Quarterly Reviews of Biophysics **32**, 2 (1999), pp. 99–187 Printed in the United Kingdom

© 1999 Cambridge University Press

# Intermediate filaments: molecular architecture, assembly, dynamics and polymorphism

David A. D. Parry[1] and Peter M. Steinert[2]

[1] Institute of Fundamental Sciences, Massey University, Private Bag 11–222, Palmerston North, New Zealand
[2] Laboratory of Skin Biology, NIAMS, National Institutes of Health, Building 6 Room 425, Bethesda, MD 20892–2752, USA

**1. Introduction 100**

**2. Molecular architecture 107**
2.1 Primary structure 108
2.1.1 Homologous regions 109
2.1.2 Chain typing 115
2.1.3 Post-translational modifications 117
2.2 Secondary structure 118
2.2.1 Central rod domain 118
2.2.2 Head and tail domains 119
2.3 Tertiary structure 123
2.3.1 Coiled-coil rod domain 123
2.3.1.1 Specificity through salt bridges 124
2.3.1.2 Specificity through apolar interactions 127
2.3.1.3 A consensus trigger sequence for two-stranded coiled-coils 128
2.3.2 Discontinuities in the rod domain 128
2.3.2.1 Links 129
2.3.2.2 Stutter 131
2.3.3 Head and tail domains 131
2.4 Electron microscope observations 133

**3. Assembly 136**
3.1 Role of the coiled-coil rod domain 137
3.2 Role of end domains 141
3.3 Experimentally induced crosslinks and modes of assembly 145
3.4 Naturally occurring crosslinks for tissue coordination 154
3.5 STEM data 154

**4. Quaternary structure 160**
4.1 Protofilaments and protofibrils 160
4.2 Head and tail domains 163
4.3 Surface lattice structure 164
4.4 Crystal studies on intermediate filament fragments 168

https://doi.org/10.1017/S0033583500003516 Published online by Cambridge University Press

**5. Polymorphism 169**
5.1 Variations on a theme 170
5.1.1 Axial structure 170
5.1.2 Lateral structure 171
**6. Keratin intermediate filament chains in diseases 172**
**7. Concluding remarks 175**
**8. Acknowledgments 176**
**9. References 176**

## 1. Introduction

Three types of intracellular filament have been identified in eukaryotic cells, and together they constitute the key elements of the cytoskeleton. They are the microtubules, the actin-containing microfilaments and the intermediate filaments. The uniqueness of the former two types of filament in cells has been well known for a long time but, in contrast, the intermediate filaments have been a relative new-comer to the scene. The microtubules and the microfilaments have always been easy to distinguish from one another on the grounds of their respective sizes (microtubules are about 25 nm in diameter and microfilaments are about 7–10 nm in diameter). Additionally, microtubules were capable of being disaggregated by the action of colchicine, and microfilaments could be disassembled by other drugs or be decorated with heavy meromyosin to generate arrowhead-like structures. Importantly, in both microtubules and microfilaments the constituent protein subunits were arranged to give the filaments a directionality, and the ability of these filaments to function *in vivo* depended crucially on this feature of their structure. Microtubules, for example, are involved in mitosis, motility and transport within the cell: each of these functions is clearly a 'directional' one. With this background the discovery and characterization of the intermediate filaments can begin.

Curiously, in retrospect, intermediate filaments had been observed in the cytoplasm and nuclei of cells from higher animals for about a century. For example, early cytologists described in fixed-stained tissues 'fibrillar netlike structures' surrounding the nucleus in the cell body of neurons which extended into axons, and elaborate arrays of 'tonofibrillae' in columnar and stratified squamous epithelia that appeared to traverse cell membranes (Wilson, 1928). However, it was only over the past 25 years or so that these were clearly distinguished from the microtubules and the microfilaments on the basis of rigorous biochemical analyses and described as a unique filament. Over the past quarter century there have been intense efforts to characterize the intermediate filaments and, in particular, to elucidate their detailed structure and to understand their modes of molecular assembly. The first intermediate filaments studied at the molecular level were those from wool and quill hard $\alpha$-keratin. Indeed, this study preceded that done on all of the other intermediate filaments by an additional 40 years though at the earlier stages of this work these filaments were believed to be distinct structures unrelated to any of the other intracellular filaments visualized by cell biologists. In the hard $\alpha$-keratins the intermediate filaments are embedded in a matrix of proteins rich in cysteine residues and also in glycine and tyrosine residues (Fig. 1*a*). The filaments were originally termed 'microfibrils'. Their diameters (about 7–8 nm) were

https://doi.org/10.1017/S0033583500003516 Published online by Cambridge University Press

measured by electron microscopy, and naturally ordered bundles of them were studied by fibre X-ray diffraction. Physical studies proved to be informative, thanks mainly to the naturally high orientation of the specimens, but the protein–chemical analysis of the constituent chains proved to be a much more difficult task. Numerous covalent disulphide bonds were present within and between the filament–matrix complex that constituted the hard $\alpha$-keratin.

The degree of chain heterogeneity that existed was an additional problem facing the protein biochemists. It was only in the late 1970s and the early 1980s that it was realized (and subsequently confirmed by sequence analysis) that the microfibrils in hard $\alpha$-keratin were members of the same family of intracellular filaments that had just been observed in various other cell types. In 1978 cell biologists termed this newly appreciated class of filaments 'intermediate filaments' because their diameters (about 10 nm) appeared to be intermediate in size between those of the microtubules and the microfilaments. This loose definition subsequently proved to be correct as more recently it has been determined that the diameter of microfilaments deduced from crystal structure studies on actin are about 10 nm, and intermediate filaments are up to 16 nm in diameter based on scanning transmission electron microscope measurements. Thus, the intermediate filaments seen initially in epithelial cells and neurons in the late 1800s and early 1900s now have been subsequently observed in the cytoplasm of virtually all other types of eukaryotic cell (see, for example, Parry & Steinert, 1995). A large number of intermediate filament chains, several hundred in total, have been sequenced and characterized. In so doing it has been established that these chains fall into a number of distinct 'types'. Type I and II chains – the keratins – occur in epithelial cells. Type Ia and IIa chains occur in the 'hard' keratinizing epithelia of hairs, nails etc., and are now commonly termed trichocyte keratins. Type Ib and IIb chains occur in other epithelia and are termed cytokeratins. Type III chains – desmin, vimentin, glial fibrillary acidic protein and peripherin – occur respectively in muscle cells, mesenchymal cells, astroglia and neurons. The Type IV chains – $\alpha$-internexin, NF-L, NF-M and NF-H – all occur in neurons. The Type V chains – the nuclear lamins – occur, as their name suggests, in the cell nucleus. Type VI chains – which include nestin and possibly such other proteins as synemin and paranemin – are also present in a variety of specialized cell types. Notably, because of their unusual features (see below), it seems possible that additional Type VI members may be identified by or at the time of the completion of the human genome project. Another common feature is that most if not all intermediate filament protein chains are capable of reconstitution either singly or together with another compatible chain into filaments *in vitro* that are indistinguishable from native intermediate filaments as judged by their respective appearances in the electron microscope (Fig. 1*b*). Although the optimal conditions for re-assembly vary somewhat, no co-factors are required, which has thereby rendered intermediate filaments particularly simple to study biochemically.

The intermediate filaments, as a general class of structures, are not cellular entities confined to the vertebrates: they are also found widely in invertebrates. Weber and colleagues have made an especial study of this latter group (see, for example, Karabinos *et al*. 1998; Riemer & Weber, 1998). In so doing, they have observed functional analogues to some of the cytoplasmic intermediate filaments but certainly not to all of them. A particular feature of the sequences exhibited by some of the intermediate filament chains from the invertebrates is that they contain a lamin-like 42 residue insert in the central region of the chains. This region is highly conserved in both length and character in all of the vertebrate chains. Examples of

https://doi.org/10.1017/S0033583500003516 Published online by Cambridge University Press

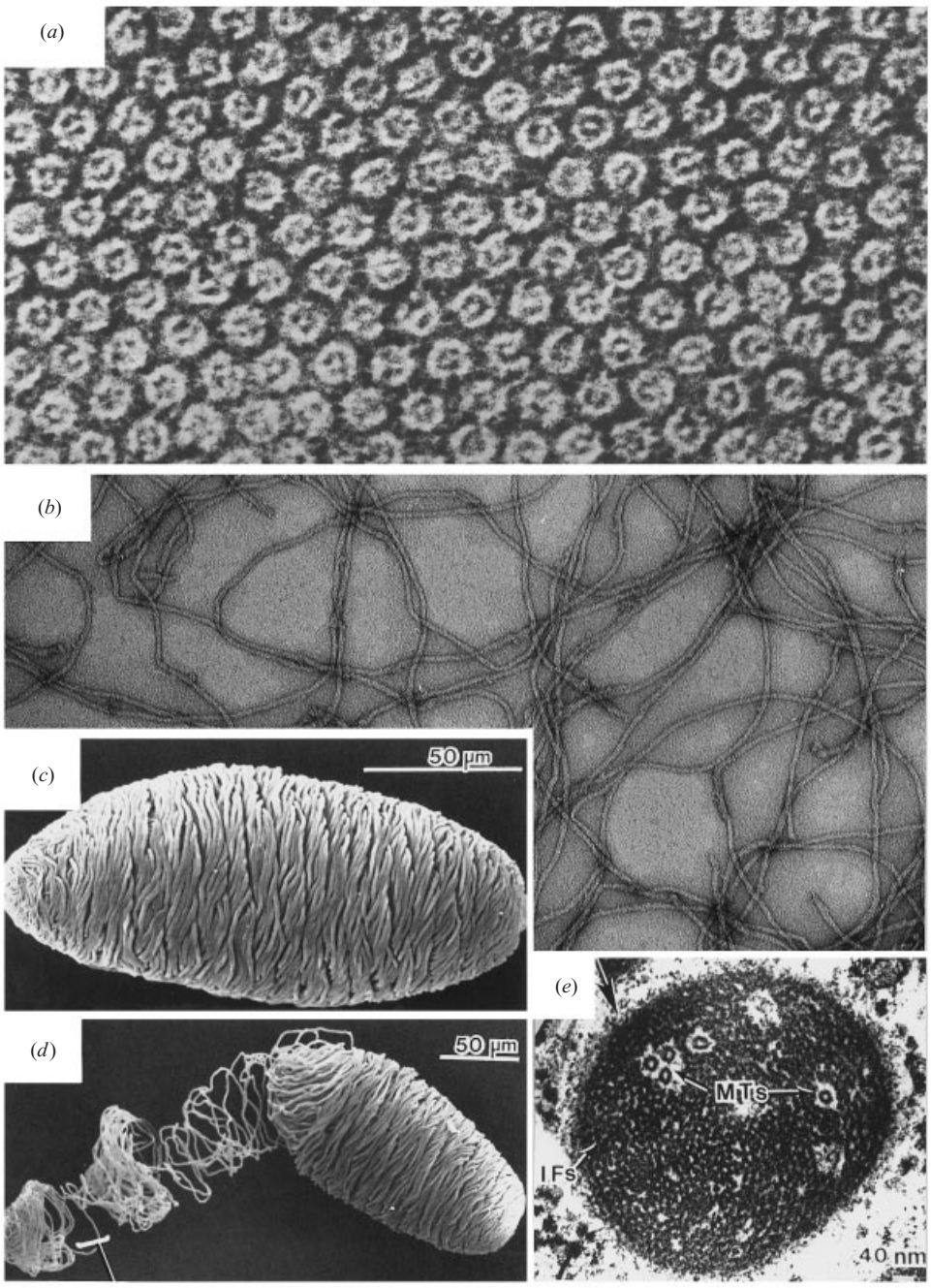



**Fig. 1.** (*a*) Electron micrograph showing a transverse section of hard α-keratin from wool. The trichocyte intermediate filaments, which are about 8 nm in diameter, are embedded in a matrix of sulphur-rich proteins. The filaments are approximately hexagonally packed in the paracortex of wool. (*b*) Reconstituted vimentin intermediate filaments. These generally follow a sinuous path and indicate that the intermediate filaments are flexible structures. (*c*) Scanning electron micrograph of a single coiled thread within the cytoplasm of the gland thread cell of the Pacific hagfish *Eptatretus stouti*. The plasma membrane was removed during the isolation of the gland thread cell from the gland. (*d*) The single thread, more than 0·5 m in overall length, has become partially unwound. A cross-section of the thread,

https://doi.org/10.1017/S0033583500003516 Published online by Cambridge University Press

invertebrates with the sequence insertion include the nematodes and the molluscs (Weber *et al*. 1988, 1989; Dodemont *et al*. 1990, 1994; Szaro *et al*., 1991; Way *et al*. 1992; Tomarev *et al*. 1993) and the annelid *Lumbricus terrestris* (Bovenschulte *et al*. 1995). In contrast, the intermediate filament proteins from the cephalocordate *Branchiostoma* (Riemer *et al*. 1998) and the urochordate *Styela plicata* (Riemer & Weber, 1998) lack the insert and thus have greater similarity to the intermediate chains from the vertebrates than to those from the invertebrates; in this way, they represent an evolutionary advance.

While intermediate filaments are often thought of as intracellular structures there is at least one recorded example where this is not the case. Intermediate filament proteins are synthesized by the gland thread cell of the hagfish (*Eptatretus stouti*) and released into an extracellular environment where they interact with mucins and sea water (Fig. 1*c–e*). The result is a sticky gelatinous mass that provides a defence mechanism for the hagfish when under stress. The two constituent proteins, known as $\alpha$ and $\gamma$ chains, have been characterized (Koch *et al*. 1994, 1995) and the intermediate filament classification confirmed. As will be shown later, however, there are again key differences between the hagfish intermediate filament proteins and those that are intracellular. This has complicated their classification into a particular sequence type. It is interesting, nonetheless, that intermediate filament proteins do have both intra- and extracellular distributions, that they can occur in both the cytoplasm and in the nucleus, and that both vertebrates and invertebrates contain this important class of structure.

For many years the roles played by intermediate filaments *in vivo* were largely undetermined. Recently, however, data have become available to allow some of these roles to be better understood. The functions of various domains in the intermediate filaments will be discussed later in this review, but it is sufficient at this point to note that all of the cytoplasmic intermediate filaments form a flexible, fairly dense, highly elaborate, and exquisitely controlled three-dimensional network that generally courses between the nucleus and the cell membrane. The intermediate filaments impact the surface of the nucleus at or near the nuclear pores, implying a direct or indirect link with the nuclear lamina and the business of the nucleus. They also impact the cell periphery at a variety of specialized junctions including gap junctions, tight junctions or desmosomes or, in the specialized case of stratified squamous epithelia, there is a direct link between the keratin intermediate filaments and the cell envelope (Meng *et al*. 1997; Steinert & Marekov, 1997; Candi *et al*. 1998; Steinert *et al*. 1998a, b). In addition, a variety of intermediate filament associated proteins (IFAPs), such as plectin (Liu *et al*. 1996; Svitkina *et al*. 1996), desmoplakin (Kouklis *et al*. 1994), BPAG1n (Yang *et al*. 1996), and presumably many others awaiting characterization, are involved in linking the filaments to each other, to actin or to microtubules. The network thus formed has the general capacity not only to provide form to the cell but also to enable the cell to withstand external stresses and strains. As discussed below, the consequences of abnormal filament proteins that arise in genetic diseases generally result in compromised cells leading to gross tissue disruption and disease. Moreover, intermediate filament networks are not static assemblies and, indeed, they display very considerable movement *in vivo*. A large body of literature has demonstrated that the organization and deployment of both the cytoplasmic intermediate

---

about 0·65 mm in diameter, indicated by the bracket in (*d*) and by the arrow in (*e*) illustrates the near parallel packing of almost 1000 intermediate filaments and about 10 microtubules.

https://doi.org/10.1017/S0033583500003516 Published online by Cambridge University Press

filaments and the nuclear lamina are grossly rearranged during mitosis. Likewise, the filaments are engaged in widespread movement in interphase cells as well. Yoon *et al*. (1998) and Ho *et al*. (1998), for example, have tagged vimentin with green fluorescent protein and demonstrated that vimentin is readily incorporated into the endogenous intermediate filament network. Furthermore, using time-lapse measurements they showed that the bundles of intermediate filaments – sometimes termed fibrils – were constantly moving. Photobleaching too revealed that the original bleached zone became translocated into several pieces, thus showing that portions of the fibrils could move in different directions. These processes are all energy-dependent and are mediated by the interactions between intermediate filaments and microtubules as well as between intermediate filaments and microfilaments. Prahlad *et al*. (1998) were also able to demonstrate a microtubule-dependent process whereby intermediate filament precursors – termed vimentin dots – moved to the peripheral regions of the cytoplasm and assembled into short fibrils. In fibroblasts, for example, kinesin-like proteins and nestin are known to be involved in these processes.

Thus with the probable exception of the trichocyte keratins, intermediate filaments are clearly not passive mechanical integrators of intracellular space, as was once thought. They are, in fact, dynamic structures capable of assembling and disassembling during normal cell cycles and, consequently, making and breaking interactions with other cellular components. Much of the dynamic behaviour of intermediate filaments seems to be under the control of kinases and phosphatases and therefore may be related to the state of phosphorylation of the intermediate filament chains. In the case of the lamin intermediate filament chains, however, it seems to be the state of isoprenylation that controls the state of organization of the filaments and their interactions with other organelles. Generally, the cytoplasmic intermediate filaments are approximately radially distributed in the cell and so link the cell membrane to the cell nucleus. Their interactions with intermediate filament-associated proteins such as desmoplakin in the desmosomes (Fig. 2; Green *et al*. 1990; 1992a, b), bullous pemphigoid antigen in the hemidesmosome (Tanaka *et al*. 1991), envoplakin and periplakin of cornified envelopes and desmosomes (Ruhrberg *et al*. 1997) and plectin throughout cytoplasmic space (Wiche *et al*. 1991), as well as with microtubules and microfilaments, govern their function. Although there are relatively few data on the mechanical properties of intermediate filaments, Leterrier *et al*. (1996) have shown that *in vitro* preparations of neurofilaments in the presence of millimolar amounts of $Mg^{2+}$ produce gels that slowly give rise to a shear modulus comparable to that seen for vimentin intermediate filaments. This was followed by a rapid rise in shear modulus that was thought to have arisen as a result of cross-bridge links formed between the neurofilaments. These cross-bridges are believed to be formed from the tail domains of the larger molecular weight neurofilament chains known as NF-M and NF-H, and determine the separation of the neurofilaments *in vivo* (see, for example, Nakagawa *et al*. 1995; Parry & Steinert, 1995). As Leterrier *et al* (1996) point out 'Dynamic and regulated interactions occur between neurofilaments to form viscoelastic networks with unique properties compared to the cytoskeletal structures'. More recent work on Type VI chains suggest a similar biomechanical role for their very long tail domains (Steinert *et al*., 1999a). Mechanically, the nuclear lamins, which are an integral part of the karyoskeleton and which are arranged tangentially in a basket weave-like structure along the inner surface of the nuclear membrane, are quite different from the cytoskeletal networks (Fig. 3; Aebi *et al*. 1986). Finally, one particular feature that distinguishes intermediate filaments in general from the other intracellular filaments is that the intermediate filaments do not have a polarity in the

https://doi.org/10.1017/S0033583500003516 Published online by Cambridge University Press

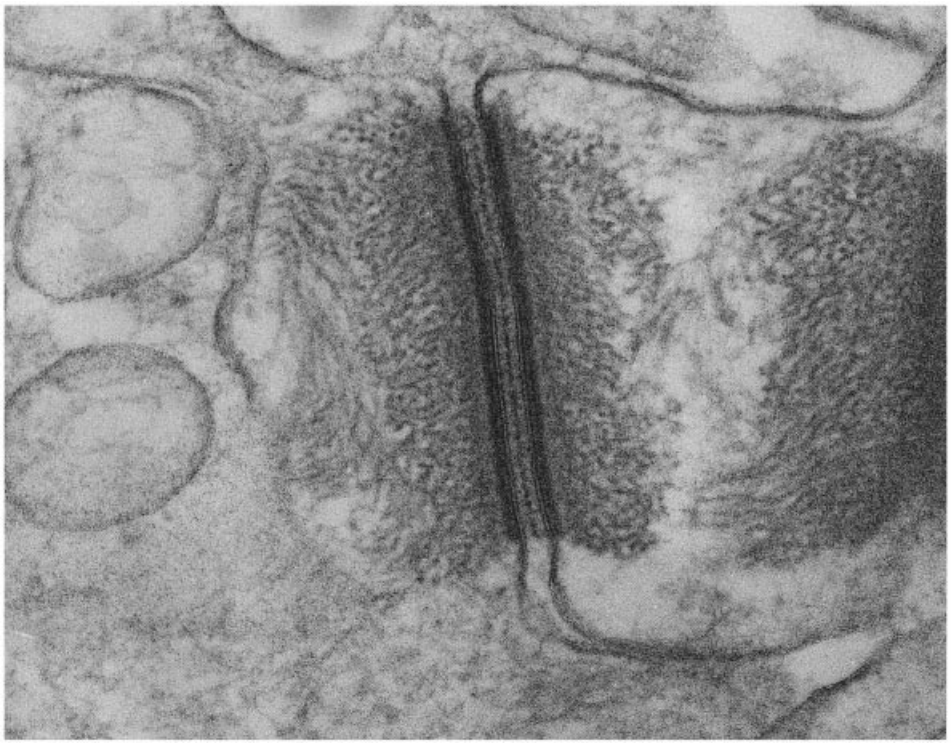


**Fig. 2.** Electron micrograph of a desmosome from bovine tongue epithelium. Within the cytoplasm the desmosome consists of an inner and outer plaque. The former includes (amongst other molecules) the rod and C-terminal domains of the desmoplakin molecules: this is the region to which the intermediate filaments are attached. The inner plaque contains (again amongst other molecules) plakoglobin and the N-terminal domain of desmoplakin. It also contains part of the transmembrane molecules known as the desmogleins and desmocollins. These molecules interact along a central line between the cells and give rise to a central dense stratum that provides the direct linkage between the cells. Figure from Kowalczyk *et al*. (1994).

same way as that exhibited by both the microtubules and the microfilaments. In this case the constituent molecules are arranged in an antiparallel manner. This observation is functionally significant.

Results from various gene knock-out experiments (which result in the non-expression of specific intermediate filament chains) have given a fascinating insight into the generic mechanical role played by the intermediate filaments as a whole. These results have been neatly summarized by Galou *et al*. (1997): knock-out of the gene encoding keratin K8 is lethal in the embryo and results in hepatic or intestinal lesions; knock-outs of the genes encoding the keratin K14 and K10 chains result in the rupture of stratified epithelia; and knock-outs of selected sheep trichocyte keratin genes generally result in matted wool (Bawden *et al*. 1998). In the case of the genes encoding Type III intermediate filament chains, their knock-out effects are as follows: loss of desmin results in the collapse of blood cell walls and the rupture of skeletal and cardiac muscle; loss of vimentin causes degeneration of the cerebellar Purkinje cells; and loss of glial fibrillary acidic protein leads to a loss of cerebral white matter. In all of these cases there is evidence that a compensating over-production of another

https://doi.org/10.1017/S0033583500003516 Published online by Cambridge University Press

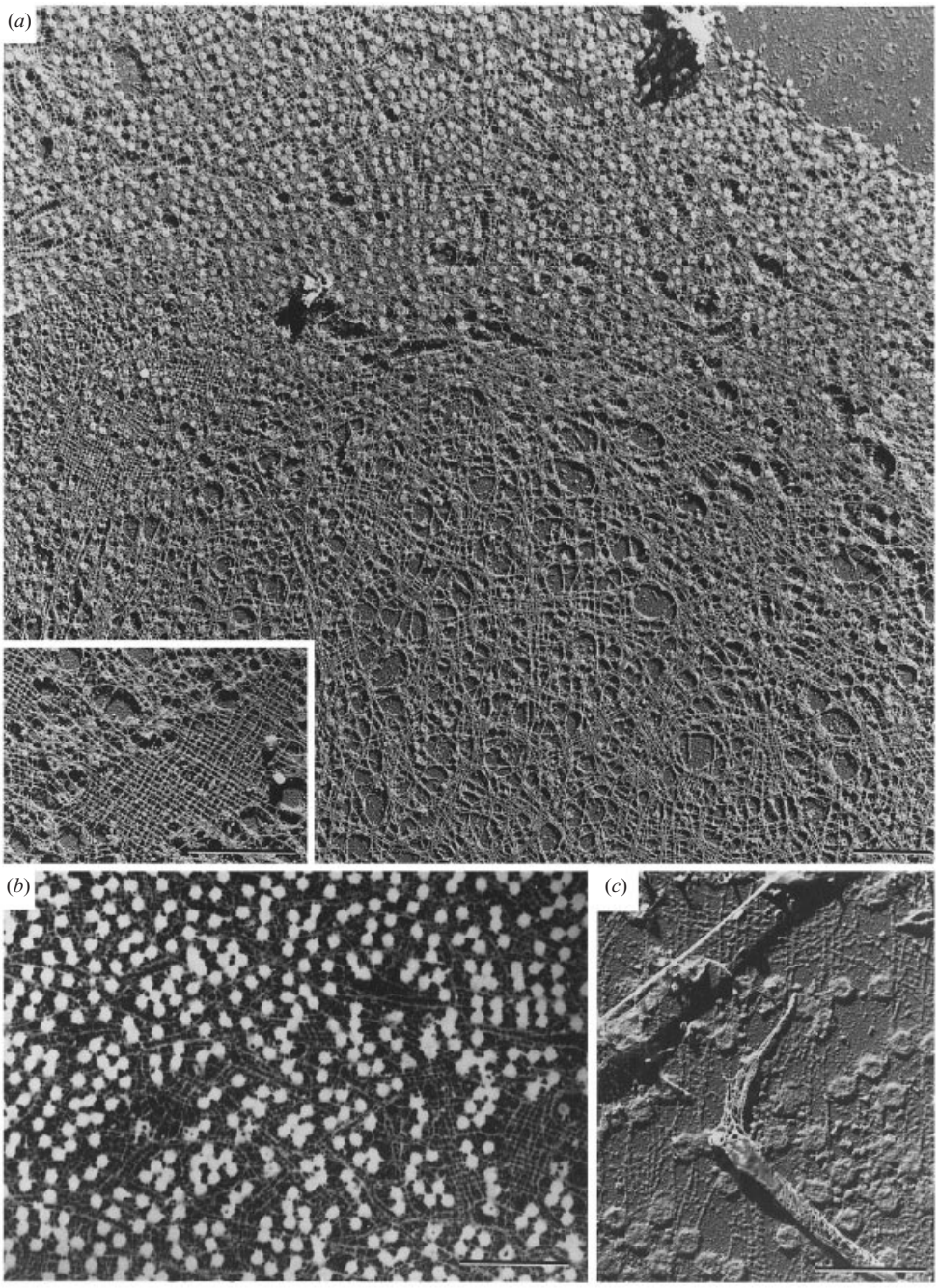



**Fig. 3.** Electron micrographs of the nuclear lamina in *Xenopus* oocytes. (*a*) Freeze-dried and metal-shadowed nuclear envelope extracted with Triton X-100. This shows the nuclear lamina covered with arrays of nuclear pore complexes. The inset reveals a highly regular portion of the network from which the pore complexes have been removed mechanically. This shows clearly the underlying meshwork of lamin filaments arranged approximately on a square lattice. (*b*) Same specimen as in (*a*) but negatively stained and imaged in dark-field mode. (*c*) Similar specimen as in (*a*) but extensively washed with low-salt buffer before freeze-drying and metal-shadowing. This reveals a 25 nm beading along the filaments. Bars are 1 $\mu$m in (*a*) and inset, as well as in (*b*). Bar in (*c*) is 0·5 $\mu$m. Figure courtesy of U. Aebi. Original in Aebi *et al*. (1986).

https://doi.org/10.1017/S0033583500003516 Published online by Cambridge University Press

intermediate filament chain does not occur. The phenotype varies considerably from severe in some cases to quite minor in others but, in general, the effects do not appear to be lethal (Chan *et al*. 1994; Colucci Guyon *et al*. 1994; Rugg *et al*. 1994; Pekny *et al*. 1995; Galou *et al*. 1996; Milner *et al*. 1996; Capetanaki *et al*. 1997). Likewise, loss by mutation or experiment of the gene for the intermediate filament associated protein plectin also resulted in a severe epithelial and muscle cell phenotype (Smith *et al*. 1996), thus indicating that such integrators of intermediate filament function are also important. It does seem clear that those cells that lack intermediate filaments are mechanically fragile and unable to withstand normal stresses and strains. Consequently, the cells tend to degenerate rapidly. It is interesting nonetheless that many of the animals that lacked an intermediate filament gene not only survived through to maturity but were able to reproduce too: however, this may be a reflection of the specialized environment of a housed animal and may not reflect their true fate in the wild. Interestingly, a null mutation in the murine NF-M chain (Elder *et al*. 1998) gives rise to marked effects on the production of both NF-L (greatly decreased levels) and NF-H chains (increased levels). Axon calibre is reduced from normal and each axon contains fewer neurofilaments but more microtubules. Curiously, however, there is no obvious phenotype. It would seem that NF-M regulates the level of NF-L expression and that the NF-M chains are a necessary prerequisite if maximal axonal diameter is to be achieved.

The scope of this review will be largely confined to a discussion of the structural hierarchy in the intermediate filament at the level of the individual chain right through to the intact functional filament. It will also deal with the various modes of assembly that have been shown to exist, and to the polymorphisms that are a prominent feature of intermediate filaments. A number of excellent reviews or books have been published in this general area. The reader is therefore recommended to contributions from Steinert & Parry (1985), Steinert & Roop (1988), Fraser *et al*. (1990), Parry (1990), Fuchs & Weber (1994), Quinlan *et al*. (1994), Parry & Steinert (1995), Parry (1997) and Herrmann & Aebi (1998).

## 2. Molecular architecture

Knowledge of the amino acid sequence of a protein does not necessarily allow its secondary and tertiary conformations to be deduced with any high degree of accuracy. This is most certainly true for globular proteins, where regions of secondary structure are generally short and where the ends of these structures are consequently less easily determined. In contrast, fibrous proteins are very commonly characterized by long regions with a common secondary structure and an underlying sequence repeat. Consequently, these proteins lend themselves more easily to investigation and subsequent characterization. As a result, the models that are produced have a much greater degree of confidence associated with them than would be the case with globular proteins. Indeed, the models proposed for a wide range of fibrous proteins have, over the years, often proved to be remarkably accurate. In the case of the intermediate filament proteins it will be seen that a major portion of all of the sequences is directly interpretable at both the secondary and tertiary levels. Some of the other regions also have unique features that are interpretable, albeit with somewhat lesser degrees of certainty. Such analyses have real value at the present time where attempts to crystallize intermediate filament chain fragments or prepare solutions suitable for NMR structural studies have proved to be

https://doi.org/10.1017/S0033583500003516 Published online by Cambridge University Press

very difficult in spite of strenuous efforts by many in the field over a considerable period of time. However, some crystal data are now becoming available and these will be discussed in Section 4.4.

## 2.1 Primary structure

Determination of the primary structure of trichocyte keratins, which was the first of the IF proteins to be tackled, proved to be a difficult technical exercise. Firstly, native intermediate filaments ('microfibrils') were embedded in a matrix of proteins rich in cysteine residues, which become oxidized *in vivo* to form a network of largely intramolecular disulphide bonds. Considerable sequence regularity occurred in some of these 'high-sulphur proteins'; this was based on a penta- and a deca-peptide repeat, especially in the distribution of the cysteine residues (Parry *et al*. 1979). Regularity of conformation seems to be a feature of many matrix proteins. Isolation of the intermediate filament proteins, however, required a reduction step prior to the use of classical protein separation methods. The intermediate filament proteins isolated were themselves members of two closely related but different families, though this was not appreciated fully at the time. Protein sequencing methods in the 1960s and 1970s were time-consuming and relied on 'clean' preparations, which were difficult to achieve with the trichocyte keratins: we now know that in human there are about ten Type Ia and ten Type IIa genes and keratin proteins (Langbein *et al*. 1999). Disappointingly (at that time) the first (partial) sequence deduced from sheep wool keratin intermediate filaments, although about 60 residues long (O'Donnell, 1969), showed little of immediate interest as the data did not lend themselves to a simple structural interpretation. However, in the 1970s two other sequences, both from sheep wool keratins were characterized. Each was about 100 residues long. They were known as Type I and Type II fragments (Crewther *et al*. 1978; Gough *et al*. 1978). In contrast to the sequence determined earlier these results were very informative as each was characterized by a heptad substructure indicative of an $\alpha$-helical structure that would, in turn, form a coiled-coil conformation (Fig. 4; see also, for example, the analysis of Dowling *et al*. (1986) on the entire sequence from which the Type I fragment was a part). It would be shown in due course that these latter fragments originated from the so-called rod domain (and Segment 1B in particular; see Section 2.2.1) and that the original sequence of O'Donnell corresponded to the N-terminal domain of one of these intermediate filament chains. Very few intermediate filament proteins were subsequently sequenced using classical protein sequencing techniques, as cDNA methods had started to come into prominence in the early 1980s. Indeed, a very large number of intermediate filament proteins originating from a wide diversity of cell types have now been sequenced using molecular biological methodology. Sequence comparisons have confirmed that all intermediate filament proteins have common features, thus linking them absolutely to a common family (Geisler *et al*. 1982a, b; Dowling *et al*. 1983). Listing the references to the 200–300 sequences of intermediate filament protein chains that have been determined would have little value here but it is worth recording the pioneering contributions in this field by Crewther *et al*. (1978), Gough *et al*. (1978), Hanukoglu & Fuchs (1982, 1983), Geisler *et al*. (1982a, b), Quax-Jeuken *et al*. (1983) and Steinert *et al*. (1983, 1984, 1985a, b).

In the following section the common features that have been observed in the sequences will be described (Section 2.1.1). It will then be shown that the sequences of the intermediate

https://doi.org/10.1017/S0033583500003516 Published online by Cambridge University Press

```
I                                                                 V C P N |
II                                                          Q N R Q C C E S N L |

          a b c d e f g a b c d e f g a b c d e f g a b c d e f g
I         Y Q S Y F R T I E E L Q Q K I L C A K S E N A R L V V Q
II        E P L F E G Y I E T L R R E A E C V E A D S G R L S S E

          a b c d e f g a b c d e f g a b c d e f g a b c d e f g
I         I D N A K L A A D D F R T K Y E T E L G L R Q L V E S D
II        L N H V Q E V L E G Y K K K Y E Q E V A L R A T A E N E

          a b c d e f g a b c d e f g a b c d e f g a b c d e f g
I         I D G L R R I L D E L T L C K S D L E A Q V E S L K E E
II        F V A L K K D V D C A Y V R K S D L E A N S E A L I Q E

          a b c d e f g a b c d e f g a b c
I         L I C L K S N H E E E V N T L     |
II        I D F L R R L Y Q E E I R V L     |
```

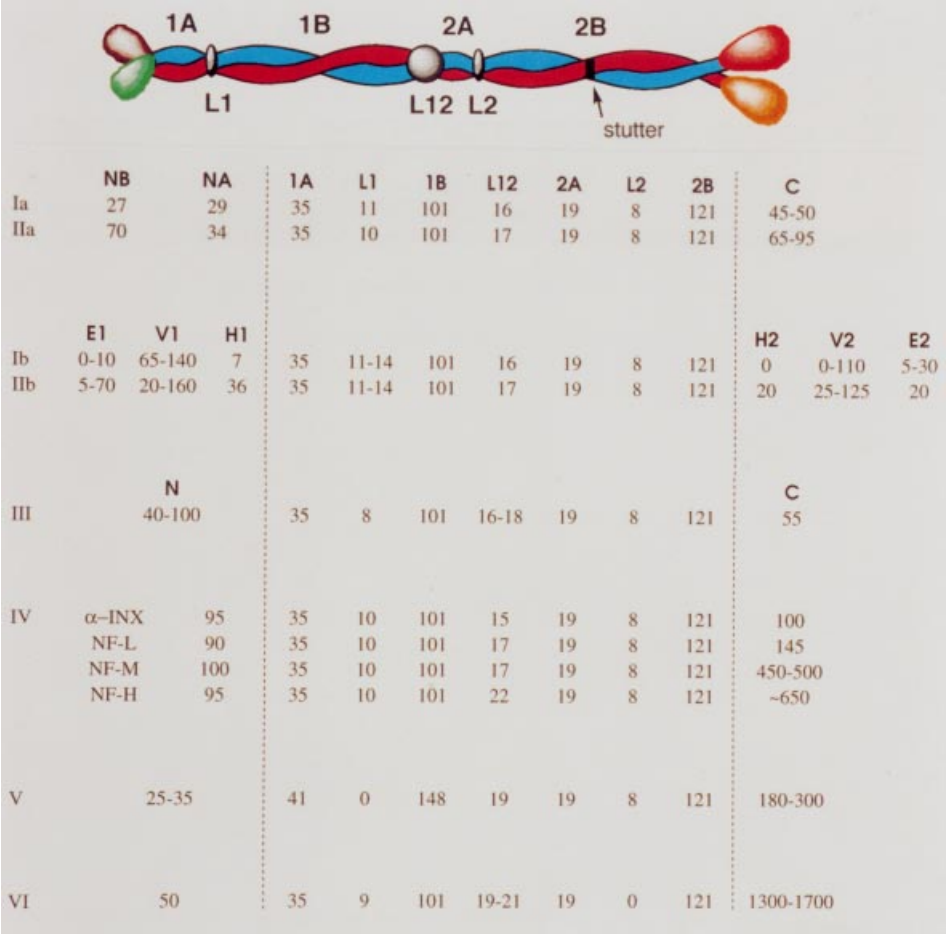
**Fig. 4.** The first sequences of trichocyte keratin fragments that showed a heptad substructure of the form (*a-b-c-d-e-f-g*) (and which could thus be attributed to a coiled-coil rod domain) were each about 100 residues long. They originated from the 8c-1 and 7c chains from sheep wool and were designated as Type I and Type II fragments respectively. Although the sequences were similar to one another they did show that significant differences were also present. At the time that this work was done in the late 1970s the four major coiled-coil segments (1A, 1B, 2A, 2B) had not been recognized so it was not clear what portion of the coiled-coil domain these sequences originated from. Subsequently, however, it proved that the sequences came from link L1 and the majority of segment 1B. In this figure the sequences have been updated to reflect any minor changes in sequence that have occurred since the initial publications of Crewther *et al*. (1978) and Gough *et al*. (1978).

filament proteins naturally group themselves into a small number of 'types' (Section 2.1.2). Such ordering is apparent not only from the amino acid sequences themselves (which was the original method used for the classification) but also from the exon/intron structure of the genes subsequently determined.

### 2.1.1 Homologous regions

One feature that is common to all intermediate filament proteins (except for some putative members of the Type VI class) is the presence of a region about 310–315 residues long which has a heptad substructure over the majority of its length. However, the nuclear lamins and invertebrate intermediate filament proteins have an additional six heptads thus generating a heptad-containing region about 355 residues in length instead. In all cases, however, the heptad has the form $(a\text{-}b\text{-}c\text{-}d\text{-}e\text{-}f\text{-}g)_n$, where *a* and *d* are occupied (to the extent of about 75%) by apolar residues such as leucine, valine and isoleucine. Furthermore, the *e* and *g* positions are often filled by charged residues. The regions of sequence that are N- and C-terminal to the heptad structure, however, are much more diverse, and large variations in size and chemical character have been observed between the various intermediate filament proteins. Nonetheless, it is possible to show (Section 2.1.2) that all intermediate filament proteins do fall into a small number of discrete groupings. Initially, characteristic differences in the 'central' heptad-containing region were sufficient to delineate these groupings (Fig. 5; see

https://doi.org/10.1017/S0033583500003516 Published online by Cambridge University Press




| | NB | NA | 1A | L1 | 1B | L12 | 2A | L2 | 2B | C | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Ia | 27 | 29 | 35 | 11 | 101 | 16 | 19 | 8 | 121 | 45-50 | | |
| IIa | 70 | 34 | 35 | 10 | 101 | 17 | 19 | 8 | 121 | 65-95 | | |
| | **E1** | **V1** | **H1** | | | | | | | **H2** | **V2** | **E2** |
| Ib | 0-10 | 65-140 | 7 | 35 | 11-14 | 101 | 16 | 19 | 8 | 121 | 0 | 0-110 | 5-30 |
| IIb | 5-70 | 20-160 | 36 | 35 | 11-14 | 101 | 17 | 19 | 8 | 121 | 20 | 25-125 | 20 |
| | **N** | | | | | | | | | **C** | | |
| III | 40-100 | | 35 | 8 | 101 | 16-18 | 19 | 8 | 121 | 55 | | |
| IV | α–INX | 95 | 35 | 10 | 101 | 15 | 19 | 8 | 121 | 100 | | |
| | NF-L | 90 | 35 | 10 | 101 | 17 | 19 | 8 | 121 | 145 | | |
| | NF-M | 100 | 35 | 10 | 101 | 17 | 19 | 8 | 121 | 450-500 | | |
| | NF-H | 95 | 35 | 10 | 101 | 22 | 19 | 8 | 121 | ~650 | | |
| V | 25-35 | | 41 | 0 | 148 | 19 | 19 | 8 | 121 | 180-300 | | |
| VI | 50 | | 35 | 9 | 101 | 19-21 | 19 | 0 | 121 | 1300-1700 | | |

**Fig. 5.** Intermediate filament molecules have a central two-stranded coiled-coil structure that is about 45 nm long. The chains are parallel to one another and in axial register. Some intermediate filament molecules, such as the keratins, display a heterodimer structure (the different chains are coloured red and blue). The coiled-coil segments (1A, 1B, 2A, 2B) are precisely maintained in length though the short linking segments (L1, L12 and L2; coloured grey) show some variation with filament type (see below). In the centre of segment 2B there is a highly conserved stutter (represented by a black line). This corresponds to a break in the continuity of the underlying heptad substructure arising from the deletion of three residues from an otherwise continuous heptad repeat.

