarity of the IF themselves. This latter point is important in view of the current uncertainty discussed earlier as to whether IF are polar or nonpolar structures based on the arrangement of their constituent four-chain complexes. On the other hand, surface lattice drawings (Figure 2*e*) deploy amino- and carboxyl-terminal end domains adjacent to each other on the periphery of IF. Irrespective of whether IF are polar or nonpolar structures, a functional polarity could be imparted to an IF by virtue of the preferences of its protruding end-domain sequences to bind to different molecules, as now demonstrated by ankyrin and lamin B.

IF-ASSOCIATED PROTEINS

In addition to the 30 or more proteins in any one mammalian species that can be demonstrated to form IF assemblies in vitro and in vivo, there is a growing list of other minor proteins that frequently coisolate and/or associate with the major IF structural proteins. These are generically termed IFAPs. While we have made the case in this review that IF end domains are often intimately involved in prescribing the functions as well as dynamic organization and distribution of IF in cells, it seems likely that many IFAPs are also involved in the supramolecular organizations that IF networks assume in cells. Table 2 lists a number of proteins so far identified that may function as IFAPs. In most cases only simple biochemical properties and tissue localization using indirect immunofluorescence techniques have been analyzed. Few data are yet available on the molecular mechanisms by which they interact with IF. Nevertheless, it seems possible to assign these IFAPs into different functional classes that may also correlate with their size.

1. IFAPs of relatively low molecular weight (10,000–45,000) that bind IF laterally into tight macrofilament aggregates. The prototypical examples are (*a*) the cysteine-rich 'high-sulfur' and high tyrosine-glycine proteins of hair and related epithelial derivatives that form an interfilamentous matrix and cross-link the keratin IF by disulfide bonds (254–256) and (*b*) filaggrin, which cross-links the keratin IF of terminally differentiating orthokeratinizing epithelia by noncovalent interactions (257–260).
2. IFAPs of generally high molecular weight that cross-link IF into loose networks. These include paranemin (261) and synemin (262), which are organized with vimentin and desmin IF networks of avian erythrocytes and muscle cells; trichohyalin, associated with the keratin IF unique to the inner root sheath cell layer of the hair follicle (263); microtubule-associated proteins (MAPs) involved in the lateral association of neurofilaments and microtubules (264); and a 300-kd protein (265, 266) [also known as plectin (267, 268)] of many cultured fibroblastic cells. Interestingly, plectin also binds to MAPs and human erythrocyte spectrin, suggesting that it is a multifunctional cross-linking IFAP (268). It has been suggested that the NF-M and NF-H chains may also be IFAPs since they apparently do not self-assemble in vitro into morphologically intact IF (73, 96, 120), but their amino acid sequences clearly show a characteristic IF rod domain. Preliminary data suggest that synemin (262) and the 300-kd protein may also possess an α-helical core (266). As with the NF-M and NF-H chains, these IFAPs may integrate into IF structures, so that their protruding end-domain sequences function as spacers in the generation of the loose IF networks.

Table 2 *Intermediate filament–associated proteins*

Protein name	Distribution	Size (kd)	Ref.
1. Lateral aggregation			
"high sulfur" family	hair and related tissues		
high glycine-tyrosine family		10–30	254–256
filaggrin family	orthokeratinizing epithelia	16–45	257–260
2. Cross-linking			
paranemin	avian erythrocytes muscle tissues	280	261
synemin		230	262
trichohyalin	inner root sheath cells of hair follicle	190	263
MAPS 1,2	neuronal tissues	300	264
300 kd[a]	fibroblasts	300	265, 266
plectin[a]		300	267, 268
NF-M, NF-H (?)	neuronal tissues	105, 135	73, 96, 120
3. Capping			
Spectrin (?)	ubiquitous	220–240	233, 234
Ankyrin			235, 236
Desmoplakin	desmosomes, hemidesmosomes	220–240	244, 250, 251
Lamin B	ubiquitous	65	238, 253, 254
4. Others			
Epinemin[b]	vimentin networks many cultured cells	45	269
p50[b]		50	270
p68[c]	vimentin networks, ubiquitous (?)	68	271, 272
β-internexin[c]	heat-shock proteins (?)	70	273
α-internexin	neuronal tissues	66	274
p95	vimentin networks	95	275

[a,b,c] These are probably the same proteins.

3. IFAPs that function as "capping" proteins. As discussed above, the proteins ankyrin (235, 236) and perhaps spectrin (233, 234) of the plasma membrane, desmoplakins of desmosomes (244, 250, 251), and the lamin B chain of the nuclear lamina complex (238, 253, 254) appear to anchor the ends of IF.
4. Finally, there are several IFAPs that do not appear at this time to conform to the above classification scheme (Table 2). Interestingly, the presumably identical proteins p68 and β-internexin seem to be heat-shock proteins

whose expression is invoked under a variety of stress stimuli. The reason for their association with the IF network remains unknown.

A large body of data has been accumulated showing the usefulness of antibodies to the principal IF structural proteins in the diagnosis of disease and of tumor types (see Refs. 7, 8, 160, 276 for detailed reviews). Some evidence exists suggesting that IFAPs may also be involved directly or indirectly in diseases, especially in keratinizing disorders of hair (277, 278) and skin (279, 280). Detailed studies will be required to understand the role of IFAPs in the organization of IF in both normal and abnormal cells.

