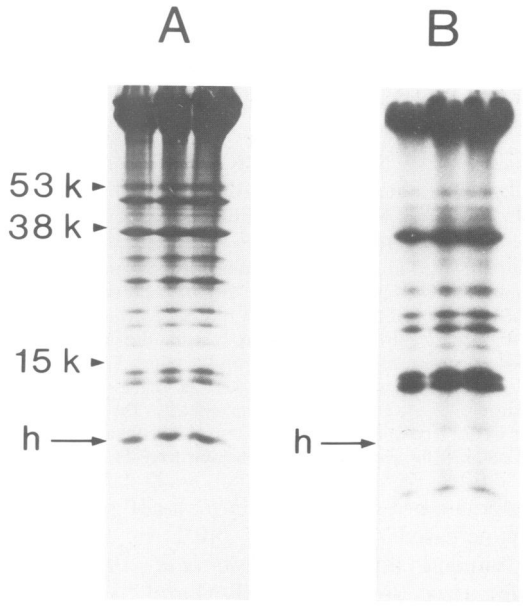
A
B
53 k
38 k
15 k
h
h

Table I. Amino acid composition for the headpiece (h) and the large fragments from the carboxy-terminal tailpiece extension

	Head-piece (h)	Tail fragments			
		CNBr I*	CNBr I	CNBr II	Ch-tail
Asx	5.3	0.8	0.4	5.1	1.8
Thr	3.7	1.2	0.8	4.9	2.2
Ser	19.0	10.7[a]	9.4	4.7[b]	8.5
Glx	4.4	23.8[c]	25.8	25.4[d]	27.3
Pro	3.3	13.6	14.3	13.7	12.6
Gly	18.4	1.3	0.6	3.5	0.9
Ala	16.8	15.1	14.4	12.8	15.2
Val	3.8	6.8	8.5	2.7	4.7
Met	n.d.	–	–	–	0.2
Ile	–	1.2	0.5	–	0.3
Leu	8.5	1.3	0.8	–	0.8
Tyr	1.4	0.1	0.1	–	–
Phe	4.5	0.4	0.3	–	–
Lys	1.5	22.3	24.5	26.3	24.9
His	2.8	–	–	–	–
Arg	6.5	0.7	0.4	–	0.3
Trp	~1	n.d.	n.d.	n.d.	n.d.

Numbers are given in mol%. For nomenclature of the tail fragments see Figure 3.
[a]About 50% of the serine residues are phosphorylated.
[b]No phosphoserine detected.
Enzymatic hydrolysis shows that >95%[c] and >85%[d] of the residues are glutamic acid rather than glutamine. Glutamic acid and phosphoserine data obtained on CNBr I* are probably very similar for CNBr I since CNBr I* exceeds CNBr I by only a few residues at the amino-terminal side (Figure 3; see text). Minor differences in composition of the Ch-tail in comparison with earlier data (Geisler *et al.*, 1983a) probably arise from the higher purity of current preparations. Dashes indicate residues not detected; n.d. stands for determination not done.

Table II. Tryptic peptides (a) isolated from the headpiece and the amino-terminal sequences of the large tail fragments (b)

(a)	T_1	F R
	T_2	K G G A G G V R
	T_3	T S V S S V S A S P S R
	T_4	S A A G S S S G F H S W A R
	T_5	G A G A A S S T D S L D T L S(1B,1Q,1P,2G)
		5 10 15
(b)	Ch-tail	X K V K S(E)E K I K V
	CNBr I	(V)K V K S(E)E K I K V V(E K)X E(I)
	CNBr II	K E E E K P Q E V K A E(K P)X K(K A E)

Results are presented in the one letter code. Less commonly used letters are Z for glutamic acid or glutamine, B for aspartic acid or asparagine, X for unidentified residues. Brackets around a residue indicate that this amino acid has not been identified unambiguously using the gas phase sequenator. In (a) note the wealth of serine, glycine and alanine. T_5 is probably the carboxy-terminal tryptic peptide since it lacks arginine or lysine. The nomenclature of the tail fragments characterized in (b) is given in Figure 3. Note that CNBr I comprises the amino-terminal part of the full Ch-tail fragment obtained by chymotryptic digestion of NF-H.