Conway & Parry, 1988, for a summary of the lengths of the sequences constituting the various structural domains) but it was clear subsequently that the N- and C-terminal regions were also characteristic of these groupings. The central heptad-containing region, however, did not consist of an unbroken heptad repeat. Four major segments were identified. These were called 1A (35 residues), 1B (101 residues in general but 143 residues for the lamins and some invertebrate intermediate filaments), 2A (19 residues) and 2B (121 residues). Segment 2B also had a discontinuity in its heptad phasing at a point close to its centre. This has been termed a 'stutter' and corresponds to a three residue deletion from what would otherwise be a continuous heptad repeat. The short regions between the heptad-containing segments, known as links L1, L12 and L2 respectively, also display high sequence conservation in each

https://doi.org/10.1017/S0033583500003516 Published online by Cambridge University Press

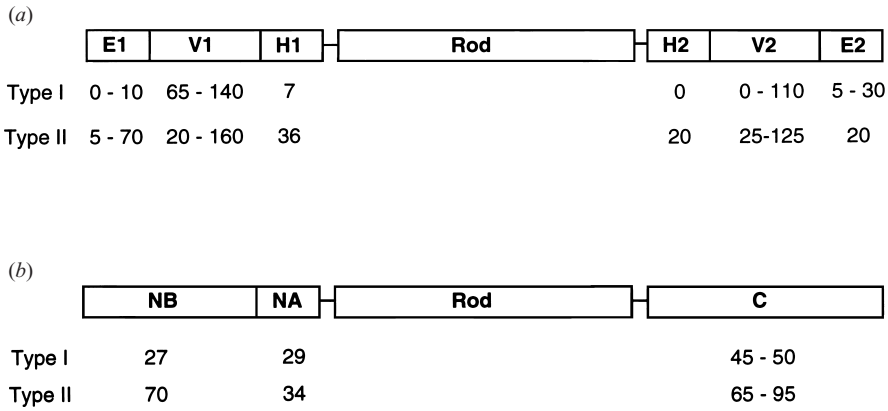



**Fig. 6.** Subdomain structure of cytokeratin and trichocyte keratins. In the Type I and II cytokeratins the head and tail terminal domains have been subdivided into end (E), variable (V) and homologous (H) subdomains. The end subdomains are short, generally basic regions. In contrast, the variable subdomains are usually quite long and have elevated contents of both glycine and serine residues. The sequences of the homologous subdomains are highly maintained and are believed to have an important role in stabilizing small oligomers during the early stages of filament formation. The head domains of the trichocyte keratins have also been subdivided, in this case into acidic regions (NA) and basic regions (NB). Parts of NA and the tail domains are homologous to the H subdomains already defined for the cytokeratins. Figure from Parry & North (1998).

of the groupings. The lengths of the links L1 and L12 show some variation with the chain but lie within the ranges 8–14 and 16–22 residues respectively. Link L2 is always eight residues long (except for the Type VI nestin protein).

A particular feature of all intermediate filament proteins is the regions of sequence located close to both ends of the heptad-containing central region. The sequence in segment 1A, sometimes known as the helix initiation motif, is given by:

K/R-X-T/Q-M/I/L-K/Q-L-N-D-R-F-A-S-F/Y-I-D/E-K-V-R-F-L-E

and that at the end of segment 2B, sometimes referred to as the helix termination motif, is given by:

E-Y-Q-X-L-L-D/N-V-K-X-R/A-L-D/E-X-E-I-A-T-Y-R-K/R-L-L-E-G-E-E/D-X-R-L/N/I

where the sequences are defined in terms of the one-letter code for amino acids. The high sequence homology present across all intermediate filament chains implies that these sequences play an important role, though whether it was primarily functional or structural was not clear until quite recently. It is also pertinent to note that the majority of the observed mutations that lead to keratinopathies have been found in or around these two regions of sequence.

Another very highly, but not absolutely, conserved feature of the rod domain sequences of intermediate filament chains is the presence of specific amino acids in certain positions (Steinert & Parry, unpublished data). In fact, there are nine such residues: four occur in

https://doi.org/10.1017/S0033583500003516 Published online by Cambridge University Press

segment 1A – leucine 7 in position *d*, asparagine 8 in position *e*, valine 18 in position *a* and glutamic acid in position *e*; one in segment 1B–aspartic acid 65 in position *b*; and four in segment 2B–leucine 103 in position *d*, glutamic acid 106 in position *g*, tyrosine 110 in position *d* and leucine 114 in position *a*. Almost all of them lie in or adjacent to the helix initiation or termination motifs. It will also be shown later that there are conserved interchain ionic interactions that may play important roles in molecular stabilization.

An interesting subdomain structure occurs in both the N- and C-terminal domains of the cytokeratins (Fig. 6; Steinert *et al*. 1985b). Immediately adjacent to and at either end of the heptad-containing region are 'homologous' regions known as H1 and H2. The sequences are preserved with high identity in each of the two chain typings that were identified (see Section 2.1.2). In turn, regions adjacent to H1 and H2, known as V1 and V2, displayed considerable variation in their sequences. Deletions and insertions in the cytokeratins expressed in the epidermis were observed by Korge *et al*. (1992a, b) (Fig. 7). These subdomains were nonetheless uniformly rich in glycine and serine residues and are thought to adopt a glycine loop-like conformation (Steinert *et al*. 1991). Finally, adjacent to these regions and at the extreme ends of the chains were short subdomains known as E1 and E2. These were generally basic. Thus the N-terminal domain of the epidermal cytokeratins consists of E1-V1-H1 (in that order) and the C-terminal domain consists of H2-V2-E2 (again in that order). The subdomains are thus arranged about the heptad-containing region with bilateral symmetry (Fig. 6).

The N-terminal domains of Type Ia and Type IIa keratins could also be divided into subdomains (Parry & North, 1998). The NA subdomain, located immediately N-terminal to the rod domain, was defined as such on the basis of its net acidity. It has lengths of 29 and 34 residues respectively for Type Ia and IIa chains. The NB subdomains, which constitute the remainder of the N-terminal domains, are net basic, and are 27 and 70 residues long respectively for Type Ia and IIa chains (Fig. 6). The C-terminal domains, on the other hand, do not obviously lend themselves to further subdivision. However, the H1 subdomains characterized for the cytokeratins have a high degree of homology with the equivalently positioned portions of the respective NA subdomains. For the Type I chains the similarity of the (short) sequences involved is as follows:

```
H1(I) G-L-L-X-G-N-E

NA(I) G-S-F-N-G-N-E
```

The seven residue sequences thus contain four identical and one homologous residues. When the longer (36 residue) H1 subdomain present in the cytokeratin Type II chains is compared with the equivalent trichocyte keratin sequences there are 21 identical residues plus several other homologous ones (see below). X represents a residue that is not highly conserved and • implies a gap in the sequence alignment. The first sequence corresponds to H1(II) and the second one to NA(II):

```
P-V-C-P-P-G-G-I-Q-E-V-T-I-N-Q-S-L-L-Q-P-L-N-V-E-I-D-P-X-I-Q-K-V-K-X-Q-E
      S                             T       L Q             R   R   E
L-•-S-P-P-•-C-I-T-T-V-S-V-N-E-S-L-L-T-P-L-N-L-E-I-D-P-N-A-Q-C-V-K-Q-E-E
P                                                                   H
```

https://doi.org/10.1017/S0033583500003516 Published online by Cambridge University Press

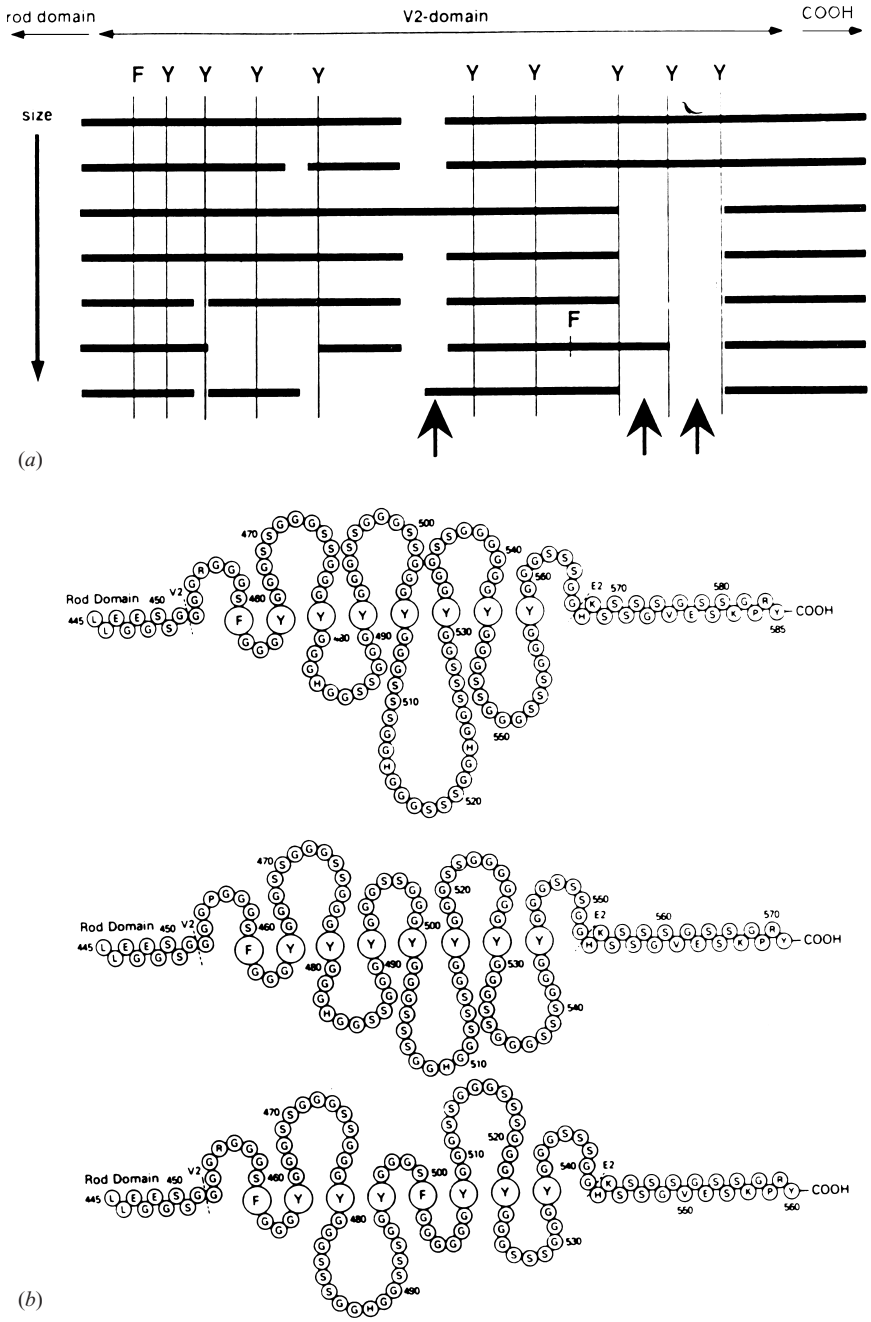



**Fig. 7.** (*a*) The sequences of the V2 domain in seven allelic variants of the human K10 epidermal cytokeratin chain display considerable differences as a result of deletions and insertions. The conclusions are that both ends of the V2 domain are conserved and that most of the observed deletions occur at three specific sites (arrowed). The positions of the tyrosine (Y) and phenylalanine (F) residues are marked. (*b*) Schematic diagram showing potential glycine loop motifs in three of the alleles. In all cases the loops are centred upon highly interacting aromatic residues. Changes in the sizes of the glycine loops are easily accommodated in a structure of this type. Figure from Korge *et al.* (1992a).

https://doi.org/10.1017/S0033583500003516 Published online by Cambridge University Press

The N-terminal parts of the C-terminal domains in the trichocyte keratins that are homologous to the 20 residue long H2 subdomain in the cytokeratins are shown below:

```
H2(II)  S-G-E-G-V-G-X-V-S-I-S-V-V-X-S-T-T-S-S-ap
        N               N             G V
                                        G
C(II)   •-C-E-G-V-G-A-V-N-V-C-V-S-S-S-R-G-G-V-V
                    S
```

Although there are fewer identities (eight) between these two consensus sequences than those that occur between the other pairs of consensus sequences, there is little doubt that the sequences H2(II) and C are homologous to one another. Since a possible function of the H subdomains in cytokeratins has been established (Section 3.2), that is, they provide intermolecular stabilization at the 2–4 molecule level during filament assembly, it must be assumed that the function of the equivalent regions in the trichocyte keratins will be the same.

The sequences of the N-terminal domain of vimentin from a variety of sources contain short regions which appear to have been highly conserved. The first, and probably the most important, is a nine residue motif of the form

```
S-S-Y-K/R-R-M-F-G-G
```

This is located very close to the N-terminal end of the chain (Herrmann *et al*. 1992). It has been proposed that the role of this motif is intimately connected to the ability of the molecules to aggregate appropriately and form normal intermediate filaments. A 12-residue stretch of sequence immediately N-terminal to the rod domain is also conserved though not with as high a degree of regularity as displayed by the nonapeptide. The entire N-terminal domain contains a quasi-regular distribution of aromatic residues with tyrosine predominating. Basic residues are similarly quasi-regularly found along the sequence. These observations give some ideas as to the possible secondary structure (Section 2.2.2)

Neurofilament and nestin chains display different homologous features. In the former case the C-terminal domains have large numbers of lys-ser-pro repeats. These are the sites at which phosphorylation occurs as a consequence of the action of various kinases. Subdomains rich in (i) glutamic acid, (ii) lysine and glutamic acid, (iii) serine and proline and also (iv) lysine, glutamic acid and proline are also apparent in the NF-M and NF-H neurofilament chains (Fig. 8). The C-terminal domain of human nestin chains, on the other hand, is rich in 22 residue repeats. These have the form:

```
Human:  E-E-D-L-E-S-L-R-X-L-E-K-E-N-Q-E-P-L-R-S-L-E
            N Q   T   K                   S   K
```

In hamster and rat nestin, however, the repeats are 44 residues long (though, not surprisingly, these are closely related to the 22 residue repeats seen in humans) (Steinert *et al*. 1999a):

```
Hamster:  E-E-D-Q-L-V-E-R-L-I-E-K-E-G-Q-E-S-L-S-S-P-E-
                    R     T
          E-E-D-Q-E-T-D-R-P-L-E-K-E-N-G-E-P-L-K-P-V-E
Rat:      E-E-D-Q-E-A-X-R-P-L-E-K-E-N-Q-E-S-L-G-Y-L-E-
              E     I                           S
          A-E-D-Q-M-L-E-R-L-L-E-K-E-S-Q-X-S-L-R-S-P-E
                                                K
```

https://doi.org/10.1017/S0033583500003516 Published online by Cambridge University Press

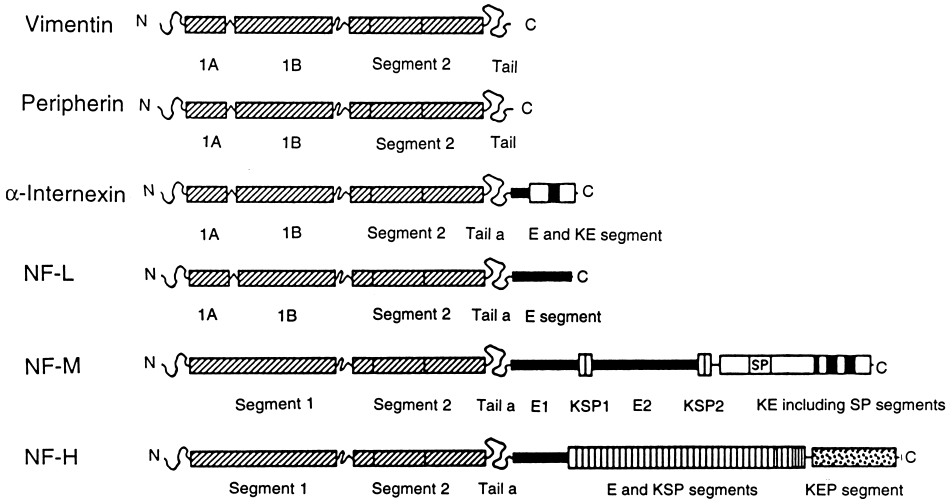



**Fig. 8.** Domain structure of the Type III (vimentin, peripherin) and the Type IV (α-internexin, the neurofilament low molecular weight chain (NF-L), the medium molecular weight chain (NF-M) and the heavy molecular weight chain (NF-H)) intermediate filament chains found in neuronal cells. Sequence features of importance in the tail domain are indicated using the one letter code. Figure from Parry & Steinert (1995) but based on an earlier one by Shaw.

Interestingly, the 22 and 44 residue motifs are composed of a underlying quasi 11 residue substructure.

### 2.1.2 Chain typing

Before describing the classification of intermediate filament chain types it is useful to make an attempt to define what actually constitutes an intermediate filament chain. This, of course, is not something on which there is unanimity. A starting point, however, was provided by us in 1995. We suggested that the following criteria should be met: there should be: (i) four heptad-containing coiled-coil segments corresponding in length to 1A (35 residues); 1B (101 or 143 residues); 2A (19 residues) and 2B (121 residues); (ii) a link L2 of length eight residues; (iii) a highly conserved sequence in segment 1A (the helix initiation motif) and an equally highly conserved sequence at the C-terminal end of segment 2B (the helix termination motif); (iv) a common period in the linear distribution of the acidic and basic residues; (v) a phase discontinuity in the heptad repeat in the middle of segment 2B, (vi) an ability to form or participate in formation of filaments 10–15 nm in diameter; and (vii) a level of homology with other IF proteins that lies well in excess of that shown by heptad-containing regions in other α-fibrous proteins.

On the basis of these criteria six unique types of IF chain have been characterized. The typing is based specifically on the degree of amino acid sequence homology within the rod domains though, more recently, the exon and intron structures of the IF genes have provided additional data on which to base the typing classification. Types I and II IF chains occur in trichocyte keratins or cytokeratins in epithelial cells. Those that originate from the trichocyte keratins have been designated Ia and IIa, and those from the cytokeratins as Ib and IIb

https://doi.org/10.1017/S0033583500003516 Published online by Cambridge University Press

(Conway & Parry, 1988). The subtyping is a result of the small but characteristic differences in rod domain sequences that have been found between the trichocyte keratins and cytokeratins. Nonetheless, both forms of keratin are clearly members of the Type I and Type II families (Hanukoglu & Fuchs, 1982, 1983; Steinert *et al*. 1984). The Type I keratins are net acidic and have molecular weights in the range 40–57 kDa. The Type II keratins are neutral-basic and have molecular weights typically in the range 52–70 kDa (slightly higher owing to the presence of H1 subdomains). Four Type III IF chains have been identified – vimentin in mesenchymal cells (mol wt 55 kDa), desmin in muscle cells (mol wt 53 kDa), glial fibrillary acidic protein in glial cells and astrocytes (mol wt 51 kDa), and peripherin in neuronal cells (mol wt 54 kDa). Four Type IV IF chains have been characterized and all are found in neurons. These are $\alpha$-internexin, neurofilament light chain (NF-L), neurofilament medium chain (NF-M) and neurofilament heavy chain (NF-H). Their molecular weights are about 56, 62–68, 102–110 and 115–130 kDa respectively. Two classes of Type V IF chains (A and B) occur in the nuclear lamina of mammalian nucleated cells. The A chains have molecular weights in the range 62–73 kDa and occur in most nucleated cells. The B chains have similar molecular weights (65–68 kDa) but are found in all nucleated cells. In both the A and B chains there is a 42 residue (six heptad) insertion in segment 1B. A single Type VI IF chain – nestin – has been added to the IF family though its inclusion as an IF chain is somewhat controversial (Lendahl *et al*. 1990). In particular, nestin lacks the L2 region found in all other IF chains. Nonetheless, it has the ability to form homodimers and homotetramers but is apparently unable to form filaments by itself (Steinert *et al*. 1999a). However, it can copolymerize with Type III vimentin and Type IV $\alpha$-internexin chains to form filaments. It would therefore seem to have assembly features which make it clearly compatible with its designation as an intermediate filament chain (Steinert *et al*. 1999a). Likewise, other proteins such as synemin (Becker *et al*. 1995; Bilak *et al*. 1998), paranemin (Hemken *et al*. 1997) and tannabin (Hemmati-Brivanluo *et al*. 1992) do possess an L2 domain, are more closely related in sequence to nestin than any other chain Type, and thus may represent members of a larger Type VI group.

In addition to those intermediate filament chains that are classified within the Type I–VI families, a host of other intermediate filament-like chains have been sequenced. These include the hagfish $\alpha$ and $\gamma$ chains (Koch *et al*. 1994, 1995), filensin (Remington, 1993; Gounari *et al*. 1993; Hess *et al*. 1998), phakinin (Merdes *et al*. 1993; Georgatos *et al*. 1997; Gounari *et al*. 1997), restin (Bilbe *et al*. 1992), tektin (Norrander *et al*. 1992), skelemin (Price *et al*. 1993), and transitin (Yuan *et al*. 1997). The differences in the sequences between these chains and 'classic' intermediate filament chains vary from quite minor to very substantial. Transitin also has greater similarity to nestin and tannabin so it too may ultimately lie within the suggested expansion of the Type VI family (Yuan *et al*. 1997). Skelemin, on the other hand, has only very limited homology with two short regions found in intermediate filament chains. Its classification within the overall family of intermediate filament chains is thus open to serious doubt. In the case of tektin the highly conserved regions at the ends of the central rod domain are absent and the lengths of segments 1A and 2A differ from normal values. Filensin (which dimerizes with phakinin) lacks 29 residues from the 'usual' rod domain sequence and also has only partial identity with the highly conserved regions at the N-terminal end of segment 1A and the C-terminal end of segment 2B. The intermediate or intermediate filament-like chains thus display many variations on a theme. Further discussion in this review, however, will concentrate on intermediate filaments that lie within the Type I–VI classification.

https://doi.org/10.1017/S0033583500003516 Published online by Cambridge University Press

### 2.1.3 Post-translational modifications

Post-translational modifications to the N- and C-terminal domains of intermediate filament chains are common whereas those within the central $\alpha$-helical domain are rare (Parry & Steinert, 1995). For lamin chains the modifications include proteolysis of the C-terminal domain plus isoprenylation at the CaaX box located at the extreme C-terminal end of the chain (Nigg, 1992; Hennekes & Nigg, 1994). These modifications have the effect of ensuring that the chains are targeted successfully to the nuclear envelope. Other intermediate filament chains, including some cytokeratin and neurofilament chains as well as lamins, are also subject to proteolysis (Eagles *et al*. 1990; Klymkowsky *et al*. 1991; Nigg, 1992). Furthermore, the N-terminal domain of neurofilament chains often contain *O*-glycosylated residues and sometimes these are found in the tail domain too (Dong *et al*. 1996). This modification affects the mechanism of assembly of molecules into intermediate filaments and thus has structural and functional importance *in vivo*. A post-translational modification that is currently unique leads to a phosphotyrosine residue at the C-terminal end of peripherin (Angelastro *et al*. 1998). Apoptosis involves participation of unique proteases that cleave cytokeratins (Ku *et al*. 1997), which presumably contributes to the orderly dissolution of the cytoskeleton of the dying cell.

The post-translational modification of greatest functional and structural importance, however, arises from the action of various kinases on serine and threonine residues embedded within specific recognition sequences. These phosphorylation events, which have a regulatory role in the assembly and turnover of most intermediate filaments *in vivo*, are confined to the N- and/or C-terminal domains of the chains (see, for example, Parry & Steinert, 1995, for a listing of pertinent references to phoshorylation events in these proteins; also Inagaki *et al*. 1994, 1996; Ku & Omary, 1994; Liao *et al*. 1995; Ando *et al*. 1996; Chou *et al*. 1996; Nakamura *et al*. 1996; Giasson & Mushynski, 1998). The phosphorylation sites of greatest significance *in vivo* seem to be those in the N-terminal rather than the C-terminal domain. Phosphorylation/dephosphorylation events in intermediate filament-associated proteins such as desmoplakin (Stappenbeck *et al*. 1994) also play an analogous functional role. Specifically, phosphorylation of a serine residue close to the C-terminus of desmoplakin is thought to negatively regulate its interaction with keratin intermediate filaments. Phosphatases, which dephosphorylate the chains, have an equally important role in governing assembly and function. In cytokeratins, vimentin and the lamins the effect of phosphorylation is to disassemble the intermediate filaments, whereas dephosphorylation allows spontaneous self-assembly to proceed. These changes typically occur during cell division. For the neurofilaments, on the other hand, the phosphorylation of the large C-terminal domains does not seem to change the state of assembly but appears to have a regulatory role on functions such as transport and supramolecular organisation in the axon. In particular, the level of phosphorylation of the neurofilament chains has been associated with axonal diameter, plasticity and stability (see, for example, Pant & Veeranna, 1995). Interestingly, the single neurofilament chain found in lamprey (NF-180) has few potential phosphorylation sites in its C-terminal domain (Jacobs *et al*. 1995) casting some doubt (in this case at least) on the ability of the lamprey neurofilament chain to regulate axonal calibre as appears to be well documented for the NF-M and NF-H chains in mammals, fish and squid.

Evidence exists that the dynamic movements of intermediate filament networks in interphase cells is also governed by rapid and continuous cycling of phosphorylation and dephosphorylation. Microinjection experiments of single tagged Type I chains or vimentin

https://doi.org/10.1017/S0033583500003516 Published online by Cambridge University Press

chains into cultured cells have demonstrated rapid assimilation of the injected protein onto the existing cytoskeleton at multiple sites throughout the cell (Miller *et al*. 1993, and references therein). Presumably the injected proteins accumulate in soluble pools before assimilation. The proteins in these pools exist as small subfilamentous units, perhaps tetramers (Soellner *et al*. 1985) and are modified on the head domains by 2–3 mol of phosphate/chain (Yatsunami *et al*. 1991; Skalli *et al*. 1992). That this is a highly dynamic phenomenon is evident when such cells are treated with phosphatase inhibitors: the entire intermediate filament cytoskeleton is lost and the proteins accumulate as phosphorylated oligomers (Eriksson *et al*. 1992). Withdrawal of the drugs allows dephosphorylation to occur and favours rapid reassembly of the elaborate intermediate filament networks.

A study of the extent and precise form of the post-translational modifications in all intermediate filament chains is not yet complete. Indeed, the field is a large and complex one that is beyond the scope of this particular review. What is clear, nonetheless, is that such modifications have real importance *in vivo* and will need to be understood if a full appreciation of the structure, assembly and function of intermediate filaments is to be gained.

## 2.2 Secondary structure

Experimental data on the secondary structure of intermediate filament proteins derived by using a variety of physical chemical techniques, including circular dichroism, and X-ray diffraction analyses on orientated fibre specimens (mainly from wool or porcupine tips), as well as secondary structure prediction techniques (Chou & Fasman, 1978; Garnier *et al*. 1978), are consistent with the idea that the intermediate filament proteins contain a 'central' $\alpha$-helix-rich region or 'rod domain'. This is about 355 residues in length for the Type V lamin proteins and some invertebrate IF proteins, and about 315 residues for all other chains (Fig. 5). The N- and C-termini which flank this central $\alpha$-helix-rich region, hereafter termed the head and tail domains, vary greatly in sequence from one IF chain type to another. In almost all cases the predicted $\alpha$-helix content is very low in these domains, and there are indications that $\beta$-structure, turns and random coil generally predominate.

### 2.2.1 Central rod domain

Although the terminology of a 'central' domain is commonplace in the literature it needs to be made clear that the head and tail terminal domains of IF proteins are rarely of equal size. Indeed, they are usually of quite different size as, for example, in the case of the heavy neurofilament chain (about 100 and 650 residues for the head and tail domains of NF-H respectively). In this respect the central domain is rarely 'central' in the amino acid sequence though it certainly is in the sense that it lies between the other two major domains in the structure and forms the middle portion of a tripartite structure.

It has already been noted that the central domain is predicted to have a very high $\alpha$-helix content. This was deduced from experimental data and from a variety of theoretical techniques, most specifically those due to Chou & Fasman (1978) and to Garnier *et al*. (1978). The coherence of the $\alpha$-helix, however, is interrupted at three positions. This was determined from the presence (or otherwise) of an underlying heptad structure that allows the ends of the $\alpha$-helical regions to be defined relatively precisely. Each of the short intervening regions generally lacks both $\alpha$-helix potential and a heptad repeat, and has been designated as a link.

https://doi.org/10.1017/S0033583500003516 Published online by Cambridge University Press

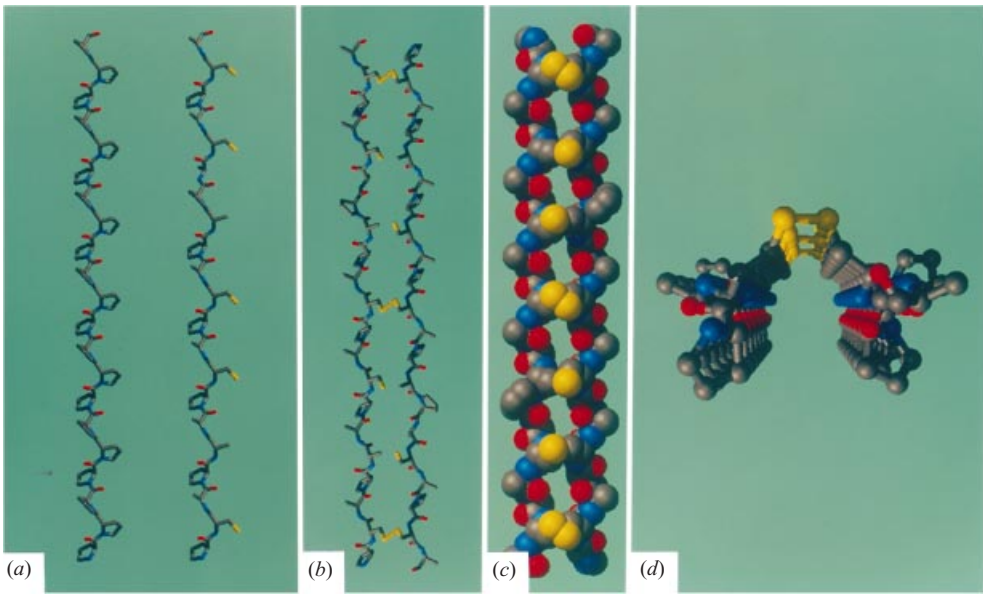



**Fig. 9.** A stretch of seven consecutive tripeptides in the tail domain of Type Ia trichocyte keratins has a consensus sequence of the form pro-cys-X. It has been proposed that this sequence will adopt the conformation of a poly-proline II helix. (*a*) Lateral view of such a structure (left) and also one with the actual sequence incorporated (right). The five cysteine residues, coloured yellow, lie along one edge of the helix. (*b*) Atom ball-and-spoke and (*c*) space-filling models illustrating the packing of an antiparallel pair of poly-proline II helices aligned axially in such a way that three disulphide bonds can be formed to stabilize the assembly. (*d*) End-on view of the structure in (*c*). Figure from Parry & North (1998).

As noted earlier the α-helical segments 1A, 1B, 2A and 2B are such that link L1 connects 1A to 1B, L12 connects 1B to 2A, and L2 connects 2A to 2B.

The Type V lamin proteins and the invertebrate IF proteins both contain an additional 42 residues (six heptads) in their 1B segments (Fisher *et al*. 1986; McKeon *et al*. 1986; Parry *et al*. 1986). Furthermore, their link regions show high α-helix potential. While clear discontinuities in the heptad phasing remain it would seem probable that the central rod domain (about 356 resides long) is virtually completely α-helical. This would correspond to a physical length of the central domain of about 52 nm.

Type VI nestin differs from the conventional model of the central domain proposed for Types I-IV intermediate filament chains in that the link L2 is absent, although an L2 link is present in other putative Type VI proteins. The other features of the substructure of the rod domain of nestin, however, are largely maintained.

### 2.2.2 Head and tail domains

In the absence of any published atomic-resolution structural studies on any intermediate filament fraction, including the head and tail domains, prediction methods have been used to assess possible features. One such study has yielded a model for the tail domain of both Type Ia and Type IIa trichocyte keratin chains. The Type Ia chains contain an eight-fold contiguous repeat of the quasi-triplet pro-cys-X. Computer modelling, based on data from

https://doi.org/10.1017/S0033583500003516 Published online by Cambridge University Press

crystallographic studies of other proteins, has indicated a high probability that this region will form a poly-proline II-like structure (Fig. 9). This would be a left-handed helix with three residues per turn and an axial rise per residue of 0·312 nm. The cysteine resides would thus lie along one edge of the helix and be in favourable positions where they may be able to form disulphide bonds with a similar structure from a different molecule. The proposed structure for the tail domain of the Type IIa chains, however, is quite different. It has a region about 55–60 residues long, immediately adjacent to the rod domain, that has a degree of sequence regularity that is quite remarkable. It was discovered by Parry & North (1998) using a method devised by Fraser & Parry (1996) in which regions of sequence compatible with a $\beta$-sheet conformation with an internally located apolar face were sought. This was delineated using hydropathy scoring on a triplet of residues each two apart. Any region of sequence with potential to form a $\beta$-sheet with an apolar face would result in alternating high and low scores of the hydropathy index. The data clearly showed that a 55–60 residue sequence immediately C-terminal to the rod domain had a near-regular two residue repeat that was postulated to correspond to four $\beta$-strands. These would be connected by two pairs of homologous $\beta$-turns (sequences S–S/G–S–R/A and A–P–C–S/G). The overall conformation would be a four-chain antiparallel $\beta$-sheet (Fig. 10), one face of which would be almost totally composed of apolar residues. The other face would be more mixed but would contain four cysteine residues and all of the charged residues that were present. These features lend themselves to a ready interpretation as regards both their function and their tertiary structure (see Section 2.3.3).

The structures of the V1 and V2 regions in Type Ib and Type IIb intermediate filament chains have been modelled by Steinert *et al*. (1991). Earlier NMR studies had shown that these regions had a high degree of flexibility (Mack *et al*. 1988), a conclusion in accordance with the theoretical studies of Conway *et al*. (1989). Although the V subdomains are rich in serine and glycine residues there are aromatic and long chain aliphatic residues scattered throughout the sequence. Steinert *et al*. (1991) proposed that these long chain aliphatic residues interact with one another to generate a series of glycine loops which maintain a degree of local secondary order but which provide very little organization at the tertiary level (Fig. 7). The nature of the loops is such that deletions/insertions of integral numbers of them, or conversely changes to the sizes of the loops, would be easy to explain. In fact, the V2 subdomains of the human keratin 1 (Type II) and especially keratin 10 (Type I) chains are polymorphic (Korge *et al*. 1992a, b: Mischke, 1993) in the normal human population. In contrast, there is no evidence that polymorphism is a feature of the V1 subdomains. It would seem that a degree of variation in the V2 subdomain can be tolerated without adversely affecting the structure or function of the epidermal intermediate filament.

The head domains of vimentin, desmin, peripherin and Type IV intermediate filament proteins (but not the glial fibrillary acidic protein) contain a highly conserved nonapeptide repeat (ser–ser–tyr–arg–arg–ile–phe–gly–gly) close to its N-terminus and another, less well-conserved repeat adjacent to the N-terminal end of the rod domain (Herrmann *et al*. 1992). Deletion experiments indicate that the former repeat is required for assembly. Subsequently, Ralton *et al*. (1994) have identified what appears to be a near equivalent motif close to the N-terminal end of glial fibrillary acidic protein. However, they propose that a second motif, called the RP-box (arg–leu–ser–leu–X–arg–met–X–pro–pro), is the important one in the nucleation and elongation phases of filament assembly. Similar motifs were found by these authors in desmin, vimentin and peripherin. The sequence of the entire head domain of

https://doi.org/10.1017/S0033583500003516 Published online by Cambridge University Press

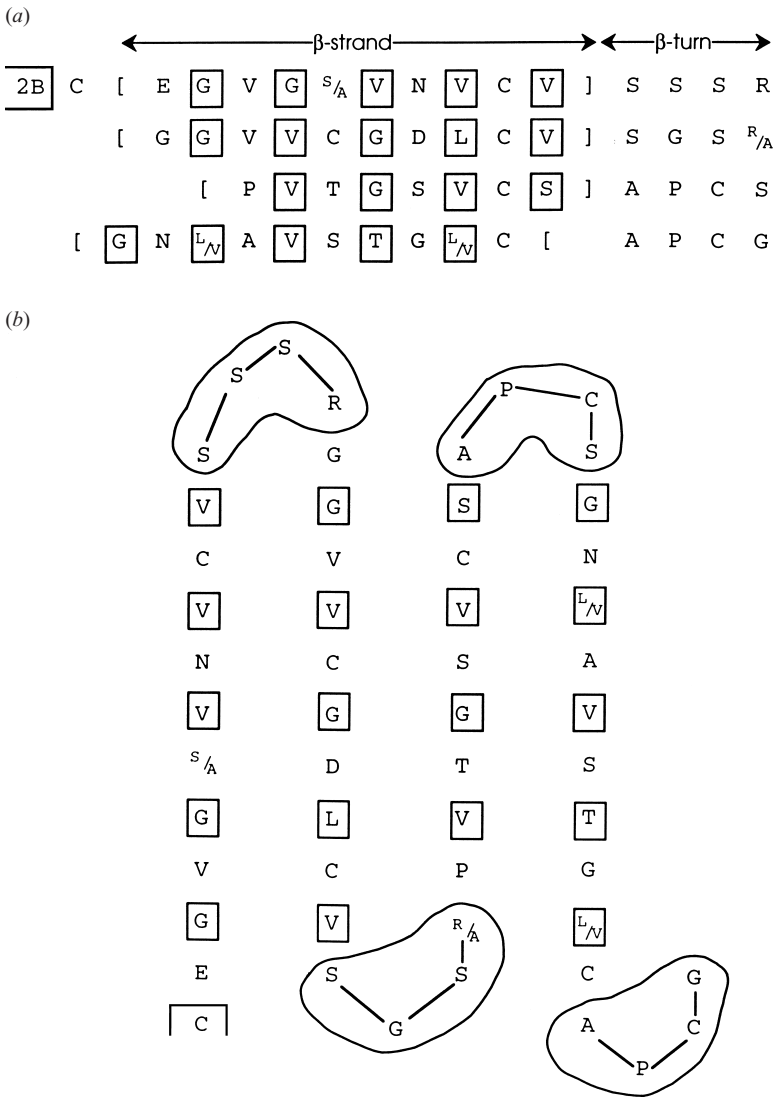



**Fig. 10.** (*a*) The consensus sequence of the first 57 residues in the tail domain of Type IIa trichocyte keratins is laid out to illustrate the four proposed $\beta$-strands and the four connecting $\beta$-turns (Parry & North, 1998). The $\beta$-strands are characterized by quasi-dipeptide repeats of the form (apolar–X) which lead to a $\beta$-sheet with one face predominantly occupied by apolar residues and the other by cysteine and charged residues (amongst others). The apolar residues are boxed. Note the high degree of homology between the sequences of the first and second $\beta$-turns and also between the third and fourth $\beta$-turns. (*b*) Schematic diagram of the proposed four-stranded antiparallel $\beta$-sheet. The $\beta$-turns are ringed. Figure modified from Parry & North (1998).

vimentin has semi-regular distributions of both aromatic and basic residues. A model for this domain (Herrmann *et al*. 1992; Herrmann & Aebi, 1998a) has similarities to the glycine loop conformation proposed for the epidermal keratins (Steinert *et al*. 1991). In the case of vimentin, it is proposed that the aromatics interact strongly and thus give rise to loops, some of which display potential phosphorylation sites.

https://doi.org/10.1017/S0033583500003516 Published online by Cambridge University Press

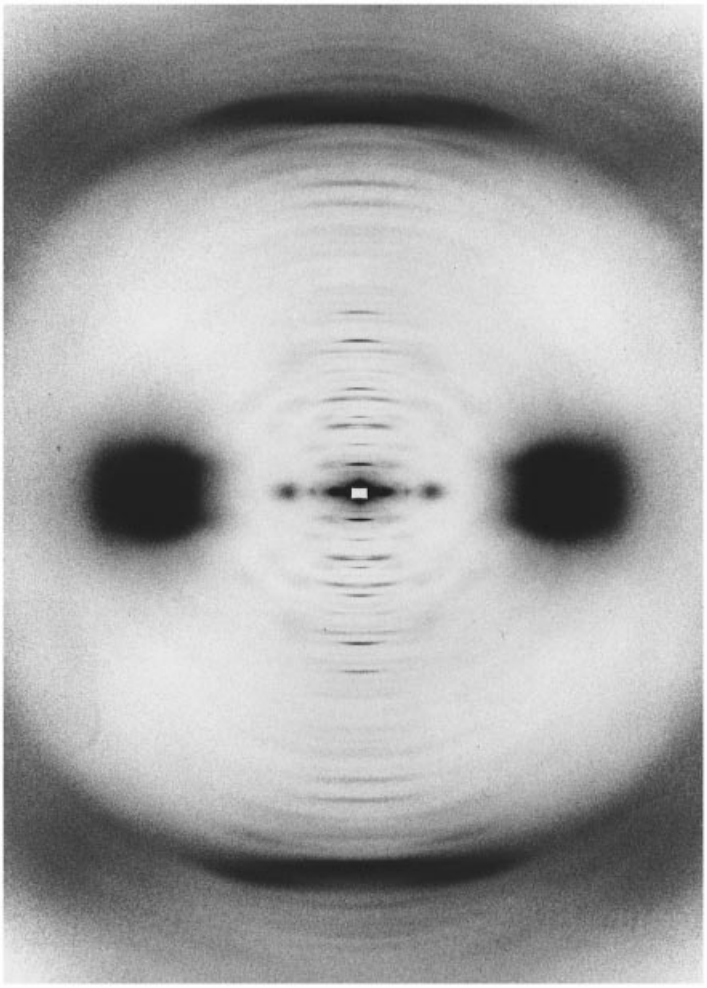


**Fig. 11.** X-ray diffraction pattern of quill. The long axis is vertical. The pattern shows a strong meridional reflection at a spacing of 0·515 nm that relates directly to the presence of a coiled-coil structure. The spacing of the near-equatorial layer line (imperfectly resolved from the equatorial maximum) is equal to the ratio of the coiled-coil pitch length to the number of strands (now known to be two). This indicates a pitch length, therefore, of about 15 nm. The meridional and near-meridional reflections, as regards both their spacings and intensities, provide the information that allows the surface lattice structure to be determined (see Section 4.3). Figure provided by R. D. B Fraser and T. P. MacRae.