CONCLUDING REMARKS

Despite considerable advances in understanding the complexity of IF proteins, their structure, patterns of gene expression, and dynamic organization in cells, a clear picture of their functions has not yet emerged. Since some cell types appear not to contain IF, and many cultured cells can still function quite well and even divide following disruption of the organization of their IF, it seems unlikely that IF have fundamentally important 'housekeeping' functions. Based on the complexity of the IF gene family, as well as their generally tissue-specific expression and diversity of the properties of their end domains, it seems probable that IF are involved in specialized functions related to the differentiation state of the cell (9, 13, 210). These may include mechanical coordination of the cytoskeleton of the cell or tissue, information transport, and signal transduction (210). On the other hand, lamins of the nuclear lamina complex seem to be ubiquitous and thus essential for the coordination of both nuclear structure and function and gene expression. They may also regulate the assembly and the state of organization of the cytoplasmic IF in each differentiated cell type. The family of IF proteins therefore remains an important and enigmatic component of cells of continuing interest to cell and molecular biologists.

ACKNOWLEDGMENTS

We are grateful to many investigators for their advice, encouragement, and valuable discussions during the course of the preparation of this review and comments on the manuscript. We especially acknowledge Drs. M. Blumenberg, M. Carden, J. Compton, I. Dawid, G. Dessev, S. Georgatos, K. Green, R. Goldman, J. Jones, J.-P. Julien, R. Lazarini, S. Lessin, D. Parry, J. Schweizer, and A. Steven for sharing with us their unpublished data.

Literature Cited

1. Lazarides, E. 1980. *Nature* 283:249–56
2. Lazarides, E. 1982. *Ann. Rev. Biochem.* 51:219–50
3. Steinert, P. M., Jones, J. C. R., Goldman, R. D. 1984. *J. Cell Biol.* 99:225–75
4. Steinert, P. M., Steven, A. C., Roop, D. R. 1985. *Cell* 42:411–19
5. Steinert, P. M., Parry, D. A. D. 1985. *Ann. Rev. Cell Biol.* 1:41–65
6. Traub, P. 1985. *Intermediate Filaments: A Review.* Berlin: Springer-Verlag
7. Wang, E., Fischman, D., Liem, R., Sun, T.-T. 1985. *Intermediate Filaments, Ann. NY Acad. Sci.* 455
8. Osborn, M., Weber, K. 1986. *Trends Biochem. Sci.* 11:469–72
9. Fraser, R. D. B., Steinert, P. M., Steven, A. C. 1987. *Trends Biochem. Sci.* 12:43–45
10. Parry, D. A. D., Fraser, R. D. B. 1985. *Int. J. Biol. Macromol.* 7:203–13
11. Conway, J. F., Parry, D. A. D. 1988. *Int. J. Biol. Macromol.* In press
12. Parysek, L. M., Goldman, R. D. 1987. *J. Neurosci.* 7:781–91
13. Franke, W. W. 1987. *Cell* 48:3–4
14. Aebi, U., Cohn, J., Buhle, L., Gerace, L. 1986. *Nature* 323:560–64
15. Gerace, L. 1986. *Trends Biochem. Sci.* 11:443–46