omous domain not found in other major IF proteins accounts for the increased mol. wt. of NF-L in comparison with desmin, vimentin and GFA (Geisler *et al.*, 1982b, 1983a). For some time it was thought that NF-M and NF-H act primarily as peripherally bound associated proteins of a filament backbone made exclusively from NF-L (Geisler and Weber, 1981a; Willard and Simon, 1981; Liem and Hutchinson, 1982; Chin *et al.*, 1983; Julien and Mushynski, 1983). However, more detailed biochemical and immunological data have suggested that the two high mol. wt. triplet components are IF proteins in their own right, i.e., neurofilament proteins contain an amino-terminal region built analogously to other non-epithelial proteins which is followed by long non-α-helical tailpiece extensions (Geisler *et al.*, 1983a; Weber *et al.*, 1983). Although this hypothesis has been fully confirmed recently for NF-M by extended amino acid sequence data (Geisler *et al.*, 1984) and by self-assembly experiments (Gardner *et al.*, 1984), no comparable information is available for NF-H. This neurofilament component is, however, particularly interesting since several immuno-electron microscopical studies (Willard and Simon, 1981; Sharp *et al.*, 1982; Debus *et al.*, 1982; Hirokawa *et al.*, 1984) indicate that NF-H or, more specifically, its putative carboxy-terminal extension (Geisler *et al.*, 1983a; Weber *et al.*, 1983), is responsible for the cross-bridges present between neighbouring neurofilaments of axons and neurites. Here we identify NF-H as a hybrid molecule by protein-chemical data and partial amino acid sequence results. It carries in its amino-terminal region the structural information typical of a non-epithelial IF protein. Its extra mass is located to a long carboxy-terminal tailpiece extension of unique amino acid composition.

Results

Characterization of a non-α-helical amino-terminal headpiece in NF-H

Porcine NF-H was subjected to chemical cleavage at cysteine residues using ^{14}C-labelled 2-nitro-5-thiocyanobenzoic acid. Polyacrylamide gel electrophoresis in the presence of SDS revealed some 12 major fragments (Figure 1A). Of the bands in the lower mol. wt. range, only one, which is marked *h*, was free of radioactivity in the corresponding autoradiograph (Figure 1B). Given the known chemistry of the reaction (Jacobsen *et al.*, 1973), fragment *h* must comprise the amino-terminal region of NF-H. Additional non-labelled fragments of higher mol. wt. probably arise from incomplete cleavage. They are expected to start at the N terminus and to extend further along the polypeptide chain to other cysteine residues. Fragment *h* with an apparent mol. wt. of ~9–11 K was isolated using preparative gel electrophoresis and chromatography on phosphocellulose, where it was eluted as the last peak. In line with the binding to phosphocellulose fragment *h* contains about six residues of arginine and one or two residues of lysine (Table I). The amino acid composition also reveals a wealth of hyroxyamino acids, particularly serine (~20%). Although these properties together with the presence of ~3% proline are in line with the headpieces of various non-epithelial IF proteins such as desmin, vimentin, GFA, NF-L and NF-M (Geisler and Weber, 1982, 1983; Geisler *et al.*, 1983a, 1983b, 1984; Quax *et al.*, 1983), the high content of glycine (18%) and particularly alanine (17%) distinguishes NF-H from these proteins. A full amino acid sequence of this fragment is not yet available but partial data on tryptic peptides are summarized in Table II. These account for ~80% of the sequence if the fragment contains ~80 residues. The partial sequence data also emphasize the wealth of serine, glycine and alanine residues indicative of a non-α-helical conformation. Currently we do not know how far the headpiece array extends beyond fragment *h*. However, the presence of at least two aspartic acid residues in peptide T_5 (Table II) suggests that this peptide might be close to the carboxyl end of the NF-H headpiece, if comparison with other sequences of non-epithelial IF proteins is made (for references see above).