Fourier transform infrared spectroscopy has been used to assess the predominant secondary structure of domains of Type III and Type IV intermediate filament proteins (Heimburg *et al*. 1996). The head domain of desmin contains a major fraction of $\beta$-sheet, whereas the tail domains of desmin, NF-L and also NF-H in their phosphorylated and dephosphorylated forms all display random structure. To date there have been no other conformational studies on either the head or the tail domain from any other intermediate filament chain. Since these regions represent much of the functional capability of intermediate filaments this is clearly an area in need of much more attention.

https://doi.org/10.1017/S0033583500003516 Published online by Cambridge University Press

### 2.3 Tertiary structure

X-ray fibre diffraction patterns of trichocyte keratins (predominantly from quill and hair) have provided virtually all of the data on which the tertiary structure of intermediate filaments is based (Fig. 11). Orientated specimens, on which this technique is heavily reliant, have not been obtained for any other intermediate filament other than for an invertebrate neurofilament from the giant axon of *Myxicola* (Day & Gilbert, 1972). Furthermore, to date no major portion of an intermediate filament chain has been crystallized, and there have been no atomic level NMR structural data deduced either (see Section 4.4, however). The very limited data, nonetheless, indicate that the intermediate filament molecules lie approximately parallel to the filament axis and that they have a high component of $\alpha$-helical coiled-coil structure. This is evident from the meridional reflections observed at spacings of 0·5165 and 0·1485 nm. The half-width of the X-ray meridional reflections show that there is an exceptionally high degree of axial ordering ($> 1\ \mu$m). This means that the molecules are very well arranged axially. Laterally and azimuthally, however, there is much greater variation from one molecule to another. The main equatorial maximum has a spacing of 1 nm and corresponds to the lateral separation of the $\alpha$-helices in the coiled-coil molecule. In contrast to the crystal-like regularity along the axis of the intermediate filament (which relates to the very precise positioning of the diffracting units) the coiled-coil molecules themselves have an axial coherence of only about 7 nm (Fraser *et al*. 1964). This can be interpreted as meaning that there are breaks in the regularity of the coiled-coil. Indeed, this value relates well to the average length of the coiled-coil segments in the rod domain. The links L1, L12 and L2 plus the stutter effectively divide the 45 nm long coiled-coil into shorter pieces that fit neatly with the observed data. The geometry of the coiled-coil itself can be deduced from the spacing of a near-equatorial layer line in the diffraction pattern. The data thus indicate that the two-stranded coiled-coil will have a pitch length of about 14–17 nm. A number of other two-stranded coiled-coils have also been shown to have pitch lengths of about 14–15 nm: GCN4 leucine zipper–14·7 nm, Seo & Cohen (1993); myosin – 14 nm, Quinlan & Stewart (1987); oothecal protein from praying mantis – 14 nm, Bullough & Tulloch (1990); tropomyosin – 13·7 nm, Parry (1975), McLachlan & Stewart (1975), Whitby *et al*. (1992); cortexillin I – 15·3 nm, Burkhard *et al*. (2000); and paramyosin – 13·6–14 nm, Elliott *et al*. (1968).

#### 2.3.1 Coiled-coil rod domain

The presence of a heptad repeat in an amino acid sequence is characteristic not only of an $\alpha$-helical secondary structure but also of a tertiary conformation in which bundles or ropes of $\alpha$-helices are formed (Cohen & Parry, 1986, 1990, 1994). This follows as a consequence of the fact that the apolar residues in the *a* and *d* positions of the heptad repeat are, on average, 3·5 residues apart whereas within the $\alpha$-helix there are about 3·6 residues per turn. Consequently, a stripe of apolar residues will wind round the axis of the helix gradually in a left-handed manner. Apolar residues generally seek the company of others so that together they can be shielded from the aqueous environment. Thus, in general two or more $\alpha$-helices assemble by coiling around one another to form a multi-stranded coiled-coil molecule stabilized by systematic close packing of apolar residues along the axis. The energy involved in distorting a straight $\alpha$-helix to one whose axis is distorted into a regular long-period coil is minimal (Parry & Suzuki, 1969a, b). The nature of the apolar residues occupying the *a* and *d* positions

https://doi.org/10.1017/S0033583500003516 Published online by Cambridge University Press

is now known to be the prime determinant of the number of α-helices that assemble to form the molecule (Section 2.3.1.2).

For intermediate filament chains there are four such heptad-containing segments (1A, 1B, 2A and 2B). While each rod domain segment has been precisely conserved in the family of intermediate filament chains, the four segments differ in length, which is strongly suggestive that the chains will aggregate so that they are parallel to one another (rather than antiparallel) since otherwise the lengths of the coiled-coil segments would not match one another. In Section 2.3.1.1 it will be shown that the charged residues in the *e* and *g* positions play an important part in specifying the chain registration and relative chain polarity. In turn, the nature of the apolar residues in positions *a* and *d* are crucial in determining that it is two chains rather than three that assemble to form the intermediate filament molecules.

### 2.3.1.1 Specificity through salt bridges

Many but not all of the residues in the *e* and *g* positions are charged ones. Because these residues lie immediately on either side of the apolar interface defined by the interlocking apolar residues in positions *a* and *d* they are in sites that would allow interchain ionic interactions (salt bridges) to be formed (Fig. 12). In addition to 1g–2e′ interactions it is also possible for 1g-2a′ and 1d-1e′ salt bridges to be formed. The symbol 1 refers to a reference heptad and 2 refers to the next heptad along the sequence. The prime symbol indicates the second chain whereas the lack of the prime symbol indicates the reference chain. Attractive interactions (which are scored +1), together with repulsive ones (which are scored −1), that is, those between two acidic residues or two basic residues, are determined and their net number calculated as a function of chain polarity and relative stagger. When all possible alignments are considered, the answer is unequivocal – the numbers of interactions have a very large peak for parallel chains that are in axial register (see also Parry *et al*. 1985). This observation is not confined to intermediate filaments but has been found for every two-stranded coiled-coil that has been studied. Conway & Parry (1990) provided a summary of these data up until that time but more recent data that have been derived confirm their conclusion in all cases. Direct experimental evidence from crosslinking of intermediate filament molecules in small assembly-competent oligomers (Section 3.3) has provided further confirmation that this is indeed the case.

Conserved features of the sequences of the intermediate filament chains are such that three key interchain interactions appear to be present for all intermediate filament molecules (Fig. 13). These are near the middle of segment 1A (basic residue 17 in a *g* position of one chain with an acidic residue 22 in an *e* position of the other chain), towards the end of segment 1B (acidic residue 84 in a *g* position with a basic residue 89 in an *e* position), and close to the end of segment 2B (acidic residue 106 in a *g* position with an acidic residue 111 in an *e* position). Notably, each conserved set lies equidistant from the end of its segment. The fact that these interchain ionic interactions seem possible in every case argues that they play an important role in stabilizing and/or specifying the molecular architecture. Indeed, mutagenesis experiments using the keratin 5 and 14 chains have documented that each of these conserved sets of ionic pairs is required for assembly *in vitro* as well as when transfected into living cells (Steinert & Parry, unpublished data). Interestingly, the sequences in the 2B segment seem to be absolutely essential, which perhaps pertains to their involvement in a coiled-coil trigger

https://doi.org/10.1017/S0033583500003516 Published online by Cambridge University Press

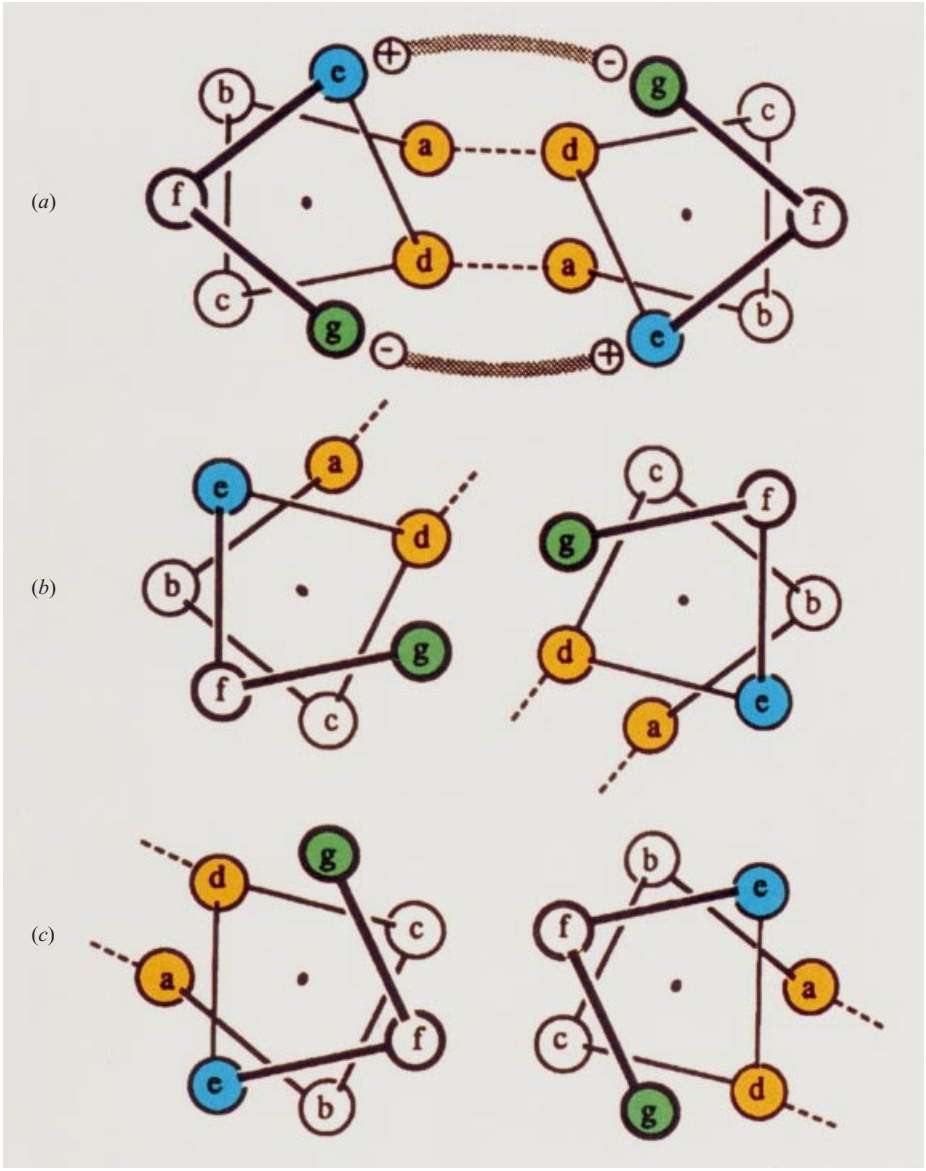



**Fig. 12.** (*a*) Axial projection of a section of a two-stranded α-helical coiled-coil showing the internal (core) positioning of the *a* and *d* residues. These positions are generally occupied by apolar residues such as leucine, valine, isoleucine and alanine. Immediately adjacent to these core residues are the *e* and *g* positions in different chains. They are ideally sited to allow interchain ionic interactions to be formed, and indeed such interactions have been seen frequently in a number of crystallographic studies on coiled-coil proteins. (*b*) The stutter at the centre of segment 2B corresponds to the deletion of three residues from an otherwise continuous heptad repeat in this region. The effect of this is to rotate the apolar seam defined by the *a* and *d* residues. In order to make this seam an internal one again at all points beyond the stutter some local structural rearrangement is required. Brown *et al*. (1996) have proposed that this can be achieved by a slight unwinding of the coiled-coil over a short axial distance. (*c*) A continuation of the heptad repeat through link L2 would effectively turn the apolar seam in each chain by 180°, thus externalizing it almost completely. Some structural rearrangement is also required to ensure that the apolar seam on both sides of the link remains internal. A model proposed for this structure is shown in Fig. 14. Figure from Conway & Parry (1990).

https://doi.org/10.1017/S0033583500003516 Published online by Cambridge University Press

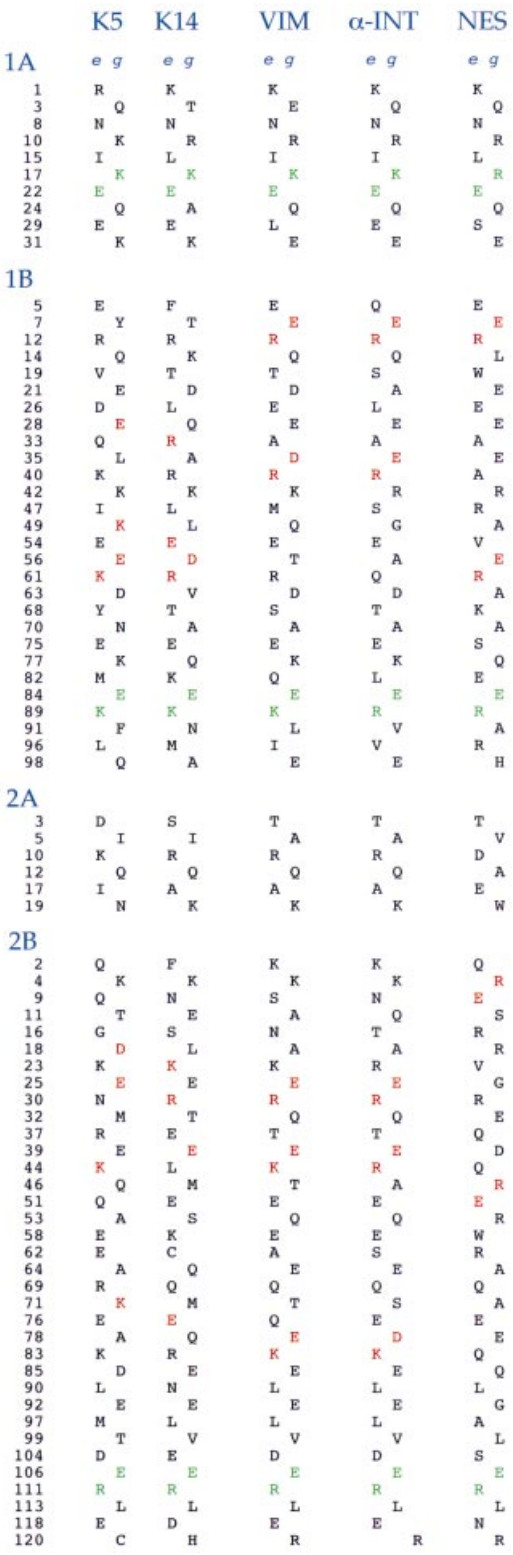



**Fig. 13.** For legend see facing page.

https://doi.org/10.1017/S0033583500003516 Published online by Cambridge University Press

motif (see below). On the other hand, loss by mutagenesis of other nonconserved charged sets of residues (Fig. 13) resulted in normal or near-normal filaments.

Data from crystal studies on coiled coils have clearly revealed the presence of the salt bridges defined by the theoretical studies (O'Shea *et al*. 1991; Harbury *et al*. 1993, 1994). Although not all of the interchain ionic interactions that are stereochemically theoretically possible do in fact form *in vivo*, the presence of at least some of them seems to be an important determinant of relative chain polarity and axial alignment in the molecule.

In addition to intramolecular ionic interactions of the type described above there are also intrachain ones arising between oppositely charged residues that lie four apart in an $\alpha$-helical structure: that is, they could form salt bonds between adjacent turns of the $\alpha$-helical strand (Marqusee & Baldwin, 1987; Perutz & Fermi, 1988; Cohen & Parry, 1990). Up to 50 such interactions are theoretically possible in the case of the intermediate filament molecules (Letai & Fuchs, 1995). These would be expected to provide further stabilization to the molecular structure. Comparable levels of intrachain ionic interactions occur, however, in the rod domains of other $\alpha$-fibrous proteins (D.A.D. Parry, unpublished data).

#### 2.3.1.2 Specificity through apolar interactions

There are now considerable crystallographic data on the GCN4 leucine zipper and on its engineered variations. From these studies it has become evident that the apolar residues in positions *a* and *d* are the prime determinants of the number of chains that co-assemble into an intact coiled-coil molecule. Early work by Parry (1982) and later studies by Conway & Parry (1990, 1991) had showed that two- and three-stranded coiled-coils had clearly distinguishable distributions of residues in these key internal positions in the coiled-coil. Detailed crystallographic studies on the coiled-coil region of the GCN4 transcription factor and variations (O'Shea *et al*. 1991; Harbury *et al*. 1993, 1994) have now shown how such residues are able to direct the number of strands involved in the molecular structure. For intermediate filaments the data were consistent with a two-stranded molecule though other more direct experimental evidence had already allowed this conclusion to be made (Steinert, 1990; Hatzfeld & Weber, 1990a). An analysis of the residues in other positions of the heptad has also provided the database which enables predictive methods to be used to determine the number of strands in any two- or three-stranded $\alpha$-fibrous protein (Woolfson & Alber, 1995; Wolf *et al*. 1997).

---

**Fig. 13.** Presence of highly conserved sets of ionic charge pairs in human intermediate filament chains. Listed are the residues located in the *e* and *g* heptad positions of two adjacent aligned chains that form a coiled-coil molecule. Shown in colour are pairs that are likely to form a salt bond between 1g-2e′ residues (and its complementary partner). Although there are several non-conserved pairs (red), three (green) have been precisely conserved in all chains: residues 17/22 of the 1A domain, residues 84/89 of the 1B domain, and residues 106/111 in the 2B domain. Mutation of either the acidic or basic residue in these conserved pairs results in diminished or failed filament formation *in vitro* and *in vivo*, whereas disruptions of other non-conserved pairs elsewhere on the chains have much less drastic effects. These may form the core of a relaxed trigger sequence motif for coiled-coil formation, and/or may be required for higher orders of molecular assembly in filaments. The abbreviations used are: K5, K14 = keratin 5, keratin 14 heterodimers; VIM = vimentin; $\alpha$-INT = $\alpha$-internexin; and NES = nestin. These three conserved pairs are also present in lamins.

https://doi.org/10.1017/S0033583500003516 Published online by Cambridge University Press

### 2.3.1.3 A consensus trigger sequence for two-stranded coiled-coils

Biophysical analyses have revealed that the slime mold protein cortexillin forms a near-perfect homodimeric two-chain coiled-coil (Steinmetz *et al*. 1998). By deletion analyses, these authors have identified a 13-residue sequence toward the end of the coiled-coil domain of this protein that is absolutely required for its assembly. Furthermore, more recent analyses of other proteins which are known to form two-chain coiled-coils also reveal the presence of this motif, which therefore suggests that such a trigger motif may be involved or even required for the assembly of two-chain coiled-coil proteins (Kammerer *et al*. 1999). Its consensus is:

*b c d e f g a b c d e f g*
x x L E x c h x c x c c x

where x may be any residue, c is charged, and h is hydrophobic. Examination of sequences of all intermediate filament chains, however, does not reveal this exact motif, although a region with a near fit is located in the highly conserved sequences two-three heptads from the end of the 2B rod domain segment (2B residues 101–113), e.g. for the keratin 14 chain:

T R L E Q E I A T Y R R L

It is not clear at this time whether the one misfit (thr in heptad position *c*) is of critical importance. Nor is it known whether it is a coincidence that the latter half of this region encompasses the helix termination sequences which have been highly conserved in all intermediate filament chains. Furthermore, sequences with two-four misfits occur two-three heptads from the ends of the 1A and 1B segments in Types I–VI chains. Thus, deletion cloning experiments will be required to test whether these sequences do in fact function as coiled-coil initiators for intermediate filament chains and whether the requirements of the consensus motif identified for other proteins should be relaxed slightly. If so, then these data suggest that two-stranded coiled-coil proteins including intermediate filament chains assemble from near the C-terminus forwards.

## 2.3.2 Discontinuities in the rod domain

There are four 'discontinuities' in the rod domain of intermediate filament proteins. Three of them – links L1, L12 and L2 – arise because the coherence of the dominant coiled-coil feature of the rod domain is disrupted, albeit to rather different extents in the three cases. The links represent points at which there is an insertion of short regions of sequence that lack both a heptad repeat and a propensity to adopt an $\alpha$-helical conformation. Details of the likely structures are not available but a model has been proposed for link L2. Some general thoughts on models for L1 and L12 have also provided a starting point. The fourth discontinuity, known as a 'stutter', represents a short region over which the phasing of the heptad repeat is changed suddenly. In this case a fairly detailed model of the structure has been proposed that is based firmly on both X-ray crystallographic data and theoretical considerations (North *et al*. 1994).

https://doi.org/10.1017/S0033583500003516 Published online by Cambridge University Press

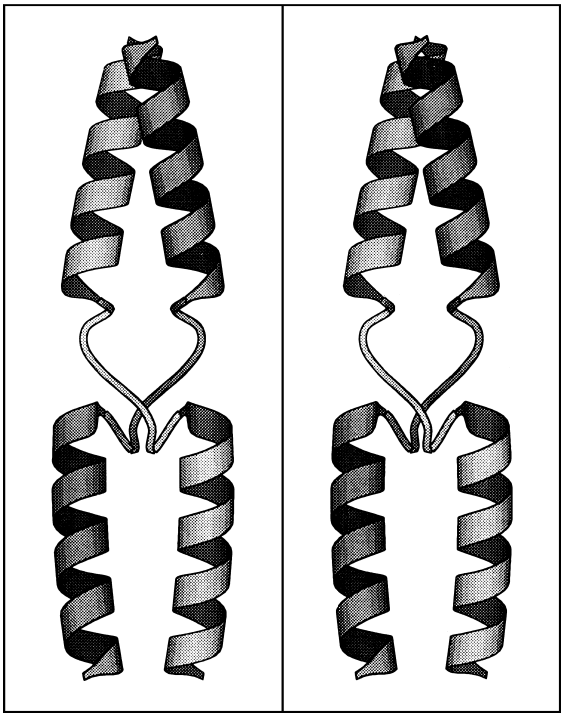


**Fig. 14.** Stereo pair showing the proposed conformation of link L2. North *et al*. (1994) used computer modelling to analyse this region and the work indicated the possibility that the C-terminal end of segment 2A and the N-terminal end of segment 2B would undergo something akin to a half-turn relative to one another. If correct, the link would be acting as a point at which the azimuth of the coiled-coil undergoes a large change over a short distance axially. Figure from Steinert *et al*. (1994).

#### 2.3.2.1 Links

Link L2 is highly conserved in length and is eight residues long for all intermediate filament proteins except nestin, where it is totally absent. The observation of conserved length would seem to argue strongly for a common structure of particular importance. What, therefore, would be such conformation be and what role would it have structurally and functionally? L2 links coiled-coil segment 2A to coiled-coil segment 2B. The phasing of these two segments is such that the internal apolar residues in the *a* and *d* positions in the coiled-coil in segment 2A would lie in external positions in segment 2B unless the intervening structure, that is link L2, reoriented the apolar seam appropriately. This must occur because the localization of numerous apolar residues in an aqueous milieu is energetically unfavourable. Therefore, it is certain that the link L2 must represent a region over which there is a significant change in the azimuthal orientation of the coiled-coil structure over a short distance axially (Fig. 14; North *et al*. 1994; Steinert *et al*. 1994). The question then arises as to the conformation adopted by the L2 sequence. Computer modelling, using the constraints of the known axial distance corresponding to L2 (North *et al*. 1994) and the connectivity to the coiled-coil segments either side of L2, lead to a family of models. The two 'extremes' will be described briefly but the actual structure might realistically therefore be expected to lie somewhere in between. The first model is one in which close contact is maintained between the two chains in link L2, and hydrogen bonds can be made between the chains to stabilize

https://doi.org/10.1017/S0033583500003516 Published online by Cambridge University Press

the structure. This results in the azimuth of the coiled-coil changing dramatically by close to 180°. The second model allows a little flexibility to the axial extent over which the apolar seam is brought back to optimum whilst simultaneously also attempting to maintain a structure something akin to an $\alpha$-helix. In this case the azimuth of the coiled-coil would change by about 30°. Either way there seems little doubt that the structural rationale for link L2 is to allow the azimuth of the coiled-coil to change significantly over a short distance axially.

Links L1 and L12 are much more difficult to model for several reasons (North *et al*. 1994). L1, for example, is quite variable in length and sequence. It is certainly possible, even probable, that L1 does not exhibit a common structure in different chain types. Certainly, its axial extent determined from crosslinking data, varies from the equivalent of about 16 residues in a coiled-coil conformation in epidermal keratin to only 11 residues for Type III and IV intermediate filaments. L12 sequences, in contrast, do show greater similarity across all intermediate filament chain types. The number of constituent residues varies between 16 and 22 but significant homology over a major portion of this region has been observed. The most recognizable feature in the sequences of L12 is an eight residue stretch in which apolar residues occur alternately with other residues. This feature is one that is usually indicative of a $\beta$-strand. The two chains in the molecule together therefore could form a minimal $\beta$-sheet of two chains with one face of the sheet being apolar. This would indicate that it is likely to co-assemble with a like sheet from a different (probably antiparallel) molecule in the filament. This model remains speculative, but nevertheless represents a reasonable starting conformation: much more data on the L1 and L12 links will be required before great confidence can be placed on the current models.

Herrmann & Aebi (1998) and Gill *et al*. (1990) have tackled the structures of L1 and L12 respectively from a different standpoint. In the former case an additional two amino acids were included in L1 and another residue was altered to make it more favourable to $\alpha$-helix formation. The vimentin filaments formed were only short and, furthermore, were only loosely associated. A second mutant, in which a pair of proline residues was inserted in L1 and another residue was replaced by a proline to generate a proline triplet, resulted in filament assembly that was indistinguishable from normal. Similarly, the insertion of additional residues in the L12 sequence of NF-L also resulted in relatively normal filament formation. The conclusion from these pieces of work is that the links convey some flexibility to the molecules which is important for higher orders of filament formation. Additional residues which do not compromise that flexibility are permitted without affecting filament formation, whereas engineered changes that give rise to less flexibility result in only imperfect filament formation. A mutation within segment L12 of keratin 5 in a patient suffering from a Kobner variant of epidermolysis bullosa simplex corresponds to a valine to alanine substitution in position 7 (Galligan *et al*. 1998). This position is absolutely conserved in Type II keratins and in almost all other intermediate filament chains too. It would seem evident that the role of this particular valine residue in L12 is an important one during filament assembly, a conclusion compatible with other evidence that the interaction of L12 segments is involved in filament assembly and that a fully overlapped arrangement of antiparallel molecules (later termed the $A_{12}$ mode: Section 3.3) leads to significant intermolecular overlap of L12 segments.

https://doi.org/10.1017/S0033583500003516 Published online by Cambridge University Press

#### 2.3.2.2 Stutter

The stutter in the centre of segment 2B, which is absolutely conserved in all IF proteins, represents a phase shift in the underlying heptad substructure. This corresponds to a deletion of three residues from an otherwise continuous heptad repeat, and is a feature seen in the sequences of many other α-fibrous proteins (Brown *et al*. 1996). Modelling by North *et al*. (1994) showed that it was possible to maintain the azimuthal relationship of the coiled-coil, but that severe distortion would occur in two or three hydrogen bonds sited one or two turns of the α-helix either side of the stutter. Consequently, it was physically conceivable that the effects of the stutter would be localized to a very short distance axially. Herrmann (1995), however, engineered an additional three residues into the sequence at the site of the stutter in vimentin. This resulted in the heptad repeat remaining in phase over the entire length of segment 2B, and generating what might be perceived as a stretch of perfect coiled-coil. The assembly of this mutant, however, while giving rise to filaments of normal diameter, was nonetheless very much shorter than normally observed, mostly corresponding to 'unit length' fragments corresponding to only the $A_{11}$ and $A_{22}$ assembly modes (see below). It appeared, therefore, that lateral aggregation was possible but that elongation of the filament was unable to proceed in the usual manner.

Later studies by Brown *et al*. (1996), using data from a wide range of α-fibrous proteins, showed that stutters were in fact a very common occurrence. Furthermore, they showed that the stammer, which corresponds to a deletion of four residues from an otherwise continuous heptad repeat, was also a feature in some sequences. All of the other observed breaks in the heptad continuity corresponded to stutters, stammers or skips. The skip residues, in which a single residue is inserted into an otherwise continuous heptad repeat, also correspond to a deletion of six residues. In other words one skip is structurally equivalent to two closely spaced stutters. In essence, therefore, only stutters and stammers interrupt heptad continuity in α-fibrous proteins, excluding those longer regions of sequence that are inserted in some cases. Evidence was presented by Brown *et al*. (1996) that stutters are likely to lead to a gradual local unwinding of the coiled-coil over a short distance axially. The result of this would be to present a different azimuthal aspect of the coiled-coil than would otherwise have occurred. Some crystallographic data from mannose binding protein and haemagglutinin are now available which support this concept. We can only speculate that this will either allow the assembly of the molecules to proceed more appropriately or that the function of the intermediate filaments (primarily via their head and tail terminal domains) would be enhanced. Stammers, on the other hand, which do not occur in intermediate filament proteins, are predicted to result in a local over-winding of the coiled-coil. So far there are no crystallographic data which pertain to this possibility.

### 2.3.3 Head and tail domains

The tertiary structure of a terminal domain from any intermediate filament chain is not known at present. Unfortunately, crystallization of these regions, or fractions thereof, has not been accomplished. Neither have NMR studies on fragments in solution yielded any insight into the structures of these functionally important domains. Needless to say this has been a source of much frustration to those in the field. The only insight so far gained arises from purely

https://doi.org/10.1017/S0033583500003516 Published online by Cambridge University Press

theoretical studies on the tail domains from Type Ia and Type IIa trichocyte keratin chains. Here, Parry & North (1998) proposed the secondary structure of a major portion of both chains. In the case of the Type I chains the proposed structure corresponded to a segment of poly-proline II helix with three residues per turn. Because the sequence quasi-repeat for this structure was of the form pro-cys-X it followed that the cysteine residues present would lie along one edge of the helix. The question then posed was what might such a structure interact with *in vivo*. The arguments put forward by Parry & North (1998) indicated that the optimum candidates that could provide an interactive element of structure were L1, L12, and the head and tail domains of both Type Ia and Type IIa chains. These regions were deduced as likely partners on the basis of their axial proximity in the surface lattice structures determined from crosslink and X-ray data (see Section 3.3). However, the proposed candidate was an identical structure from another Type I chain emanating from a (different) antiparallel molecule. Using computational molecular modelling it was shown that the cysteine residues in the poly-proline II helix would indeed be in positions where covalent disulphide bonds could be formed between the two domains. If this hypothesis is correct then the role of the terminal domains would be to provide additional means of stabilizing and specifying the molecular packing.

In the case of the Type IIa tail domain a very similar line of reasoning was used. It too indicated that its likely interaction partner would be the same domain, again originating from a different and antiparallel molecule. However, in this case the structure formed would consist of a pair of $\beta$-sheets stabilized by apolar residues lying along the interface between them. The individual sheets themselves would be twisted and right-handed, assuming that such this feature, already seen in many crystallized globular proteins, would occur here also. In summary, the model was proposed that a major fragment of the tail domain of a Type Ia chain would interact only with another identical segment from another molecule, and that a major fragment of a Type IIa chain would interact only with another identical segment from another molecule. It remains to be seen whether this observation is a general one for all intermediate filaments or whether it relates specifically to the trichocyte keratins.

The neurofilaments display long arms (approximately 60 nm in length and 3–5 nm in diameter) emanating from the core of the filaments at about 20–50 nm axial intervals (Hirokawa *et al*. 1984; Hisanaga & Hirokawa, 1988). These correspond to the tail domains of the NF-M and NF-H chains. Both regions are highly charged and are thus certain to have an elongated structures. Electron microscope observations indicate that these arms interact with microtubules and with those proteins associated with microtubules (Miyasaka *et al*. 1993). There are no direct data relating to the tertiary structure of these domains other than that relating to its overall shape.

The cytoplasmic intermediate filament proteins from cephalocordate *Branchiostoma* fall into four sub-families known as A, B, C and D (Riemer *et al*. 1998). The C-terminal domains of the C1 and C2 chains both contain a pair of regions with a heptad substructure extending over a distance of about 64 residues. A heptad repeat is, of course, a feature characteristic of a coiled-coil structure. In this way the tail domains of these two particular chains may be involved in the formation of an additional stretch of coiled-coil about 9·5 nm long. The role of this piece of coiled-coil is unclear, but it is unlikely to participate in a coiled-coil of the rod domain which forms the core of the intermediate filament. Rather, it may be designed to position appropriately other functional groups in the tail domain.

https://doi.org/10.1017/S0033583500003516 Published online by Cambridge University Press

## 2.4 Electron microscope observations

Transmission electron microscopy (TEM) and scanning transmission electron microscopy (STEM) have both provided important data on the structure of intermediate filaments. The STEM analyses, which relate essentially to the determination of the mass-per-unit-length of the intermediate filaments, will be discussed in Section 3.4. TEM, on the other hand, has yielded important information on the physical dimensions and axial regularities of the intermediate filaments, the possible subfibrillar structure of intermediate filaments and the ability of intermediate filaments to reassemble *in vitro* under various conditions.

The earliest high-resolution observations on intermediate filaments were those made on trichocyte keratins in stained sections of wool (Fig. 1*a*; Birbeck & Mercer, 1957; Rogers, 1964). The filaments appeared to be approximately circular in cross-section and had diameters of about 7–8 nm. Each of the filaments had the appearance of a central core surrounded by an annular ring of density, though the interpretation of these data was controversial. The filaments themselves were quasi-hexagonally packed in the paracortex of wool and quill, and were embedded in a matrix of other proteins that preferentially took up stain. Longitudinal sections confirmed that the filaments had a near constant diameter. Many years later intermediate filaments were also observed in various cell types. These entities were usually measured to be about 10 nm in diameter (and even somewhat larger in many cases). As more detail on the intermediate filaments was gleaned from all sources it has become apparent that the central rod domains pack closely together to form the annular ring that defines the filament and that the head and tail domains lie essentially (though not totally) on the surface of the intermediate filament where they interact with other proteins in the cell. Thus, the perceived (and measured) variations in filament diameter are likely to be a reflection of the different sizes of the terminal domains of the molecules in the intermediate filaments rather than to major changes in the modes of packing of the molecules themselves. It must be pointed out though that the trichocyte keratin intermediate filament molecules appear to differ in their packing in such a way that the molecular head-to-tail overlap characteristic of all of the other intermediate filaments is absent (Parry, 1995). This would be expected to result in a slightly smaller observed diameter for trichocyte intermediate filaments compared to that for intermediate filaments from other sources.

Axial beading is commonly but not always clearly observed on intermediate filaments (Fig. 15), and arises from a combination of the two structural features just described. The first, and probably the most important, is the regular disposition of the head and tail domains of the molecules on the filament surface. The second factor, and one likely to be of lesser importance, is the head-to-tail overlap that occurs between parallel molecules in the filament. The precise values of the axial repeats in the intermediate filaments have not been easy to ascertain for a variety of reasons, and a wide range of values has been reported. Sometimes this has arisen as a consequence of the magnification of the electron microscope not being determined experimentally either by taking an electron micrograph of a diffraction grating replica in the same series of micrographs as those taken of the intermediate filaments, or by including specimens of known dimensions on the same grid as the intermediate filaments. As a result it has been very difficult to know whether the variations in period that have been reported are a genuine reflection of different axial structures for different types of intermediate filaments or whether the range of values is merely a result of experimental variability. One study that has been carried out in a much more sophisticated manner was that undertaken by

https://doi.org/10.1017/S0033583500003516 Published online by Cambridge University Press

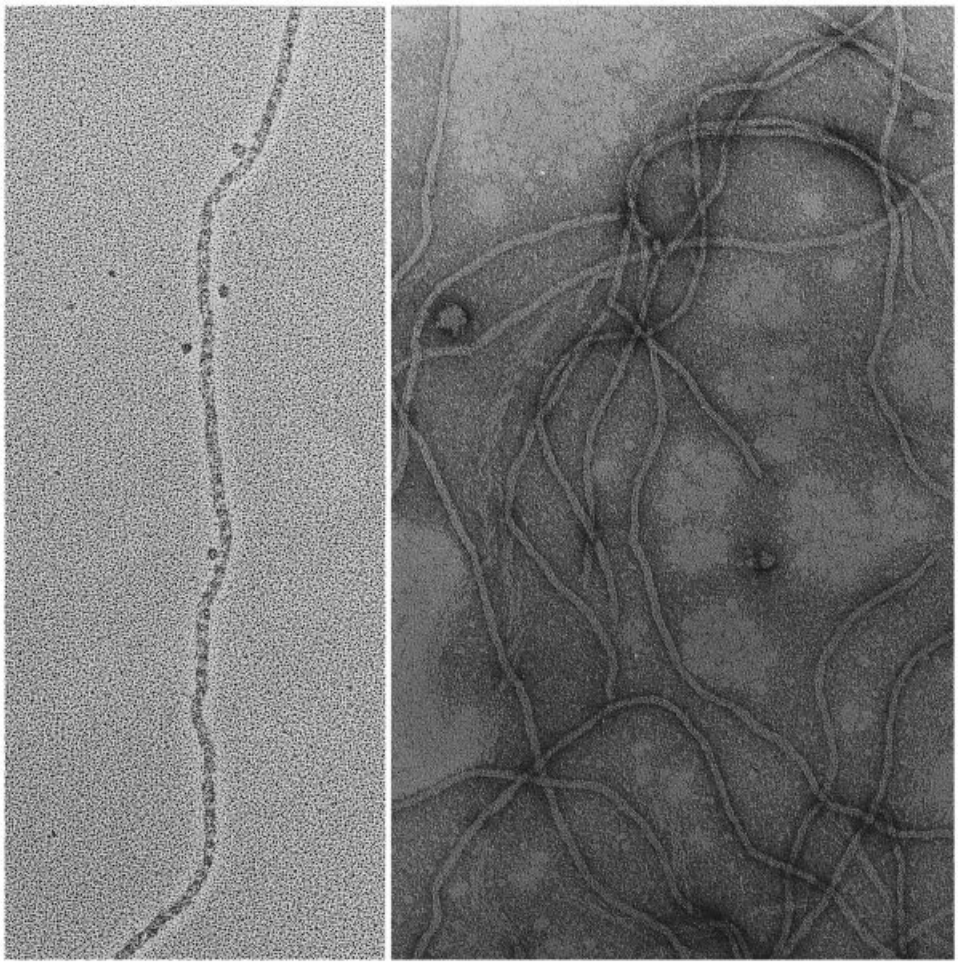


**Fig. 15.** (Left) Electron micrograph showing the axial beading of bovine epidermal cytokeratin intermediate filaments. The axial period deduced from a Fourier analysis of the density of the filaments after the latter were computationally straightened was 22·7 nm (Steven *et al*. 1985). (Right) Reconstituted trichocyte intermediate filaments. These were assembled from equimolar mixtures of bacterially expressed mouse Type Ia and IIa chains (Wang & Steinert, unpublished observations). By negative staining these filaments are morphologically similar to vimentin although they appear much smoother and lack any obvious axial beading.

Steven *et al*. (1985). Negatively stained micrographs displaying *in vitro* reassembled epidermal cytokeratin filaments were treated computationally, firstly to 'straighten' the filaments in order to minimize artifacts and, secondly, to allow the axial period to be determined accurately using Fourier techniques (Steven *et al*. 1985). The period thus determined was 22·7 nm. This value corresponds almost exactly to half the period (i.e. 22·6 nm) deduced from crosslinking studies reported in Section 3.3. The close correspondence between values deduced using fundamentally different techniques gives additional credence to both studies. Unfortunately, there have been no comparable studies on either Type III or Type IV intermediate filaments. Type V intermediate filaments (lamin) and some invertebrate cytoplasmic intermediate filaments are composed of longer molecules than other intermediate filaments, and the axial beading observed has a period of about 24·5 nm (Geisler *et al*. 1998).

https://doi.org/10.1017/S0033583500003516 Published online by Cambridge University Press

This additional 2 nm can be explained by the 42 residue insertion in their 1B rod domain segments. Furthermore, the axial period in trichocyte keratin intermediate filaments, as measured from X-ray diffraction, is different from that measured for epidermal cytokeratin intermediate filaments using electron microscopy, and that the period calculated for vimentin from crosslinking data is different yet again from the others. Thus altogether, it appears that the intermediate filaments adopt at least four unique structures, albeit closely related ones.

The existence or otherwise of subfibrillar structure of intermediate filaments has remained a controversial topic. Certainly, electron microscope observations have been made of structures 3–4.5 nm in diameter occurring in degraded or partially reconstituted intermediate filaments (see, for example, some early reports from Steinert, 1977; Krishnan *et al*. 1979; Aebi *et al*. 1983; Sauk *et al*. 1984; Eichner *et al*. 1985). The difficulty, however, has been the lack of reproducibility and the uncertainty that the observations genuinely relate to a defined unit of structure as distinct from non-specific aggregates. At the electron microscope level this issue has not been satisfactorily resolved though evidence of the existence of such subfilamentous structures does exist from other experimental techniques, such as crosslinking. Notably, however, the subfilaments seen in the electron microscope appear to contain the same axial beading as observed in the intact intermediate filament. The beading would therefore result not from the packing of subfilaments within the filament but from the structure of the subfilaments themselves.