16. Zackroff, R. V., Goldman, A. E., Jones, J. C. R., Steinert, P. M., Goldman, R. D. 1984. *J. Cell Biol.* 98:1231–37
17. Goldman, A. E., Maul, G., Steinert, P. M., Yang, H.-Y., Goldman, R. D. 1986. *Proc. Natl. Acad. Sci. USA* 83:3839–43
18. Parry, D. A. D., Conway, J. F., Goldman, A. E., Goldman, R. D., Steinert, P. M. 1987. *Int. J. Biol. Macromol.* 9:137–45
19. Shelton, K. R., Guthrie, V. H., Cochran, D. L. 1982. *J. Biol. Chem.* 257:4328–32
20. Krohne, G., Benavente, R. 1986. *Exp. Cell Res.* 162:1–10
21. Krohne, G., Dabauvalle, M., Franke, W. W. 1981. *J. Mol. Biol.* 151:121–41
22. McKeon, F. D., Kirschner, M. W., Caput, D. 1986. *Nature* 319:463–68
23. Fisher, D. Z., Chaudhary, N., Blobel, G. 1986. *Proc. Natl. Acad. Sci. USA* 33:6450–54
24. Parry, D. A. D., Conway, J. F., Steinert, P. M. 1986. *Biochem. J.* 238:305–8
25. Bartnik, E., Osborn, M., Weber, K. 1986. *J. Cell Biol.* 102:2033–41
26. Dawson, P. J., Hulme, J. S., Lloyd, C. W. 1985. *J. Cell Biol.* 100:1793–98
27. Soll, D. R., Mitchell, L. H. 1983. *J. Cell Biol.* 96:486–93
28. Koury, S. T., Eckert, B. S. 1984. *J. Cell Biol.* 99:320a
29. Linck, R. W., Langevin, G. L. 1982. *J. Cell Sci.* 58:1–22
30. Linck, R. W., Amos, L. A., Amos, W. B. 1985. *J. Cell Biol.* 100:126–35
31. Linck, R. W., Stephens, R. E. 1987. *J. Cell Biol.* 104:1069–75
32. Chang, X.-J., Piperno, G. 1987. *J. Cell Biol.* 104:1563–68
33. Lasek, R., Phillips, L., Katz, M. J., Autillo-Gambetti, A. 1985. *Ann. NY Acad. Sci.* 455:462–78
34. Pruss, R. M., Mirsky, R., Raff, M. C., Thorpe, R., Dowding, A. J., Anderton, B. J. 1981. *Cell* 27:419–28
35. Weber, K. 1986. *Nature* 320:402
36. Jackson, A. P., Seow, H.-F., Holmes, N., Drickamer, K., Parham, P. 1987. *Nature* 326:154–59
37. Bader, B. L., Magin, T. M., Hatzfeld, M., Franke, W. W. 1986. *EMBO J.* 5:1865–75
38. Steinert, P. M., Parry, D. A. D., Idler, W. W., Johnson, L. D., Steven, A. C., Roop, D. R. 1985. *J. Biol. Chem.* 260:7142–49
39. Jorcano, J. L., Rieger, M., Franz, J. K., Schiller, D. L., Moll, B. L., Franke, W. W. 1984. *J. Mol. Biol.* 179:257–81
40. Roop, D. R., Cheng, C. K., Titterington, L., Meyers, C. A., Stanley, J. R., et al. 1984. *J. Biol. Chem.* 259:8037–40
41. Dowling, L. M., Crewther, W. G., Inglis, A. S. 1986. *Biochem. J.* 236:695–703
42. Powell, B. C., Cam, G. R., Fietz, M. J., Rogers, G. E. 1986. *Proc. Natl. Acad. Sci. USA* 83:5048–52
43. Geisler, N., Plessmann, U., Weber, K. 1985. *FEBS Lett.* 182:475–78
44. Geisler, N., Fischer, S., Vandekerkhove, J., Plessmann, U., Weber, K. 1985. *EMBO J.* 4:57–63
45. Julien, J.-P., Ramachandran, K., Grosveld, F. 1985. *Biochim. Biophys. Acta* 825:398–404
46. Lewis, S. A., Cowan, N. J. 1986. *Mol. Cell. Biol.* 6:1529–34
47. Napolitano, E. N., Chin, S. S. M., Colman, D. R., Liem, R. K. H. 1987. *J. Neurosci.* 7:2590–95
48. Myers, M. W., Lazarini, R. A., Lee, V. M.-Y., Schlaepfer, M. W., Nelson, D. L. 1987. *EMBO J.* 6:1617–26
49. Julien, J.-P., Grosveld, F., Vazdan-