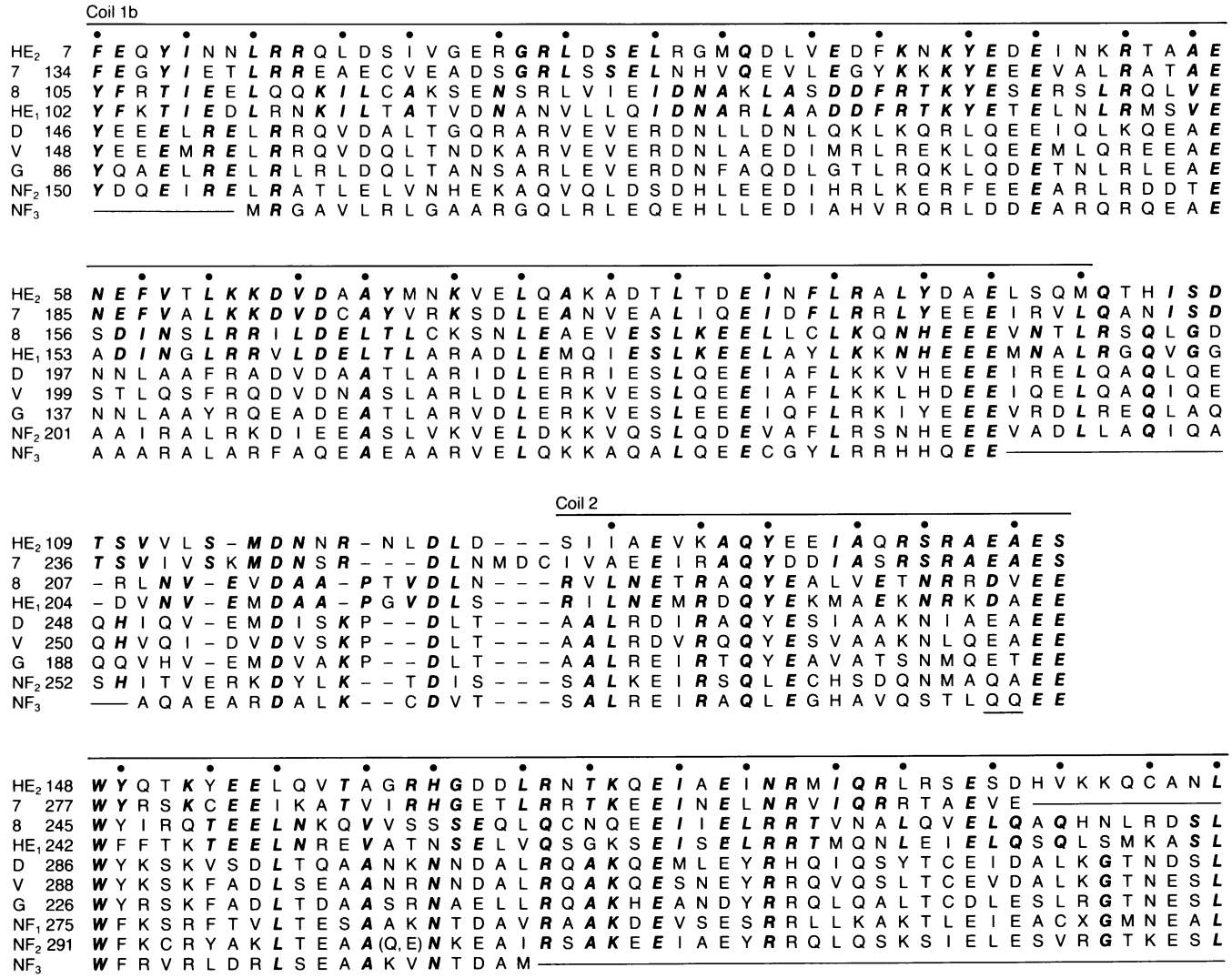

Fig. 2. Sequence relationship between NF-H and other IF proteins. Alignment is based on previous arguments (Geisler and Weber, 1982). Since only a 17-kd fragment of NF-H is known, the figure provides only a part of the 'common' α-helical middle domain (for a full documentation, see for instance Geisler *et al.*, 1984, and references given there). For primary sequence data see the following references (Geisler *et al.*, 1982a, 1983a; Geisler and Weber, 1982, 1983, 1984; Hanukoglu and Fuchs, 1982, 1983; Lewis *et al.*, 1984; Quax-Jeuken *et al.*, 1983; Quax *et al.*, 1983). Abbreviations are HE_1 and HE_2, human epidermal keratins 50 kd and 56 kd, respectively; 8 and 7, sheep wool α-keratins 8c-1 and 7c; D, chicken desmin; V, hamster vimentin; G, murine GFAP; NF_1, NF_2 and NF_3 are porcine neurofilament proteins NF-L, NF-M and NF-H, respectively. α-Keratin sequences (Crewther *et al.*, 1980; Sparrow and Inglis, 1980) are arranged as proposed (Geisler *et al.*, 1982, 1983a; for supporting evidence see Dowling *et al.*, 1983; Crewther *et al.*, 1983). Horizontal lines indicate as yet unestablished sequences. X is an arginine or lysine residue in NF_1. The predominantly hydrophobic residues *a* and *d* in the consecutive heptades of the proposed coiled-coils are indicated by dots above the top line. Note that the presentation used here for space saving reasons does not cover the amino-terminal 50 residues of the domain (coil Ia followed by a non-α-helical region) and the carboxy-terminal 70 residues of coil II (Geisler and Weber, 1982). Bold letters indicate identical residues among the different members of each of the three prototype sequences: i.e., non-epithelial IF proteins, keratins I and keratins II (Hanukoglu and Fuchs, 1983; Weber and Geisler, 1984). Deletions (dashes) allow for better alignment in the short non-α-helical spacer between coil Ib and II. Here a horizontal line indicates a short stretch of ~12 residues for which the sequence of NF-H remains undetermined (see text). Additionally, in the NF-H sequence, the two underlined glutamine residues 131 and 132 have not been identified unambiguously using the gas phase sequenator.