There have been many observations of individual intermediate filament molecules or of fully overlapped pairs of molecules in the electron microscope using such techniques as negative staining, shadowing and antibody labelling (see, for example, Geisler *et al*. 1982a, 1985; Quinlan *et al*. 1984, 1986; Ip *et al*. 1985). In all cases there is general consistency in the reported values of the molecular length (about 45 nm) though small variations of up to 5 nm have been recorded. Transient electric birefringence measurements on desmin, vimentin and glial fibrillary acidic protein have also revealed the presence of structures with a strong dipole moment and a length of about 45 nm (Kooijman *et al*. 1995a, 1997). It is almost certain that these entities represent individual molecules. A molecular length of about 45 nm corresponds almost exactly to that predicted from both the crosslink data (Section 3.3) and the molecular modelling of the coiled-coil rod domain. In addition, there is a wealth of other electron microscope observations and transient electric birefringence data on *in vitro* assemblies of various intermediate filament proteins which indicate the presence of particles with lengths of about 65–70 nm and 90 nm. These must correspond respectively to a half-staggered arrangement of an antiparallel pair of molecules and to a consecutive half-staggering of three molecules arranged antiparallel with respect to one another.

Reconstitution of intermediate filament proteins into filamentous structures that are virtually indistinguishable from native intermediate filaments have been studied using electron microscopy. The first structural intermediate seems to be the 'full-length half-filaments' which contain about 16 chains (eight molecules) in cross-section. These are about 65–70 nm long (Herrmann & Aebi, 1998; Abumuhor *et al*. 1998; Geisler *et al*. 1998; Herrmann *et al*. 1999). Electric birefringence measurements also revealed particles derived from various intermediate filament proteins having no permanent dipole moment and with lengths of 63–68 nm (Kooijmann *et al*. 1995a, b, 1997). The observed length corresponds to antiparallel pairs of molecules that are approximately half-staggered with respect to one another. The electric birefringence data suggested further that the half-staggered arrays were formed using the antiparallel interaction between 1B segments (the $A_{11}$ mode of interaction,

https://doi.org/10.1017/S0033583500003516 Published online by Cambridge University Press

see below) rather than an antiparallel interaction between 2B segments (the $A_{22}$ mode, again see later). The next stage in filament formation results in 'full-length' filaments, presumably the result of the aggregation of two half-filaments. The measured diameter is greater than that observed in native intermediate filaments but there is believed to be a subsequent condensation phase where the loose packing becomes close packing (Herrmann & Aebi, 1998). The last stage of filament assembly results in longitudinal growth which proceeds by annealing of full-length filaments.

## 3. Assembly

The first step in the assembly of an intermediate filament is the formation of the two-chained coiled-coil molecule. Having specified the six types of IF chain the question arises as to the possible manners in which the chains may dimerize to form the molecule. Although all combinations have not been experimentally studied there is now a considerable body of data on this subject. In the case of the trichocyte keratins and cytokeratins it was clear that both Type I and Type II chains were involved in the molecular structure but it was not certain initially whether the molecules in the filament were both Type I/Type I homodimers and Type II/Type II homodimers or, conversely, whether only Type I/Type II heterodimers existed. Early theoretical data indicated that the heterodimer solution was more likely (Parry *et al*. 1977, 1985; Crewther *et al*. 1983) but definitive experimental evidence took some years to obtain. It was ultimately provided by a number of groups more or less simultaneously (Parry *et al*. 1985; Steinert, 1990; Hatzfeld & Weber, 1990b; Herrling & Sparrow, 1991). Evidence from several lines of work indicates that a Type I chain can assemble with a Type II chain (Hatzfeld & Weber, 1990b; Steinert, 1990; Herrling & Sparrow, 1991); that is, the molecules are heterodimers and the keratins are obligate copolymers. However, it is not yet clear whether trichocyte keratins and cytokeratins can coassemble *in vivo* or *in vitro*. In contrast, Type III chains can homodimerize. Alternatively, they may dimerize with a different Type III chain (Steinert *et al*. 1981; and see Parry & Steinert, 1995, and Rogers *et al*. 1996, for early references), with Type IV α-internexin (Steinert *et al*. 1999b), with each of the NF-L, NF-M and NF-H chains (summarized in Quinlan *et al*. 1994) (Table 1), or with the Type VI nestin chain (Steinert *et al*. 1999a). Type IV α-internexin and NF-L chains can homodimerize, or form heterodimers with each other and either NF-M or with NF-H (Ching & Liem, 1993; Lee *et al*. 1993; Athlan & Mushynski, 1997; Balin & Miller, 1995; Steinert *et al*. 1999b; Abumuhor *et al*. 1998). The latter chains, however, are apparently unable to form homodimers and exist only with NF-L as heterodimers (Carpenter & Ip, 1996; Leung & Liem, 1996). Likewise, the α-internexin chains can form heterodimers with Type VI nestin chains (Steinert *et al*. 1999a). Finally, of the likely Type VI chains, nestin is also capable of forming homodimers with itself and heterodimers with some other proteins *in vitro* (Steinert *et al*. 1999a) and probably *in vivo* as well (Marvin *et al*. 1998). The mechanism whereby heterodimers are specified remains unknown at this time, but appears to lie predominantly within the terminal domains rather than in the rod domain. The mode of Type V chain aggregation is unclear but evidence indicates that both homodimers (A/A and B/B) and heterodimers (A/B) may exist. Copolymerization with other chain types does not occur as a result of their different rod domain structures (that is, the difference in lengths of the 1B segments).

https://doi.org/10.1017/S0033583500003516 Published online by Cambridge University Press

**Table 1.** *Observed copolymerisation of chains into intermediate filaments*

| | D | V | GFAP | PN | INX | NF-L | NF-M | NF-H | NES |
|---|---|---|---|---|---|---|---|---|---|
| Desmin (D) | × | × | | | | | | | |
| Vimentin (V) | | × | × | × | × | × | × | × | × |
| Glial fibrillary acidic protein (GFAP) | | | × | | | | | | |
| Peripherin (PN) | | | | × | | × | × | × | |
| α-Internexin (INX) | | | | | × | × | × | × | × |
| Neurofilament light chain (NF-L) | | | | | | × | × | × | |
| Neurofilament medium chain (NF-M) | | | | | | | No | No | |
| Neurofilament heavy chain (NF-H) | | | | | | | | No | |
| Nestin (NES) | | | | | | | | | × |

### 3.1 Role of the coiled-coil rod domain

It had been observed from the X-ray diffraction patterns of wool and quill that the α-helices of the trichocyte keratins must lie approximately parallel to the long axis of the intermediate filament. Of course, such a conclusion is also the only one compatible with molecules about 45 nm long constituting filaments only 10 nm in diameter. What information lies within the coiled-coil rod domain that specifies the molecular packing? The answer in part lies in the regular linear distribution of acidic and basic residues that occur in segments 1B and 2B. The acidic residues (aspartic acid and glutamic acid) have a highly significant periodicity of about 9·54 residues in segment 1B and a slightly longer period of about 9·84 residues in segment 2B (Parry *et al*. 1977; McLachlan & Stewart, 1982; Parry & Fraser, 1985). Similarly, the basic residues (arginine and lysine) have virtually identical periods in each of the two segments. Furthermore, the periods in the acidics and basics are out of phase with one another, which means that there are alternating bands of acidic and basic residues along both segment 1B and segment 2B. These observations immediately indicate that molecular assembly is likely to be specified to some extent by the maximisation of intermolecular ionic interactions. Thus, bands of acidic residues in one molecule would be expected to be axially aligned with bands of basic residues in an adjacent molecule. In order to confirm the role of charge interactions between molecules in the formation of the filament, Meng *et al*. (1994) used deletion mutagenesis to prepare mutant vimentin lacking two, three and four heptads in segment 1B. The assembly of these proteins was studied using crosslinking, *in vitro* assembly and by DNA transfection into intermediate filament null cells. The results provided strong support for the hypothesis that charge interactions between molecules are indeed important and that they not only specify the assembly but, additionally, provide stability to the intermediate filament. The yeast two-hybrid system has also provided strong experimental support for the concept of strongly interacting 1B segments and strongly interacting 2B segments in vimentin (Meng *et al*. 1996).

An analysis of the (multiple) modes of aggregation permitted via intermolecular interactions has been studied in detail by Fraser *et al*. (1985, 1986). The data do not clearly distinguish between parallel and antiparallel assemblies, neither do they unequivocally delineate which of the family of intermolecular staggers will correspond to the ones found

https://doi.org/10.1017/S0033583500003516 Published online by Cambridge University Press

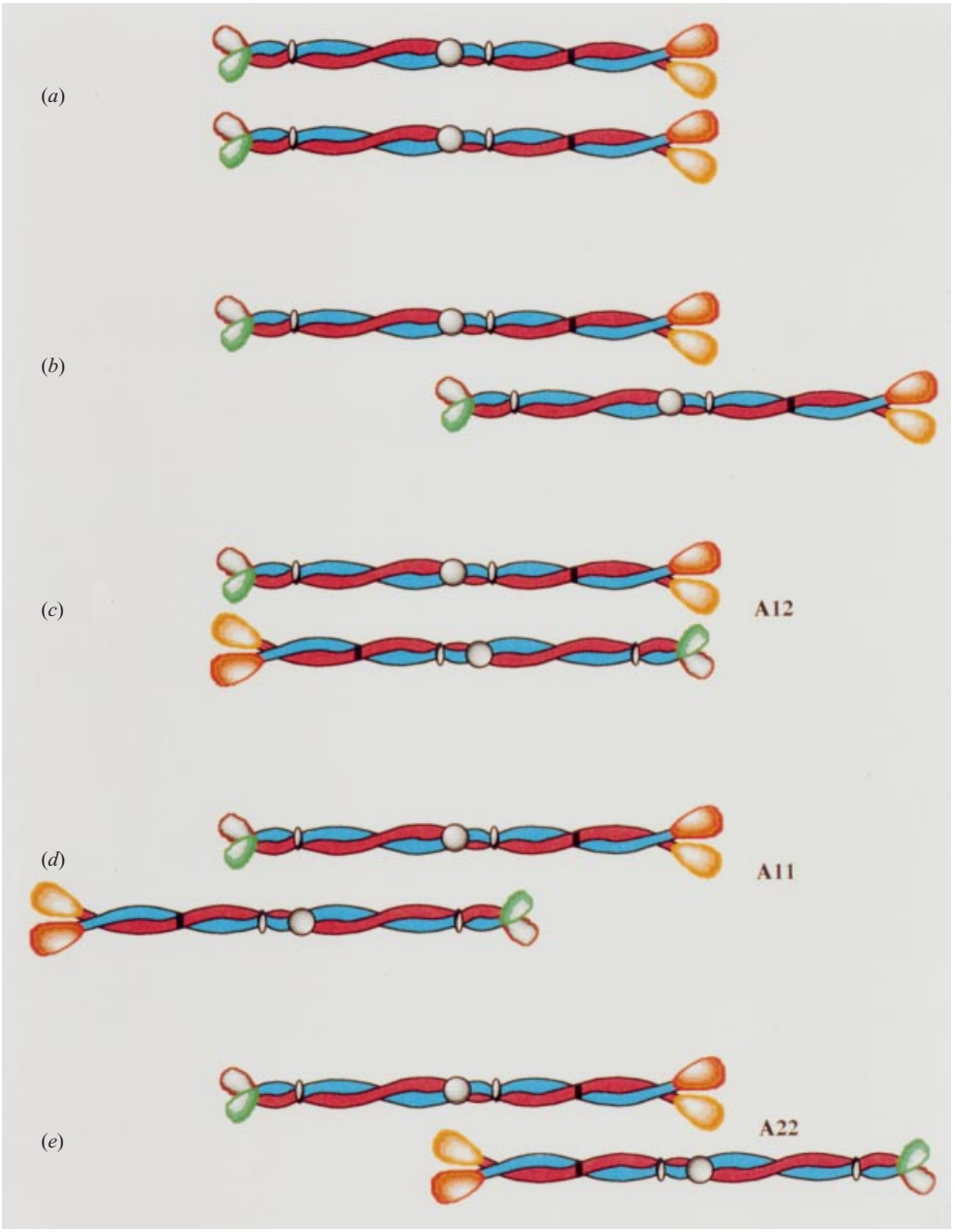



**Fig. 16.** On the assumption that the observed periodicity in the linear distribution of the acidic residues and the basic residues in segments 1B and 2B will lead to a maximization of intermolecular ionic interactions between either two 1B segments, two 2B segments or between a 1B and a 2B segment, only five generic models for the packing of intermediate filament molecules are possible. Two of these relate to parallel modes (*a*, *b*) and three lead to antiparallel modes (*c*, *d*, *e*). Each of the five cases is represented by a family of closely related structures in which the relative axial staggers between the molecules will differ by a multiple of 9·54–9·85 residues. Figure after Crewther *et al.* (1983).

https://doi.org/10.1017/S0033583500003516 Published online by Cambridge University Press

*in vivo*. Nonetheless, they have provided powerful constraints in the model building process for the five generic modes of assembly deemed possible (Fig. 16) (Crewther *et al*. 1983). These corresponded to three antiparallel arrangements (one with the 1B segments overlapped, one with the 2B segments overlapped and one with the entire molecules overlapped) and two parallel arrangements (one with the entire molecules overlapped and one with segments 1B and 2B overlapped). As will be shown later, only the first three antiparallel molecular arrangements have been subsequently confirmed by crosslink analysis (Fig. 17) (Section 3.3). Experimental data from Steinert (1991a, b), obtained prior to the crosslinking studies, had also provided strong evidence that the intermediate filaments were constructed from assemblies of half-staggered antiparallel molecules. Further, it was shown that a pair of fully overlapped antiparallel molecules could not lead to intermediate filament formation, though later occurrence of these forms did lead to incorporation into intermediate filaments during the growth phase.

Another insight into the role of the coiled-coil domain on intermediate formation has been provided by Munoz Marmol *et al*. (1998). They characterized a mutation in the desmin gene that resulted in the deletion of seven residues from the coiled-coil portion of the rod domain. The mutant protein was unable to form intermediate filaments *in vitro* or intermediate filament networks in cultured muscle cells. The patient suffered from severe generalized myopathy and had aberrant accumulation of desmin intermediate filaments within the sarcoplasm. This work illustrates the necessity of a full coiled-coil rod domain for normal intermediate filament formation to proceed.

While discussing the features of the coiled-coil rod domain that are important in the aggregation of molecules into filaments it is appropriate to mention that peptides derived from the rod domain are capable of causing filament disaggregation *in vivo* and *in vitro*. Two sets of peptides have been identified. One set corresponds to sequences of the helix termination motif toward the end of segment 2B (Steinert *et al*. 1993a; Hatzfeld & Weber, 1992; Kouklis *et al*. 1993) and the other corresponds to sequences encompassing the helix initiation motif near the beginning of segment 1A (Goldman *et al*. 1996). Peptides from either region have essentially the same effect in that they are capable of disassembling preformed intermediate filaments into oligomers and individual chains, or inhibiting filament assembly. Upon microinjection into living fibroblasts or keratinocytes, the peptides disassemble the intermediate filament networks, resulting in dramatic changes in cell shape, apparently owing to secondary effects on microtubule and microfilament organization. The process is completely reversible: withdrawal of the peptides restores filaments and their elaborate cytoplasmic arrays within an hour. Such observations provide evidence on the important role of intermediate filaments in determination of the mechanical properties and shape of cells. More specifically, it highlights the importance of the highly conserved helix initiation and termination motifs. How do these peptides cause disassembly in the manner observed? It is believed that the peptides competitively inhibit the intermediate filament molecules for binding sites which facilitate filament formation and/or stability. It is also worth mentioning here in passing that most mutations in keratin diseases also occur in or near the helix initiation or termination regions, indicating that the precise three-dimensional structure of these regions is critical for intermediate filament structure and function (and see Section 6).

https://doi.org/10.1017/S0033583500003516 Published online by Cambridge University Press

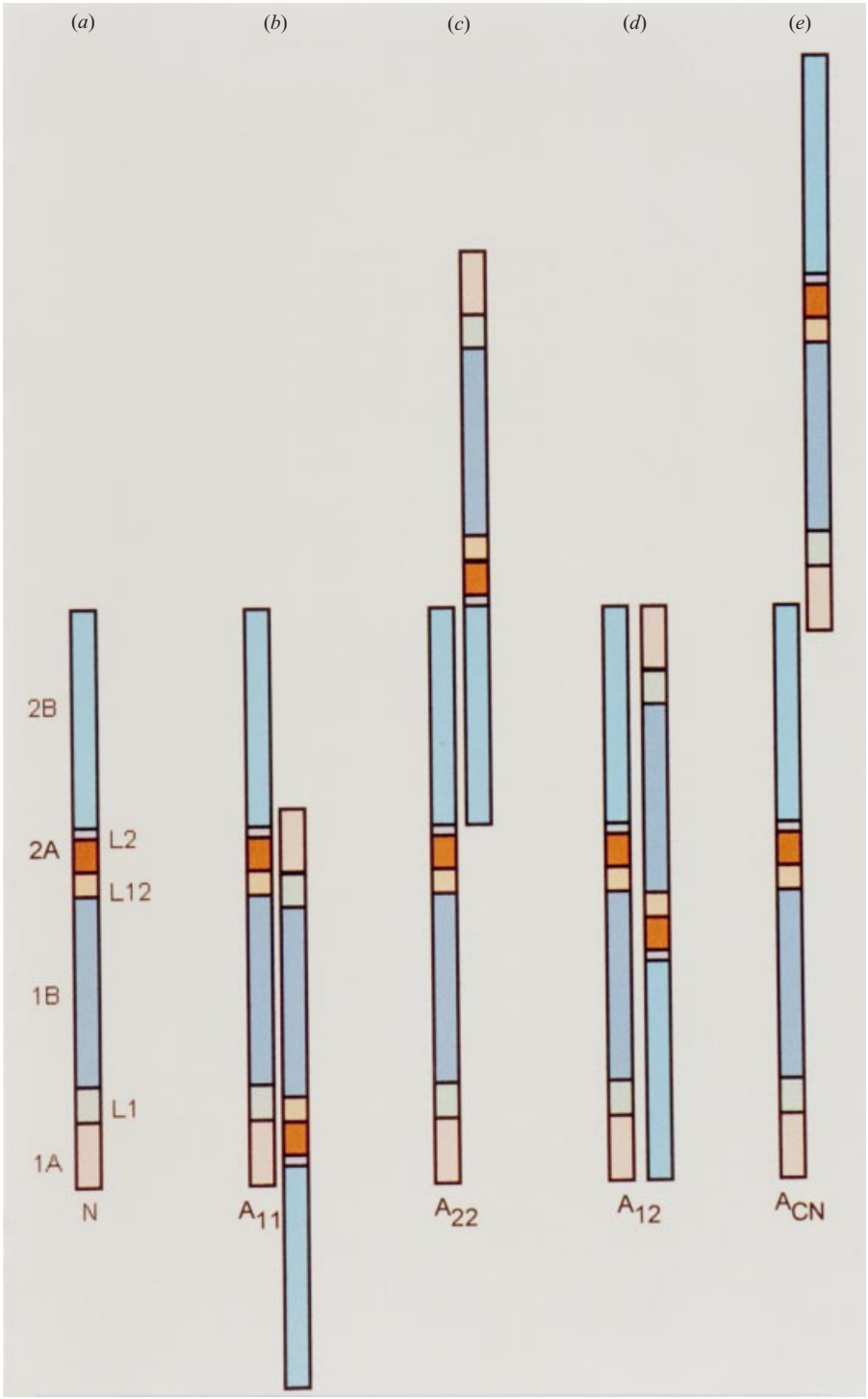



**Fig. 17.** Schematic figure illustrating the basic domain structure of an intermediate filament molecule (N, left), the three modes of aggregation involving antiparallel molecules ($A_{11}$, $A_{22}$ and $A_{12}$), and the mode of aggregation involving a head-to-tail overlap between parallel molecules ($A_{CN}$).

https://doi.org/10.1017/S0033583500003516 Published online by Cambridge University Press

### 3.2 Role of end domains

Evidence has been presented that proteolytic deletion of the head or tail domains of intermediate filament chains does not prevent formation of either a coiled-coil dimer (that is, a molecule) or tetramer (that is, a pair of molecules) (Kaufmann *et al*. 1985). Additionally, the heterodimeric structure of keratin molecule is maintained (see, for example, Woods & Inglis, 1984; Parry *et al*. 1985; Hatzfeld *et al*. 1987; Magin *et al*. 1987). However, filament formation is compromised and this implies that at least a portion of the terminal domains is required for stabilizing small oligomers and/or possibly larger assemblies of molecules (Geisler *et al*. 1982a). Using molecular cloning techniques, Kaufmann *et al*. (1985) showed that the deletion of the head domain of desmin prevented filament formation. Similarly, deletion of the head domains of vimentin (Herrmann *et al*. 1996a) or cytokeratins (Hatzfeld & Burba, 1994), allowed the formation of single molecules and molecular pairs, but discrete entities that were larger than tetramers did not form. Likewise, Sun *et al*. (1997) showed that NF-H chains with deleted head domains could no longer assemble with vimentin, NF-L or NF-M. More extensive related studies with α-internexin and various neurofilament chains also lacking various portions of their terminal domains have generated similar data (Ching & Liem, 1998). Thus all of these studies have shown a consistent picture that a major portion of the head domain must be involved in filament formation, whilst its absence apparently does not affect molecular or tetramer formation.

Similar studies by Quinlan *et al*. (1989), in which the entire tail domain of glial fibrillary acidic protein was deleted, also resulted in no discernible filament formation. However, when only the latter half of the tail domain of desmin was deleted, filaments could be produced. There would seem to be a special role in filament formation for those regions located immediately adjacent to both ends of the rod domain (cf. the role of H1 and H2 in Type II epidermal keratins). Further studies have been undertaken by Hatzfeld & Weber (1990a) in which they produced a mutant of K8 (a Type II chain) that had just six residues in its tail domain. Filament formation with its usual Type I partner (K18) occurred nonetheless, which indicates either that the tail domain as a whole is not important in filament formation or that the key region involved is very short indeed, possibly as few as six residues. In fact, one of the smallest Type I keratin chains (K19) normally has a very short tail domain (13 residues), but it can assemble with many Type II chains into normal-appearing filaments, albeit somewhat shorter than usual (Fradette *et al*, 1998). Other studies (Rogers *et al*. 1995; Herrmann *et al*. 1996a) showed that various tail domain truncated mutants of vimentin did not affect filament formation but that they did have an effect on cellular organization. Comparable studies have been undertaken for the neurofilament chains NF-L, NF-M and NF-H (Sun *et al*. 1997). Aggregation studies of NF-L and NF-M chains with mutants of NF-H chains which lack either their head or their tail domains, in the presence or absence of vimentin, have indicated that NF-H prefers NF-L to either NF-M or vimentin as its heterodimer partner. Furthermore, the head domain of NF-H was shown to have a defining role in the production of NF-L/NF-H filaments, whereas the tail domains were intimately involved in allowing an extensive network of NF-L/NF-H filaments to be formed (see also the work of Ching & Liem, 1998, on the assembly of mutants of α-internexin with mutants of neurofilament chains). Interestingly, in about 5% of the human population a polymorphic variant or mutant Type I hair keratin hHa1 is found which lacks the complete tail domain. However, both the mutant protein and the wild type assemble with a Type II partner and

https://doi.org/10.1017/S0033583500003516 Published online by Cambridge University Press

produce normal intermediate filaments *in vitro* and presumably *in vivo* as well (Winter *et al*. 1997c).

In addition, there is evidence on the structural/functional roles played by portions (or subdomains) within the end domains in Types I/II keratin intermediate filament chains. These include: the highly conserved H1 and H2 subdomains in cytokeratins (though primarily in the Type II chains); the tail domains in the Type Ia and Type IIa trichocyte keratin chains; and regions in the head domains of the trichocyte keratins. In the Type II cytokeratins, the H1 and H2 subdomains are 36 and 20 residues long respectively. From their size and high degree of homology, it seems likely that they are designed to play an important part *in vivo*. In contrast, the H1 subdomain in Type I cytokeratins is only seven residues long and the H2 subdomain does not seem to exist at all as a finite entity. Consequently, it would seem less likely that H1 will play a significant role in the aggregation process (if indeed it has any role at all). Three pieces of evidence have been presented which provide direct experimental evidence for the importance of the H1 subdomain (and perhaps by analogy the H2 subdomain also) in Type IIb keratin chains. Based on formation of disulphide crosslinks, it was concluded that these sequences specify the axial alignment of the first two or three molecules during keratin intermediate filament assembly (Steinert & Parry, 1993; Steinert *et al*. 1993a; Hatzfeld & Burba, 1994). It would seem also that the interactions involving H1 provides important stabilization for this oligomer. Second, it was observed that a synthetic peptide with a sequence corresponding to that of H1 chains was able to disassemble pre-existing keratin intermediate filaments, or prevent filament assembly *in vitro* (Chipev *et al*. 1992; Steinert & Parry, 1993; Steinert *et al*. 1993a). Also, when micro-injected into living epithelial cells, the H1 peptide caused reversible disassembly of the keratin intermediate filament network, but not vimentin networks of fibroblasts (Steinert *et al*. 1993d). As for the helix initiation or termination synthetic peptides, the most likely explanation of this phenomenon is competition between the peptide and the H1 subdomains in the intact epidermal keratin molecules for binding sites in the filament. Furthermore, this conclusion was substantiated by Hatzfeld & Burba (1994) who tackled the issue by studying the effect on filament formation of deletion mutants involving the H1 subdomain. Finally, several cases of keratin diseases have been attributed to mutations in the H1 subdomain of Type IIb cytokeratins, but interestingly, no cases have been found yet to involve the H2 subdomain or H1 subdomain of Type Ib cytokeratins (see Section 6).

Note that the trichocyte keratins possess similar H1 and H2 subdomains in equivalent positions (Section 2.1.1). Presumably, therefore, they have similar functions *in vivo*, namely to specify and stabilise the axial alignment of the first two or three molecules during intermediate filament assembly.

There is another region in the head domain of Type IIa trichocyte keratins that has a nine residue quasi-repeat that occurs tandemly four times:

G-I-S-C-Y-R-G-L-T

G-G-F-G-S-R-S-V-C

G-G-F-R-A-G-S-C-G

R-S-F-G-Y-R-S-G-G

There is no direct evidence, however, as to the conformation adopted by the consensus sequence G-G-F-G-Y-R-S-X-G, but a region rich in aromatic and glycine residues might

https://doi.org/10.1017/S0033583500003516 Published online by Cambridge University Press

**Table 2.** *Equations derived from crosslink data for intermediate filament proteins*

| Crosslink | $A_{12}$ | $A_{11}$ | $A_{22}$ | L1 | L12 | L2 | Value |
|---|---|---|---|---|---|---|---|
| K5/K14 | | | | | | | |
| K5:L1-6/K5:2B-83 | 1 | 0 | 0 | −0·5 | 0 | 0 | −4 |
| K5:1B-40/K14:2B-23 | 1 | 0 | 0 | −1 | 0 | 0 | −24 |
| K5:1B-49/K5:2B-23 | 1 | 0 | 0 | −1 | 0 | 0 | −15 |
| K5:B-66/K14:1B-6 | 1 | 0 | 0 | −1 | 0 | 0 | −15 |
| K5:1A-10/K5:2B-112 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| K5:2A-10/K14:1B-90 | 1 | 0 | 0 | −1 | 0 | 1 | −6 |
| K5:1B-49/K14:2B-23 | 1 | 0 | 0 | −1 | 0 | 0 | −15 |
| K14:1B-82/K14:2A-19 | 1 | 0 | 0 | −1 | 0 | 1 | −5 |
| K5:2B-71/K14:1B-6 | 1 | 0 | 0 | −1 | 0 | 0 | −10 |
| K5:1A-10/K14:2A-19 | 0 | 1 | 0 | 0 | 0 | 1 | −112 |
| K5:1B-49/K5:1B-49 | 0 | 1 | 0 | −1 | 1 | 1 | −109 |
| K5:1B-40/K5:1B-61 | 0 | 1 | 0 | −1 | 1 | 1 | −106 |
| K14:1B-6/K14:1B-89 | 0 | 1 | 0 | −1 | 1 | 1 | −112 |
| K5:1B-61/K5:1B-42 | 0 | 1 | 0 | −1 | 1 | 1 | −104 |
| K14:1B-6/K14:1B-90 | 0 | 1 | 0 | −1 | 1 | 1 | −111 |
| K5:1B-61/K14:1B-42 | 0 | 1 | 0 | −1 | 1 | 1 | −104 |
| K5:1A-17/K5:2A-10 | 0 | 1 | 0 | 0 | 0 | 1 | −114 |
| K14:1B-14/K14:1B-82 | 0 | 1 | 0 | −1 | 1 | 1 | −111 |
| K5:1B-89/K14:1B-6 | 0 | 1 | 0 | −1 | 1 | 1 | −112 |
| K5:2B-44/K5:2B-83 | 0 | 0 | 1 | −1 | −1 | −1 | 160 |
| K5:2B-23/K5:2B-100 | 0 | 0 | 1 | −1 | −1 | −1 | 156 |
| K5:2B-44/K14:2B-81 | 0 | 0 | 1 | −1 | −1 | −1 | 158 |
| K5:2B-45/K5:2B-83 | 0 | 0 | 1 | −1 | −1 | −1 | 161 |
| K1/K10 | | | | | | | |
| K1:1A-10/K1:2B-111 | 1 | 0 | 0 | −1 | 0 | 0 | −1 |
| K10:1B-6/K1:2B-63 | 1 | 0 | 0 | −1 | 0 | 0 | −18 |
| K1:1B-49/K1:2B-14 | 1 | 0 | 0 | −1 | 0 | 0 | −24 |
| K1:1B-49/K1:2B-23 | 1 | 0 | 0 | −1 | 0 | 0 | −15 |
| K1:1B-61/K10:2B-4 | 1 | 0 | 0 | −1 | 0 | 0 | −22 |
| K1:1B-77/K1:L2-2 | 1 | 0 | 0 | −1 | 0 | 0·750 | −10 |
| K1:1B-77/K10:2A-19 | 1 | 0 | 0 | −1 | 0 | 1 | −9 |
| K1:1A-10/K10:2B-4 | 0 | 1 | 0 | 0 | 0 | 0 | −108 |
| K10:1A-17/K1:2B-4 | 0 | 1 | 0 | 0 | 0 | 0 | −101 |
| K10:1B-6/K10:1B-89 | 0 | 1 | 0 | −1 | 1 | 1 | −112 |
| K1:1B-19/K1:1B-77 | 0 | 1 | 0 | −1 | 1 | 1 | −111 |
| K1:1B-49/K1:1B-49 | 0 | 1 | 0 | −1 | 1 | 1 | −109 |
| K1:1B-61/K10:1B-42 | 0 | 1 | 0 | −1 | 1 | 1 | −104 |
| K1:1A-10/K1:L2-2 | 0 | 1 | 0 | 0 | 0 | 0·750 | −112 |
| K10:1A-17/K1:L2-2 | 0 | 1 | 0 | 0 | 0 | 0·750 | −105 |
| K10:2B-4/K1:2B-111 | 0 | 0 | 1 | −1 | −1 | −1 | 148 |
| K1:2B-14/K1:2B-111 | 0 | 0 | 1 | −1 | −1 | −1 | 158 |
| K10:2B-23/K1:2B-100 | 0 | 0 | 1 | −1 | −1 | −1 | 156 |
| K1:2B-62/K1:2B-62 | 0 | 0 | 1 | −1 | −1 | −1 | 157 |
| K1:2B-43/K1:2B-82 | 0 | 0 | 1 | −1 | −1 | −1 | 158 |
| K1:2B-43/K1:2B-86 | 0 | 0 | 1 | −1 | −1 | −1 | 162 |
| Desmin (D) | | | | | | | |
| D:1A-17/D:L12-15 | 0 | 1 | 0 | 0 | 0·063 | 1 | −124 |
| D:1A-1/D:2A-19 | 0 | 1 | 0 | 0 | 0 | 1 | −121 |
| D:2A-19/D:2B-112 | 0 | 0 | 1 | −1 | −1 | 0 | 145 |
| Vimentin (V) | | | | | | | |
| V:2B-83/V:L1-5 | 1 | 0 | 0 | −0·625 | 0 | 0 | −4 |
| V:1B-90/V:L1-1 | 0 | 1 | 0 | −0·125 | 1 | 1 | −117 |
| V:1B-90/V:2A-19 | 1 | 0 | 0 | −1 | 0 | 1 | 3 |

[*continued overleaf*

https://doi.org/10.1017/S0033583500003516 Published online by Cambridge University Press

**Table 2** (*cont.*)

| Crosslink | $A_{12}$ | $A_{11}$ | $A_{22}$ | L1 | L12 | L2 | Value |
|---|---|---|---|---|---|---|---|
| V:1B-42/V:1B-42 | 0 | 1 | 0 | −1 | 1 | 1 | −123 |
| V:1A-17/V:L12–15 | 0 | 1 | 0 | 0 | 0·063 | 1 | −124 |
| V:1A-1/V:2A-19 | 0 | 1 | 0 | 0 | 0 | 1 | −121 |
| V:1B-89/V:L1–5 | 0 | 1 | 0 | −0·625 | 1 | 1 | −118 |
| V:1B-89/V:L1–1 | 0 | 1 | 0 | −0·125 | 1 | 1 | −118 |
| V:1B-90/V:L1–1 | 0 | 1 | 0 | −0·125 | 1 | 1 | −117 |
| V:2A-19/V:2B-112 | 0 | 0 | 1 | −1 | −1 | 0 | 145 |
| V:2B-83/V:2B-23 | 0 | 0 | 1 | −1 | −1 | −1 | 139 |
| V:2B-100/V:2B-2 | 0 | 0 | 1 | −1 | −1 | −1 | 135 |
| α-Internexin (INX) | | | | | | | |
| INX:1A-17/INX:2B-112 | 1 | 0 | 0 | 0 | 0 | 0 | 7 |
| INX:1B-76/INX:2B-04 | 1 | 0 | 0 | −1 | 0 | 0 | −7 |
| INX:1B-77/INX:2B-04 | 1 | 0 | 0 | −1 | 0 | 0 | −6 |
| INX:1B-77/INX:2B-02 | 1 | 0 | 0 | −1 | 0 | 0 | −8 |
| INX:1B-76/INX:2B-02 | 1 | 0 | 0 | −1 | 0 | 0 | −9 |
| INX:1A-01/INX:2A-19 | 0 | 1 | 0 | 0 | 0 | 1 | −121 |
| INX:1A-17/INX:L12-18 | 0 | 1 | 0 | 0 | 0·053 | 1 | −124 |
| INX:2B-04/INX:2B-100 | 0 | 0 | 1 | −1 | −1 | −1 | 137 |
| INX:2B-02/INX:2B-100 | 0 | 0 | 1 | −1 | −1 | −1 | 135 |
| INX:2A-19/INX:2B-112 | 0 | 0 | 1 | −1 | −1 | 0 | 145 |
| α-Internexin/Vimentin (INX/V) | | | | | | | |
| INX:1A-17/V:2B-112 | 1 | 0 | 0 | 0 | 0 | 0 | 7 |
| INX:2B-100/V:1A-26 | 1 | 0 | 0 | 0 | 0 | 0 | 4 |
| INX:1B-77/V:2B-04 | 1 | 0 | 0 | −1 | 0 | 0 | −6 |
| INX:1B-77/V:2B-02 | 1 | 0 | 0 | −1 | 0 | 0 | −8 |
| INX:1B-76/V:2B-02 | 1 | 0 | 0 | −1 | 0 | 0 | −9 |
| INX:1B-58/V:2B-23 | 1 | 0 | 0 | −1 | 0 | 0 | −6 |
| INX:2A-19/V:1B-89 | 1 | 0 | 0 | −1 | 0 | 1 | 2 |
| INX:2A-19/V:1B-90 | 1 | 0 | 0 | −1 | 0 | 1 | 3 |
| INX:2B-83/V:L1-05 | 1 | 0 | 0 | −0·625 | 0 | 0 | −4 |
| INX:1A-01/V:2A-19 | 0 | 1 | 0 | 0 | 0 | 1 | −121 |
| INX:1A-17/V:L12-15 | 0 | 1 | 0 | 0 | 0·063 | 1 | −124 |
| INX:1B-58/V:1B-22 | 0 | 1 | 0 | −1 | 1 | 1 | −127 |
| INX:2B-100/V:2B-04 | 0 | 0 | 1 | −1 | −1 | −1 | 137 |
| INX:2B-100/V:2B-02 | 0 | 0 | 1 | −1 | −1 | −1 | 135 |
| INX:2B-83/V:2B-23 | 0 | 0 | 1 | −1 | −1 | −1 | 139 |
| INX:2B-112/V:2A-19 | 0 | 0 | 1 | −1 | −1 | 0 | 145 |
| Unique data from vimentin, α-internexin and vimentin/α-internexin | | | | | | | |
| VINX:1A-17/VINX:2B-112 | 1 | 0 | 0 | 0 | 0 | 0 | 7 |
| VINX:1B-77/VINX:2B-04 | 1 | 0 | 0 | −1 | 0 | 0 | −6 |
| VINX:1B-77/VINX:2B-02 | 1 | 0 | 0 | −1 | 0 | 0 | −8 |
| VINX:1B-76/VINX:2B-02 | 1 | 0 | 0 | −1 | 0 | 0 | −9 |
| VINX:1B-58/VINX:2B-23 | 1 | 0 | 0 | −1 | 0 | 0 | −6 |
| VINX:1A-26/VINX:2B-100 | 1 | 0 | 0 | 0 | 0 | 0 | 4 |
| VINX:1B-89/VINX:2A-19 | 1 | 0 | 0 | −1 | 0 | 1 | 2 |
| VINX:1B-90/VINX:2A-19 | 1 | 0 | 0 | −1 | 0 | 1 | 3 |
| VINX:L1–05/VINX:2B-83 | 1 | 0 | 0 | −0·625 | 0 | 0 | −4 |
| VINX:1B-76/VINX:2B-04 | 1 | 0 | 0 | −1 | 0 | 0 | −7 |
| VINX:1A-17/VINX:L12-18 | 0 | 1 | 0 | 0 | 0·053 | 1 | −124 |
| VINX:1B-89/VINX:L1-05 | 0 | 1 | 0 | −0·625 | 1 | 1 | −118 |
| VINX:1B-89/VINX:L1-01 | 0 | 1 | 0 | −0·125 | 1 | 1 | −118 |
| VINX:1B-90/VINX:L1-01 | 0 | 1 | 0 | −0·125 | 1 | 1 | −117 |
| VINX:1B-42/VINX:1B-42 | 0 | 1 | 0 | −1 | 1 | 1 | −123 |
| VINX:1A-01/VINX:2A-19 | 0 | 1 | 0 | 0 | 0 | 1 | −121 |

https://doi.org/10.1017/S0033583500003516 Published online by Cambridge University Press

**Table 2** (*cont.*)

| Crosslink | $A_{12}$ | $A_{11}$ | $A_{22}$ | L1 | L12 | L2 | Value |
|---|---|---|---|---|---|---|---|
| VINX:1A-17/VINX:L12-15 | 0 | 1 | 0 | 0 | 0·063 | 1 | −124 |
| VINX:1B-58/VINX:1B-22 | 0 | 1 | 0 | −1 | 1 | 1 | −127 |
| VINX:2B-04/VINX:2B-100 | 0 | 0 | 1 | −1 | −1 | −1 | 137 |
| VINX:2B-02/VINX:2B-100 | 0 | 0 | 1 | −1 | −1 | −1 | 135 |
| VINX:2A-19/VINX:2B-112 | 0 | 0 | 1 | −1 | −1 | 0 | 145 |
| VINX:2B-83/VINX:2B-23 | 0 | 0 | 1 | −1 | −1 | −1 | 139 |

adopt a flexible structure (courtesy of the glycine residues) that would allow the aromatic residues to interact strongly amongst themselves or with aromatics from a similar region originating from a different molecule. The region is free of proline residues which might influence chain mobility. Alternatively, there is the possibility that the motif will form a $\beta$-strand and $\beta$-turn, with glycine residues located predominantly towards the centre of the strand and the larger aromatic and apolar residues sited immediately adjacent to the $\beta$-turn.

Part of the tail domain of the trichocyte keratins have predicted structures (Parry & North, 1998) based on the poly-proline II structure (Type I chains) and a twisted four-strand antiparallel $\beta$-sheet with an apolar face (Type II chains). Arguments have been presented that such structures are capable of aggregating only with the same element of structure from a different (antiparallel) molecule. If this argument can be substantiated from other experimental data then it would seem clear that a major role of the tail domains of these keratins is to specify and stabilize particular oligomeric assemblies, possibly in the same way as achieved by the H subdomains.

In summary, the bulk of the data currently available support the concept that the end domains are not absolutely required for the earliest steps in intermediate filament assembly, that is, formation of molecules or pairs of molecules (tetramers). However, considerable portions of the head domains are required for the higher-order levels of assembly of molecules into filaments. On the other hand, only a small portion (at the very most) of the tail domains are involved in filament formation, but where there is a contribution, the vital segments, especially in Type I, II and III chains, would probably be those that lie immediately adjacent to the rod domain.