kaksh, K., Flavell, D., Meijer, D., Mushynski, W. 1987. *Biochim. Biophys. Acta* 909:10–20
50. Mack, J. W., Torchia, D. A., Steinert, P. M. 1987. *Biophys. J.* 51:860
51. Leszczynski, J. N., Rose, G. D. 1986. *Science* 234:849–55
52. Crewther, W. G., Dowling, L. A., Steinert, P. M., Parry, D. A. D. 1983. *Int. J. Biol. Macromol.* 5:267–74
53. Fraser, R. D. B., Suzuki, E., Parry, D. A. D. 1985. *Int. J. Biol. Macromol.* 7:257–73
54. Fraser, R. D. B., MacRae, T. P., Parry, D. A. D., Suzuki, E. 1986. *Proc. Natl. Acad. Sci. USA* 83:1179–83
55. Steinert, P. M., Parry, D. A. D., Racoosin, E. L., Idler, W. W., Steven, A. C., et al. 1984. *Proc. Natl. Acad. Sci. USA* 81:5709–13
56. Woods, E. F., Inglis, A. S. 1984. *Int. J. Biol. Macromol.* 6:277–83
57. Parry, D. A. D., Steven, A. C., Steinert, P. M. 1985. *Biochem. Biophys. Res. Commun.* 127:1012–18
58. Quinlan, R. A., Franke, W. W. 1982. *Proc. Natl. Acad. Sci. USA* 79:3452–56
59. Quinlan, R. A., Cohlberg, J. A., Schiller, D. L., Hatzfeld, M., Franke, W. W. 1984. *J. Mol. Biol.* 178:365–88
60. Carden, M. J., Eagles, P. A. M. 1983. *Biochem. J.* 215:227–37
61. Carden, M. J., Eagles, P. A. M. 1986. *Biochem. J.* 234:587–60
62. Lam, K. S., Kasper, C. B. 1979. *J. Biol. Chem.* 254:11713–20
63. Steinert, P. M., Idler, W. W., Zimmerman, S. B. 1976. *J. Mol. Biol.* 108:547–67
64. Hatzfeld, M., Franke, W. W. 1985. *J. Cell Biol.* 101:1826–41
65. Steinert, P. M., Idler, W. W., Aynardi-Whitman, M., Zackroff, R. V., Goldman, R. D. 1982. *Cold Spring Harbor Symp. Quant. Biol.* 46:465–74
66. Steinert, P. M. 1986. *J. Invest. Dermatol.* 86:508
67. Quinlan, R. A., Hatzfeld, M., Franke, W. W., Lustig, A., Schultes, T., Engel, J. 1986. *J. Mol. Biol.* 192:337–52
68. Steinert, P. M., Idler, W. W., Cabral, F., Gottesman, M. M., Goldman, R. D. 1981. *Proc. Natl. Acad. Sci. USA* 78:3692–96
69. Cabral, F., Gottesman, M. M., Zimmerman, S. B., Steinert, P. M. 1981. *J. Biol. Chem.* 256:1428–31
70. Quinlan, R. A., Franke, W. W. 1983. *Eur. J. Biochem.* 132:477–84
71. Moon, H. M., Wisniewski, T., Merz, P., DeMartini, J., Wisniewski, H. M. 1981. *J. Cell Biol.* 89:560–67
72. Geisler, N., Weber, K. 1981. *J. Mol. Biol.* 151:565–71
73. Liem, R. K. H., Hutchinson, S. B. 1982. *Biochemistry* 21:3221–26
74. Zackroff, R. V., Idler, W. W., Steinert, P. M., Goldman, R. D. 1982. *Proc. Natl. Acad. Sci. USA* 79:754–57
75. Zackroff, R. V., Goldman, R. D. 1980. *Science* 208:152–55
76. Eagles, P. A. M., Gilbert, D. S., Maggs, A. 1981. *Biochem. J.* 199:89–100
77. Kaufmann, E., Weber, K., Geisler, N. 1985. *J. Mol. Biol.* 185:733–42
78. Traub, P., Vorgias, C. 1983. *J. Cell Sci.* 63:43–67
79. Ahmardi, B., Speakman, P. T. 1978. *FEBS Lett.* 94:365–67
80. Woods, E. F., Gruen, L. C. 1981. *Aust. J. Biol. Sci.* 34:515–26
81. Steinert, P. M. 1986. *J. Invest. Dermatol.* 86:508
82. Ward, W. S., Schmidt, W. N., Schmidt, L. A., Hnilica, L. S. 1985. *Biochemistry* 24:4429–34
83. Ip, W., Hartzer, M. K., Pang, Y.-Y. S., Robson, R. M. 1985. *J. Mol. Biol.* 183:365–75
84. Geisler, N., Weber, K. 1982. *EMBO J.* 1:1649–56
85. Soellner, P., Quinlan, R. A., Franke, W. W. 1985. *Proc. Natl. Acad. Sci. USA* 82:7929–33
86. Crewther, W. G., Harrap, B. S. 1967. *J. Biol. Chem.* 242:4310–19
87. Skerrow, D., Matoltsy, M. N., Matoltsy, A. G. 1973. *J. Biol. Chem.* 248:4820–26
88. Steinert, P. M. 1978. *J. Mol. Biol.* 123:49–70
89. Steinert, P. M., Idler, W. W., Goldman, R. D. 1980. *Proc. Natl. Acad. Sci. USA* 77:4534–38
90. Steinert, P. M., Steven, A. C. 1985. *Nature* 316:767
91. Geisler, N., Kaufmann, E., Weber, K. 1985. *J. Mol. Biol.* 182:173–77
92. Potschka, M. 1986. *Biophys. J.* 49:129–30
93. Steven, A. C., Wall, J. S., Hainfeld, J. F., Steinert, P. M. 1982. *Proc. Natl. Acad. Sci. USA* 79:3101–5
94. Steven, A. C., Hainfeld, J. F., Trus, B. L., Wall, J. S., Steinert, P. M. 1983. *J. Biol. Chem.* 258:8323–29
95. Steven, A. C., Hainfeld, J. F., Trus, B. L., Wall, J. S., Steinert, P. M. 1983. *J. Cell Biol.* 97:1939–44
96. Steven, A. C., Trus, B. L., Hainfeld, J. T., Wall, J. S., Steinert, P. M. 1985. *Ann. NY Acad. Sci.* 455:371–80
97. Eichner, R., Rew, P., Engel, A., Aebi,