Amino acid sequence of a 17-K CNBr fragment covering half of the coiled-coil array in non-epithelial IF proteins

Since cysteine cleavage leads to blocked amino termini (Jacobsen *et al.*, 1973) a different approach was taken to obtain partial sequence information on the remaining regions of NF-H. Cleavage of NF-H with CNBr provided a complex pattern of fragments when analyzed by gel electrophoresis. In addition to several bands in the low mol. wt. region (≤10 K) there were two large mol. wt. bands (see below) and a pronounced fragment at 17 K. Amino acid composition indicated that the latter fragment could cover a highly α-helical region, whereas the large fragments (Table I) were clearly related to a non-α-helical region (see below). We have therefore developed the amino acid sequence of the 17-K fragment using enzymatic cleavage with trypsin, thermolysin and protease V8. Peptides were separated by two-dimensional paper fingerprints and h.p.l.c. They were analyzed by amino acid composition and stepwise Edman degradation. The combined results allow us to propose the amino acid sequence given in Figure 2. This proposal is supported by two automatic sequenator runs covering residues 2–23 and residues 119–137, respectively. The sole region of the 17 K fragment which remains unsolved corresponds to the residues tentatively given as 86–97. The peptides corresponding to these residues

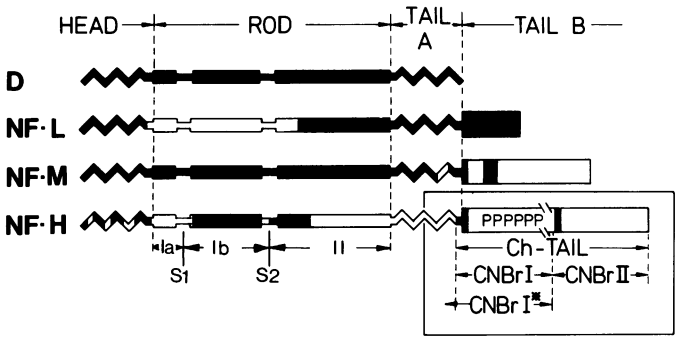

Fig. 3. Schematic representation of neurofilament proteins NF-L, NF-M and NF-H in comparison with muscle desmin (D). The common rod is divided by spacer 1 (S_1) and spacer 2 (S_2) into helices Ia, Ib and II (for details see Geisler and Weber, 1982). It is flanked at the amino-terminal side by the non-α-helical headpiece and at the carboxy-terminal side by the tailpiece and its extension. Tail part A is common to all non-epithelial IF proteins but has not yet been demonstrated for NF-H (see text). Part B, the extension, is specific for the neurofilament proteins (Geisler *et al.*, 1983a, 1984, and text). It increases in size from NF-L to NF-H. The Ch-tail of NF-H can be isolated by chymotryptic digestion of native NF-H (Geisler *et al.*, 1983a) and yields two CNBr fragments: CNBr I and CNBr II. Cleavage of intact NF-H by CNBr provides CNBr II and CNBr I*. The latter fragment exceeds CNBr I probably by only a few residues situated at the amino-terminal side. The mol. wts. of CNBr I and CNBr II are uncertain (see text). P in CNBr I indicate that this peptide contains many phosphoserine residues. The filled out stretches indicate regions of known amino acid sequence (Geisler and Weber, 1982; Geisler *et al.*, 1983a, 1984, and this paper). The sequence of the NF-H headpiece is only partially known (Table II).

could not be obtained in sufficient purity and amount to allow sequence determination (see below).