### 3.3 Experimentally induced crosslinks and modes of assembly

X-ray diffraction techniques have been employed to determine the helix parameters for intermediate filaments from the trichocyte keratins (Fraser & MacRae, 1971, 1983, 1985). These data refer specifically, however, to the surface lattice on which groupings of molecules are positioned. Determination of the spatial arrangement of the molecules constituting such groupings in the intermediate filament is also important if the entire structure is to be described with a high degree of precision. Because crystallization of the filament is currently not a feasible option, an alternative method has been devised in which covalent crosslinks have been deliberately introduced into the structure and their position determined by sequence analyses. In particular, it has been possible to induce covalent crosslinks between (i) cysteine residues in small assembly-competent oligomers of intermediate filament chains using the copper–phenanthroline reaction (Steinert & Parry, 1993), and (ii) lysine residues

https://doi.org/10.1017/S0033583500003516 Published online by Cambridge University Press

that are in close proximity (up to about 1.2–1.5 nm apart) using the cleavable reagent disulphosuccinimidyl tartrate (DST) (Steinert *et al*. 1993a). The portions of the sequences connected by the crosslinks were then sequenced, hence allowing the residues involved in the crosslink to be identified. Consequently, those portions of the intermediate filament molecules that are spatially adjacent to one another can be specified. The DST reaction has proved to be particularly effective and important in the study of the aggregation of intermediate filament molecules in the small assembly-competent oligomers that have been studied to date. Depending on the numbers and siting of the lysine residues significant numbers of data have been obtained (Table 2): epidermal cytokeratins K1/K10 (Fig. 18) – Steinert *et al*. 1993a; epidermal cytokeratins K5/K14 (Fig. 19) – Steinert *et al*. 1993*b*; desmin – Geisler *et al*. 1992; vimentin (Fig. 20) – Steinert *et al*. 1993c; $\alpha$-internexin/$\alpha$-internexin (Fig. 21) – Steinert *et al*. 1999b; and $\alpha$-internexin/vimentin (Fig. 22) – Steinert *et al*. 1999*b*. Inspection of these data revealed that all of the crosslinks can be accounted for in terms of three modes of intermolecular aggregation (Fig. 16 and 17). All three occur between antiparallel molecules. The first of these modes, known as $A_{11}$, results from a near half-stagger between molecules such that the 1B segments of adjacent molecules are largely overlapped. The second mode ($A_{22}$) again corresponds to a near half-stagger between antiparallel molecules, but in this case the 2B segments are almost completely overlapped. The third mode is one in which antiparallel molecules are almost totally overlapped ($A_{12}$). It is important to emphasize that these modes (with one exception about to be described) account for all of the data obtained and, furthermore, have been found for all of the different IF systems investigated thus far. Combination of the $A_{11}$ and $A_{22}$ modes leads to the situation where a small (about eight residues) head-to-tail overlap occurs between parallel molecules ($A_{CN}$). This feature is common to Type Ib/IIb cytokeratins, Type III vimentin molecules, and Type IV $\alpha$-internexin molecules. There is some doubt, however, that it is also present in the oxidized form of Type Ia/IIa trichocyte keratins where, alternatively, it is believed that the presence of multiple disulphide bonds may lead to a small head-to-tail gap rather than to an overlap (Parry, 1995). A head-to-tail overlap, however, would still be expected to be a feature of the reduced form of the trichocyte keratins in the living cells.

The exception alluded to above refers to a small number of crosslinks that are intramolecular (rather than intermolecular). In other words there are usually a small number of lysine–lysine crosslinks that occur between identically located residues. This means not only that the crosslinks occur between the constituent chains in the intermediate filament molecule but that these chains must necessarily be in axial register and also parallel to one another. This confirms earlier theoretical studies. Some of the crosslinks formed between identical residues are in positions where they could correspond to arrangements of antiparallel molecules. These possibilities are only included in the least-squares analyses where there is other evidence that supports this assignment.

As noted earlier, the $A_{11}$ and $A_{22}$ modes of assembly in murine vimentin have received direct experimental support from the yeast two-hybrid technique (Meng *et al*. 1996). Their work suggested that segment 1B–segment 1B and segment 2B–segment 2B interactions were true protein–protein interactions. However, they found no evidence of a significant segment 1B–segment 2B interaction, expected for the $A_{12}$ mode of assembly, and they suggested therefore that this interaction occurred between adjacent but non-interacting molecules. This technique would likely be unable to detect an interaction based on only 5–10 residues expected for the $A_{CN}$ mode.

https://doi.org/10.1017/S0033583500003516 Published online by Cambridge University Press

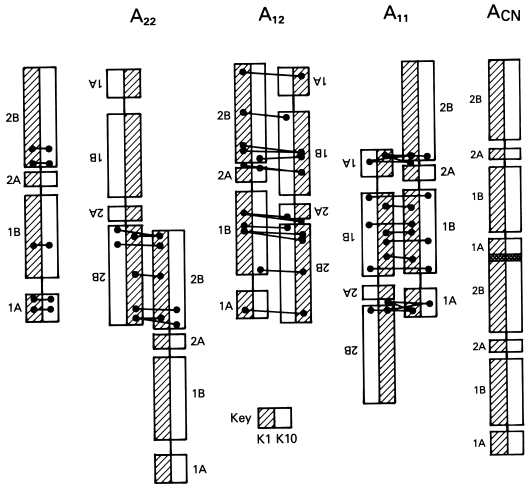



**Fig. 18.** Disulphosuccinimidyl tartrate (DST) has been used to induce covalent crosslinks between lysine residues that are spatially adjacent in murine epidermal K1/K10 cytokeratin intermediate filaments. The links have been characterized and the residues involved have been specified by protein sequencing. All of the links can be explained in terms of (i) chains that are parallel and in axial register within a molecule, (ii) an antiparallel arrangement of molecules with their 2B segments largely overlapped ($A_{22}$), (iii) an antiparallel arrangement of molecules with their 1B and 2B segments largely overlapped ($A_{12}$), and (iv) an antiparallel arrangement of molecules with their 1B segments largely overlapped ($A_{11}$). Least squares analysis allows the axial staggers to be deduced as well as the projected lengths of the links L1, L12 and L2. The $A_{CN}$ mode arises from the $A_{11}$ and $A_{22}$ modes, and corresponds to a small head-to-tail overlap between parallel molecules. Figure from Parry & Steinert (1995).

Having such crosslink data available, it is then possible to relate them directly to six parameters (the relative axial staggers $A_{11}$, $A_{22}$, $A_{12}$ and the axial lengths of the links L1, L12 and L2). The equations thus deduced are then subjected to a least-squares analysis to determine the above parameters and, provided that there are sufficient numbers of independent equations available (and this depends not only on the number of data but also on the distribution of the crosslinked residues involved), then each can be derived with some confidence. A summary of the parameters determined using this method is given in Table 3. A related analysis for trichocyte keratins has also been undertaken (Parry, 1995) but the data here are very limited. These data will be described after those relating to the other intermediate filaments.

Firstly, the two epidermal cytokeratin systems (K1/K10 and K5/K14) gave rise to almost identical axial alignments of the molecules. This is not surprising, of course, since it is known that during differentiation and development the intermediate filaments are reconstructed via a dynamic exchange process in which one set of epidermal keratin chains/molecules are replaced by another. One particular example of this is where intermediate filaments composed largely of K5/K14 molecules expressed in basal stratified squamous epithelia are replaced during terminal differentiation by other differentiation-specific pairs, such as those composed largely of K1/K10 molecules in terminally differentiating epidermis. It follows that molecular exchange would be a simple process because both K1/K10 and K5/K14 molecules have identical lengths and axial assembly parameters. The parameters deduced for vimentin (Steinert *et al.* 1993*c*) were based on a much smaller number of crosslinks but indicative

https://doi.org/10.1017/S0033583500003516 Published online by Cambridge University Press

https://doi.org/10.1017/S0033583500003516 Published online by Cambridge University Press



**Table 3.** *Structural parameters for IF proteins deduced from crosslinking studies and theoretical predictions**

| Structural feature | Epidermal keratin (Type I/II) | Hard α-keratin (Type I/II) | Vimentin (Type III) | α-Internexin (Type IV) | Vimentin (Type III), α-internexin (Type IV) and α-internexin/vimentin Type (III/IV) |
|---|---|---|---|---|---|
| | Crosslinks | Theoretical/ crosslinks | Crosslinks | Crosslinks | Crosslinks |
| $A_{12}$ | 0·43 | 6·38 | −0·45 | 7·00 | 4·19 |
| $A_{11}$ | −111·62 | −126·92 | −131·05 | −130·00 | −132·24 |
| $A_{22}$ | 192·82 | 189·68 | 156·58 | — | 165·68 |
| L1 | 16·27 | 12·51 | 5·24 | 14·50 | 10·99 |
| L12 | 13·94 | 13·94 | 6·06 | — | 8·99 |
| L2 | 5·06 | 5·06 | 8·41 | 9·00 | 8·94 |
| Length | 311·27 | 307·51 | 295·72 | — | 304·92 |
| 1BU, 1BD | −3·89 | −15·43 | −16·81 | — | −20·30 |
| 2BU, 2BD | 2·56 | 3·17 | −18·14 | −19·00 | −18·24 |
| 1BU, 2D | −50·84 | −41·13 | −40·69 | −42·50 | −41·80 |
| Repeat | 304·44 | 316·60 | 287·63 | — | 297·93 |
| Overlap | 6·83 | −9·09 | 8·09 | 10·00 | 6·99 |
| No. of data | 39 | | 11 | 10 | 22 |

* Axial data are measured in multiples of the mean axial rise in a coiled-coil conformation, $\bar{h}_{cc}$ (= 0.1485 nm).

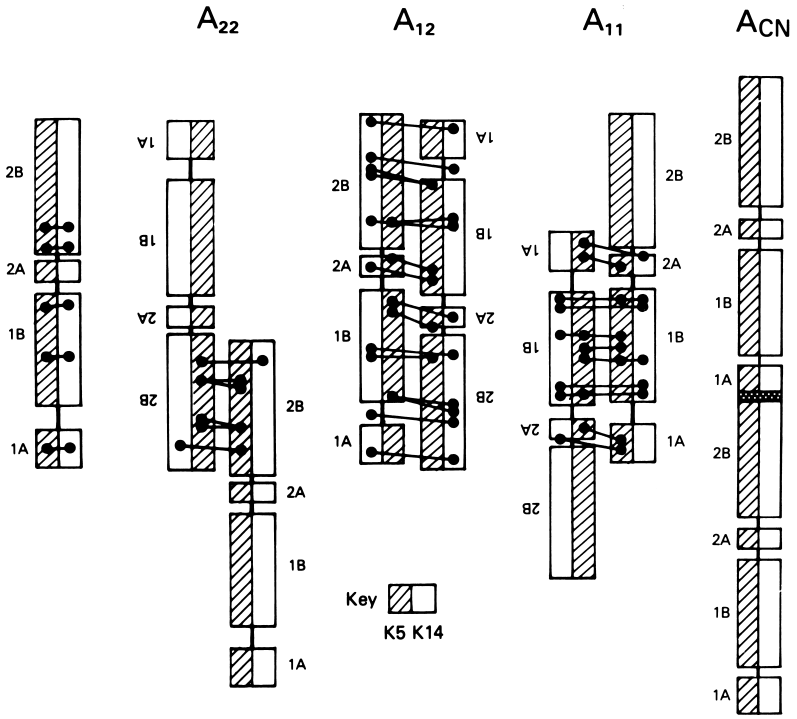



**Fig. 19.** DST-induced crosslinks for human epidermal K5/K14 cytokeratin intermediate filaments. The same modes of interaction ($A_{22}$, $A_{12}$, $A_{11}$ and $A_{CN}$) as those deduced for K1/K10 intermediate filaments were found. Furthermore, the axial staggers and the link lengths were closely maintained. It follows that K1/K10 and K5/K14 molecules could be interchanged in the intermediate filament with little difficulty as is indeed observed during differentiation. Figure from Steinert *et al.* (1993*b*).

parameters were nonetheless determined. The important point was that these values were quite clearly different from those derived for the epidermal keratins, indicating (as is observed) that Type III chains do not copolymerise *in vitro* or *in vivo* with either Type I or Type II cytokeratins. More data were collected subsequently for α-internexin homodimers and for α-internexin/vimentin heterodimers (Steinert *et al.* 1999b). The parameters thus deduced in the latter two cases were nearly identical and very similar to those found earlier from the limited data on vimentin homodimers. Using all of the extended data, a more highly refined set of parameters has now been calculated for vimentin intermediate filaments. This differs only in two (related) parameters from those values previously published. Parameters for Type III intermediate filaments, for Type IV intermediate filaments and for Type III/IV copolymers are thus now all known.

What can we deduce about the epidermal keratins, the Type III and the Type IV intermediate filaments from these data? The first key point is that while the same modes of aggregation are used in all cases, the exact values of the axial alignments ($A_{11}$, $A_{22}$ and $A_{12}$) and the lengths of the links L1, L12 and L2 are different. The effect of this would be to prevent copolymerization of the Types I and/or II proteins with the Types III and/or IV proteins. The molecular lengths are predicted to differ slightly (46·2 nm for Type I/II epidermal keratin and 45·3 nm for Types III and IV intermediate filaments) as are the repeat distances in the filament (45·2 and 44·2 nm respectively). A 22·6 nm axial repeat in isolated epidermal keratin filaments (which corresponds exactly to half of the predicted repeat length)

https://doi.org/10.1017/S0033583500003516 Published online by Cambridge University Press

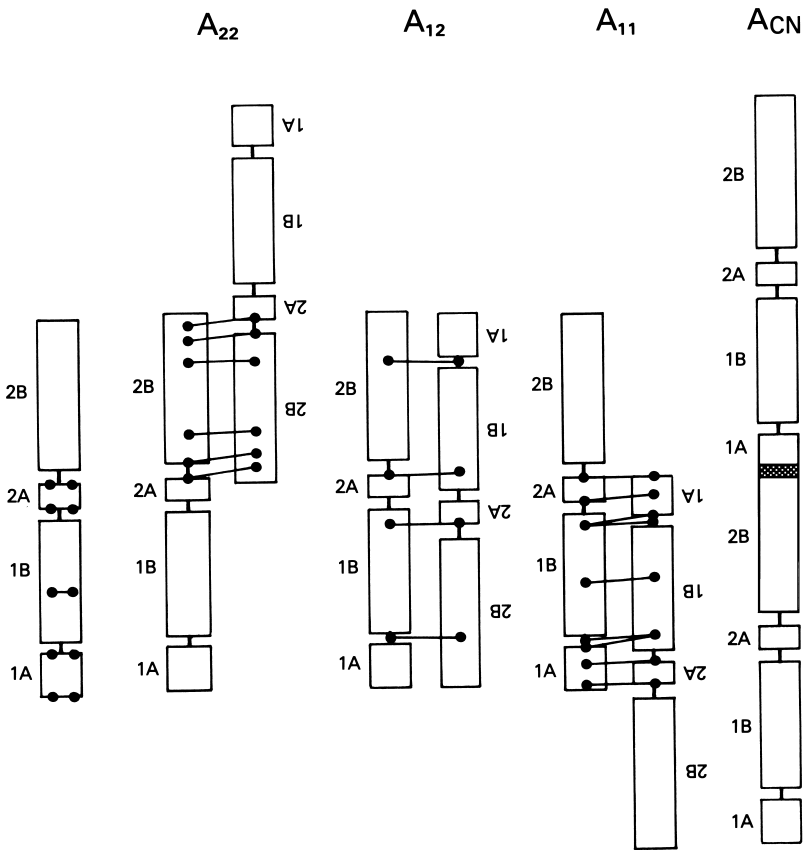



**Fig. 20.** DST-induced crosslinks for vimentin intermediate filaments. Although fewer crosslink data were available for vimentin than for the epidermal cytokeratins the observed data refine satisfactorily on the basis of the same set of modes of aggregation characterized for the cytokeratins (Table 3). Additional data, subsequently derived from a vimentin/α-internexin copolymer (Fig. 22), allowed a more complete analysis but only small changes in two parameters resulted. Note that the exact axial staggers between the antiparallel molecules and the link lengths are not identical to those derived for the cytokeratins. This indicates that vimentin and the cytokeratins are most unlikely to copolymerise *in vivo*. Figure from Steinert *et al*. (1993c).

has been observed directly using electron microscopy (Section 2.4). This correspondence in the predicted and observed values gives confidence in the idea that the crosslinks characterized do indeed occur in functional and assembly-competent oligomers as initially believed. The expected axial repeat in vimentin intermediate filaments would be about 22·1 nm. So far there are no experimental data available to confirm or disprove this proposal, but paracrystals of neurofilaments (which are believed to have the same molecular dimensions and packing characteristics as vimentin) have an axial period a little below this value (21·25 nm; Heins & Aebi, 1994). The true period in vimentin/neurofilaments would therefore, if anything, be expected to be larger rather than smaller or the same. A report by Carter *et al*. (1997) has suggested that cytokeratins might copolymerize with NF-L chains in cells. However, the axial parameters of the keratins and the Type IV intermediate filaments, as epitomised by α-internexin, are very clearly different from one another. It would seem more likely that the filaments observed represent separate populations of cytokeratin and NF-L filaments rather than a copolymerized mixture of the two chain types.

https://doi.org/10.1017/S0033583500003516 Published online by Cambridge University Press

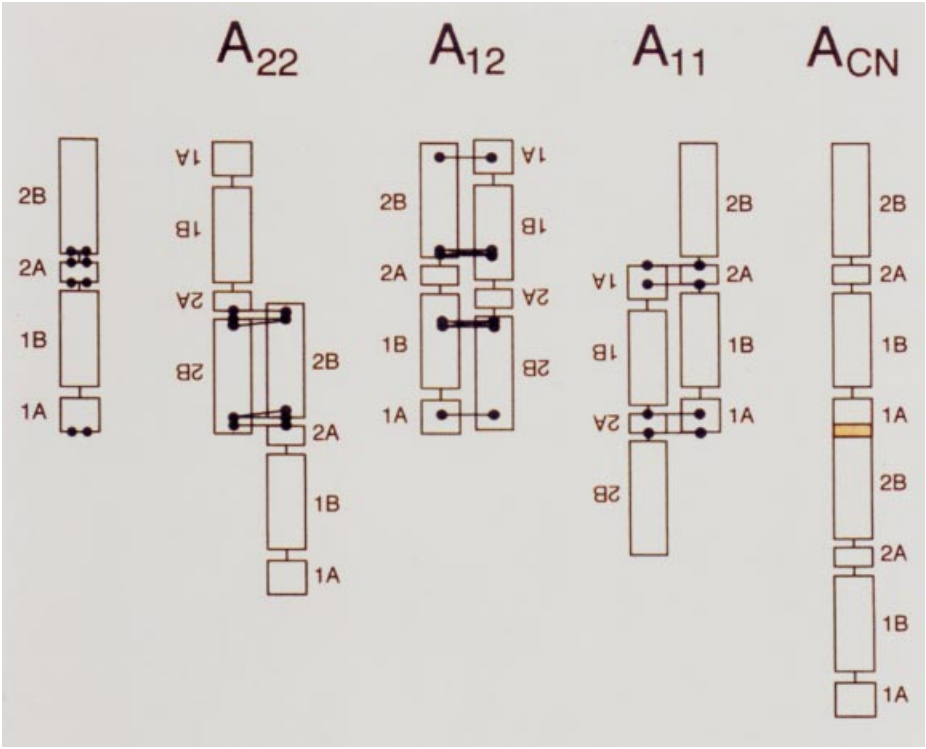



**Fig. 21.** DST-induced crosslinks for Type IV $\alpha$-internexin intermediate filaments. Insufficient data were available to refine all of the parameters satisfactorily though the observed crosslinks were all compatible with the $A_{22}$, $A_{12}$, $A_{11}$ and $A_{CN}$ modes of molecular aggregation. The calculated parameters were nonetheless similar to those derived for vimentin suggesting that they may have a common intermediate filament structure. Figure from Steinert *et al.* (1999b).

The second key point is that the head-to-tail overlap between parallel molecules can be predicted from the dimensions of $A_{11}$ and $A_{22}$. The deduced value ($A_{CN}$) is equivalent to an axial rise of seven residues in an $\alpha$-helical coiled-coil conformation for the cytokeratins, to eight residues in Type III vimentin, to ten residues in Type IV $\alpha$-internexin, and to seven residues in the combined Type III and Type IV data. Within the accuracy of the data it is likely that this value is the same in all cases. As already discussed, the first 10–15 residues of segment 1A and the last 10–15 residues of segment 2B are exceptionally highly conserved in all intermediate filament chains. The predicted head-to-tail overlap involves precisely these sequences or those which are immediately adjacent. Thus the role of these conserved sequences is to provide an important interaction to specify and stabilise molecular assembly in the intermediate filament. The fact that these overlaps are so highly conserved in three quite different intermediate filament types (as judged on the basis of their axial alignment parameters: Table 3) and also in lamins (Stuurman *et al.* 1996)) is testament to their importance. Yet further evidence for this conclusion lies in the locations of mutations that have been characterized in many different keratin chains and which give rise to various keratinopathies: most mutations have been confined very largely to these same highly conserved regions of sequence at either end of the rod domain (see Section 6).

A third point of possible importance is that a small number of the observed crosslinks are consistent not only with the mode of assembly already attributed to them but also with

https://doi.org/10.1017/S0033583500003516 Published online by Cambridge University Press

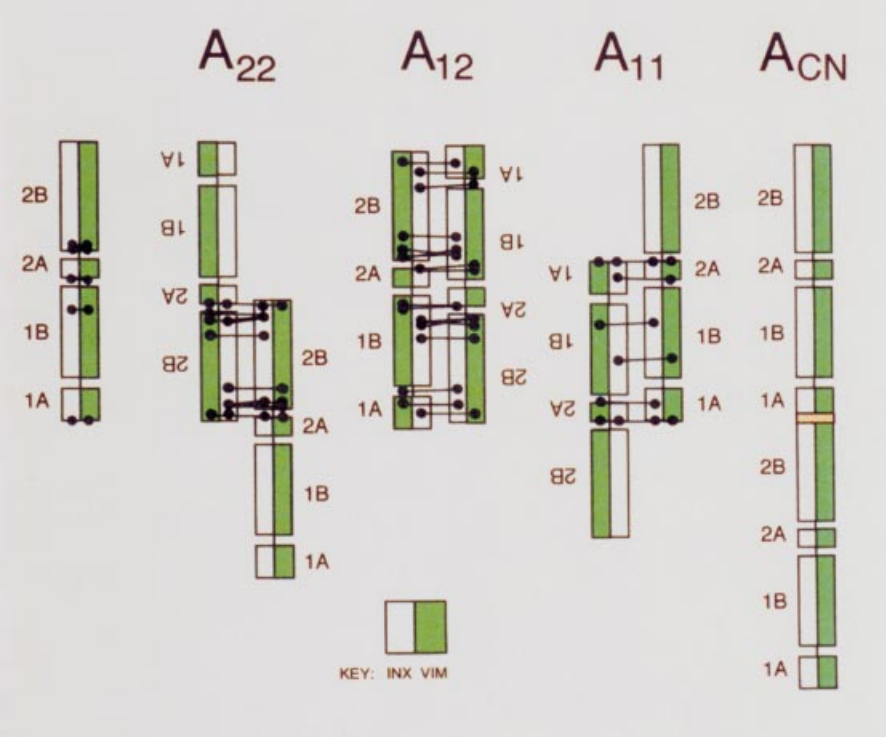



**Fig. 22.** DST-induced crosslinks for α-internexin/vimentin copolymers. These data refine well and show that vimentin and α-internexin molecules have the same molecular parameters. The link lengths and axial parameters for vimentin/α-internexin and for the epidermal cytokeratins are different (Table 3). The crosslink data strongly imply that at least two different but related structures must exist for intermediate filaments. With knowledge also of the lamin structure and the trichocyte keratin axial repeats, it is clear that at least four unique structures exist for the class of structures known as the intermediate filaments. Figure from Steinert *et al*. (1999b).

crosslinking between parallel molecules two apart on the surface lattice (Steinert *et al*. 1994). This could be related to the existence of a sub-filamentous structure, such as a protofilament or a protofibril (see Section 2.4). Possible crosslinks between molecules three or more apart on the surface lattice, however, have not yet been identified.

So far no discussion has been made of the trichocyte keratins. The very recent discovery of methods for *in vitro* assembly of bacterially expressed Type Ia/IIa chains into intermediate filaments (Fig. 15) should now facilitate crosslinking studies of the sort reported for several other forms of filaments as described above. Nevertheless, a different way of tackling the problem has been attempted. In this case six equations were derived for the six parameters and while this represents a minimalist approach (to say the very least) it is the best that can be done at present. Two equations were based on the possible conservation of the axial lengths of L1 and L12 between trichocyte keratins and cytokeratins, as deduced from the high identity or homology that exists between the pertinent sequences. Two other equations were derived from the dimensions of the surface lattice measured from fibre X-ray diffraction patterns. The fifth equation was determined using the data that a transpeptidation event led to the covalent linking of 1B segments from antiparallel molecules after the action of chymotrypsin on the so-called low-sulphur fraction of wool. The sixth constraint arose from

https://doi.org/10.1017/S0033583500003516 Published online by Cambridge University Press

considerations that the axial alignment of antiparallel 1B segments could physically take only one of two possible values deduced earlier from other studies on the role of the intermolecular ionic interactions (Fraser *et al*. 1985, 1986). Naturally, the limited data available allow less confidence in the conclusions than was possible for the other intermediate filaments that were studied. The results are interesting, nonetheless, and would seem to differ radically in some respects from those previously described. While the modes of interaction are absolutely maintained there is no head-to-tail overlap. Indeed, there is a gap of nine residues between the end of one molecule and the beginning of the next parallel molecule. It is difficult to understand why this should be so when the sequences at the ends of the rod domain are so similar to those in all other intermediate filament chains. An explanation, however, was proposed by Parry (1995, 1996). It was suggested that the assembly is stabilized preferentially by a network of disulphide bonds. The cysteine residues present in the rod domain of trichocyte keratin molecules lie in precisely the correct axial positions for disulphide bond formation to occur. There are no data, however, to show that the azimuthal positions of these residues are also appropriate to allow these bonds to form. It is likely that the trichocyte keratin molecules have the same head-to-tail overlap during their early stages of development (i.e. when they exist in the reduced form in living cells) as seen for the other intermediate filaments. It is possible, however, that subsequent oxidation and possible interaction with cysteine-rich matrix proteins forces a molecular rearrangement to the one that has been proposed for wool and quill. Further data will be required to clarify these ideas. It must be noted here though that mutations in the trichocyte keratins that give rise to hair/nail disorders also occur in the same regions of the intermediate filament chains as those that result in diseases in cytokeratins; that is, in and around the helix initiation motif (N8H and N8D in segment 1A) and the termination motif (E117K, E117D and E106K in segment 2B) (Korge *et al*. 1998, 1999; Winter *et al*. 1997a, b, 1998; Zlotogorski *et al*. 1998). Notably, the effects of these abnormal keratin chains is evident early in the hair follicle or nail bed, prior to the time the chains become oxidized and crosslinked by matrix proteins.

Less detail is known about the packing of the nuclear lamin proteins. However, Parry *et al*. (1987) have shown that the packing of the nuclear lamin molecules in paracrystalline arrays, filamentous structures (Zackroff *et al*. 1984; Goldman *et al*. 1986) and the lattice seen in *Xenopus* oocytes (Aebi *et al*. 1986) can all be explained in terms of a single model in which an array of approximately half-staggered antiparallel molecules is the key element. This form of packing has major similarities to that deduced using crosslinking data derived for other intermediate filament types. Geisler *et al*. (1998) have come to a similar conclusion from their work on the assembly of invertebrate cytoplasmic intermediate filaments. They showed that these filaments had assembly characteristics that were common to both the nuclear lamins and to the vertebrate cytoplasmic intermediate filaments. Stuurman *et al*. (1998) have recently summarized studies on nuclear lamin assemblies, including those originating from lamin mutants. The latter (in particular) provide some insight into the particular roles played by the various domains in the lamin molecule. Stuurman *et al*. (1996) also showed a common propensity of all intermediate filaments to display a short head-to-tail overlap between similarly directed (parallel) molecules. Their experimentally derived conclusion provides further data in support of the conclusion obtained using the crosslink approach for other intermediate filament proteins.

https://doi.org/10.1017/S0033583500003516 Published online by Cambridge University Press

### 3.4 Naturally occurring crosslinks for tissue coordination

So far the discussion has been confined to the characterization of experimentally induced crosslinks and to their interpretation in terms of the molecular assembly mechanisms existing within the intermediate filament. The $N^{\varepsilon}$-($\gamma$-glutamyl)lysine isopeptide crosslink catalysed by transglutaminases is formed naturally between intermediate filament proteins and other proteins in stratified squamous epithelia. Such epithelia form a specialized structure, termed the cell envelope, located immediately under their plasma membrane which is critically involved in barrier function. In most epithelia, the cell envelope affords primarily a physical barrier for protection against the environment. In mammalian epidermis and the cuticle of the hair, the 'cornified' cell envelope also participates in water barrier function. Many such isopeptide crosslinks have been found between cytokeratins and several of the cell envelope proteins, most notably loricrin, small proline-rich proteins (in the epidermis), and involucrin, desmoplakin and envoplakin (in all epithelia examined so far) (Steinert & Marekov, 1997; Candi *et al*. 1998; Steinert *et al*. 1998a, b). Most crosslinks involved a single lysine residue that is highly conserved within the V1 subdomain of Type IIb keratins. Loss of this lysine residue in the K1 chain by mutation results in an ichthyosis skin disease which, interestingly, is largely restricted to the palms and soles (non-epidermolytic palmar–plantar keratoderma, or Ulna–Thost disease), although it also occurs in other thickened portions of the skin in other body sites (Kimonis *et al*. 1994). In this case, the tonofibril bundles of keratin intermediate filaments are normal, but the associations of such bundles with the cell periphery and desmosomes of the cells are abnormal. Also, the intercellular lipids required for normal water barrier function are inappropriately organized. It was concluded based on the peptide sequencing and these morphological findings that firm attachment of the keratin intermediate filament cytoskeleton by transglutaminase crosslinking to the cell periphery is essential for coordination of cell structure and function in terminally differentiating stratified squamous epithelial cells, which in turn affects the physical and water barrier functions of the entire epithelium (Candi *et al*. 1998).

### 3.5 STEM data

STEM has been able to provide key data on the structure of intermediate filaments. Experimentally, it is a technique with the capability of allowing the mass per unit length of the filaments to be determined. Such data places significant restrictions on the possible number of molecules in cross-section and thus higher orders of alignments of molecules. When allied with electron energy loss spectroscopy it is also possible to measure the state of phosphorylation of the constituent proteins (Leapman *et al*. 1997). STEM is also capable of allowing the radial mass density of the filaments to be determined. The results will be discussed firstly in terms of the axial mass distributions for native intermediate filaments, and then for *in vitro* assemblies, and secondly for the radial mass distributions.

Three native intermediate filaments have been studied by using STEM. These are desmin/vimentin copolymers (Steven *et al*. 1982), human vimentin (Herrmann *et al*. 1996a) and trichocyte keratins (Jones *et al*. 1997). In each instance the mass per unit length distributions were unimodal (Fig. 23) and thus characteristic of an homogeneous population of filaments. In order to calculate the number of chains in cross-section both the molecular

https://doi.org/10.1017/S0033583500003516 Published online by Cambridge University Press

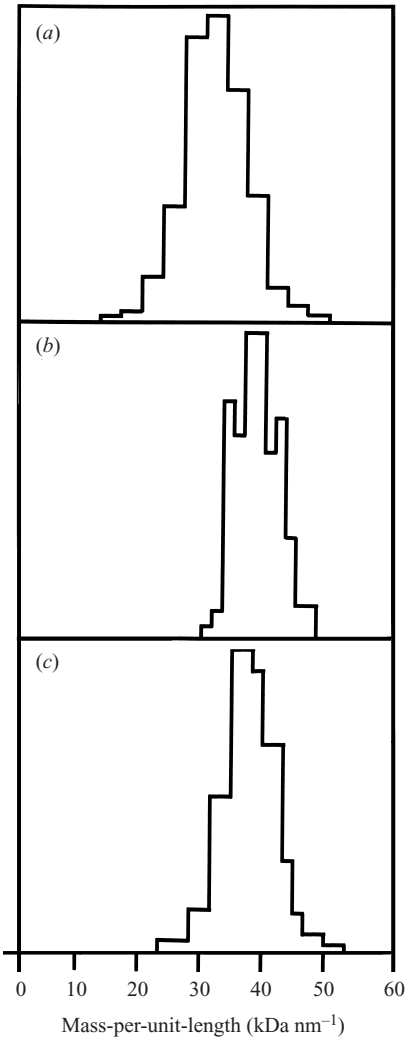



**Fig. 23.** Mass-per-unit-length distributions determined by using STEM. The data pertain to native intermediate filaments and originate from: (*a*) rat vibrissae (Jones *et al*. 1997, and unpublished data from the same authors); (*b*) desmin/vimentin copolymers from BHK-21 cells (Steven *et al*. 1982); and (*c*) human vimentin (Herrmann *et al*. 1996a). The distributions are unimodal and have mean values of 31·2 (and 31·8 unpublished), 36·2 and 37 kDa respectively. In each case the mass values correspond to about 30 chains in section. Distribution (*b*) does not show a very small peak that appeared in the original published distribution as this was attributed by the authors to measurements that were made either from very short intermediate filaments or from one end of a filament that was clearly discontinuous. Figure redrawn from the originals cited above.

weights of vimentin/desmin, vimentin and trichocyte keratins as well as their axial repeat lengths are needed. With the exception of the repeat lengths in desmin/vimentin and in vimentin, these data have all been experimentally determined. The missing data for desmin/vimentin and vimentin, however, have been assessed from crosslinking data but will ultimately require verification from X-ray diffraction studies on orientated fibres. The results show that in each case the number of chains in cross-section is close to 32 (Table 4). It remains to be seen whether all native intermediate filaments contain the same number of chains in

https://doi.org/10.1017/S0033583500003516 Published online by Cambridge University Press

https://doi.org/10.1017/S0033583500003516 Published online by Cambridge University Press



**Table 4.** *STEM data for intermediate filaments*

| Source | Reference | Mass/length (kDa $nm^{-1}$) | Number of chains | | | | | Molecular weight (kDa) | Repeat (nm) |
|---|---|---|---|---|---|---|---|---|---|
| Bovine epidermis | Steven *et al.* 1983a | 26·6, 38·0, 48·3 | — | 22·9 | 32·7 | 41·6 | — | (105) | (45·2) |
| Rat vibrissae | Jones *et al.* 1997 | 31·2 | — | — | 29·3 | — | — | 100·2 | 47·0 |
| Rat vibrissae | Watts *et al.* unpublished | 31·8 | — | — | 29·8 | — | — | 100·2 | 47·0 |
| Human hair follicle | Watts *et al.* unpublished | 16.0, 32·1 | 15·0 | — | 30·1 | — | — | 100·2 | 47·0 |
| Human vimentin | Herrmann *et al.* 1996 | 37 | — | — | 30.5 | — | — | 107·2 | (44·2) |
| Bovine vimentin | Steven *et al.* 1983a | 25·0, 36·5, 51·0 | — | 20·8 | 30·4 | 42·5 | — | (106) | (44·2) |
| Bovine vimentin/ desmin | Steven *et al.* 1983a | 28·6, 36·2, 44·1 | — | 23·9 | 30·2 | 36·8 | — | (106) | (44·2) |
| *Xenopus* vimentin | Herrmann *et al.* 1996 | 39, 50, 61 | — | — | 32·8 | 42·0 | 51·3 | 105·2 | (44·2) |
| Mouse neurofilament NF-L | Heins *et al.* 1993 | 47 | — | — | 33·8 | — | — | 122.8 | (44·2) |

Best estimates are given in parentheses.

section but it is currently a reasonable hypothesis to assume that this is indeed the case. It is interesting to note, however, that STEM data for neurofilaments from an invertebrate (the squid giant axon) contain only 16 chains in cross-section and that these are near maximally phosphorylated (Leapman *et al*. 1997). The squid neurofilaments, like those listed above, arise from a near homogeneous population. It remains to be seen whether the observed difference is maintained across the vertebrate/invertebrates divide. It is possible, maybe even likely, that the squid data have originated in some way from 'normal' intermediate filaments that have split into two or, conversely, have yet to assemble into full filaments. Time will tell which alternative is the real one.

In passing it should be noted that measurements of the diameter of the trichocyte intermediate filaments together with data on the number of chains in cross-section deduced from the structure of wool (Fraser *et al*. 1969) both suggested that this type of intermediate filament would contain only about 26 chains. Reinterpretation of the low-angle X-ray diffraction data, however, led to the conclusion that the trichocyte intermediate filaments need not be considered as solid cylinders and could have a core with a different density to that of the annular region defining the outside of the filament. In other words the X-ray data are consistent both with a solid cylinder of diameter 7.5 nm and with a larger diameter cylinder with a less dense core. The latter structure would be compatible with intermediate filaments containing about 32 chains in cross-section (Parry, 1996).

Intermediate filament proteins (or mutants thereof) reassemble *in vitro* into structures that are generally similar to native filaments as judged by their electron microscope appearances (e.g. Steinert *et al*. 1976, 1981). However, they do give rise to STEM data that show that the mass distributions are, with one exception, not unimodal, as seen for native preparations, but multimodal (Fig. 24; data summarized in Table 4; Steven *et al*. 1983a, b; Engel *et al*. 1985; Heins *et al*. 1993; Herrmann *et al*. 1996a). The main peak in all of the distributions, nonetheless, does correspond to about 32 chains in section for the observed filaments, the same value reported for native intermediate filaments. Independent measurements of the number of chains in cross-section in intermediate filaments *in vitro* have been made by Kooijman *et al*. (1997) using the technique of transient electric birefringence. Their data too are consistent with 32 chains in section for both vimentin and glial fibrillary acidic protein intermediate filaments. The remaining peaks in the mass distributions determined using STEM are not random but tend to occur close to values of about 16, 24, 40 and 48 chains. This has been interpreted as implying that the intermediate filament *in vivo* contains four protofibrils each with eight chains, but that *in vitro* assemblies are able to form from two, three, four, five or even six such protofibrils. This suggests that the specificity of assembly *in vitro* is lacking, though why this should be so is not understood. The key implication of these results is not so much the lack of specificity of assembly that appears to be a feature *in vitro* but the fact that an eight chain protofibrillar structure is likely to be an important element of intermediate filaments *in vivo*. TEM observations have been made of sub-filamentous structures that could correspond to these protofibrils. However, the inconsistency of the appearances of these structures has not resulted in general acceptance of their existence. It does seem, nonetheless, that the STEM data show that protofibrils may indeed be a real feature of intermediate filaments.

STEM data from intermediate filaments have also shown that the mass may vary along the length of an individual filament. The most likely explanation of this phenomenon derives from the current view that intermediate filaments assemble *in vitro* from full-length half-mass

https://doi.org/10.1017/S0033583500003516 Published online by Cambridge University Press

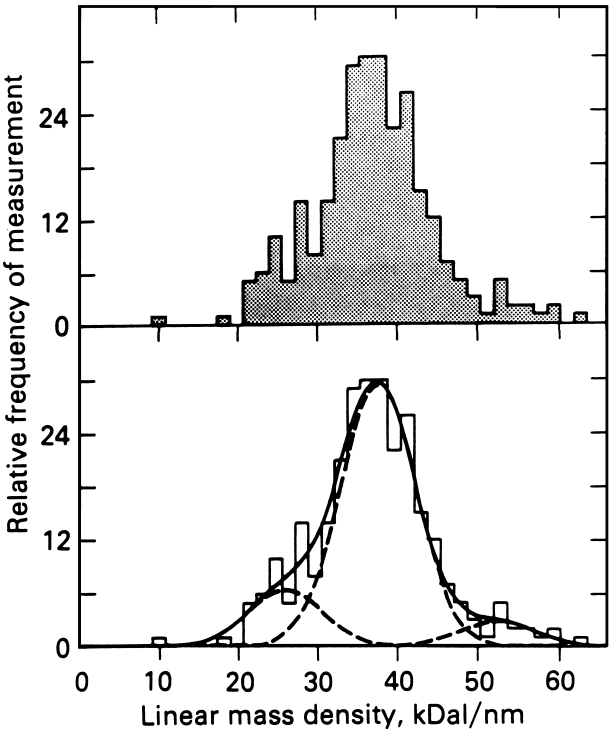



**Fig. 24.** Mass-per-unit-length distributions determined using STEM. In this case the data pertain to bovine vimentin intermediate filaments that have been assembled *in vitro*. Virtually all intermediate filaments assembled *in vitro* display multimodal distributions such as the one shown here (Steven *et al*. 1982). Furthermore, successive peaks in the distributions tend to differ from one another by the equivalent of about eight chains. This has been interpreted as indicating that an eight chain protofibril may be an inherent element of substructure of the intermediate filament. Figure from Steven *et al*. (1982).

particles, and that this is perhaps an uneven event, at least *in vitro* (Herrmann *et al*. 1999). It remains to be seen whether such axial mass variations also occur in native intermediate filaments.