U. 1985. *Ann. NY Acad. Sci.* 455:381–402
98. Aebi, U., Engel, A., Eichner, R. 1985. *J. Ultrastruct. Res.* 90:323–35
99. Fraser, R. D. B., MacRae, T. P. 1973. *Polymer* 14:61–67
100. Fraser, R. D. B., MacRae, T. P. 1983. *Biosci. Rep.* 3:517–25
101. Ip, W., Heuser, J. E. 1985. *Ann. NY Acad. Sci.* 455:185–99
102. Henderson, D., Geisler, N., Weber, K. 1982. *J. Mol. Biol.* 155:173–76
103. Milam, L., Erickson, H. P. 1982. *J. Cell Biol.* 94:592–96
104. Steven, A. C., Stall, R., Steinert, P. M., Trus, B. L. 1986. In *Electron Microscopy and Alzheimer's Disease,* ed. J. Metuzals, pp. 31–33. San Francisco: San Francisco Press Inc.
105. Steven, A. C., Trus, B. L., Steinert, P. M. 1987. *Proc. 45th Ann. EM Soc.*, pp. 802–5
106. Ip, W. 1986. See Ref. 104, pp. 46–49
107. Steinert, P. M. 1977. In *Biochemistry of Cutaneous Epidermal Differentiation,* ed. M. Seiji, I. A. Bernstein, pp. 444–66. Tokyo: Tokyo Univ. Press
108. Krishnan, N., Kaiserman-Abramof, I. R., Lasek, R. J. 1979. *J. Cell Biol.* 82:323–35
109. Aebi, U., Fowler, W. E., Rew, P., Sun, T.-T. 1983. *J. Cell Biol.* 97:1131–43
110. Sauk, J. T., Krumweide, M., Cocking-Johnson, D., White, J. G. 1984. *J. Cell Biol.* 99:1590–97
111. Aebi, U., Glararis, E. C., Eichner, R. 1986. See Ref. 104, pp. 9–12
112. Lieska, N., Maisel, H., Romero-Herrera, A. E. 1981. *Curr. Eye Res.* 1:339–50
113. Fraser, R. D. B., MacRae, T. P., Suzuki, E. 1976. *J. Mol. Biol.* 108:435–52
114. Renner, W., Franke, W. W., Schmid, E., Geisler, N., Weber, K., Mandelkow, E. 1981. *J. Mol. Biol.* 149: 285–306
115. Steinert, P. M., Rice, R. H., Roop, D. R., Trus, B. L., Steven, A. C. 1983. *Nature* 302:794–800
116. Eagles, P. A. M., Maggs, A. 1985. *Ann. NY Acad. Sci.* 455:779–81
117. Eagles, P. A. M. 1986. See Ref. 104, pp. 6–9
118. Eagles, P. A. M., Hopkins, J. M., Rao, R., Chin, T. K. 1986. See Ref. 104, pp. 42–45
119. Carden, M. J., Eagles, P. A. M. 1983. *Biochem. J.* 215:239–52
120. Hirokawa, N., Glicksman, M. A., Willard, M. B. 1984. *J. Cell Biol.* 98:1523–36
121. Steinert, P. M. 1979. *Biochemistry* 18:5664–69
122. Vidrich, A., Gilmartin, M., Zimmerman, J., Freedberg, I. M. 1982. *J. Cell Biol.* 96:237a
123. Robinson, P. A., Wion, D., Anderton, B. H. 1986. *FEBS Lett.* 209:203–8
124. Gerace, L., Blobel, G. 1980. *Cell* 19:277–87
125. Ottaviano, Y., Gerace, L. 1985. *J. Biol. Chem.* 260:624–32
126. Miake-Lye, R., Kirschner, M. 1985. *Cell* 41:165–75
127. Dessev, G., Iocheva, C., Tasheva, B., Goldman, R. 1987. *Proc. Natl. Acad. Sci. USA.* In press
128. Evans, R. M. 1984. *J. Biol. Chem.* 259:5372–75
129. Zieve, G. W. 1985. *Ann. NY Acad. Sci.* 455:720–23
130. Inagaki, M., Nishi, Y., Nishizawa, K., Matsuyama, M., Sato, C. 1987. *Nature* 328:649–52
131. Krieg, T. M., Schafer, M. P., Cheng, C. K., Filpula, D., Flaherty, P., et al. 1985. *J. Biol. Chem.* 260:5867–70
132. Rueger, M., Jorcano, J. L., Franke, W. W. 1985. *EMBO J.* 4:2261–67
133. Marchuk, D., McCrohon, S., Fuchs, E. 1984. *Cell* 39:491–98
134. RayChaudhury, A., Marchuk, D., Lindhurst, M., Fuchs, E. 1986. *Mol. Cell Biol.* 6:539–48
135. Johnson, L. D., Idler, W. W., Zhou, X.-M., Roop, D. R., Steinert, P. M. 1985. *Proc. Natl. Acad. Sci. USA* 82:1896–900
136. Klinge, E. M., Sylvestre, Y., Freedberg, I. M. 1987. *J. Mol. Evol.* 24:319–29
137. Tyner, A. L., Eichman, M. J., Fuchs, E. 1985. *Proc. Natl. Acad. Sci. USA* 82:4683–87
138. Quax, W., Egberts, W. V., Hendriks, W., Quax-Jeuken, Y., Bloemendal, H. 1983. *Cell* 35:215–23
139. Quax, W., van der Brock, L., Egberts, W. V., Ramaekers, F., Bloemendal, H. 1985. *Cell* 43:327–38
140. Balcarek, J., Cowan, N. J. 1985. *Nucleic Acids Res.* 173:159–76
141. Miyatani, S., Winkles, J. A., Sargent, T. D., Dawid, I. B. 1986. *J. Cell Biol.* 103:1957–65
142. Ferrari, S., Battini, R., Kaczmarek, L., Rittling, S., Calabretta, B., et al. 1986. *Mol. Cell Biol.* 6:3614–20
143. Zehner, Z., Li, Y., Roe, R. A., Paterson, B. M., Sax, C. M. 1987. *J. Biol. Chem.* 262:8112–20
144. Levy, E., Liem, R. K. H., D'Eustachio, P., Cowan, N. J. 1987. *Eur. J. Biochem.* 166:71–77
145. Roop, D. R., Steinert, P. M. 1986. In *Cell and Molecular Biology of the Cy-*