Inspection of the sequence shows that it approximately covers the middle region of the structurally preserved coiled-coil domain of 310 residues found in all IF proteins (Geisler and Weber, 1982). It can be directly aligned with the chicken desmin sequence if start and stop correspond to desmin residues 153 and 305, respectively (Figure 2). Thus, the 17-K fragment covers, in the general scheme of the rod domain, nearly the entire coil 1b, spacer 2 and the amino-terminal part of coil 2 (Figures 2,3). In the α-helical regions involved in coiled-coil formation (Geisler and Weber, 1982) the 17-K fragment of NF-H reveals the typical heptade repeat pattern with *a* and *d* positions being primarily hydrophobic in nature (see also Discussion). As expected from the coiled-coil forming ability, addition or deletion of even a single residue is not detected in these arrays. The as yet undocumented rather short array of residues 86–97 corresponds to the amino-terminal region of the non-α-helical spacer 2 and is in a region which in previous studies on other IF proteins has also posed some difficulties. The tentative estimate of ~12 residues for this gap is based on amino acid compositional data. Any small deviation from this number, if it were to be found in future studies, does not interfere with the major conclusions drawn so far. The 17-K fragment covers about half of the conserved coiled-coil array characteristic of IF proteins, it reveals the non-epithelial type sequence (see Discussion), and can be aligned with other IF sequences without difficulty.

Characterization of a long non-α-helical extension present at the carboxyl end of NF-H

When NF-H is cleaved by CNBr, two large fragments of apparent mol. wt. 110 kd (I*) and 45 kd (II) show an amino acid composition highly related to that of the non-α-helical tailpiece extension (Ch-tail) obtained by mild chymotryptic digestion of NF-H (Geisler *et al.*, 1983a; Table I, Figure 3). Furthermore, when the isolated Ch-tail is cleaved by CNBr, again two fragments are obtained – CNBr I and CNBr II (Figure 3). These are very similar, if not identical, in apparent mol. wt. and amino acid composition, with the corresponding CNBr fragments derived from the whole protein (Table I). Their low content of large hydrophobic amino acids and arginine is in line with their derivation from regions outside the coiled-coil array (Table I). These fragments reveal a wealth of alanine (15%), proline (14%) and particularly lysine (25%) and glutamic acid (25%) in standard acid hydrolysis. Enzymatic hydrolysis performed on fragments I* and II show that 95% and 85%, respectively, of the glutamic acid value is indeed the acid rather than the amide. In spite of a generally similar composition, CNBr fragments I* and II differ distinctly in the following points. First, fragment II seems to lack, or contain only trace amounts of, the following common amino acids: isoleucine, leucine and arginine, whereas fragment I* shows noticeable but very small amounts of these residues. In addition I* has a slightly higher content of serine and a lower content of aspartic acid and threonine than II. Second, in spite of its high glutamic acid content, fragment II, probably because of the wealth of lysine residues, does not bind to DEAE-cellulose. This is not the case for fragment I*. The anionic character of fragment I* seems to arise from the presence of many serine phosphates noticeably absent in fragment II (see Table II). Approximate estimates indicate values of ~50 mol of serine phosphate per apparent mol. wt. of 110 kd for CNBr I* (see, however, Discussion).

The relative order of the tail fragments CNBr I and CNBr II could be determined by matching the amino-terminal sequences obtained for CNBr I with the corresponding sequence obtained for the chymotryptic Ch-tail (Table II). CNBr I is therefore the amino-terminal fragment whereas CNBr II covers the carboxy-terminal array of the Ch-tail domain. The amino-terminal sequences obtained for CNBr I and CNBr II are in line with the prediction made from the compositional data, i.e., they reveal a wealth of lysine and glutamic acid residues.