The radial distributions of mass in intermediate filaments can also be determined from STEM data. Steven *et al*. (1984), using dark-field STEM micrographs of freeze-dried, unstained intermediate filaments reassembled from bovine cytokeratins have shown that these filaments have a relatively uniformly dense core (8–10 nm in diameter) and that this is surrounded by an annular region over which the density drops to zero (outer diameter 13–15 nm). The interpretation of these data is that the rod domains of the intermediate filament molecules pack closely to form the core of the filament and that the head and tail domains lie on the surface. This is consistent with the STEM measurement that 65% of the mass of the filament lies in the core region, and that the mass of the rod domain of the molecules in this particular type of intermediate filament is also about 65% of the total mass of the molecule. Interestingly, however, electron cryomicroscopy data (Watts *et al*., unpublished) have shown that the trichocyte intermediate filaments from rat whiskers and human hair follicles have a hollow core (Fig. 25). This observation is compatible with, although not clearly obvious from, the radial mass profiles described above. While the statement of a

https://doi.org/10.1017/S0033583500003516 Published online by Cambridge University Press

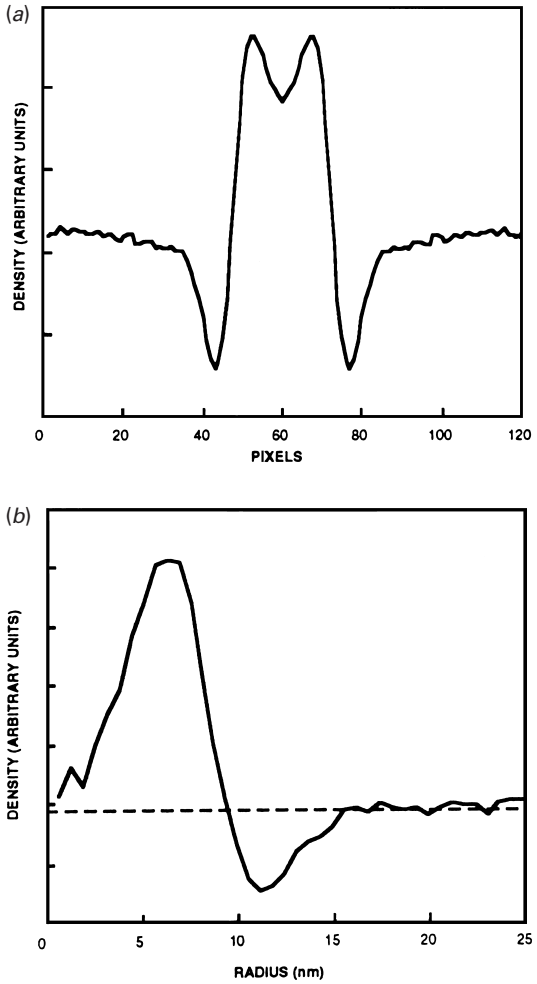



**Fig. 25.** (*a*) Transverse density profile of an undecorated trichocyte keratin intermediate filament from hair and (*b*) the corresponding radial mass profile. These data indicate that these keratin intermediate filament may have a 'hollow' core. Most likely, the core would not be occupied by intermediate filament molecules but by water and the sidechains of residues in the molecules that occupy the positions around the annulus of the intermediate filament. Figure redrawn from the unpublished data of Steven, Jones and colleagues (with permission).

hollow core indicates the total absence of matter this is almost certainly not the case in practice. At the very least the core would be hydrated and would contain many of the sidechains of the molecules that form the region occupied by the intermediate filament molecules. The limited resolution of the STEM data would not easily allow the expected drop in mass density to be identified at the very centre of the filament. It remains to be seen whether this conclusion is one that applies to all types of intermediate filament but it is not unlikely that this will turn out to be the case. Such a model has a degree of symmetry (attractive to biophysicists) that is not present in earlier models in which a centrally located molecule was surrounded by (say) seven others.

https://doi.org/10.1017/S0033583500003516 Published online by Cambridge University Press

## 4. Quaternary structure

The quaternary structures of the intermediate filaments, of which at least four distinct but related forms have now been shown to exist *in vivo*, remain largely undetermined in detail though there are several clear indications as to the main features that are likely to be present in all cases. The data on which the modelling has been based come from four major sources: (a) X-ray diffraction (of wool/quill, which remain the only sources of intermediate filaments from which high resolution and high quality data have been obtained); (b) induction of crosslinks in oligomers of various intermediate filament proteins (cytokeratins, vimentin, α-internexin) to generate intermolecular axial staggers and projected link lengths; (c) STEM measurements of mass-per-unit-length of the intermediate filaments; and (d) detailed and high resolution electron microscopy on primarily *in vitro* aggregates. As discussed above, it seems very probable that all native intermediate filaments contain 16 molecules in cross-section and that the $A_{11}$, $A_{22}$ and the $A_{12}$ modes of interaction are those that generate the fully functional intermediate filament. In all cases (with the likely exception of the trichocyte keratins in the oxidized state), the $A_{CN}$ mode (corresponding to a short head-to-tail overlap between parallel molecules) is a common interaction that both specifies aggregation and stabilizes the assemblies so formed. Specific roles at the quaternary level have been assigned to subfilamentous structures (protofilaments, protofibrils) and to the terminal domains. The quaternary models for the intermediate filament remain incomplete but it is hoped that crystal structure data may become available in the not-too-distant future to provide the hard evidence that is needed to substantiate each level of the filament hierarchy that has been proposed.

### 4.1 Protofilaments and protofibrils

Protofilaments have been defined as those long subfilamentous structures that in cross-section contain two molecules (four chains). The structure, if in fact it exists as a real entity, is proposed to consist of an antiparallel arrangement of molecular strands. Each strand consists of parallel molecules linked head-to-tail with an overlap of about 1 nm ($A_{CN}$ mode). The two antiparallel strands would be aligned such that the constituent molecules are approximately half-staggered ($A_{11}$ and $A_{22}$ modes). Although some electron microscope data are consistent with such a structure (its diameter would be about 3 nm) (Aebi *et al.* 1983) this hypothetical structure remains only a possibility and has only limited experimental support. It is, nonetheless, compatible with the intermolecular staggers determined from the crosslinking studies (Section 3.3).

Protofibrils on the other hand, while still requiring additional experimental verification, look to be a much more likely subfilamentous component of the intermediate filament. Again, there is some electron microscope support from the appearance of degraded or reconstituted filaments, but the main line of evidence comes from STEM observations on both *in vivo* and *in vitro* assemblies. Consistently, observations of the masses-per-unit-length of *in vitro* assemblies of intermediate filaments indicate that polymorphism exists and that mass variants exist that differ from one another by units containing eight chains. The protofibril, therefore, is proposed to be that structure which contains eight chains in cross-section (i.e. four molecules). It can most easily be thought of as a pair of protofilaments staggered axially in

https://doi.org/10.1017/S0033583500003516 Published online by Cambridge University Press

such a way that the $A_{12}$ mode becomes involved in stabilising the structure (Parry & Steinert, 1995). Additional indirect supportive data come from two sources. Waseem *et al*. (1997), using a specific monoclonal antibody (Mab LE65), have mapped the epitope to the L12 segment of keratins K18 and K8. The results they obtained were interpreted to show that the L12 segments interact during filament assembly in a manner expected both of a parallel in-register chain structure for the molecule and the $A_{12}$ mode of molecular interaction. In apparent support of this, mutations in L12 in several keratin chains are known to lead to skin blistering diseases and could thus have their origins in the disruption to the filament structure caused by imperfections in the intermolecular interactions between L12 segments.

There is now considerable evidence from other fibrous proteins that the handedness of supercoiling that characterises this group of proteins alternates at successive levels of the hierarchy. Since $\alpha$-helices are right-handed and the coiled-coil structure is left-handed it would seem clear that the next level of supercoiling should be right-handed (Fig. 26). If protofilaments do exist then the molecular strands would be expected to coil about one another in a right-handed manner. In turn, if the protofibril also exists then the protofilaments would then probably coil about the axis of the protofibril in a left-handed manner. Conversely, if the protofilament does not exist as a distinct subfilamentous structure then the molecular strands would probably coil about the axis of the protofibril in a right-handed manner and the protofibrils would coil about the axis of the intermediate filament in a left-handed manner. Clearly, more data are required before these options can be distinguished from one another.

So far the discussion has centred on the subfilamentous structure of the vertebrate intermediate filaments. Does a related structural hierarchy exist also for the invertebrate intermediate filaments, especially those with the extended rod domain? Many of the answers to these questions have been forthcoming from the work of Geisler *et al*. (1998). They studied the assembly characteristics of the two chains from esophagus epithelium of the snail *Helix pomatia* and the two chains from the nematode *Ascaris suum*. Firstly, they observed molecules with lengths of about 56 nm, a value about 3 nm longer than that expected of structures with coiled-coil rod domains containing about 356 residues. The molecules were parallel in-register dimers as judged by the two large terminal domains that were observed at only one end of the molecule. (The head domains of these molecules are small and the tail domains are large.) The next level of hierarchy was characterized by rod-like structures 72 nm in length with the large globular domains at the extreme ends. These were therefore believed to be tetramers formed from the N-terminal overlap of antiparallel half-staggered dimers. These tetramers then appear to anneal longitudinally into protofilaments by further assembly via the C-terminal overlap of antiparallel half-staggered molecules. These modes of assembly are essentially identical to those discussed above for the shorter rod domain structures characteristic of the vertebrate cytoplasmic intermediate filament molecules. Geisler *et al*. (1998) then postulated that radial growth occurs whereby pairs of protofilaments aggregate laterally. These in turn further aggregate into full size intermediate filaments. The axial period of the intermediate filaments generated in essence by the axial disposition of the large tail domains is 24·5 nm. Related studies by Stuurman *et al*. (1996) have shown that the head-to-tail overlap between parallel molecules in *Drosophila* lamin $Dm_0$ involves the same highly conserved helix initiation and termination regions at either end of the rod domain as used by vertebrate cytoplasmic intermediate filament molecules. In addition, they showed that two point mutations in $Dm_0$ which are similar to those which cause keratinopathies in human

https://doi.org/10.1017/S0033583500003516 Published online by Cambridge University Press

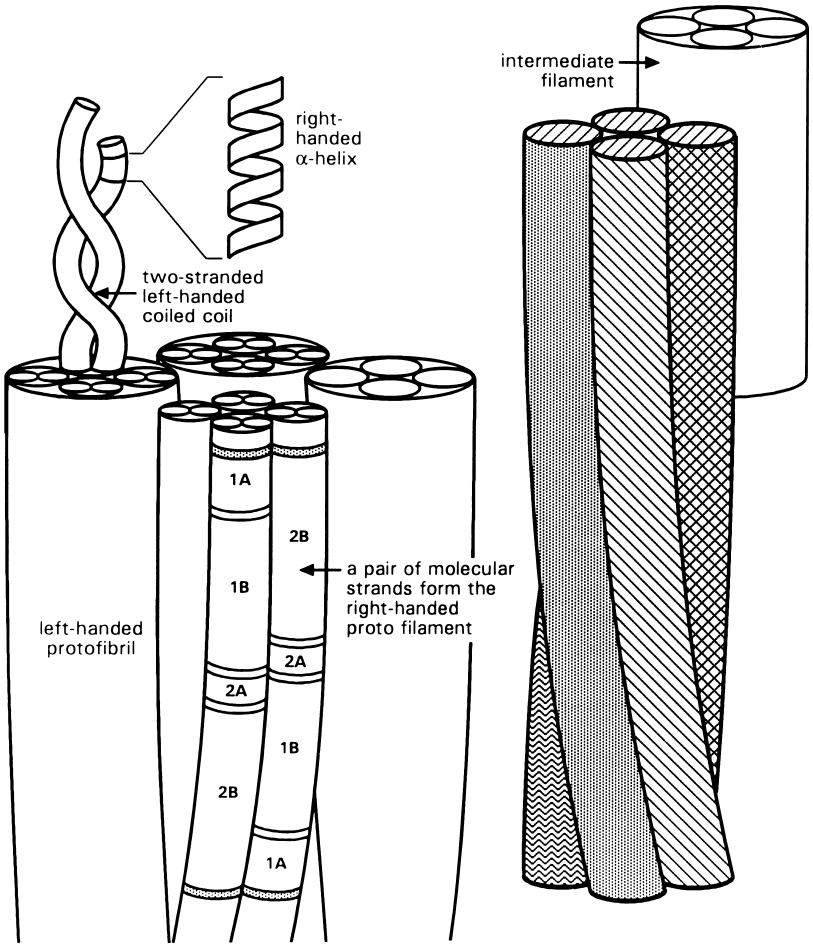



**Fig. 26.** Schematic of the possible structural hierarchy of an intermediate filament. Individual chains adopt a right-handed α-helical conformation, and two parallel chains aggregate in axial register to form a left-handed coiled-coil molecule. The molecules aggregate further, firstly by an antiparallel arrangement involving a strong 1B–1B interaction ($A_{11}$) and then by an antiparallel arrangement involving a 2B–2B interaction ($A_{22}$). As a consequence of these interactions, head-to-tail overlaps occur between parallel molecules ($A_{CN}$) generating two antiparallel molecular strands each composed of parallel molecules overlapped in a heat-to-tail manner. The structure thus formed has four chains in section and is termed a protofilament. Evidence for its existence as a discrete entity, however, is not unequivocal. It is likely, but not proven, that two pairs of molecular strands (that is, two protofilaments with eight chains) interacting via the $A_{12}$ mode of interaction will wind around a common axis in a right-handed manner; that is, with a handedness opposite to that in the preceding level of the structural hierarchy. This generates the structure known as a protofibril. Evidence for this entity arises primarily from STEM data (see text). Finally, the four protofibrils (32 chains in section as determined from STEM data) would likely coil around one another in a left-handed manner to generate the intact intermediate filament. Some of the higher levels of the putative structural hierarchy given here remain speculative and will require further experimental verification. From Parry & Steinert (1995).

keratins, interfere with head-to-tail associations. Sasse *et al*. (1997) developed this work further by investigating the assembly of recombinant *Drosophila* lamin proteins that lacked their head domains, their tail domains or both. The absence of the head domain seemed to inhibit axial assembly whereas the absence of the tail domain resulted in the accelerated

https://doi.org/10.1017/S0033583500003516 Published online by Cambridge University Press

production of filamentous structures. Sasse *et al*. (1998) showed that these particular structures formed protofilaments not from head-to-tail assemblies of parallel molecules but by the production of half-staggered arrays of molecules. They also revealed that a 41 residue segment (at most) in the tail domain of $Dm_0$ lamin is involved in altering the assembly from predominantly a lateral process to a predominantly axial one. The conclusion to be drawn from these four pieces of work is that the general modes of assembly of all intermediate filaments are the same irrespective of the rod domain length or the presence/absence of link segments.

## 4.2 Head and tail domains

The role of the head and tail domains in specifying the quaternary structure of intermediate filaments has not been easy to determine, although experimental evidence using increasingly sophisticated techniques has now begun to provide strong indications as to their likely contributions. Using electron microscopy and STEM, Heins *et al*. (1993) studied fragments of the mouse NF-L chain that had been expressed in bacteria, purified and then allowed to assemble into filaments under specified conditions. The first observation made was that genuine intermediate filaments (or the closely related but unglycosylated structures observed in the electron microscope) required not only the rod domain but both the head and tail domains as well. Heins *et al*. (1993) proposed that the head domain promoted the lateral assembly of protofilaments into protofibrils and that the tail domain controlled the lateral assembly of the protofilaments so that a filamentous structure of the correct diameter was formed. In addition, the tail domain was thought to play an important part in facilitating interactions with other filaments in the cell, and thus in defining the networks so formed (Heins & Aebi, 1994). The point has also been made by Heins & Aebi (1994) that minor changes in assembly conditions *in vitro* may result in ultrastructural changes in the filaments. Certainly, the possibility that *in vitro* assemblies differ meaningfully from *in vivo* ones cannot be ignored. A consistent finding also is that *de novo* filament formation permits less variation in the structure of the intermediate filament chain than does incorporation of these chains into an existing intermediate filament (Heins & Aebi, 1994). Studies on cytokeratins (Hatzfeld & Burba, 1994), desmin (Kaufmann *et al*. 1985) and vimentin (Herrmann *et al*. 1996a) in which the head domains of the chains were largely deleted have shown that no oligomers beyond dimers and tetramers exist under normal filament-forming conditions. Herrmann & Aebi (1998a, b) suggested that either the dimers formed associate in the fully overlapped antiparallel $A_{12}$ mode (which is known to be an unproductive one; see Steinert, 1991a, b), or that the head domains play an important part in allowing the accretion of additional dimers.

More recent work by Herrmann *et al*. (1998b) included a study of the assembly of human and frog vimentin that lacked either the head or tail domains. They showed that the deletion of the head domain prevented the formation of intermediate filaments. In marked contrast, however, there was little discernible effect on filament assembly when the tail domain was deleted, though STEM data indicated that alternative structural forms existed depending on the initial conditions. This indicates that the diameters of the structures were no longer under strict control, and it is reasonable to assume that this could be attributed (in part at least) to the absence of the tail domains. However, the axial growth of the intermediate filaments for human and frog, as well as for vimentin lacking its tail domain, proceeded normally through the intermediary of unit length filaments. These structures were flexible, about 65 nm long,

https://doi.org/10.1017/S0033583500003516 Published online by Cambridge University Press

thicker in their central regions and contain as many as 16 pairs of antiparallel molecules approximately half-staggered with respect to one another (Herrmann & Aebi, 1998). These unit length filaments then undergo axial annealing to generate the full intermediate filament (Herrmann *et al*. 1996a). In contrast to the conclusion that the tail domains of the intermediate filament chains control and limit filament diameters, Winter *et al*. (1997c) have shown that the Type I trichocyte keratin chain hHa1 (which lacks a tail domain *in vivo*) and its Type II partner hHb3 form what are apparently completely normal intermediate filaments. It remains a matter of conjecture whether or not the tail domain in one rather than both chains in the molecule is sufficient to control lateral assembly, but the evidence of Winter *et al*. (1997c) would suggest that this might be so. Further data, however, will be required to confirm this point.

### 4.3 Surface lattice structure

The X-ray diffraction data necessary to elucidate the surface lattice structure of vertebrate intermediate filaments have proved difficult to obtain, primarily because the specimens do not naturally have (or cannot be induced to have) sufficiently high orientations to allow high quality and high resolution data to be obtained. Nonetheless, the trichocyte keratins (and they alone) have been amenable to such analyses, primarily as a consequence of both the very high natural orientation that specimens such as porcupine quills exhibit *in vivo* and to the small X-ray beam size used in these experiments.

The term 'surface lattice' is not one that many find easy to understand and so some explanation may be helpful at this point. Suppose that the filament structure of an intermediate filament is not only known but is so large that it can be seen and touched. A (large) piece of transparent paper could be wrapped around the structure so that the edges of the paper parallel to the long axis of the filament just touch one another. In other words the paper would just wrap round the structure once without overlapping anywhere. Looking through the paper a mark could be made on it at every point where some recurring structural motif appeared. After all of these motifs were marked the paper could be opened up and the pattern of marks examined. This is a two-dimensional representation of a surface lattice structure. It provides knowledge about the axial and azimuthal coordinates on which identical structural motifs are positioned. In practice, the problem must be tackled very differently, of course. It is necessary to determine not only the surface lattice using X-ray diffraction data but also the structure of the motif itself.

The meridional data from wool keratin were initially believed to be accounted for by an axial period of 19·8 nm (Bear, 1944), but specimens with higher orientation used in conjunction with X-ray beams of smaller sizes revealed that the true period was 47·0 nm (Fraser & MacRae, 1971). A similar value of 46·8 nm has been determined from baboon hair by using a synchrotron source and an imaging plate detector (James *et al*. 1995). Curiously, the majority of the expected meridional reflections are either absent or do not appear to have any significant intensity. The challenge posed by these observations was solved by Fraser & MacRae (1983, 1985), who showed that these data could be accounted for by an infinite axial lattice of spacing 47·0 nm convoluted with a structure having an axially projected period of 19·8 nm and containing about eight repeats. This model would naturally put intensity along the meridional region of the diffraction pattern at about 19·8 nm axial intervals (as observed). The sampling of these data by the 47·0 nm axial period then accounts not only for the absence of many reflections but also the presence of those that are indeed observed. In addition there

https://doi.org/10.1017/S0033583500003516 Published online by Cambridge University Press

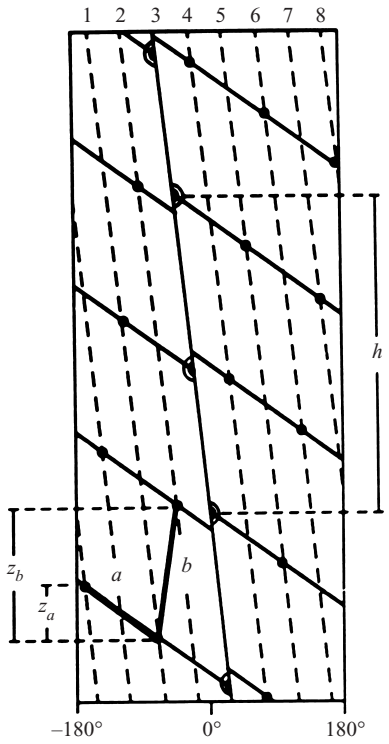



**Fig. 27.** The surface lattice structure of a wool trichocyte keratin intermediate filament showing eight protofilaments per repeat (Fraser *et al*. 1976, 1990; Fraser & MacRae, 1983, 1985). The experimentally determined values of $h$, $z_a$ and $z_b$ are 47·0, 7·42 and 19·79 nm respectively. The physical significance of $z_a$ and $z_b$ is that they correspond respectively to the axial distances between consecutive lattice points and between adjacent protofilaments. The lattice, however, displays a helical discontinuity along the line which forms the boundary when the lattice is rolled up in to a tube that corresponds, topologically, to the intermediate filament. The helix thus formed has a unit height of 47 nm, a unit twist of 49·1° and a pitch length of 344·7 nm. Figure from Parry & Steinert (1995).

were many off-meridional reflections in the diffraction patterns. These, together with the meridional information, allowed the dimensions of the surface lattice to be determined. The lattice sides had axially projected lengths $z_a$ and $z_b$ equal to 7·42 and 19·79 nm respectively (Fig. 27). Both left- and right-handed surface lattice structures are possible and further data will be required before the two options can be distinguished. Rather unusually (but not uniquely) the proposed surface lattice contains a helical dislocation. Its form is described more fully below in the context of the cytokeratins. Having ascertained the surface lattice structure for wool keratin the next step was to determine the common structural motif that exists on each lattice point. It was thus necessary to relate the X-ray data to the postulated number of chains in section (32) and to the modes of aggregation ($A_{11}$, $A_{22}$ and $A_{12}$) elucidated from the crosslinking studies. In this regard, Parry (1995) had undertaken an analysis of the likely molecular arrangements in trichocyte keratins based on the same approach used in the crosslink study undertaken for other intermediate filaments. Further, the mass data are relatively easy to interpret and are compatible with each lattice point being associated with a tetramer. Thus, using the $A_{11}$ and $A_{22}$ modes of aggregation to generate protofilaments, and the $A_{12}$ mode to facilitate the aggregation of two protofilaments into a four-molecule-in-

https://doi.org/10.1017/S0033583500003516 Published online by Cambridge University Press

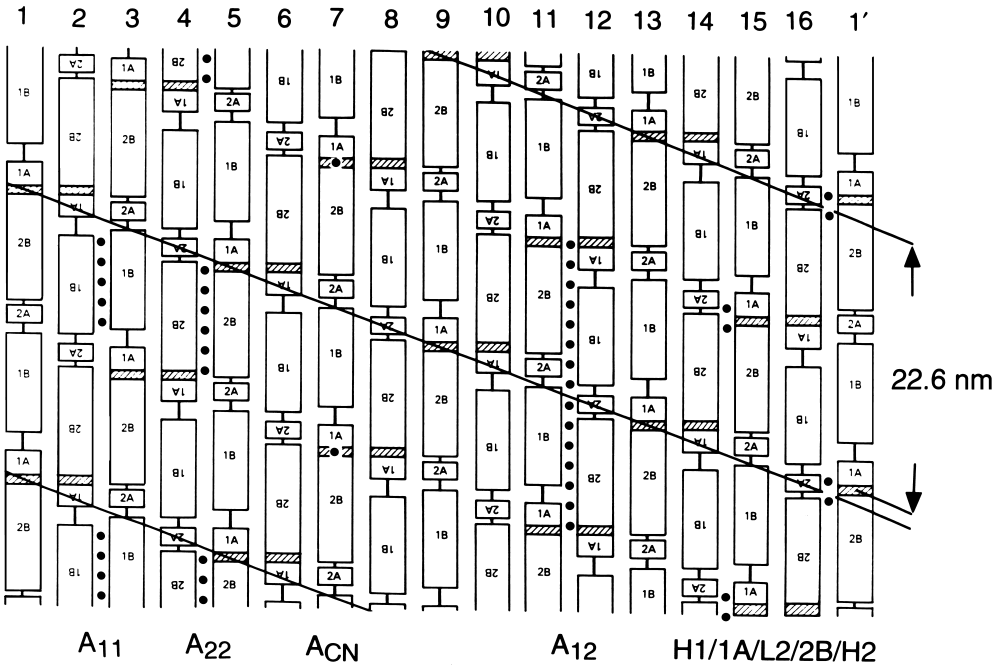



**Fig. 28.** Surface lattice structure proposed for epidermal keratin (Steinert *et al*. 1993a). In section the model contains 32 chains (i.e. 16 molecules as marked) as indicated by the STEM data. Eight protofilaments are formed as a consequence of $A_{11}$ and $A_{22}$ interactions (indicated by large dots). Neighbouring protofilaments are linked to one another by $A_{12}$ interactions (large dots) to form either a sheet or, more likely, an eight-chain protofibril. Other interactions of importance that have been established by Steinert & Parry (1993) are those involving the segments or subdomains H1/1A/L2/2B/H2. These too are indicated by large dots. Note that a seventeenth molecule (labelled 1′) is directly translated from the first one (labelled 1) to illustrate that the lattice does not match up exactly and that a helical dislocation will occur. The extent of the lattice mismatch is indicated by the arrows. The pitch length of the structure as drawn is 22·6 nm, an axial value observed experimentally in electron micrographs by Steven *et al*. (1985). The repeat length in the protofilaments (45·2 nm) is twice the pitch length. Figure from Steinert *et al*. (1993a).

section protofibril, it is possible to account for all of the observed data, including those derived from X-ray diffraction (Fig. 27).

Surface lattice structures can be proposed for other intermediate filaments on the basis of that derived for trichocyte keratin. It must be emphasised again, however, that independent X-ray data are not available for these cases and that the derived surface lattices require independent verification. With that proviso, consider firstly the surface lattice proposed for K1/K10 and K5/K14 cytokeratin filaments (Fig. 28; Steinert *et al*. 1993a). This is based on the assumption that these keratin molecules are arranged regularly and in a manner defined by the $A_{11}$, $A_{12}$, $A_{22}$, $A_{CN}$ modes of interaction described earlier. In section the model contains 16 molecules to agree with the STEM data. Protofilaments could be formed through $A_{11}$ and $A_{22}$ interactions, and neighbouring protofilaments could be linked to one another by $A_{12}$ interactions to form either a sheet or four eight-chain protofibrils. Other established interactions of importance are those involving the segments or subdomains H1/1A/L2/2B/H2. A seventeenth molecule (labelled 1′ in Fig. 28) is directly translated from the first one (labelled 1) to illustrate that the lattice does not match up exactly and that a helical dislocation will occur. The extent of the lattice mismatch is indicated by the arrows. The pitch

https://doi.org/10.1017/S0033583500003516 Published online by Cambridge University Press

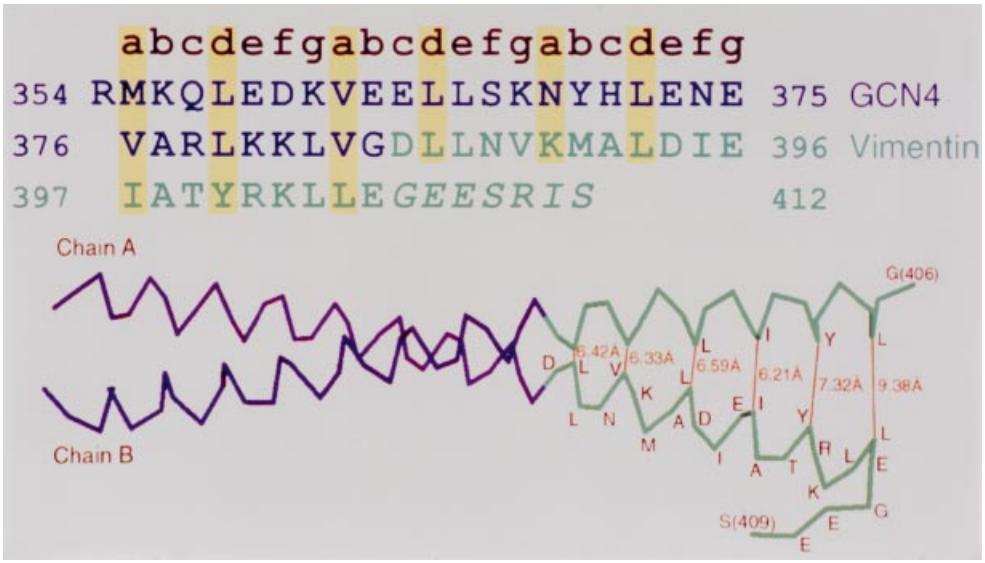



**Fig. 29.** The amino acid sequence of GCN4 linked covalently to the highly conserved C-terminal end of segment 2B (from human vimentin) is shown at top, together with the heptad designation of the residues. The crystal structure of this peptide has been solved to 1·9 Å resolution (Herrmann, Strelkov, Feja, Rogers, Brettel, Lustig, Häner, Parry, Steinert, Burkhard & Aebi, unpublished) and shows that the backbone of the coiled-coil begins to splay out a little from Ile397. At Leu 404 one chain bends away from the coiled-coil and folds back on itself. Figure courtesy of Ueli Aebi.

length of the structure as drawn is 22·6 nm, an axial value observed experimentally in electron micrographs of epidermal cytokeratin intermediate filaments by Steven *et al*. (1985). The repeat length in the protofilaments (45·2 nm) is twice this pitch length. The predicted values of $z_a$ and $z_b$ for the cytokeratin surface lattice are 11·93 and 16·64 nm respectively.

An identical approach has been used to assess the common surface lattice structure for Type III, Type IV and Type III/IV intermediate filaments. As before, the model will require independent verification and that is likely to rely on the acquisition of high resolution X-ray diffraction data. The surface lattice would have predicted $z_a$ and $z_b$ values of 3·72 and 20·26 nm respectively. In this case the pitch length of the structure would be 22·1 nm. There are no direct electron microscope data from these filaments to allow this value to be corroborated independently, but paracrystals of neurofilaments do have an axial period of about 21·35 nm. This would represent a minimum value and is hence reasonably compatible with the predicted value. Interestingly, the $z_b$ values determined for trichocyte keratins, cytokeratins and vimentin are, whilst different, still relatively similar to one another (19·79, 16·64 and 20·26 nm respectively). Much greater variation is evident with the $z_a$ values (7·42, 11·93 and 3·72 nm respectively), though whether or not this has any significance is matter of speculation.

X-ray diffraction patterns generated by the highly oriented neurofilaments in the giant axon from the invertebrate *Myxicola* have been recorded and these display a different axial period from that seen for trichocyte keratins (25·2 nm versus 47·0 nm; Wais-Steider *et al*. 1983). This is not surprising, of course, since the molecular length in some invertebrate intermediate filaments is considerably longer than that found for the vertebrates (approximately 52 nm versus 46 nm) as a result of the insertion of an additional six heptads in the 1B segment of the invertebrate chains. There have been no further detailed studies of this structure but the

https://doi.org/10.1017/S0033583500003516 Published online by Cambridge University Press

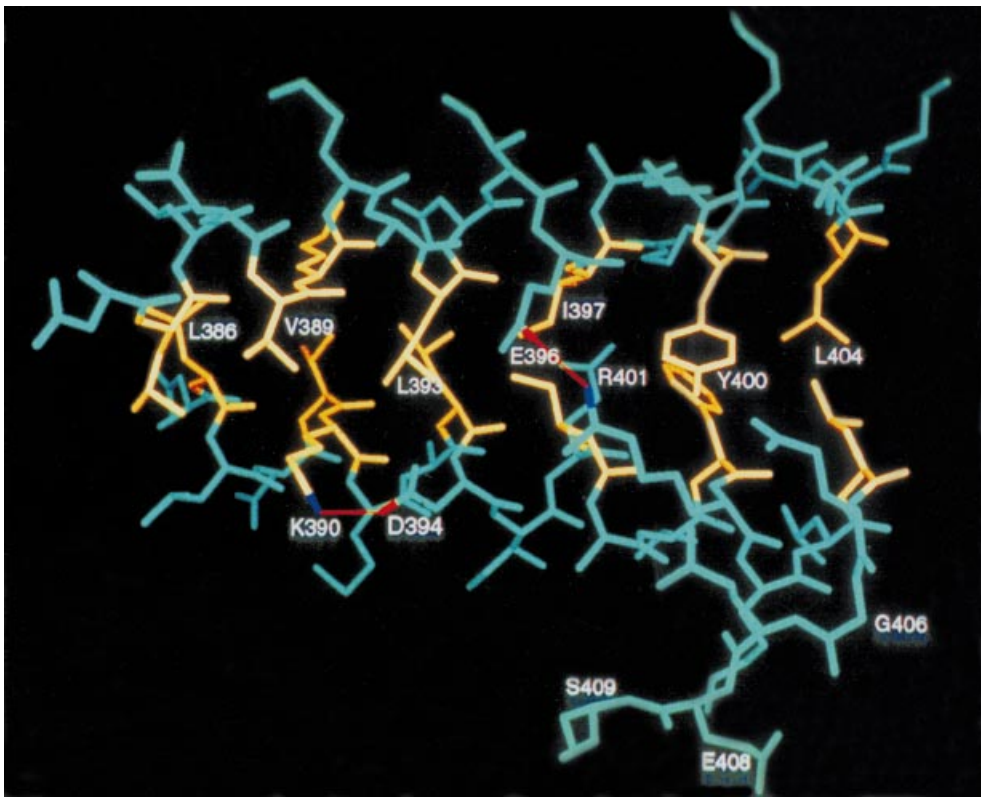



**Fig. 30.** As in Fig. 29 but showing the packing of all of the atoms along the axis of the coiled-coil for the terminal portion of segment 2B. Figure courtesy of Ueli Aebi.

expectation is, nonetheless, that general features deduced from other intermediate filaments will prove to be closely related to those present in *Myxicola* neurofilaments.

### 4.4 Crystal studies on intermediate filament fragments

Considerable efforts have been devoted to crystallizing intermediate chains or molecular fragments thereof, or preparing solutions suitable for NMR study. The results so far have been disappointing though much work continues to be done in this area. The recent discovery that a trigger motif appears to be a necessary prerequisite in a heptad sequence for the formation of a two chain coiled-coil (Kammerer *et al*. 1998; Steinmetz *et al*. 1998) may represent a significant step forward in designing peptide fragments for crystallization. Towards that approach our latest attempts, in conjunction with Herrmann, Strelkov, Feja, Rogers, Brettel, Lustib, Häner, Burkhard and Aebi, have centred on efforts to crystallize or co-crystallize peptides consisting of segment 1A or segment 2B linked directly to the GCN4 leucine zipper sequence, which does contain such a trigger motif. This has resulted in some real progress being made, all of which occurred during the time that this review was being written. The GCN4–2B peptide designed and prepared by us has been crystallized and solved to 1·9 Å resolution in the laboratory of Ueli Aebi in Basel. The structure is indeed a coiled-coil but it does diverge a little over the last heptad predicted from the sequence data. Indeed, one chain folds back on itself at this point (Figs. 29, 30). The 1A-GCN4 peptide has not been

https://doi.org/10.1017/S0033583500003516 Published online by Cambridge University Press

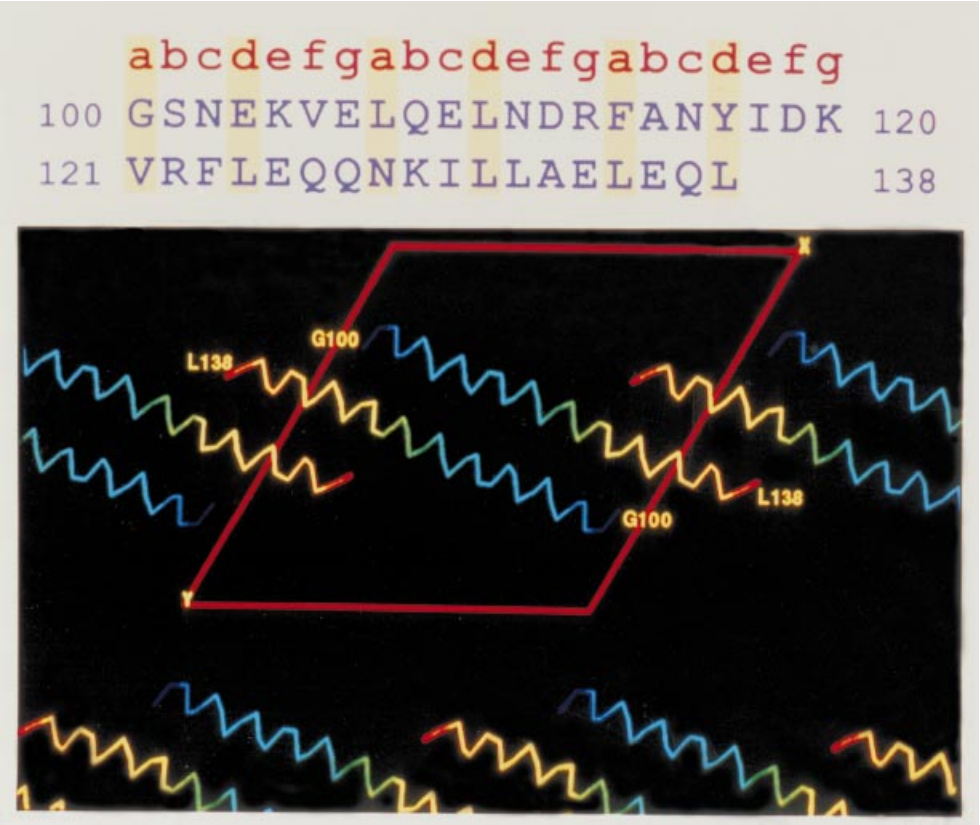



**Fig. 31.** Segment 1A from human vimentin (sequence at top in its heptad designation) has been crystallized and the structure solved to a resolution of 1·4 Å (Strelkov, Aebi *et al.*). Only monomeric α-helices have been observed. Figure courtesy of Ueli Aebi.

crystallized yet, however, but the structure of segment 1A in human vimentin (which apparently lacks a trigger motif) has been solved to 1·4Å resolution by Aebi, Strelkov and colleagues in Basel. The conformation observed for this peptide is not a coiled-coil but monomeric α-helices (Fig. 31). It is too early to comment meaningfully on the significance of both sets of observations but these crystal studies do represent a major step forward in our understanding of the structure of intermediate filaments. It is to be hoped and expected that many more high resolution structures will be solved over the next 5–10 years.

## 5. Polymorphism

Polymorphism is a feature of many proteins, but most especially the fibrous proteins. The latter almost always contain long, sometimes extremely long, regions with an underlying sequence repeat consistent with a particular secondary structure. The sequence repeat naturally generates periodicities in the linear distributions of acidic residues, basic residues and/or apolar residues. Because fibrous protein aggregation is largely a consequence of

https://doi.org/10.1017/S0033583500003516 Published online by Cambridge University Press

intermolecular apolar interactions and/or intermolecular ionic interactions, it follows that the sequence regularities will naturally give rise to a family of packing arrangements with closely related axial alignments. The energy differences between these variants will often be extremely small, and minor changes in assembly conditions may easily result in different forms being adopted *in vitro*. In the intermediate filaments the rod domain displays a very regular linear distribution of acidic residues and also basic residues. The periodicity is about 9·54 residues in segment 1B for both residue groupings, and about 9·84 residues in segment 2 (and 2B in particular), also for both residue groupings. Early models of Crewther *et al*. (1983) were thus based on the proposal that intermolecular ionic interactions would be maximized and, therefore, that the intermediate filaments would contain half-staggered arrays of antiparallel molecules with their 1B segments overlapped, with their 2B segments overlapped, and with the whole molecules overlapped. The crosslinking studies (detailed in Section 3.3) provided strong support for these models and gave details of the relative molecular staggers and the projected lengths of the links L1, L12 and L2, which enabled even more detailed structures for the intermediate filament to be proposed (Steinert *et al*. 1993a–c). Axial data derived from X-ray diffraction patterns and electron microscopy have confirmed the general packing scheme and shown that the axial periods of at least four groups of intermediate filaments do differ slightly but significantly from one another. The remainder of this section will describe these conclusions in more detail.

## 5.1 Variations on a theme

The similarity at electron microscope level of intermediate filaments from diverse sources led many in the field to the (erroneous) conclusion that all intermediate filaments had the same structure and, hence, that this was independent of cell type origin. There is a wealth of data now, however, to indicate that there are at least four unique intermediate filament structures. These are epitomized by: (i) the trichocyte keratins (Type Ia/IIa); (ii) the cytokeratins (Type Ib/IIb); (iii) the class represented by vimentin, $\alpha$-internexin, neurofilaments and possibly Type VI (Type III/III, Type III/IV, Type IV/IV, and possibly III/VI and IV/VI); and (iv) the nuclear lamins (Type V/V). These forms, nonetheless, bear remarkable similarities to one another as they all appear to contain the same number of chains in cross-section (approximately 32) and, furthermore, they utilize the same three general modes of axial interactions. The differences lie almost totally in relatively small axial shifts between molecules, which result in changes to the axial period and hence to the surface lattice structure (Section 4.4). The best estimate (or in the case of wool trichocyte keratin the experimentally measured value) of the axial period in the four unique forms of intermediate filaments seen *in vivo* are 47·0, 45·2, 44·2 and 50 nm respectively. Section 5.1.1 will consider the variation that exists between the axial assembly parameters of the four unique intermediate filament structures thus far characterized, and Section 5.1.2 will deal with the observed variation that occurs in the number of chains in cross-section in *in vitro* assemblies and in non-viable *in vivo* ones.

### 5.1.1 Axial structure

The DST crosslinking methodology has been described earlier, as have the results on the molecular packing parameters and the projected link lengths. The conclusions are simply

https://doi.org/10.1017/S0033583500003516 Published online by Cambridge University Press

stated: the parameters for the Type Ib/IIb cytokeratins are different from the parameters common to Types III/III, III/IV and IV/IV intermediate filaments. Hence there must be at least two unique intermediate filament structures. The nuclear lamins contain an additional 42 residues in their rod domain and hence must be different from the two structures just described. Trichocyte keratins, in turn, have a measured axial period that is different from that deduced for all of the other intermediate filaments. Thus, there must be a minimum of four unique structures within the form known generically as the intermediate filament. Data derived from the trichocyte keratins have been subjected to a least squares analysis roughly equivalent to that undertaken for the cytokeratins, vimentin and $\alpha$-internexin. The results are less secure since the number of data are equal to the number of parameters. With that major proviso, however, it seems that the parameters governing the axial assembly of the molecules in trichocyte keratin intermediate filaments are a hybrid of those deduced for the cytokeratins and those for the Type III intermediate filaments. A key difference lies in the head-to-tail overlap of about eight residues that is found in all intermediate filaments except the trichocyte keratins. In the latter case, when oxidized, these intermediate filaments have a head-to-tail gap of about eight residues between similarly directed molecules in the filament instead (Parry, 1995).