toskeleton, ed. J. W. Shay, pp. 69–83. New York: Plenum
146. Savtchenko, E. S., Freedberg, I. M., Blumenberg, M. 1987. *J. Invest. Dermatol.* 88:516
147. Vasseur, M., Duprey, P., Brulet, P., Jacob, F. 1985. *Proc. Natl. Acad. Sci. USA* 82:1155–59
148. Mischke, D., Wild, G. 1987. *J. Invest. Dermatol.* 88:191–97
149. Yamada, Y., Avvedimento, V. E., Mudryji, M., Ohkubo, H., Vogeli, G., et al. 1980. *Cell* 22:887–92
150. Ohno, S., Matsunaga, T., Wallace, R. B. 1982. *Proc. Natl. Acad. Sci. USA* 77:1999–2002
151. Alexander, F., Young, P. R., Tilghman, S. M. 1984. *J. Mol. Biol.* 173:159–76
152. Nishioda, Y., Leder, A., Leder, P. 1980. *Proc. Natl. Acad. Sci. USA* 77:2806–9
153. Karin, M., Richards, R. I. 1982. *Nature* 299:797–802
154. Lee, M. G. S., Lewis, S. A., Wilde, C. D., Cowan, N. J. 1983. *Cell* 33:477–87
155. Julien, J.-P., Meyer, D., Flavell, D., Harold, J., Grosveld, F. 1986. *Mol. Brain Res.* 1:243–50
156. Bowden, P. E., Blanchard, A., Steinert, P. M. 1987. *J. Invest. Dermatol.* 88: 478
157. Blessing, M., Zentgraf, H., Jorcano, J. L. 1987. *EMBO J.* 6:567–75
158. Quax, W., Khan, P. M., Quax-Jeuken, Y., Bloemendal, H. 1985. *Gene* 38:189–96
159. Moll, R., Franke, W. W., Schiller, D. L., Geiger, B., Krepler, R. 1982. *Cell* 31:11–24
160. Heid, H. W., Werner, E., Franke, W. W. 1986. *Differentiation* 32:101–19
161. Quinlan, R. A., Schiller, D. L., Hatzfeld, M., Achtstatter, T., Moll, R., et al. 1985. *Ann. NY Acad. Sci.* 455:282–306
162. Sun, T.-T., Eichner, R., Schermer, A., Cooper, D., Nelson, W. G., Weiss, R. A. 1984. In *Cancer Cells, The Transformed Phenotype,* 1:169–76. NY: Cold Spring Harbor Lab.
163. Lynch, M. H., O'Guin, W. M., Hardy, C., Mak, L., Sun, T.-T. 1986. *J. Cell Biol.* 103:2593–606
164. Sun, T.-T., Tseng, S. C. G., Huang, A. J.-W., Cooper, D., Schermer, A., et al. 1985. *Ann. NY Acad. Sci.* 455:307–29
165. Cooper, D., Schermer, A., Sun, T.-T. 1985. *Lab. Invest.* 52:234–56
166. Banks-Schlegel, S. P. 1982. *J. Cell Biol.* 93:551–59
167. Schweizer, J., Kinjo, M., Furstenberger, G., Winter, H. 1984. *Cell* 37:159–70
168. Moll, R., Moll, I., Wiest, W. 1982. *Differentiation* 23:170–78
169. Weiss, R. A., Eichner, R., Sun, T.-T. 1984. *J. Cell Biol.* 98:1397–406
170. Nakazawa, H., Mebrel, T., Cheng, C. K., Yuspa, S. H., Roop, D. R. 1987. Submitted for publication
171. Fisher, C., Jones, A., Roop, D. R. 1987. *J. Cell Biol.* 104: In press
172. Franz, J. K., Gall, L., Williams, M. A., Picheral, B., Franke, W. W. 1983. *Proc. Natl. Acad. Sci. USA* 80:6254–58
173. Ellison, T. R., Mathisen, P. M., Miller, L. 1985. *Dev. Biol.* 112:329–37
174. Jonas, E., Sargent, T. D., Dawid, I. B. 1985. *Proc. Natl. Acad. Sci. USA* 82:5413–17
175. Winkles, J. A., Sargent, T. D., Parry, D. A. D., Jonas, E., Dawid, I. B. 1985. *Mol. Cell Biol.* 5:2575–81
176. Sargent, T. D., Dawid, I. B. 1983. *Science* 222:135–39
177. Jamrich, M., Sargent, T. D., Dawid, I. B. 1987. *Genes Dev.* 1:124–32
178. Benevente, R. 1985. *Cell* 41:177–90
179. Benevente, R., Krohne, G. 1985. *Proc. Natl. Acad. Sci. USA* 82:6176–80
180. Stick, R., Hausen, P. 1985. *Cell* 41:191–200
181. Winter, H., Schweizer, J. 1983. *Proc. Natl. Acad. Sci. USA* 80:6480–84
182. Toftgard, R., Yuspa, S. H., Roop, D. R. 1985. *Cancer Res.* 45:5845–50
183. Roop, D. R., Krieg, T. M., Mehrel, T., Cheng, C. K., Yuspa, S. H. 1987 Submitted for publication
184. Tyner, A. L., Fuchs, E. 1986. *J. Cell Biol.* 103:1945–55
185. Knapp, B., Rentrop, M., Schweizer, J., Winter, H. 1987. *J. Biol. Chem.* 262:938–45
186. Doran, T. I., Vidrich, A., Sun, T.-T. 1980. *Cell* 22:17–25
187. Connell, N. D., Rheinwald, J. G. 1983. *Cell* 34:245–53
188. Breitkreutz, D., Bohnert, A., Herzmann, E., Bowden, P. E., Boukamp, P., Fusenig, N. E. 1984. *Differentiation* 26:154–69
189. Asselineau, D., Bernhard, B., Bailly, C., Darmon, M. 1985. *Exp. Cell Res.* 159:536–39
190. Breitkreutz, D., Harnung, J., Pohlmann, J., Broen-Biermann, L., Bohnert, A., et al. 1986. *Eur. J. Cell Biol.* 42:255–67
191. Ben-Ze'ev, A. 1986. *Trends Biochem. Sci.* 11:478–81
192. Rentrop, M., Nischt, R., Knapp, B., Schweizer, J., Winter, H. 1987. *Differentiation.* In press
193. Fuchs, E., Green, H. 1981. *Cell* 25:617–25

194. Eckert, R. L., Green, H. 1984. *Proc. Natl. Acad. Sci. USA* 81:4321–25
195. Tseng, S. C. G., Hatchell, D., Tierney, N., Huang, A. J.-W., Sun, T.-T. 1984. *J. Cell Biol.* 99:2279–86
196. Gilfix, B. M., Eckert, R. L. 1985. *J. Biol. Chem.* 260:14026–29
197. Huang, F. L., Roop, D. R., DeLuca, L. M. 1986. *In Vitro* 22:223–30
198. Kapan, R., Traska, G., Fuchs, E. 1987. *J. Cell Biol.* 105:427–40
199. Schweizer, J., Furstenberger, G., Winter, H. 1987. *J. Invest. Dermatol.* 88: In press
200. Roop, D. R., Toftgard, R., Yuspa, S. H., Kronenberg, M. S., Clark, J. H. 1984. In *The Molecular Biology of the Cytoskeleton*, ed. G. G. Borisy, D. W. Cleveland, D. B. Murphy, pp. 409–14. NY: Cold Spring Harbor Lab.
201. Kronenberg, M. S., Clark, J. H. 1985. *Endocrinology* 117:1469–79
202. Kronenberg, M. S., Clark, J. H. 1985. *Endocrinology* 117:1480–89
203. Laskin, J. D., Mufson, A. R., Piccini, L., Engelhardt, D. L., Weinstein, I. B. 1981. *Cell* 25:441–49
204. Nelson, K. G., Stephenson, K. B., Slaga, T. J. 1982. *Cancer Res.* 42:4164–75
205. Nelson, K. G., Slaga, T. J. 1982. *Cancer Res.* 42:4176–81
206. Schweizer, J., Winter, H. 1982. *Cancer Res.* 42:1517–29
207. Roop, D. R., Huitfeldt, H., Kilkenney, A., Yuspa, S. H. 1987. *Differentiation.* In press
208. Palmiter, R. D., Brinster, R. L. 1986. *Ann. Rev. Genet.* 20:465–91
209. Julien, J.-P., Tretjakoff, I., Beaudet, L., Peterson, A. 1987. *Genes Dev.* In press
210. Goldman, R. D., Goldman, A. E., Green, K., Jones, J. C. R., Lieska, N., Yang, S.-Y. 1985. *Ann. NY Acad. Sci.* 455:1–17
211. Goldman, R. D., Knipe, D. M. 1972. *Cold Spring Harbor Symp. Quant. Biol.* 34:523–34
212. Celis, J. E., Small, J. V., Larsen, P. M., Fey, S. J., DeMay, J., Celis, A. 1984. *Proc. Natl. Acad. Sci. USA* 81:1117–21
213. Knapp, L. W., O'Guin, W. M., Sawyer, R. H. 1983. *J. Cell Biol.* 91:1788–94
214. Klymkowsky, M. W., Miller, R. H., Lane, E. B. 1983. *J. Cell Biol.* 96:494–509
215. Lane, E. B., Klymkowsky, M. W. 1982. *Cold Spring Harbor Symp. Quant. Biol.* 46:387–402
216. Eckert, B. S. 1986. *Cell Motility Cytoskel.* 6:15–24
217. Wang, E., Choppin, P. W. 1981. *Proc. Natl. Acad. Sci. USA* 78:2363–67
218. Thomas, G. P., Welch, W. J., Mathews, M. B., Feramisco, J. R. 1982. *Cold Spring Harbor Symp. Quant. Biol.* 46:985–96
219. Jones, J. C. R., Goldman, A. E., Steinert, P. M., Yuspa, S. H., Goldman, R. D. 1982. *Cell Motility* 2:197–213
220. Jones, J. C. R., Goldman, A. E., Yang, S.-Y., Goldman, R. D. 1985. *J. Cell Biol.* 100:93–102
221. Celis, J. E., Larsen, P. M., Fey, S. J., Celis, A. S. 1983. *J. Cell Biol.* 97:1429–34
222. Franke, W. W., Schmid, E., Grund, C., Geiger, B. 1982. *Cell* 30:103–13
223. Goldman, R. D., Follet, E. A. C. 1970. *Science* 169:286–88
224. Fey, E. G., Capco, D. G., Krochmalnic, G., Penman, S. 1984. *J. Cell Biol.* 99:203s–208s
225. Green, K. J., Goldman, R. D. 1986. *Cell Motility Cytoskel.* 6:389–405
226. Eckert, B. S., Daley, R. A., Parysek, L. M. 1982. *J. Cell Biol.* 92:575–78
227. Ben-Ze'ev, A., Duerr, A., Solomon, F., Penman, S. 1979. *Cell* 17:859–67
228. Burridge, K., Kelly, T., Mangeur, P. 1982. *J. Cell Biol.* 95:478–86
229. Granger, B. L., Lazarides, E. 1982. *Cell* 30:263–75
230. Marchesi, V. T. 1985. *Ann. Rev. Cell Biol.* 1:531–61
231. Hirokawa, N., Tilney, L. B., Fujiwara, K., Heuser, J. E. 1982. *J. Cell Biol.* 94:425–43
232. Hirokawa, N., Cheney, R. E., Willard, M. 1983. *Cell* 32:953–65
233. Granger, B. L., Lazarides, E. 1984. *Mol. Cell Biol.* 4:1943–50
234. Mangeat, P. H., Burridge, K. 1984. *J. Cell Biol.* 98:1363–77
235. Georgatos, S. D., Marchesi, V. T. 1985. *J. Cell Biol.* 100:1955–61
236. Georgatos, S. D., Weaver, D. C., Marchesi, V. T. 1985. *J. Cell Biol.* 100:1962–67
237. Georgatos, S. D., Blobel, G. 1987. *J. Cell Biol.* 105:105–15
238. Georgatos, S. D., Weber, K., Geisler, N., Blobel, G. 1987. *Proc. Natl. Acad. Sci. USA.* In press
239. Skerrow, C. J., Skerrow, D. 1980. In *Cell Adhesion and Motility*, ed. A. G. Curtis, J. D. Pitts, pp. 445–64. Cambridge: Cambridge Univ. Press
240. Cohen, S. M., Gorbsky, G., Steinberg, M. S. 1983. *J. Biol. Chem.* 258:2621–27
241. Cowin, P., Garrod, D. R. 1983. *Nature* 302:148–50