These results have important connotations for intermediate filament assembly and dynamics in cells. Many examples are known where several chain types co-exist within a single cell (Type I, Type II and Type III chains, for example). Clearly, some mechanism is required to prevent intermediate filaments being formed from functionally inappropriate combinations of chains. Nature, it seems, has devised a particularly simple way of achieving this end. Molecules with different lengths and differing axial assembly parameters are extremely unlikely to copolymerize into viable intermediate filaments and, indeed, this seems to be the case. When expressed after transfection or upon micro-injection, Type I/II and III chains sort themselves onto their respective separate networks in living cells without apparent interference. Conversely, there are numerous examples in which similar or compatible chain types replace each other during development and differentiation. For example, neuronal cells consecutively express Type VI, Type III, and then several Type IV chains, all by an apparent dynamic replacement scheme. Similarly, in keratinocytes, keratin molecules containing one particular chain pair will ultimately be replaced by molecules comprising a different chain pair as differentiation proceeds. In each system this is possible because the molecules have compatible lengths and axial assembly characteristics as demonstrated by the crosslinking data. In this regard, it is interesting that trichocyte keratins also appear to begin life by a dynamic replacement mechanism of existing cytokeratin intermediate filaments, which supports our ideas outlined above that the trichocyte keratin molecules become rearranged following oxidation as the tissues harden into hair, nails, etc.

### 5.1.2 Lateral structure

Transmission electron microscopy is unable to provide diameter data for the intermediate filaments that is sufficiently accurate to allow the number of molecules in cross-section to be determined except within rather wide limits. Even accurately measured X-ray diffraction data derived from the low-angle equatorial patterns (Fraser & MacRae, 1958) and which relate to

https://doi.org/10.1017/S0033583500003516 Published online by Cambridge University Press

the intermediate filament diameter are capable of more than a single interpretation. If the available data are assumed to arise from a uniform solid cylinder then its diameter would necessarily be about 7.45 nm in order to explain the positions of the observed X-ray diffraction maxima and minima. If, on the other hand, it is assumed that the intermediate filament has a core that may be more or less dense than the annular region then it is possible for the intermediate filament to have a larger external diameter and still be consistent with the observed X-ray data. Hence, the numbers of chains that could fit within the envelopes defined by the two structures can (and do) differ significantly from about 23 to 32 (Parry, 1996). The best approach to determining the number of intermediate filament chains in section is by using the mass-per-unit-length measurements derived from a STEM analysis. Even these data have some variation but by analysing many different intermediate filaments (Table 4) very clear trends emerge, as have already been shown. Firstly, an intermediate filament with 32 chains in section is common to both *in vivo* and *in vitro* assemblies. The most direct interpretation of these results is that all intermediate filaments contain the same number of chains in section (approximately 32) but that *in vitro* it is possible for other well-defined (but artefactual) polymorphs to exist. Such polymorphs differ from one another in containing differing multiples of eight chains – 16, 24, 32, 40 and 48 chains. As discussed above, the eight-chain unit has been termed the protofibril. It would seem that *in vitro* it is possible for two, three, four, five or even six protofibrils to aggregate. The simplest interpretation of these observations is that some degree of specificity present *in vivo* has been lost *in vitro*. It remains a matter of speculation as to what this might be.

## 6. Keratin intermediate filament chains in diseases

One of the most exciting developments in the field of intermediate filament research in recent years has been the discovery that mutations in the genes that encode the various chains often result in disease. Most mutations that have been observed to date involve simple single nucleotide changes, resulting in inappropriate amino acid substitutions in an otherwise perfect chain. A minority of mutations have involved nucleotide insertions or deletions, or other more rare occurrences such as intron/exon splice errors etc., which together usually result in premature chain terminations and a truncated protein. To date, the vast majority of mutations have been observed in keratin genes which cause a variety of epithelial disorders or 'keratinopathies'. Almost all mutations are transmitted as a dominant negative trait, which means that only one allele is affected and that the abnormal protein chain expressed from that allele is sufficient to cause disease. In virtually all cases of disease, the abnormal protein accumulates as insoluble clumps inside the keratinocytes, and interferes with the normal dynamic and elaborate three-dimensional organization of the filaments in cells. This results in a physical weakening of the cell, interferes with the normal cell–cell interactions through desmosomes and other junctional complexes, so that the entire tissue is weakened and breaks, often manifesting as blisters. Some diseases also involve hyperkeratosis and/or reduced skin barrier function, which may be secondary effects of the disturbed epithelium. Several reviews have been written recently which have summarized the history of discovery, genetics, complexity, pathophysiology and locations of the disease-causing changes, of which some include: Parry & Steinert (1995), McLean & Lane (1995), Korge & Krieg (1996) and Irvine

https://doi.org/10.1017/S0033583500003516 Published online by Cambridge University Press

**Table 5.** *Keratin intermediate filaments and diseases*

| | Keratin chain | |
|---|---|---|
| Disease | I | II |
| Epidermolysis bullosa simplex | K14 | K5 |
| Epidermolytic hyperkeratosis | K10 | K1 |
| Epidermolytic palmoplantar keratoderma | K9 | |
| Diffuse non-epidermolytic palmoplantar keratoderma | | K1 |
| Focal non-epidermolytic palmoplantar keratoderma | K16 | K6c, K6d, K6e, K6f |
| Ichthyosis bullosa of Siemens | | K2e |
| Pachyonychia congenita type I | K16 | K6a |
| Pachyonychia congenita type II | K17 | K6b |
| Steatocystoma multiplex | K17 | |
| White sponge nevus | K13 | K4 |
| Cryptogenic cirrhosis | K18 | K8 |
| Meesmann's corneal dystrophy | K12 | K3 |
| Monilethrix | Hhb6, Hhb1 | |

Not yet found: K2p, K7, K19, K20, K21; 18 other trichocyte keratins.

& McLean (1999). Suffice to say that we will not duplicate these reviews here, but will instead comment on the structural implications of some of these mutations. Table 5 lists the keratin chains which have been shown to cause diseases.

It turns out that the vast majority of mutations involve the various rod domain segments and the linker L12. A few have been described for the H1 region of Type IIb chains, and only a couple for distal head domain sequences (Fig. 32). No mutations have been reported yet in large sections of the 1B or 2B rod domain segments, or in the tail domain. To a first approximation, the observed distribution and gross numbers of reported mutations would seem to reflect the importance of these domains in intermediate filament structure, as inappropriate amino acid changes might be expected to alter structure in unacceptable ways. However, the details are not quite as simple. The phenomenon of 'ascertainment bias' is particularly evident in the skin diseases: usually only the more severe cases of disease are brought to the attention of physicians and scientists, so that these tend to be studied first and in more detail. Some of the other non-epidermal diseases are not terribly troubling to the patients, and so are not frequently brought to the attention of researchers. Indeed, it is possible that some of the more informative mutations, from a structural point of view, may cause only mild disease. Despite these limitations, closer inspection of the body of mutation data reveals that the first 15 residues of the 1A segment and the last 10 residues of the 2B segment are scored most frequently. In general, very similar disease presentation occurs when any one of the first 10 residues of the 1A domain, or last 10 residues of the 2B domain, are affected. These of course correspond to the most highly conserved residues across the entire spectrum of intermediate filament chains, and occupy the residue region where molecules overlap in the $A_{12}$ and $A_{CN}$ mode as discussed above. The most often scored residue is an arginine residue in position 10 of the 1A domain of Type Ib cytokeratins, because the dinucleotide CpG island is somewhat more susceptible to mutation at the gene level, and this arginine residue apparently is absolutely critical for the correct three-dimensional orientation of molecules.

https://doi.org/10.1017/S0033583500003516 Published online by Cambridge University Press

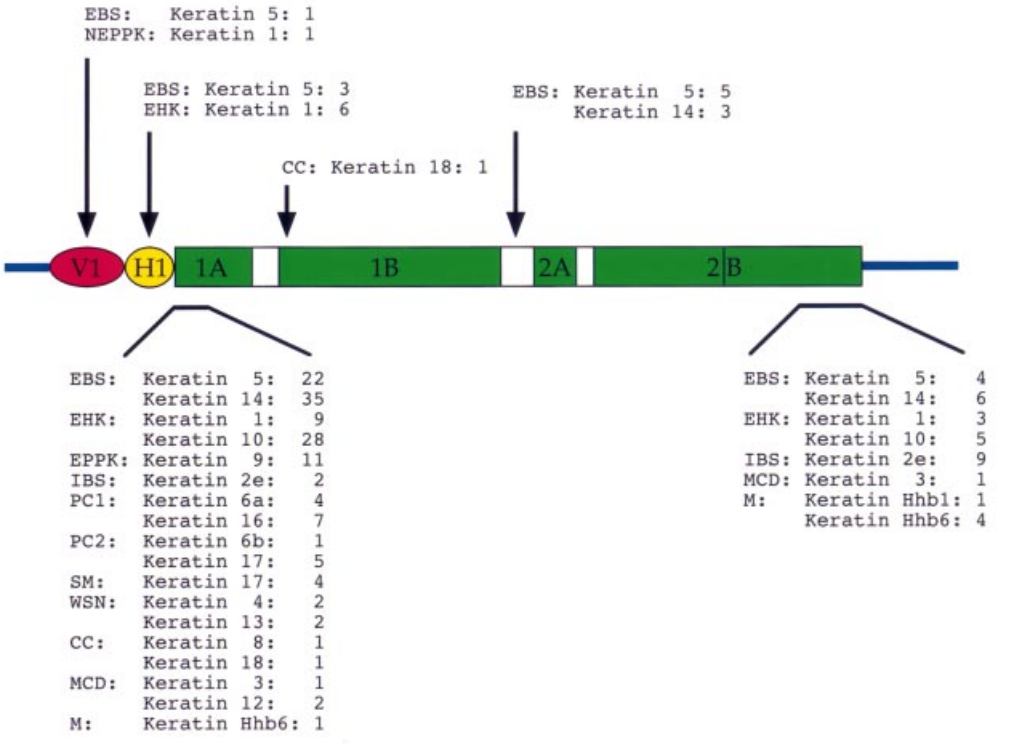



**Fig. 32.** A drawing of a generic Type I and II chain illustrating the locations and numbers of known keratin mutations. The vast majority occur in the vicinity of the helix initiation region at the beginning of segment 1A or in the vicinity of the helix termination region toward the end of segment 2B. In most diseases, either the Type I or II chain can cause equal disease. However, there is some asymmetry in the distribution of occurrence of mutations: for example, most cases of Ichthyosis bullosa of Siemens involve 2B mutations; and generally more mutations have been found in the Type I partner, particularly in position 10 of the 1A segment, but this is likely to be due to the phenomenon of ascertainment bias. All of these mutations represent single nucleotide substitutions resulting in insertion of another amino acid instead. To date, only six other types of mutations (deletions, premature stop codons, etc.) have been reported, and are not listed here. Abbreviations used are: EBS, epidermolysis bullosa simplex; EHK, epidermolytic hyperkeratosis; EPPK, epidermolytic palmar-plantar keratoderma; IBS, Ichthysois bullosa of Siemens; PC1, pachyonychia congenita type 1; PC2, pachyonychia congenita type 2; SM, steatocytoma multiplex; WSN, white sponge nevus syndrome; CC, cryptogenic cirrhosis; MCD, Meesman's corneal dystrophy; and M, monilethrix. List of July 1999, courtesy of John Compton.

To date, the structural/functional consequences of a few of these mutations have been explored. Synthetic peptides bearing the observed mutations in the conserved H1, 1A and 2B regions have been made and used in filament interference studies (e.g. Chipev *et al*. 1992; Steinert *et al*. 1993a, b). In many cases, the abnormal peptide was unable to disassemble preformed filaments in the same way as for the wildtype peptide, thus suggesting indirectly that the mutation in the intact chain had changed the structure and interactions of molecules that stabilize higher orders of filament structure. In one curious case involving an E106D substitution of the 2B domain segment of the K1 chain, which caused a moderately severe case of epidermolytic hyperkeratosis, the mutant synthetic peptide did not interfere with

https://doi.org/10.1017/S0033583500003516 Published online by Cambridge University Press

preformed filaments. However, filaments assembled from a mutant form of the K5 or K14 chain bearing this mutation, together with a wildtype partner chain, were markedly shorter than normal (Yang *et al*. 1999). Similarly, when transfected into living epithelial cells, the mutant proteins behaved abnormally. In this case, molecules and tetramers which formed were normal in terms of their alignments and molecular stability in concentrated urea solutions, but higher orders of filament assembly were affected. Apparently, therefore, a glutamic residue in position 106 is required to stabilize the alignments, whereas a slightly shorter aspartic acid residue instead cannot (P. M. Steinert, unpublished). A similar study has revealed that the arginine residue in position 10 of the 1A rod domain segment is also absolutely required to stabilize $A_{12}$ and $A_{CN}$ molecular interactions. These few data seem to emphasize the extreme importance of the conserved residues at or in the vicinity of the helix initiation and termination motifs of the intermediate filament protein chains. We look forward to the opportunity of examining the atomic level structural organization of these sequences. Indeed, such information eventually may offer useful strategies for management of the diseases in patients afflicted by the mutations.

One final comment should be made concerning the consequences of the one mutation characterized to date located in distal sequences of the head domain of the K1 chain. This mutation involved loss of a lysine residue in the V1 subdomain of the head domain. This residue is required for transglutaminase crosslinking in order to link the keratin intermediate filaments to the cell periphery of stratified squamous epithelia (Candi *et al*. 1998). In this case, the mutant does not cause accumulation of abnormal protein which ultimately results in a blistering presentation, but rather affects the supramolecular organization of the filaments in cells and their association with the cell periphery. Thus this represents a cogent example of where mutations in different domains of a protein chain may result in quite different functional consequences for the filaments and cells.

## 7. Concluding remarks

This review has been devoted to a study of the family of closely related structures known as the intermediate filaments. In doing so attempts have been made to describe just one part of what is an increasingly large and diverse field of study. While efforts have concentrated on detailing the molecular architecture, the modes of assembly of the chains and the molecules, and the polymorphism that is so characteristic of these filaments, it will be clear to the reader that much more detail is still required if we are to approach the level of understanding that we have for many globular proteins, such as enzymes, where in many cases crystal structures have been derived at atomic resolution. Notwithstanding that, much progress has been made and even more is likely to be made over the next 10 years. Although crystals of intermediate filament chains or molecular fragments have remained frustratingly elusive some real progress has been made in recent months. When this facet of the research reaches fruition there can be little doubt that many of the structural and functional enigmas currently faced by researchers will become much clearer. Many of the keratinopathies, for example, can be related back to mutations in the helix initiation and termination motifs. Co-crystallization of appropriate fragments of these regions allied to an X-ray crystallographic analysis could provide important data that would allow a deeper understanding to be gained of how the

https://doi.org/10.1017/S0033583500003516 Published online by Cambridge University Press

disease manifests itself at the molecular level. It would also help to show why the small head-to-tail overlap between parallel molecules is so important during normal assembly of intermediate filaments. Moreover, it may offer a glimmer of hope for treatment for those who are afflicted with these diseases.

It is well known that the functions of the intermediate filaments, as well as part of their ability to assemble, resides within the structures of the head and tail domains of the molecules. The conformations of these regions, however, remain generally unexplored territory though low resolution models have been proposed for some of them. Again, crystal studies or possibly NMR studies will be required to unlock the data so much desired and needed by the intermediate filament community. Surely such analyses will be completed over the next decade.

## 8. Acknowledgments

Writing a review on any subject invariably relies on the assistance and contributions of so many friends, collaborators and colleagues in the field. We thank them all for allowing us access to their recent data and for providing figures. Ueli Aebi in particular must be mentioned in this regard. Any omissions in content remain our responsibility. In some cases, however, difficult decisions were made to limit the text in order to keep the review within specified bounds. This necessitated some selection of material, which we hope will not have detracted from the value of the review.

## 9. References

Abumuhor, I. A., Spencer, P. H. & Cohlberg, J. A. (1998). The pathway of assembly of intermediate filaments from recombinant α-internexin. *J. Struct. Biol.* **123**, 187–198.

Aebi, U., Fowler, W. E., Rew, P. & Sun, T.-T. (1983). The fibrillar substructure of keratin filaments unraveled. *J. Cell Biol.* **97**, 1131–1143.

Aebi, U., Cohn, J., Buhle, L. & Gerace, L. (1986). The nuclear lamina is a meshwork of intermediate-type filaments. *Nature* (*London*) **323**, 560–564.

Ando, S., Tokui, T., Yano, T. & Inagaki, M. (1996). Keratin 8 phosphorylation *in vitro* by cAMP-dependent protein kinase occurs within the amino- and carboxy-terminal end domains. *Biochem. Biophys. Res. Commun.* **221**, 67–71.

Angelastro, J. M., Ho, C. L., Frappier, T., Liem, R. K. H. & Greene, L. A. (1998). Peripherin is tyrosine-phosphorylated at its carboxy-terminal tyrosine. *J. Neurochem.* **70**, 540–549.

Athlan, E. S. & Mushynski, W. E. (1997). Heterodimeric associations between neuronal intermediate filament proteins. *J. Biol. Chem.* **272**, 31073–31078.

Balin, B. J. & Miller, M. E. (1995). Reassembly of the 66 kD neurofilament protein *in vitro* following isolation and purification from bovine spinal cord. *J. Neurosci. Res.* **40**, 79–88.

Bawden, C. S., Powell, B. C., Walker, S.K, & Rogers, G. E. (1998). Expression of a wool intermediate filament keratin transgene in sheep fibre alters structure. *Transgenic Res.* **7,** 273–287.

Bear, R. S. (1944). X-ray diffraction studies on protein fibers. II Feather rachis, porcupine quill and clam muscle. *J. Am. Chem. Soc.* **66**, 2043–2050.

Becker, B., Bellin, R. M., Sernett, S. W., Huiatt, T. W. & Robson, R. M. (1995). Synemin contains the rod domain of intermediate filaments. *Biochem. Biophys. Res. Commun.* **213**, 796–802.

Bilak, S. R., Sernett, S. W., Bilak, M. M., Bellin, R. M., Stromer, M. H., Huiatt, T. W. & Robson, R. M. (1998). Properties of the novel intermediate filament protein synemin and its identification in mammalian muscle. *Arch. Biochem. Biophys.* **355**, 63–76.

Bilbe, G., Delabie, J., Bruggen, J., Richener, H., Asselbergs, F. A. M., Cerletti, N., Sorg C., Odink, K., Tarcsay, L., Wiesendanger, W., DeWolf-Peeters, C. & Shipman, R. (1992). Restin:

https://doi.org/10.1017/S0033583500003516 Published online by Cambridge University Press

a novel intermediate filament-associated protein highly expressed in the Reed–Sternberg cells of Hodgkin's Disease. *EMBO J.* **11**, 2103–2113.

Birbeck, M. S. C. & Mercer, E. H. (1957). The electron microscopy of the human hair follicle. *J. Biophys. Biochem. Cytol.* **3**, 203–214.

Bovenschulte, M., Riemer, D. & Weber, K. (1995). The sequence of a cytoplasmic intermediate filament (IF) protein from the annelid *Lumbricus terrestris* emphasizes a distinctive feature of protostomic IF proteins. *FEBS Lett.* **360**, 223–226.

Brown, J. H., Cohen, C, & Parry, D. A. D. (1996). Heptad breaks in α-helical coiled coils: stutters and stammers. *Proteins: Structure, Function and Genetics* **26**, 134–145.

Bullough, P. A. & Tulloch, P. A. (1990). High-resolution spot-scan electron microscopy of micro-crystals of an α-helical coiled-coil protein. *J. Mol. Biol.* **215**, 161–173.

Burkhard, P., Kammerer, R. A., Steinmetz, M. O., Bourenkov, G. P. & Aebi, U. (2000). The coiled-coil trigger site of the red domain of cortexillin I unveils a distinct network of interhelical salt bridges. *Structure* **8**, 223–230.

Candi, E., Tarcsa, E., DiGiovanna, J. J., Compton, J. G., Elias, P. M., Marekov, L. N. & Steinert, P. M. (1998). A highly conserved lysine residue on the head domain of Type II keratins is essential for the attachment of keratin intermediate filaments to the cornified cell envelope through isopeptide cross-linking by transglutaminases. *Proc. Natl. Acad. Sci. USA.* **95**, 2067–2072.

Capetanaki, Y., Milner, D. J. & Weitzer, G. (1997). Desmin in muscle formation and maintenance: knock-outs and consequences. *Cell Struct. Funct.* **22**, 103–116.

Carpenter, D. A. & Ip, W. (1996). Neurofilament triplet protein interactions: evidence for the preferred formation of NF-L-containing dimers and a putative function for the end domains. *J. Cell Sci.* **109**, 2493–2498.

Carter, J. E., Robertson, J., Anderton, B. H. & Gallo, J. M. (1997). Incorporation of NF-L into keratin filaments in transfected epithelial cells. *Neuroreport* **8**, 2225–2228.

Chan, Y. M., Anton Lamprecht, I., Yu, Q. C., Jäckel, A., Zabel, B., Ernst, J. P. & Fuchs, E. (1994). A human keratin 14 'knockout': the absence of K14 leads to severe epidermolysis bullosa simplex and a function for an intermediate filament protein. *Genes Develop.* **8**, 2574–2587.

Ching, G. Y. & Liem, R. K. H. (1993). Assembly of Type IV neuronal intermediate filaments in non-neuronal cells in the absence of preexisting cytoplasmic intermediate filaments. *J. Cell Biol.* **122**, 1323–1335.

Ching, G. Y. & Liem, R. K. H. (1998). Roles of head and tail domains in α-internexin's self-assembly and coassembly with the neurofilament triplet proteins. *J. Cell Sci.* **111**, 321–333.

Chipev, C. C., Korge, B. P., Markova, N.G, Bale, S. J., DiGiovanna, J. J., Compton, J. G. & Steinert, P. M. (1992). A leucine–proline mutation in the H1 subdomain of keratin 1 causes epidermolytic hyperkeratosis. *Cell* **70**, 821–828.

Chou, P. Y. & Fasman, G. D. (1978). Empirical predictions of protein conformation. *Annu. Rev. Biochem.* **47**, 251–276.

Chou, Y. H., Opal, P., Quinlan, R. A. & Goldman, R. D. (1996). The relative roles of specific N- and C-terminal phosphorylation sites in the disassembly of intermediate filament in mitotic BHK-21 Cells. *J. Cell. Sci.* **109**, 817–826.

Cohen, C. & Parry, D. A. D. (1986). α-Helical coiled-coils–a widespread motif in proteins. *Trends Biochem. Sci.* **11**, 245–248.

Cohen C. & Parry D. A. D. (1990). α-Helical coiled-coils and bundles: how to design an a-helical protein. *Proteins: Structure, Function and Genetics* **7**, 1–15.

Cohen C. & Parry D. A. D. (1994). α-Helical coiled coils: more facts and better predictions. *Science* **263**, 488–489.

Colucci Guyon, E., Portier, M. M., Dunia, I., Paulin, D., Pournin, S. & Babinet, C. (1994). Mice lacking vimentin develop and reproduce without an obvious phenotype. *Cell* **79**, 679–694.

Conway, J. F. & Parry, D. A. D. (1988). Intermediate filament structure. 3. Analysis of sequence homologies. *Int. J. Biol. Macromol.* **10**, 79–98.

Conway, J. F. & Parry, D. A. D. (1990). Structural features in the heptad substructure and longer range repeats of two-stranded α-fibrous proteins. *Int. J. Biol. Macromol.* **12**, 328–334.

Conway, J. F. & Parry, D. A. D. (1991). Three-stranded α-fibrous proteins: the heptad repeat and its implications for structure. *Int. J. Biol. Macromol.* **13**, 14–16.

Conway, J. F., Fraser, R. D. B., MacRae, T. P. & Parry, D. A. D. (1989). Protein chains in wool and epidermal keratin IF: structural features and spatial arrangement. In: *The Biology of Wool and Hair* (eds. G. E. Rogers, P. J. Reis, K. A. Ward & R. C. Marshall), pp. 127–144. Chapman and Hall: London and New York.

Crewther, W. G., Inglis, A. S. & McKern, N. M. (1978). Amino acid sequences of a-helical segments from S-carboxymethylkerateine A: complete sequence of a Type II segment. *Biochem. J.* **173**, 365–371.

Crewther, W. G., Dowling, L. M., Steinert, P. M. & Parry, D. A. D. (1983). Structure of intermediate filaments. *Int. J. Biol. Macromol.* **5**, 267–274.

Day, W. A. & Gilbert, D. S. (1972). X-ray diffraction

https://doi.org/10.1017/S0033583500003516 Published online by Cambridge University Press

pattern of axoplasm. *Biochim. Biophys. Acta* **285**, 503–506.

Dodemont, H., Riemer, D. & Weber, K. (1990). Structure of an invertebrate gene encoding cytoplasmic intermediate filament (IF) proteins: implications for the origin and the diversification of IF proteins. *EMBO J.* **9**, 4083–4094.

Dodemont H., Riemer D., Ledger N. & Weber K. (1994). Eight genes and alternative RNA processing pathways generate an unexpectedly large diversity of cytoplasmic intermediate filament proteins in the nematode *Caenorhabditis elegans*. *EMBO J.* **13**, 2625–2638.

Dong, D. L. Y., Xu, Z. S., Hart, G. W. & Cleveland, D. W. (1996). Cytoplasmic O-GlcNac modification of the head domain and the KSP repeat motif of the neurofilament protein neurofilament-H. *J. Biol. Chem.* **271**, 20845–20852.

Dowling, L. M., Parry, D. A. D. & Sparrow, L. G. (1983). Structural homology between hard α-keratin and the intermediate filament proteins, desmin and vimentin. *Biosci. Rep.* **3**, 73–78.

Dowling, L. M., Crewther, W. G. & Parry, D. A. D. (1986). Secondary structure of component 8c-1 of α-keratin: an analysis of the amino acid sequence. *Biochem. J.* **236**, 705–712.

Eagles, P. A. M., Pant, H. C. & Gainer, H. (1990). Neurofilaments. In: *Cellular and Molecular Biology of Intermediate Filaments* (eds. R. D. Goldman & P. M. Steinert), pp 37–94. New York: Plenum Press.

Eichner, R., Rew, P., Engel, A. & Aebi, U. (1985). Human epidermal keratin filaments: studies on their structure and assembly. In: *Intermediate Filaments* (eds. E. Wang, D. Fischman, R. K. H. Liem & T.-T. Sun), pp. 381–402. New York: New York Academy of Sciences.

Elder, G. A., Friedrich, V. L. Jr., Bosco, P., Kang, C., Gourov, A., Tu, P. H., Lee, V. M. Y. & Lazzarini, R. A. (1998). Absence of the mid-sized neurofilament subunit decreases axonal calibers, levels of light neurofilament (NF-L), and neurofilament content. *J. Cell Biol.* **141**, 727–739.

Elliott, A., Lowy, J., Parry, D. A. D. & Vibert, P. J. (1968). Puzzle of the coiled-coils in the paramyosin filaments of molluscan muscles. *Nature* (*London*) **218**, 656–659.

Engel, A., Eichner, R. & Aebi, U. (1985). Polymorphism of reconstituted human epidermal keratin filaments: determination of their mass-per-unit-length and width by scanning transmission electron microscopy (STEM). *J. Ultrastruct. Res.* **90**, 323–335.

Eriksson, J. E., Brautigan, D. L., Vallee, R., Olmsted, J., Fujiki, H. & Goldman, R. D. (1992). Cytoskeletal integrity in interphase cells requires protein phosphatase activity. *Proc. Natl. Acad. Sci. USA* **89**, 11093–11097.

Fisher, D. Z., Chaudhary, N. & Blobel, G. (1986). cDNA sequencing of nuclear lamins A and C reveals primary and secondary structural homology to intermediate filament proteins. *Proc. Natl. Acad. Sci. USA* **83**, 6450–6454.

Fradette, J., Germain, L., Seshaiah, P. & Coulombe, P. A. (1998). The Type I keratin 19 possesses distinct and context-dependent assembly properties. *J. Biol. Chem.* **273**, 35176–35184.

Fraser, R. D. B. & MacRae, T. P. (1958). Structural implications of the equatorial X-Ray diffraction pattern of α-keratin. *Biochim. Biophys. Acta* **29**, 229–240.

Fraser, R. D. B. & MacRae, T. P. (1971). Structure of α-keratin. *Nature* (*London*) **233**, 138-140.

Fraser, R. D. B. & MacRae, T. P. (1983). The structure of the α-keratin microfibril. *Biosci. Rep.* **3**, 517–525.

Fraser, R. D. B. & MacRae, T. P. (1985). Intermediate filament structure. *Biosci. Rep.* **5**, 573–579.

Fraser, R. D. B. & Parry, D. A. D. (1996). The molecular structure of reptilian keratin. *Int. J. Biol. Macromol.* **19**, 207–211.

Fraser, R. D. B., MacRae, T. P. & Miller, A. (1964). The coiled-coil model of α-keratin structure. *J. Mol. Biol.* **10**, 147–156.

Fraser, R. D. B., MacRae, T. P., Parry, D. A. D. & Suzuki, E. (1969). The structure of β-keratin. *Polymer* **10**, 810–826.

Fraser, R. D. B., MacRae, T. P. & Suzuki, E. (1976). Structure of the α-keratin microfibril. *J. Mol. Biol.* **108**, 435–452.

Fraser, R. D. B., MacRae, T. P., Suzuki, E. & Parry, D. A. D. (1985). Intermediate filament structure. 2. Molecular interactions in the filament. *Int. J. Biol. Macromol.* **7**, 258–274.

Fraser, R. D. B., MacRae, T. P., Parry, D. A. D. & Suzuki, E. (1986). Intermediate filaments in α-keratins. *Proc. Natl. Acad. Sci. USA* **83**, 1179–1183.

Fraser, R. D. B., MacRae, T. P., Sparrow, L. G. & Parry, D. A. D. (1988). Disulphide bonding in α-keratin. *Int. J. Biol. Macromol.* **10**, 106–112.

Fraser, R. D. B., MacRae, T. P. & Parry, D. A. D. (1990). The three-dimensional structure of IF. In: *Cellular and Molecular Biology of Intermediate Filaments* (eds. R. D. Goldman & P. M. Steinert) Ch 8, pp 205–231. New York: Plenum Press.

Fuchs, E. & Weber, K. (1994). Intermediate filaments: structure, dynamics, function, and diseases. *Ann. Rev. Biochem.* **63**, 345–382.

Galligan, P., Listwan, P., Siller, G. M. & Rothnagel, J. A. (1998). A novel mutation in the L12 domain of keratin 5 in the Kobner variant of epidermolysis bullosa simplex. *J. Invest. Dermat.* **111**, 524–527.

Galou, M., Colucci Guyon, E., Ensergueix, D.,

https://doi.org/10.1017/S0033583500003516 Published online by Cambridge University Press

Ridet, J. L., Gimenez, Y., Ribotta, M., Privat, A., Babinet, C. & Dupouey, P. (1996). Disrupted glial fibrillary acidic protein network in astrocytes from vimentin. *J. Cell Biol.* **133**, 853-863.

Galou, M., Gao, J., Humbert, J., Mericskay, M., Li, Z., Paulin, D. & Vicart, P. (1997). The importance of intermediate filaments in the adaptation of tissues to mechanical stress: evidence from gene knockout studies. *Biol. Cell* **89**, 85–97.

Garnier, J., Osguthorpe, D. J. & Robson, B. (1978). Analysis of the accuracy and implications of simple methods for predicting the secondary structure of globular proteins. *J. Mol. Biol.* **120**, 97–120.

Geisler, N., Kaufmann, E. & Weber, K. (1982a). Protein-chemical characterization of three structurally distinct domains along the protofilament unit of desmin 10 nm filaments. *Cell* **30**, 277–286.

Geisler, N., Plessmann, U. & Weber, K. (1982b). Related amino acid sequences in neurofilaments and non-neuronal intermediate filaments. *Nature* (*London*) **296**, 448–450.

Geisler, N., Kaufmann, E. & Weber, K. (1985). Antiparallel orientation of the two double-stranded coiled-coils in the tetrameric protofilament unit of intermediate filaments. *J. Mol. Biol.* **182**, 173–177.

Geisler, N., Schünemann, J. & Weber, K. (1992). Chemical cross-linking indicates a staggered and antiparallel protofilament of desmin intermediate filaments and characterizes one higher-level complex between protofilaments. *Eur. J. Biochem.* **206**, 841–852.

Geisler, N., Schünemann, J., Weber, K., Häner, M. & Aebi, U. (1998). Assembly and architecture of invertebrate cytoplasmic intermediate filaments reconcile features of vertebrate cytoplasmic and nuclear lamin-type intermediate filaments. *J. Mol. Biol.* **282**, 601–617.

Georgatos, S. D., Gounari, F., Goulielmos, G. & Aebi, U. (1997). To bead or not to bead? Lens-specific intermediate filaments revisited. *J. Cell Sci.* **110**, 2629–2634.

Giasson, B. I. & Mushynski, W. E. (1998). Intermediate filament disassembly in cultured dorsal root ganglion neurons is associated with amino-terminal head domain phosphorylation of specific subunits. *J. Neurochem.* **70**, 1869–1875.

Gill, S. R., Wong, P. C., Monteiro, M. J. & Cleveland, D. W. (1990). Assembly properties of dominant and recessive mutations in the small mouse neurofilament (NF-L) subunit. *J. Cell Biol.* **111**, 2005–2019.

Goldman, A. E., Maul, G., Steinert, P. M., Yang, H.-Y. & Goldman, R. D. (1986). Keratin-like proteins that co-isolate with intermediate filaments of BHK-21 cells are nuclear lamins. *Proc. Natl. Acad. Sci. USA* **83**, 3839–3843.

Goldman, R. D., Khuon, S., Chou, Y. H., Opal, P. & Steinert, P. M. (1996). The function of intermediate filaments in cell shape and cytoskeletal integrity. *J. Cell Biol.* **134**, 971–983.

Gough, K. H., Inglis, A. S. & Crewther, W. G. (1978). Amino acid sequences of a-helical segments from S-carboxymethylkerateine A: complete sequence of a Type I segment. *Biochem. J.* **173**, 373–385.

Gournari, F., Merdes, E., Quinlan, R., Hess, J., FitzGerald, P. G., Ouzounis, C. A. & Georgatos, S. D. (1993). Bovine filensin possesses primary and secondary structure similarity to intermediate filament proteins. *J. Cell Biol.* **121**, 847–853.

Gounari, F., Karagianni, N., Mincheva, A., Lichter, P., Georgatos, S. D. & Schirrmacher, V. (1997). The mouse filensin gene: structure and evolutionary relation to other intermediate filament genes. *FEBS Lett.* **413**, 371–378.

Green, K. J., Parry, D. A. D., Steinert, P. M., Virata, M. L. A., Wagner, R. M., Angst, B. D. & Nilles, L. A. (1990). Structure of the human desmoplakins: implications for function in the desmosomal plaque. *J. Biol. Chem.* **265**, 2603–2612.

Green, K. J., Virata, M. L. A., Elgart, G. W., Stanley, J. R. & Parry, D. A. D. (1992a). A comparative structural analysis of desmoplakin, bullous pemphigoid antigen and plectin: members of a new gene family involved in organization of intermediate filaments. *Int. J. Biol. Macromol.* **14**, 145–153.

Green, K. J., Stappenbeck, T. S., Parry, D. A. D. & Virata, M. L. A. (1992b). Structure of desmoplakin and its association with intermediate filaments. *J. Dermatol.* **19**, 765–769.

Hanukoglu, I. & Fuchs, E. (1982). The cDNA sequence of a human epidermal keratin: divergence of sequence but conservation of structure among intermediate filament proteins. *Cell* **31**, 243–252.

Hanukoglu, I. & Fuchs, E. (1983). The cDNA sequence of a Type II cytoskeletal keratin reveals constant and variable structural domains among keratins. *Cell* **33**, 915–924.

Harbury, P. B., Zhang, T., Kim P. S. & Alber, T. (1993). A switch between two- , three-, and four-stranded coiled coils in GCN4 leucine zipper mutants. *Science* **262**, 1401–1407.

Harbury, P. B., Kim, P. S. & Alber, T. (1994). Crystal structure of an isoleucine-zipper trimer. *Nature* (*London*) **371**, 80–83.

Hatzfeld, M. & Burba, M. (1994). Function of Type I and Type II keratin head domains: their role in dimer, tetramer and filament formation. *J. Cell Sci.* **107**, 1959–1972.

Hatzfeld, M. & Weber, K. (1990a). Tailless keratins assemble into regular intermediate filaments *in vitro*. *J. Cell Sci.* **97**, 317–324.

Hatzfeld, M. & Weber, K. (1990b). The coiled-coil of

https://doi.org/10.1017/S0033583500003516 Published online by Cambridge University Press

*in vitro* assembled keratin filaments is a heterodimer of Type I and II keratin: use of site-specific mutagenesis and recombinant protein expression. *J. Cell Biol.* **110**, 1199–1210.

Hatzfeld, M. & Weber, K. (1992). A synthetic peptide representing the consensus sequence motif at the carboxy-terminal end of the rod domain inhibits intermediate filament assembly and disassembles preformed filaments. *J. Cell Biol.* **116**, 157–166.

Hatzfeld, M., Maier, G. & Franke, W. W. (1987). Cytokeratin domains involved in heterotypic complex formation determined by *in vitro* binding assays. *J. Mol. Biol.* **197**, 237–255.

Heimburg, T., Schånemann, J., Weber, K. & Geisler, N. (1996). Specific recognition of coiled coils by infrared spectroscopy: analysis of the three structural domains of Type III intermediate filaments. *Biochem.* **35**, 1375–1382.

Heins, S. & Aebi, U. (1994). Making heads and tails of intermediate filaments assembly, dynamics and networks. *Curr. Opin. Cell Biol.* **6**, 25-33.

Heins, S., Wong, P. C., Müller, S., Goldie, K., Cleveland, D. W. & Aebi, U. (1993). The rod domain of NF-L determines neurofilament architecture, whereas the end domains specify filament assembly and network formation. *J. Cell Biol.* **123**, 1517–1533.

Hemken, P. M., Bellin, R. M., Sernett, S. W., Becker, B., Huiatt, T. W. & Robson, R. M. (1997). Molecular characteristics of the novel intermediate filament protein paranemin. Sequence reveals EAP-300 and IFAPa-400 are highly homologous to paranemin. *J. Biol. Chem.* **272**, 32489–32499.

Hemmati-Brivanluo, A., Mann, R. W. & Harland, R. M. (1992). A protein expressed in the growth cones of embryonic vertebrate neurons defines a new class of intermediate filament protein. *Neuron* **9**, 417–428.

Hennekes, H & Nigg, E. A. (1994). The role of isoprenylation in membrane attachment of nuclear lamins. A single point mutation prevents proteolytic cleavage of the lamin A precursor and confers membrane binding properties. *J. Cell Sci.* **107**, 1019–1029.

Herrling, J. & Sparrow, L. G. (1991). Interactions of intermediate filament proteins from wool. *Int. J. Biol. Macromol.* **13**, 115–119.

Herrmann, H. (1995). Strukturbildende und Topogene Domanen von Intermediar-filamentproteinen, Habilitationsschrift, University of Heidelberg, Germany.

Herrmann, H. & Aebi, U. (1998a). Intermediate filament assembly: fibrillogenesis is driven by decisive dimer-dimer interactions. *Curr. Opin. Struct. Biol.* **8**, 177–185.

Herrmann, H. & Aebi, U. (1998b). In *Subcellular Biochemistry*, *Vol.* 31 (eds. H. Herrmann & R. Harris), pp. 319–362. New York: Plenum Press.

Herrmann, H., Hofmann, I. & Franke, W. W. (1992). Identification of a nonapeptide motif in the vimentin head domain involved in intermediate filament assembly. *J. Mol. Biol.* **223**, 637–650.

Herrmann, H., Häner, M., Brettel, M., Müller, S. A., Goldie, K. N., Fedtke, B., Lustig, A., Franke, W. W. & Aebi, U. (1996a). Structure and assembly properties of the intermediate filament protein vimentin: the role of its head, rod and tail domains. *J. Mol. Biol.* **264**, 933–953.

Herrmann, H., Münick, M. D., Brettel, M., Fouquet, B. & Markl, J. (1996b). Vimentin in a cold-water fish, the rainbow trout: highly conserved primary structure but unique assembly properties. *J. Cell Sci.* **109**, 569–578.

Herrmann, H., Häner, M., Brettel, M., Ku, N.-O. & Aebi, U. (1999). Characterization of distinct early assembly units of different intermediate filament proteins. *J. Mol. Biol.* **286**, 1403–1420.

Hess, J. F., Caselman, J. T., King, A. P. & Fitzgerald, P. G. (1998). Primary sequence, secondary structure, gene structure, and assembly properties suggests that the lens-specific cytoskeletal protein filensin represents a novel class of intermediate filament protein. *Exp. Eye Res.* **66**, 625–644.

Hirokawa, N., Glicksman, M. A. & Willard, M. B. (1984). Organization of mammalian neurofilament polypeptides within the neuronal cytoskeleton. *J. Cell Biol.* **98**, 1523–1536.

Hisanaga, S.-I. & Hirokawa, N. (1988). Structure of the peripheral domains of neurofilaments revealed by low angle rotary shadowing. *J. Mol. Biol.* **202**, 297–305.