242. Geiger, B., Schmid, E., Franke, W. W. 1983. *Differentiation* 23:189–205
243. Cowin, P., Kapprell, H.-P., Franke, W. W. 1985. *J. Cell Biol.* 101:1442–54
244. Jones, J. C. R., Goldman, R. D. 1985. *J. Cell Biol.* 101:506–17
245. Thornell, L. E., Eriksson, A., Johansson, B., Kjorell, U., Franke, W. W., et al. 1985. *Ann. NY Acad. Sci.* 455:213–40
246. Green, K. J., Geiger, B., Jones, J. C. R., Talian, J. C., Goldman, R. D. 1987. *J. Cell Biol.* 104:1389–402
247. Hennings, H., Michael, D., Cheng, C. K., Steinert, P. M., Holbrook, K. A., Yuspa, S. Y. 1980. *Cell* 19:245–54
248. Hennings, H., Holbrook, K. A. 1983. *Exp. Cell Res.* 143:127–42
249. Watt, F. M., Mattey, D. L., Garrod, D. R. 1984. *J. Cell Biol.* 99:2211–15
250. Kartenbeck, J., Schweicheimer, K., Moll, R., Franke, W. W. 1984. *J. Cell Biol.* 98:1072–81
251. Jones, J. C. R., Yokoo, K. M., Goldman, R. D. 1986. *Cell Motility Cytoskel.* 6:560–69
252. Kapprell, H. P., Cowin, P., Franke, W. W. 1985. *Eur. J. Cell Biol.* 36:217–29
253. Georgatos, S. D., Blobel, G. 1987. *J. Cell Biol.* 105:117–25
254. Fraser, R. D. B., MacRae, T. P., Rogers, G. E. 1972. *Keratins, Their Structure Composition and Biosynthesis.* Springfield, Ill: Thomas
255. Powell, B. C., Rogers, G. E. 1986. In *Biology of the Integument*, ed. J. Bereiter-Hahn, A. G. Matoltsy, K. S. Richards, 2:695–781. Berlin: Springer-Verlag
256. Rogers, G. E. 1985. *Ann. NY Acad. Sci.* 455:403–25
257. Dale, B. A., Holbrook, K. A., Steinert, P. M. 1978. *Nature* 276:729–31
258. Steinert, P. M., Cantieri, J. S., Teller, D. C., Lonsdale-Eccles, J. D., Dale, B. A. 1981. *Proc. Natl. Acad. Sci. USA* 78:4097–101
259. Dale, B. A., Resing, K. A., Lonsdale-Eccles, J. D. 1985. *Ann. NY Acad. Sci.* 455:330–42
260. Rothnagel, J. A., Mehrel, T., Idler, W. W., Roop, D. R., Steinert, P. M. 1987. *J. Biol. Chem.* 262:15643–48
261. Price, M. G., Lazarides, E. 1983. *J. Cell Biol.* 97:1860–74
262. Granger, B. L., Lazarides, E. 1981. *Cell* 22:727–38
263. Rothnagel, J. A., Rogers, G. E. 1986. *J. Cell Biol.* 102:1419–29
264. Williams, R. C., Aamodt, E. J. 1985. *Ann. NY Acad. Sci.* 455:509–24
265. Yang, H.-Y., Lieska, N., Goldman, A. E., Goldman, R. D. 1985. *J. Cell Biol.* 100:620–31
266. Lieska, N., Yang, H.-Y., Goldman, R. D. 1985. *J. Cell Biol.* 101:802–13
267. Wiche, G., Krepler, R., Artlieb, U., Pytela, R., Denk, H. 1983. *J. Cell Biol.* 97:887–901
268. Herman, H., Wiche, G. 1987. *J. Biol. Chem.* 262:1320–25
269. Lawson, D. 1983. *J. Cell Biol.* 97:1891–905
270. Wang, E., Cairncross, J. G., Yung, W. K. A., Garber, E. A., Liem, R. K. H. 1983. *J. Cell Biol.* 97:1507–14
271. Wang, E., Asai, D. J., Lazarides, E. 1980. *Proc. Natl. Acad. Sci. USA* 77:1541–45
272. Wang, E., Gomer, R. H., Lazarides, E. 1981. *Proc. Natl. Acad. Sci. USA* 78:3531–35
273. Napolitano, E. W., Pachter, J. S., Chin, S. S. M., Liem, R. K. H. 1984. *J. Cell Biol.* 101:1323–31
274. Pachter, J. S., Liem, R. K. H. 1984. *J. Cell Biol.* 101:1316–22
275. Lin, J.-J. C., Feramisco, J. R. 1981. *Cell* 24:185–93
276. Osborn, M., Altmannsberger, M., Debus, E., Weber, K. 1985. *Ann. NY Acad. Sci.* 455:649–68
277. Raphael, K. A., Marshall, R. C., Pennycuik, P. R. 1984. *Genet. Res. Cambridge* 44:29–38
278. Marshall, R. C. 1986. See Ref. 255, pp. 722–38
279. Dale, B. A., Gown, A. M., Fleckman, P., Kimball, J. R., Resing, K. A. 1987. *J. Invest. Dermatol.* 88:306–13
280. Holbrook, K. A., Dale, B. A., Smith, L. T., Foster, C. A., Williams, M. L., et al. 1987. In *Current Problems in Dermatology. Prenatal Diagnosis of Heritable Skin Disorders*, ed. T. Gedde-Dahl, K. D. Wuepper, 16:94–108. Basel: Karger