Ho, C. L., Martys, J. L., Mikhailov, A., Gundersen, G. G. & Liem, R. K. H. (1998). Novel features of intermediate filament dynamics revealed by green fluorescent protein chimeras. *J. Cell Sci.* **111**, 1767–1778.

Inagaki, M., Nakamura, Y., Takeda, M., Nishimura, T. & Inagaki, N. (1994). Glial fibrillary acidic protein: dynamic property and regulation by phosphorylation. *Brain Path.* **4**, 239–243.

Inagaki, M., Matsuoka, Y., Tsujimura, K., Ando, S., Tokui, T., Takahashi, T. & Inagaki, N. (1996). Dynamic property of intermediate filaments: regulation by phosphorylation. *Bioessays* **18**, 481–487.

Ip, W., Hartzer, M. K., Pang, Y.-Y. S. & Robson, R. M. (1985). Assembly of vimentin *in vitro* and its implications concerning the structure of intermediate filaments. *J. Mol. Biol.* **183**, 365–375.

Irwin, A. D. & McLean. W. H. I. (1999). Human keratin diseases: the increasing spectrum of disease and sublety of the phenotype-genotype correlation. *Br. J. Dermatol.* **140**, 815–828.

James, V. J., Wilk, K. E., McConnell, J. F., Baranov, E. P. & Amemiya, Y. (1995). Intermediate

https://doi.org/10.1017/S0033583500003516 Published online by Cambridge University Press

filament structure of α-keratin in baboon hair. *Int. J. Biol. Macromol.* **17**, 99–104.

Jacobs, A. J., Kamholz, J. & Selzer, M. E. (1995). The single lamprey neurofilament subunit (NF-180) lacks multiphosphorylation repeats and is expressed selectively in projection neurons. *Mol. Brain Res.* **29**, 43–52.

Jones, L. N., Simon, M., Watts, N. R., Booy, F. P., Steven, A. C., & Parry, D. A. D. (1997). Intermediate filament structure: hard α-keratin. *Biophys. Chem.* **68**, 83–93.

Kammerer, R. A., Schulthess, T., Landwehr, R., Lustig, A., Engel, J., Aebi, U. & Steinmetz, M. O. (1998). An autonomous folding unit mediates the assembly of two-stranded coiled coils. *Proc. Natl. Acad. Sci. USA* **95**, 13419–13424.

Karabinos, A., Riemer, D., Erber, A. & Weber, K. (1998). Homologues of vertebrate Type I, II and III intermediate filament (IF) proteins in an invertebrate: the IF multigene family of the cephalochordate *Branchiostoma*. *FEBS Lett.* **437**, 15–18.

Kaufmann, E., Weber, K. & Geisler, N. (1985). Intermediate filament forming ability of desmin derivatives lacking either the amino-terminal 67 or the carboxy-terminal 27 residues. *J. Mol. Biol.* **185**, 733–742.

Kimonis, V., DiGiovanna, J. J., Yang, J.-M., Doyle, S. Z., Bale, S. J. & Compton, J. G. (1994). A mutation in the V1 end domain of keratin 1 in non-epidermolytic palmar-plantar keratoderma. *J. Invest. Dermatol.* **103**, 764–769.

Klymkowsky, M. W., Maynell, L. A. & Nislow, C. (1991). Cytokeratin phosphorylation, cytokeratin filament severing and the solubilization of the maternal mRNA *Vg*1. *J. Cell Biol.* **114**, 787–797.

Koch,, E. A., Spitzer, R. H., Pithawalla, R. B. & Parry, D. A. D. (1994). An unusual intermediate filament subunit from the cytoskeletal biopolymer released extracellularly into seawater by the primitive hagfish (*Eptatretus stouti*). *J. Cell Sci.* **107**, 3133–3144.

Koch, E. A., Spitzer, R. H., Pithawalla, R. B., Castillos, F. A. & Parry, D. A. D. (1995). Hagfish biopolymer: a Type I/Type II homologue of epidermal keratin intermediate filaments. *Int. J. Biol. Macromol.* **17**, 283–292.

Kooijman, M., Bloemendal, M., Traub, P., van Grondelle, R. & van Amerongen, H. (1995a). Hydrodynamic and electrical characterization of T-vimentin dimers and tetramers by transient electric birefringence measurements. *J. Biol. Chem.* **270**, 2931–2937.

Kooijman, M., van Amerongen, H., Traub, P., van Grondelle, R. & Bloemendal, M. (1995b). The assembly state of the intermediate filament proteins desmin and glial fibrillary acidic protein at low ionic strength. *FEBS Lett.* **358**, 185–188.

Kooijman, M., Bloemendal, M., Traub, P., van Grondelle, R. & van Amerongen, H. (1997). Transient electric birefringence study of intermediate filament formation from vimentin and glial fibrillary acidic protein. *J. Biol. Chem.* **272**, 22548–22555.

Korge, B. P. & Krieg, T. (1996). The molecular basis for inherited bullous diseases. *J. Molec. Med.* **74**, 59–70.

Korge, B. P., Gan, S-Q., McBride, O. W., Mischke, D. & Steinert, P. M. (1992a). Extensive size polymorphism of the human keratin 10 chain resides in the C-terminal V2 subdomain due to variable numbers and sizes of glycine loops. *Proc. Natl. Acad. Sci. USA* **89**, 910–914.

Korge, B. P., Compton, J. G., Steinert, P. M. & Mischke, D. (1992b). The two size alleles of human keratin 1 are due to a deletion in the glycine-rich carboxy-terminal V2-subdomain. *J. Invest. Dermatol.* **99**, 697–702.

Korge, B. P., Healy, E., Munro, C. S., Punter, C., Birch-Machin, M., Holmes, S. C., Darlington, S., Hamm, H., Messenger, A. G., Rees, J. L. & Traupe, H. (1998). A mutational hotspot in the 2B domain of human hair basic keratin 6 (hHb6) in monilethrix patients. *J. Invest. Dermatol.* **111**, 896–899.

Korge, B. P., Hamm, H., Jury, C. S., Traupe, H., Irvine, A. D., Healy, E., Birch-Machin, M., Rees, J. L., Messenger, A. G., Holmes, S. C., Parry, D. A. D. & Munro, C. S. (1999). Identification of novel mutations in basic hair keratins hHb1 and hHb6 in monilethrix: implications for protein structure and clinical phenotype. *J. Invest. Dermatol.* **113**, 607–612.

Kouklis, P. D., Traub, P. & Georgatos, S. D. (1993). Involvement of the consensus motif at coil 2B in the assembly and stability of vimentin intermediate filaments. *J. Cell Sci.* **102**, 31–41.

Kouklis, P. D., Hutton, E. & Fuchs, E. (1994). Making a connection: direct binding between keratin intermediate filaments and desmosomal proteins. *J. Cell Biol.* **127**, 1049–1060.

Kowalczyk, A. P., Stappenbeck, T. S., Parry, D. A. D., Palka, H. L., Virata, M. L. A., Bornslaeger, E. A., Nilles, L. A. & Green, K. J. (1994). Structure and function of desmosomal transmembrane core and plaque molecules. *Biophys. Chem.* **50**, 97–112.

Krishnan, N., Kaiserman-Abramof, I. R. & Lasek, R. J. (1979). Helical substructure of neurofilaments isolated from *Myxicola* and squid giant axons. *J. Cell Biol.* **82**, 323–335.

Ku, N.-O. & Omary, M. B. (1994). Identification of the major physiologic phosphorylation site of human keratin 18: potential kinases and a role in filament organization. *J. Cell Biol.* **127**, 161–171.

Ku, N.-O., Liao, J. & Omary, M. B. (1997). Apoptosis generates stable fragments of human type I keratins. *J. Biol. Chem.* **272,** 33197–33203.

Langbein, L., Rogers, M. A., Winter, H., Praetzel,

https://doi.org/10.1017/S0033583500003516 Published online by Cambridge University Press

S., Beckhaus, U., Rackwitz, H. R. & Schweizer, J. (1999). The catalog of human hair keratins. I. Expression of the nine Type I members in the hair follicle. *J. Biol. Chem*. **274**, 19874–19884.

Leapman, R. D., Gallant, P. E., Reese, T. S. & Andrews, S. B. (1997). Phosphorylation and subunit organization of axonal neurofilaments determined by scanning transmission electron microscopy. *Proc. Natl. Acad. Sci. USA* **94**, 7820–7824.

Lee, M.-K., Xu, Z., Wong, P.-C. & Cleveland, D. W. (1993). Neurofilaments are obligate heteropolymers *in vivo*. *J. Cell Biol*. **122**, 1337–1350.

Lendahl, U., Zimmerman, L. B. & McKay R. D. G. (1990). CNS stem cells express a new class of intermediate filament protein. *Cell* **60**, 585–595.

Letai, A. & Fuchs, E. (1995). The importance of intramolecular ion pairing in intermediate filaments. *Proc. Natl. Acad. Sci. USA* **92**, 92–96.

Leterrier, J. F., Kas, J., Hartwig, J., Vegners, R. & Janmey, P. A. (1996). Mechanical effects of neurofilament cross-bridges: Modulation by phosphorylation, lipids, and interactions with F-actin. *J. Biol. Chem*. **271**, 15687–15694.

Leung, C. L. & Liem, R. K. H. (1996). Characterization of interactions between the neurofilament triplet proteins by the yeast two-hybrid system. *J. Biol. Chem*. **271**, 14041–14044.

Liao, J., Lowthert, L. A., Ku, N. O. & Fernandez, R. (1995). Dynamics of human keratin 18 phosphorylation: polarized distribution of phosphorylated keratins in simple epithelial tissues. *J. Cell Biol*. **131**, 1291–1301.

Liu, C.-G., Märcker, C., Castanon, M. J., Hauptmann, R. & Wiche, G. (1996). Human plectin: organization of the gene, sequence analysis, and chromosome localization (8q24). *Proc. Natl. Acad. Sci. USA* **93**, 4278–4283.

Mack, J. W., Torchia, D. A. & Steinert, P. M. (1988). Solid-state NMR studies of the dynamics and structure of mouse keratin intermediate filaments. *Biochemistry* **27**, 5418–5426.

Magin, T. M., Hatzfeld, M. & Franke, W. W. (1987). Analysis of cytokeratin domains by cloning and expression of intact and deleted polypeptides in *Escherichia coli*. *EMBO J*. **6**, 2607–2615.

Marqusee, S. & Baldwin, R. L. (1987). Helix stabilization by $glu^{-}...lys^{+}$ salt bridges in short peptides of *de novo* design. *Proc. Natl. Acad. Sci. USA* **84**, 8898–8902.

Marvin, M. J., Dahlstrand, J., Lendahl, U. & McKay, R. D. G. (1998). A rod end deletion in the intermediate filament protein nestin alters its subcellular localization in neuroepithelial cells of transgenic mice. *J. Cell Sci*. **111**, 1951–1961.

McKeon, F. D., Kirschner M. W. & Caput D. (1986). Homologies in both primary and secondary structure between nuclear envelope and intermediate filament proteins. *Nature* (*London*) **319**, 463–468.

McLachlan, A. D. & Stewart, M. (1975). Tropomyosin coiled-coil interactions: evidence for an unstaggered structure. *J. Mol. Biol*. **98**, 293–304.

McLachlan, A. D. & Stewart, M. (1982). Periodic charge distribution in the intermediate filament proteins desmin and vimentin. *J. Mol. Biol*. **162**, 693–698.

McLean, W. H. I. & Lane, E. B. (1995). Intermediate filaments in disease. *Curr. Opinion Cell Biol*.**7**, 118–125.

Meng, J. J, Khan, S. & Ip, W. (1994). Charge interactions in the rod domain drive formation of tetramers during intermediate filament formation. *J. Biol. Chem*. **269**, 18679–18685.

Meng, J. J, Khan, S. & Ip, W. (1996). Intermediate filament protein domain interactions as revealed by two-hybrid screens. *J. Biol. Chem*. **271**, 1599–1604.

Meng, J. J., Bornsläger, E. A., Green, K. J., Steinert, P. M. & Ip, W. (1997). Two-hybrid analysis reveals fundamental differences in direct interactions between desmoplakin and cell-specific intermediate filaments. *J. Biol. Chem*. **272**, 21495–21503.

Merdes, A., Gounari, F. & Georgatos, S. D. (1993). The 47-kD lens-specific protein phakinin is a tailless intermediate filament protein and an assembly partner of filensin. *J. Cell Biol*. **123**, 1507–1516.

Miller, R. K., Khuon, S. & Goldman, R. D. (1993). Dynamics of keratin assembly: exogenous type I keratin rapidly associates with type II keratin *in vivo*. *J. Cell Biol*. **122**, 123–135.

Milner, D. J., Weitzer, G., Tran, D., Bradley, A. & Capetanaki, Y. (1996). Disruption of muscle architecture and myocardial degeneration in mice lacking Desmin. *J. Cell Biol*. **134**, 1255–1270.

Mischke, D. (1993). Frequencies of human keratin 10 alleles. *Hum. Mol. Genet*. **2**, 618.

Miyasaka, H., Okabe, S., Ishiguro, K., Uchida, T. & Hirokawa, N. (1993). Interaction of the tail domain of high molecular weight subunits of neurofilaments with the COOH-terminal region of tubulin and its regulation by tau protein kinase II. *J. Biol. Chem*. **268**, 22695–22702.

Munoz Marmol, A. M., Strasser, G., Isamat, M., Coulombe, P. A., Yang, Y., Roca, X., Vela, E., Mate, J. L., Coll, J., Fernandez Figuera, M. T., Navas Palacios, J. J., Ariza, A. & Fuchs E. (1998). A dysfunctional desmin mutant in a patient with severe generalized myopathy. *Proc. Natl. Acad. Sci. USA* **95**, 11312–11317.

Nakagawa, T., Chen, J., Zhang, Z., Kanai, Y. & Hirokawa, N. (1995). Two distinct functions of the carboxy-terminal tail domain of NF-M upon neurofilament assembly: cross-bridge formation and longitudinal elongation of filaments. *J. Cell Biol*. **129**, 411–429.

https://doi.org/10.1017/S0033583500003516 Published online by Cambridge University Press

Nakamura, Y., Takeda, M. & Nishimura, T. (1996). Dynamics of bovine glial fibrillary acidic protein phosphorylation. *Neurosci. Lett.* **205**, 91–94.

Nigg, E. A. (1992). Assembly-disassembly of the nuclear lamina. *Curr. Opin. Cell Biol.* **4**, 105–109.

Norrander, J. M., Amos, L. A. & Linck, R. W. (1992). Primary structure of tektin A1: comparison with intermediate-filament proteins and a model for its association with tubulin. *Proc. Natl. Acad. Sci. USA* **89**, 8567–8571.

North, A. C. T., Steinert, P. M. & Parry, D. A. D (1994). Coiled-coil stutter and link segments in keratin and other intermediate filaments: a computer modeling study. *Proteins: Struct. Funct. Gen.* **20**, 174–184.

O'Donnell, I. J. (1969) Studies on reduced wool. IX. The N-terminal sequence of a fragment produced by cleavage of component 8 with cyanogen bromide *Aust. J. Biol. Sci.* **22**, 471–488.

O'Shea, E. K., Klemm, J. D., Kim, P. S. & Alber, T. (1991). X-ray structure of the GCN4 leucine zipper, a two-stranded, parallel coiled coil. *Science* **254**, 539–544.

Pant, H. C. & Veeranna, X. (1995). Neurofilament phosphorylation. *Biochem. Cell Biol.* **73**, 575–592.

Parry, D. A. D. (1975). Analysis of the primary structure of α-tropomyosin. *J. Mol. Biol.* **98**, 519–535.

Parry, D. A. D. (1982). Coiled-coils in α-helix-containing proteins: analysis of the residue types within the heptad repeat and the use of these data in the prediction of coiled-coils in other proteins. *Biosci. Rep.* **2**, 1017–1024.

Parry, D. A. D. (1990). Primary and secondary structure of IF protein chains modes of aggregation. In: *Cellular and Molecular Biology of Intermediate Filaments* (eds. R. D. Goldman & P. M. Steinert), Ch. 7, pp. 175–204. New York: Plenum Press.

Parry, D. A. D. (1995). Hard α-keratin IF: a structural model lacking a head-to-tail molecular overlap but having hybrid features characteristic of both epidermal keratin and vimentin IF. *Proteins: Struct., Funct. Gen.*, **22**, 267–272.

Parry, D. A. D. (1996). Hard α-keratin intermediate filaments: an alternative interpretation of the low-angle equatorial X-ray diffraction pattern, and the axial disposition of putative disulphide bonds in the intra- and inter-protofilamentous networks. *Int. J. Biol. Macromol* **19**, 45–50.

Parry, D. A. D. (1997). Protein chains in hair and epidermal keratin IF: structural features and spatial arrangements. In: *Formation and Structure of Human Hair* (eds. P. Jollés, H. Zahn & H. Höcker) pp. 177–207. Birkhäuser: Basel.

Parry, D. A. D. & Fraser, R. D. B. (1985). Intermediate filament structure. 1. Analysis of IF protein sequence data. *Int. J. Biol. Macromol.* **7**, 203–213.

Parry, D. A. D. & North, A. C. T. (1998). Hard α-keratin intermediate filament chains: substructure of the N- and C-terminal domains and the predicted structure and function of the C-terminal domains of Type I and Type II chains. *J. Struct. Biol.* **122**, 67–75.

Parry, D. A. D. & Steinert, P. M. (1995). *Intermediate Filament Structure*, pp. 1–183. Springer-Verlag: Heidelberg.

Parry, D. A. D. & Suzuki, E. (1969a). Intrachain potential energy of the α-helix and of a coiled-coil strand. *Biopolymers* **7**, 189–197.

Parry, D. A. D. & Suzuki, E. (1969b). Interchain packing energies of multistranded α-helical ropes. *Biopolymers* **7**, 199–206.

Parry, D. A. D., Crewther, W. G., Fraser, R. D. B. & MacRae, T. P. (1977). Structure of α-keratin: structural implication of the amino acid sequences of the Type I and Type II chain segments. *J. Mol. Biol.* **113**, 449–454.

Parry, D. A. D., Fraser, R. D. B. & MacRae, T. P. (1979). Repeating patterns of amino acid residues in the sequences of some high-sulphur proteins from α-keratin. *Int. J. Biol. Macromol.* **1**, 17–22.

Parry, D. A. D., Steven, A. C. & Steinert, P. M. (1985). The coiled-coil molecules of intermediate filaments consist of two parallel chains in exact axial register. *Biochem. Biophys. Res. Commun.* **127**, 1012–1018.

Parry, D. A. D., Conway, J. F. & Steinert, P. M. (1986). Structural studies on lamin: similarities and differences between lamin and intermediate filament proteins. *Biochem. J.* **238**, 305–308.

Parry, D. A. D., Conway, J. F., Goldman, A. E., Goldman, R. D. & Steinert, P. M. (1987). Nuclear lamin proteins: common structures for paracrystalline, filamentous and lattice forms. *Int. J. Biol. Macromol.* **9**, 137–145.

Pekny, M., Leveen, P., Pekna, M., Eliasson, C., Berthold, C. H., Westermark, B. & Betsholtz, C. (1995). Mice lacking glial fibrillary acidic protein display astrocytes devoid of intermediate filaments but develop and reproduce normally. *EMBO J.* **14**, 1590–1598.

Perutz, M. F. & Fermi, G. (1988). Stereochemistry of salt-bridge formation in α-helices and β-strands. *Proteins: Struct. Funct. Genet.* **4**, 294–295.

Prahlad, V., Yoon, M., Moir, R. D., Vale, R. D. & Goldman, R. D. (1998). Rapid movements of vimentin on microtubule tracks: kinesin-dependent assembly of intermediate filament networks. *J. Cell Biol.* **143**, 159–170.

Price, M. G., Gomer, R. H. & Skelemin, A. (1993). Cytoskeletal M-disc periphery protein, contains motifs of adhesion/recognition and intermediate filament proteins. *J. Biol. Chem.* **268**, 21800–21810.

Quax-Jeuken, Y. E. F. M., Quax, W. J. & Bloemendal, H. (1983). Primary and secondary structure

https://doi.org/10.1017/S0033583500003516 Published online by Cambridge University Press

of hamster vimentin predicted from the nucleotide sequence. *Proc. Natl. Acad. Sci. USA* **80**, 3548–3552.

Quinlan, R. A. & Stewart, M. (1987). Crystalline tubes of myosin long subfragment-2 that show the coiled-coil pitch and molecular interaction geometry. *J. Cell Biol.* **105**, 403–415.

Quinlan, R. A., Cohlberg, J. A., Schiller, D. L., Hatzfeld, M. & Franke, W. W. (1984). Heterotypic tetramer ($A_2D_2$) complexes of non-epidermal keratins isolated from cytoskeletons of rat hepatocytes and hepatoma cells. *J. Mol. Biol.* **178**, 365–388.

Quinlan, R. A., Hatzfeld, M., Franke, W. W., Lustig, A., Schulthess, T. & Engel, J. (1986). Characterization of dimer subunits of intermediate filament proteins. *J. Mol. Biol.* **192**, 337–349.

Quinlan, R. A., Moir, R. D. & Stewart, M. (1989). Expression in *Escherichia coli* of fragments of glial fibrillary acidic protein: characterization, assembly properties and paracrystal formation. *J. Cell Sci.* **93**, 71–83.

Quinlan, R., Hutchison, C. & Lane, B. (1994). Intermediate filament proteins. *Protein Profile* **1** (8), 779–911.

Ralton, J. E., Lu, X., Hutcheson, A. M. & Quinlan, R. A. (1994). Identification of two N-terminal non-$\alpha$-helical domain motifs important in the assembly of glial fibrillary acidic protein. *J. Cell Sci.* **107**, 1935-1948.

Remington, S. G. (1993). Chicken filensin: a lens fiber cell protein that exhibits its sequence similarity to intermediate filament proteins. *J. Cell Sci.* **105**, 1057–1068.

Riemer, D. & Weber, K. (1998). Common and variant properties of intermediate filament proteins from lower chordates and vertebrates; two proteins from the tunicate *Styela* and the identification of a Type III homologue. *J. Cell Sci.* **111**, 2967–2975.

Riemer, D., Karabinos, A. & Weber, K. (1998). Analysis of eight cDNAs and six genes for intermediate filament (IF) proteins in the cephalochordate *Branchiostoma* reveals differences in the IF multigene families of lower chordates and the vertebrates. *Gene* **211**, 361–373.

Rogers, G. E. (1964). Structural and biochemical features of the hair follicle. In: *The Epidermis* (eds. W. Montagna & W. C. Lobitz), pp. 179–236. New York: Academic Press.

Rogers, K. R., Eckelt, A., Nimmrich, V., Jannsen, K. P., Schliwa, M., Herrmann, H. & Franke, W. W. (1995). Truncation mutagenesis of the non-helical carboxyterminal tail domain of vimentin reveals contributions to cellular localization but not to filament assembly. *Eur. J. Cell Biol.* **66**, 136–150.

Rogers, K. R., Herrmann, H. & Franke, W. W. (1996). Characterization of disulfide crosslink formation of human vimentin at the dimer, tetramer, and intermediate filament levels. *J. Struct. Biol.* **117**, 55–69.

Ruhrberg, C., Hajibagheri, M. A. N., Parry, D. A. D. & Watt, F. M. (1997). Periplakin, a novel component of cornified envelopes and desmosomes that belongs to the plakin family and forms complexes with envoplakin. *J. Cell Biol.* **139**, 1835–1849.

Rugg, E. L., McLean, W. H. I., Lane, E. B., Pitera, R., McMillan, J. R., Dopping Hepenstal, P. J. C., Navsaria, H. A., Leigh, I. M. & Eady, R. A. J. (1994). Functional 'knockout' of human keratin 14. *Genes Develop.* **8**, 2563–2573.

Sasse, B., Lustig, A., Aebi, U. & Stuurman, N. (1997). *In vitro* assembly of Drosophila lamin $Dm_0$: lamin polymerization properties. *Eur. J. Biochem.* **250**, 30–38.

Sasse, B., Aebi, U. & Stuurman, N. (1998). A tailless Drosophila $Dm_0$ fragment reveals lateral associations of dimers. *J. Struct. Biol.* **123**, 56–66.

Sauk, J. J., Krumweide, M., Cocking-Johnson, D. & White, J. G. (1984). Reconstitution of cytokeratin filaments *in vitro*: further evidence for the role of non-helical peptides in filament assembly. *J. Cell Biol.* **99**, 1590–1597.

Seo, J. & Cohen, C. (1993). Pitch diversity in $\alpha$-helical coiled-coils. *Proteins: Struct. Funct. Genet.* **15**, 223–234.

Skalli, O., Chou, Y.-H. & Goldman, R. D. (1992). Cell cycle-dependent changes in the organization of an intermediate filament-associated protein: correlation with phosphorylation by p34cdc2. *Proc. Natl. Acad. Sci. USA.* **89**, 11959–11963.

Smith, F. J., Eady, R. A. Leigh, I. M., McMillan, J. R., Rugg, E. L., Kelsell, D. P., Bryant, S. P., Spurr N. K., Geddes, J. F., Kirtschig, G., Milana, G., de Bono, A. G., Owaribe, K., Wiche, G., Pulkkinen, L., Uitto, J., McLean, W. H. & Lane, E. B. (1996). Plectin deficiency results in muscular dystrophy with epidermolysis bullosa. *Nat. Genet.* **13**, 450–457.

Soellner, P., Quinlan, R. A. & Franke, W. W. (1985). Identification of a distinct soluble subunit of an intermediate filament protein: tetrameric vimentin from living cells. *Proc. Natl. Acad. Sci. USA.* **82**, 7929–7933.

Stappenbeck, T. S., Lamb, J. A., Corcoran, C. M. & Green, K. J. (1994). Phosphorylation of the desmoplakin COOH terminus negatively regulates its interaction with keratin intermediate filament networks. *J. Biol. Chem.* **269**, 29351–29354.

Steinert, P. M. (1977). The mechanism of assembly of bovine epidermal keratin filaments *in vitro*. In: *Biochemistry of Cutaneous Epidermal Differentiation* (eds. M. Seiji and I. A. Bernstein), pp. 444–464. Tokyo: University of Tokyo Press.

Steinert, P. M. (1990). The two-chain coiled-coil molecule of native epidermal keratin intermediate

https://doi.org/10.1017/S0033583500003516 Published online by Cambridge University Press

filaments is a type I–type II heterodimer. *J. Biol. Chem.* **265**, 8766–8774.

Steinert, P. M. (1991a). Organisation of coiled-coil molecules in native keratin 1/keratin 10 intermediate filaments: evidence for alternating rows of antiparallel in-register and antiparallel staggered molecules. *J. Struct. Biol.* **107**, 157–174.

Steinert, P. M. (1991b). Analysis of the mechanism of assembly of mouse keratin 1/keratin 10 intermediate filaments *in vitro* suggests that intermediate filaments are built from multiple oligomeric units rather than a unique tetrameric building block. *J. Struct. Biol.* **107**, 175–188.

Steinert, P. M. & Marekov, L. N. (1997). Direct evidence that involucrin is a major early isopeptide cross-linked component of the keratinocyte cornified cell envelope. *J. Biol. Chem.* **272**, 2021–2030.

Steinert, P. M. & Parry, D. A. D. (1985). Intermediate filaments: conformity and diversity of expression and structure. *Ann. Rev. Cell Biol.* **1**, 41–65.

Steinert, P. M. & Parry, D. A. D. (1993). The conserved H1 domain of the Type II keratin 1 chain plays an essential role in the alignment of nearest-neighbor molecules in mouse and human keratin 1/keratin 10 intermediate filaments at the two- to four-molecule level of structure. *J. Biol. Chem.* **268**, 2878–2887.

Steinert, P. M. & Roop, D. R. (1988). Molecular and cellular biology of intermediate filaments. *Annu. Rev. Biochem.* **57**, 593–625.

Steinert, P. M., Idler, W. W. & Zimmerman, S. B. (1976). Self-assembly of bovine epidermal keratin filaments *in vitro*. *J. Mol. Biol.* **108**, 547–567.

Steinert, P. M., Idler, W. W., Cabral, F., Gottesman, M. M. & Goldman, R. D. (1981). *In vitro* assembly of homopolymer and copolymer filaments from intermediate filament subunits of muscle and fibroblastic cells. *Proc. Natl. Acad. Sci. U. S. A.* **78**, 3692–3696.

Steinert, P. M., Rice, R. H., Roop, D. R., Trus, B. L. & Steven, A. C. (1983). Complete amino acid sequence of a mouse epidermal keratin subunit and implications for the structure of intermediate filaments. *Nature* (*London*) **302**, 794–800.

Steinert, P. M., Parry, D. A. D., Racoosin, E. L., Idler, W. W., Steven, A. C., Trus, B. L. & Roop, D. R. (1984). The complete cDNA and deduced amino acid sequence of a Type II mouse epidermal keratin of 60000 Da: analysis of sequence differences between Type I and Type II keratins. *Proc. Natl. Acad. Sci. USA* **81**, 5709–5713.

Steinert, P. M., Idler, W. W., Zhou, X. M., Johnson, L. D., Parry, D. A. D., Steven, A. C. & Roop, D. R. (1985a). Structural and functional implications of amino acid sequences of keratin intermediate filament subunits. *Proc. N. Y. Acad. Sci.* **455**, 451–461.

Steinert, P. M., Parry, D. A. D., Idler, W. W., Johnson, L. D., Steven, A. C. & Roop, D. R. (1985b). Amino acid sequences of mouse and human epidermal Type II keratins of $M_r$ 67000 provide a systematic basis for the structural and functional diversity of the end domains of keratin intermediate filament subunits. *J. Biol. Chem.* **260**, 7142–7149.

Steinert, P. M., Mack, J. W., Korge, B. P., Gan, S.-Q., Haynes, S. R. & Steven, A. C. (1991). Glycine loops in proteins: their occurrence in certain intermediate filament chains, loricrins and single-stranded RNA binding proteins. *Int. J. Biol. Macromol.* **13**, 130–139.

Steinert, P. M., Marekov, L. N., Fraser, R. D. B. & Parry, D. A. D. (1993a). Keratin intermediate filament structure: crosslinking studies yield quantitative information on molecular dimensions and mechanism of assembly. *J. Mol. Biol.* **230**, 436–452.

Steinert, P. M., Marekov, L. N. & Parry, D. A. D. (1993b). Conservation of the structure of keratin intermediate filaments: molecular mechanism by which different keratin molecules integrate into pre-existing keratin intermediate filaments during differentiation. *Biochemistry* **32**, 10046–10056.

Steinert, P. M., Marekov, L. N. & Parry, D. A. D. (1993*c*). Diversity of intermediate filament structure: evidence that the alignment of coiled-coil molecules in vimentin is different from that in keratin intermediate filaments. *J. Biol. Chem.* **268**, 24916–24925.

Steinert, P. M., Khuon, S. & Goldman, R. D. (1993d). Use of synthetic peptides to probe the highly dynamic structures of keratin intermediate filaments. *J. Invest. Dermatol.* **100**, 500.

Steinert, P. M., North, A. C. T. & Parry, D. A. D. (1994). Structural features of keratin intermediate filaments. *J. Invest. Dermatol.* **103**, 19S-24S.

Steinert, P. M., Candi, E., Kartasova, T. & Marekov, L. (1998a). Small proline-rich proteins are cross-bridging proteins in the cornified cell envelopes of stratified squamous epithelia. *J. Struct. Biol.* **122**, 76–85.

Steinert, P. M., Kartasova, T. & Marekov, L. N. (1998b). Biochemical evidence that small proline-rich proteins and trichohyalin function in epithelia by modulation of the biomechanical properties of their cornified cell envelopes. *J. Biol. Chem.* **273**, 11758–11769.

Steinert, P. M., Chou, P-Y., Prahlad, V., Parry, D. A. D., Marekov, L. N., Wu, K. C., Jang, S-I. & Goldman, R. D. (1999a). A high molecular weight intermediate filament-associated protein in BHK-21 cells is nestin, a Type VI intermediate filament protein: limited co-assembly *in vitro* to form heteropolymers with Type III vimentin and Type IV a-internexin. *J. Biol. Chem.* **274**, 9881–9890.

Steinert, P. M., Marekov, L. N. & Parry, D. A. D.

https://doi.org/10.1017/S0033583500003516 Published online by Cambridge University Press

(1999b). Molecular parameters of Type IV a-internexin and Type IV–Type III a-internexin-vimentin copolymer intermediate filaments. *J. Biol. Chem.* **274**, 1657–1666.

Steinmetz, M. O., Stock, A., Schulthess, T., Landwehr, R., Lustig, A., Faix, J., Gerisch, G., Aebi, U. & Kammerer, R. A. (1998). A distinct 14 residue site triggers coiled-coil formation in cortexillin I. *EMBO J.* **17**, 1883–1891.

Steven, A. C., Wall, J. S., Hainfeld, J. F. & Steinert, P. M. (1982). Structure of fibroblastic intermediate filaments: analysis by scanning transmission electron microscopy. *Proc. Natl. Acad. Sci. USA* **79**, 3101-3105.

Steven, A. C., Hainfeld, J. F., Trus, B. L., Wall, J. S. & Steinert, P. M. (1983a). The distribution of mass in heteropolymer intermediate filaments assembled *in vitro*: STEM analysis of vimentin/desmin and bovine epidermal keratin *J. Biol. Chem.* **258**, 8323–8329.

Steven, A. C., Hainfeld, J. F., Trus, B. L., Wall, J. S. & Steinert, P. M. (1983b). Epidermal keratin filaments assembled *in vitro* have masses-per-unit-length that scale according to average subunit mass: structural basis for homologous packing of subunits in intermediate filaments. *J. Cell Biol.* **97**, 1939–1944.

Steven, A. C., Hainfeld, J. F., Trus, B. L., Wall, J. S. & Steinert, P. M. (1984). Radial distributions of density within macromolecular complexes determined from dark-field electron micrographs. *Proc. Natl. Acad. Sci. USA* **81**, 6363–6367.

Steven, A. C., Stall, R., Steinert, P. M. & Trus, B. L. (1985). Computational straightening of images of curved macromolecular helices by cubic spline interpolation facilitates structural analysis by Fourier methods. *Proc. 43rd Annu. Meet. EMSA* **7**, 738–739.

Stuurman, N., Sasse, B. & Fisher, P. A. (1996). Intermediate filament protein polymerization: molecular analysis of *Drosophila* nuclear lamin head-to-tail binding. *J. Struct. Biol.* **117**, 1–15.

Stuurman, N., Heins, S. & Aebi, U. (1998). Nuclear lamins: their structure, assembly, and interactions. *J. Struct. Biol.* **122**, 42–66.

Sun, D., Macioce, P., Chin, S. S. M. & Liem, R. K. H. (1997). Assembly properties of amino- and carboxy-terminally truncated neurofilament NF-H proteins with NF-L and NF-M in the presence and absence of vimentin. *J. Neurochem.* **68**, 917–926.

Svitkina, T. M., Verkhovsky, A. B. & Borisy, G. G. (1996). Plectin sidearms mediate interaction of intermediate filaments with microtubules and other components of the cytoskeleton. *J. Cell Biol.* **135**, 991–1007.

Szaro, B. G., Pant, H. C., Way, J. & Battey, J. (1991). Squid low molecular weight neurofilament proteins are a novel class of neurofilament protein. A nuclear lamin-like core and multiple distinct proteins formed by alternative RNA processing. *J. Biol. Chem.* **266**, 15035–15041.

Tanaka, T., Parry, D. A. D., Klaus-Kovtun, V., Steinert, P. M. & Stanley, J. R. (1991). Comparison of molecularly cloned bullous pemphigoid antigen to desmoplakin I confirms that they define a new family of cell adhesion junction plaque proteins. *J. Biol. Chem.* **266**, 12555–12559.

Tomarev, S. I., Zinovieva, R. D. & Piatigorsky, J. (1993). Primary structure and lens-specific expression of genes for an intermediate filament protein and a $\beta$-tubulin in cephalopods. *Biochim. Biophys. Acta* **1216**, 245–254.

Wais-Steider, C., Eagles, P. A. M., Gilbert, D. S. & Hopkins, J. M. (1983). Structural similarities and differences amongst neurofilaments. *J. Mol. Biol.* **165**, 393–400.

Waseem, A., White, K. & Waseem, N. H. (1997). Identification of a novel keratin epitope: evidence for associations between non-helical sub-domains L12 during filament assembly. *Int. J. Biochem. Cell Biol.* **29**, 971–983.

Watts, N. R., Jones, L. N., Cheng, N., Parry, D. A. D. & Steven, A. C. Intermediate filaments of hard alph$\alpha$-keratin are hollow (unpublished).

Way, J., Hellmich, M. R., Jaffe, H., Szaro, B., Pant, H. C., Gainer, H. & Battey, J. (1992). A high molecular weight squid neurofilament protein contains a lamin-like rod domain and a tail domain with Lys-Ser-Pro repeats. *Proc. Natl. Acad. Sci. USA* **89**, 6963–6967.

Weber, K., Plessmann, U., Dodemont, H. & Kossmagk-Stephan, K. (1988). Amino acid sequences and homopolymer-forming ability of the intermediate filament proteins from an invertebrate epithelium. *EMBO J.* **7**, 2995–3001.

Weber, K., Plessmann, U. & Ulrich, W. (1989). Cytoplasmic intermediate filament proteins of invertebrates are closer to nuclear lamins than are vertebrate intermediate filament proteins: sequence characterization of two muscle proteins of a nematode. *EMBO J.* **8**, 3221–3227.

Whitby, F. G., Kent, H., Stewart, F., Stewart, M., Xie, X., Hatch, V., Cohen, C. & Phillips, G. N. Jr (1992). Structure of tropomyosin at 9 angstroms resolution. *J. Mol. Biol.* **227**, 441–452.

Wiche, G., Becker, B., Luber, K., Weitzer, G, Castanon, M. J., Hauptaman, R., Stratowa, C., & Stewart, M. (1991). Cloning and sequencing of rat plectin indicates a 466kDa polypeptide chain with a three-domain structure based on a central alpha helical coiled coil. *J. Cell Biol.* **114**, 83–99.

Wilson, E. (1928) *The Cell in Development and Heredity*. Macmillan Co., New York.

Winter, H., Rogers, M. A., Langbein, L., Stevens,

https://doi.org/10.1017/S0033583500003516 Published online by Cambridge University Press

H. P., Leigh, I. M., Labrèze, C., Roul, S., Taieb, A., Krieg, T. & Schweizer, J. (1997a). Mutations in the hair cortex keratin hHb6 cause the inherited hair disease monilethrix. *Nature Genet*. **16**, 372–374.

Winter, H., Rogers, M. A., Gebhardt, M., Wollina, U., Boxall, L., Chitayat, D., Babul-Hirji, R., Stevens, H. P., Zlotogorski, A. & Schweizer, J. (1997*b*). A new mutation in the Type II hair cortex keratin hHb1 involved in the inherited hair disorder monilethrix. *Hum. Genet*. **101**, 165–169.

Winter, H., Hofmann, I., Langbein, L., Rogers, M. A. & Schweizer, J. (1997c). A splice site mutation in the gene of the human Type I hair keratin hHa1 results in the expression of a tailless keratin isoform. *J. Biol. Chem*. **272**, 32345–32352.

Winter, H., Labrèze, C., Chapalain, V., Surlève-Bazeille, J. E., Mercier, M., Rogers, M. A. Taieb, A. & Schweizer, J. (1998). A variable monilethrix phenotype associated with a novel mutation, Glu-402Lys in the helix termination motif of the Type II hair keratin hHb1. *J. Invest. Dermatol*. **111**, 169–172.

Wolf, E., Kim, P. S. & Berger, B. (1997). MultiCoil: a program for predicting two- and three-stranded coiled coils. *Protein Science* **6**, 1179–1189.

Woods, E. F. & Inglis, A.S (1984). Organization of the coiled-coils in the wool microfibril. *Int. J. Biol. Macromol*. **6**, 277–283.

Woolfson, D. N. & Alber, T. (1995). Predicting oligomerization states of coiled coils. *Protein Science* **4**, 1596–1607.

Yang, Y., Dowling, J., Yu, Q. C., Kouklis, P., Cleveland, D. W. & Fuchs, E. (1996). An essential cytoskeletal linker protein connecting actin microfilaments to intermediate filaments. Cell **86**, 655–665.

Yang, J.-M., Nam, K., Kim, H.-C., Lee, J.-H., Park, J.-K., Wu, K., Lee, E.-S. & Steinert, P. M. (1999). A novel glutamic acid to aspartic acid mutation near the end of the 2B rod domain in the keratin 1 chain in epidermolytic hyperkeratosis. *J. Invest. Dermatol*. **112**, 376–379.

Yatsunami, J., Fujiki, H., Suganuma, M., Yoshizawa, S., Eriksson, J. E., Olson, M. O. & Goldman, R. D. (1991). Vimentin is hyperphosphorylated in primary human fibroblasts treated with okadaic acid. *Biochem. Biophys. Res. Commun*. **177**, 1165–1170.

Yoon, M., Moir, R. D., Prahlad, V. & Goldman, R. D. (1998). Motile properties of vimentin intermediate filament networks in living cells. *J. Cell Biol*. **143**, 147–157.

Yuan, Y., Lee, J. A., Napier, A. & Cole, G. J. (1997). Molecular cloning of a new intermediate filament protein expressed by radial glia and demonstration of alternative splicing in a novel heptad repeat region located in the carboxy-terminal tail domain. *Mol. Cell. Neurosci*. **10**, 71–86.

Zackroff, R. V., Goldman, A. E., Jones, J. C. R., Steinert, P. M. & Goldman, R. D. (1984). Isolation and characterization of keratin-like proteins from cultured cells with fibroblastic morphology. *J. Cell Biol*. **98**, 1231–1237.

Zlotogorski, A., Horev, L. & Glaser, B. (1998). Monilethrix: a keratin hHb6 mutation is co-dominant with variable expression. *Exp. Dermatol*. **7**, 268–272.

https://doi.org/10.1017/S0033583500003516 Published online by Cambridge University Press