Discussion

The amino acid sequence and compositional data developed here identify the largest mammalian neurofilament component as a hybrid molecule. NF-H carries, in its amino-terminal region, the sequence information currently thought to identify a non-epithelial IF protein. The extra mass of NF-H arises from a large non-α-helical extension most likely situated at the carboxyl end. These results on NF-H, together with a detailed sequence characterization of NF-M (Geisler *et al.*, 1984), fully confirm our previous hypothesis that the large mol. wt. neurofilament proteins are IF proteins in their own right co-polymerized in the filament with the major component NF-L, which has a much lower mol. wt. (Geisler *et al.*, 1983a). This interaction most likely occurs *via* the extended coiled-coil forming arrays thought to cover most of the conserved middle region of 310 residues found so far in all IF proteins. In addition, our results allow a better understanding of the carboxy-terminal extension of NF-H which, from immuno-electron microscopical studies, seems the prime candidate for the thin cross-bridge structure connecting neighbouring neurofilaments of axons and neurites in well preserved specimens analyzed by electron microscopy (Willard and

Simon, 1981; Sharp *et al.*, 1982; Geisler *et al.*, 1981a; Weber *et al.*, 1983; Hirokawa *et al.*, 1984).

Mammalian neurofilament proteins have apparent mol. wts. of ~200 K (NF-H), 160 K (NF-M) and 68 K (NF-L) when analyzed by SDS-gel electrophoresis (Geisler and Weber, 1981a; Liem and Hutchinson, 1982; Julien and Mushynski, 1982, 1983; Chin *et al.*, 1983). However, recent studies using gel filtration and sedimentation equilibrium centrifugation in 6 M guanidine-HCl point to lower values, i.e., 110–140 K, 107 K and 62 K and indicate that the aberrant behaviour in gel electrophoresis stems from the unusual tailpiece extensions (Kaufmann *et al.*, 1984). Although the extended sequence data on NF-L (Geisler *et al.*, 1983a) support this prediction, a firm conclusion as to the actual mol. wts. of NF-M and NF-H will require more chemical data. Nevertheless, they are very large molecules and the determination of the full amino acid sequences by protein chemistry alone is still a formidable task. Over the last 3 years we have developed the full amino acid sequence of chicken desmin and characterized porcine desmin, porcine vimentin and porcine glial fibrillary acidic protein (GFA) to 38%, 59% and 50%, respectively (Geisler and Weber, 1981b, 1982, 1983; Geisler *et al.*, 1982b, 1983b). This approach yielded a detailed description of the topographical organization of IF proteins in general (Geisler and Weber, 1982; see also Hanukoglu and Fuchs, 1982) and non-epithelial IF proteins in particular. The conclusions were confirmed when DNA technology provided a complete and a nearly complete sequence for hamster vimentin and murine GFA, respectively (Quax *et al.*, 1983; Lewis *et al.*, 1984). Neurofilament data have so far only been developed by protein sequence. Given the presence of three large polypeptides, we have been more concerned with general structural principles than with complete sequences. Thus NF-L and NF-M are known only to 65 and 52%, respectively. Currently missing are the amino-terminal part of the rod domain of NF-L (Geisler *et al.*, 1983a) and nearly the entire carboxy-terminal tailpiece extension of NF-M (Geisler *et al.*, 1984) (Figure 3). In the case of NF-H, sequence data are less advanced, but given the established general framework they are sufficient to prove the predicted hybrid character of a non-epithelial IF protein domain followed by a very large non-α-helical extension. Thus we have characterized a 9–10 K non-α-helical headpiece unambiguously spanning the amino-terminal region of NF-H. Although its high content of alanine is unusual, its wealth of hyroxyamino acids together with the presence of several arginine and proline residues points to the features of the variable headpieces found in non-epithelial IF proteins (reviewed in Geisler and Weber, 1983). Previous proteolytic studies using chymotrypsin on NF-H allowed the isolation of a highly α-helical 40-K fragment which was assumed to span the rod domain documented in other IF proteins (Geisler *et al.*, 1983a). This proposal is now directly verified by the amino acid sequence of a 17-K fragment obtained by CNBr cleavage. It reveals extensive α-helical arrays with the characteristic heptade repeat pattern indicative of coiled-coil forming ability. As expected the sequence can be aligned unambiguously with the middle part of the rod domain of other IF proteins (Figure 2). That the conserved rod domain will also be ~35 K (i.e., 310 residues) for NF-H is predicted by the isolation of the α-helical 40-K fragment and its reactivity with a general monoclonal antibody to IF proteins whose epitope lies in the consensus type sequence at the carboxyl end of the presumptive coiled-coil array (Geisler *et al.*, 1983a).

We have also analyzed the carboxy-terminal extension of NF-H by amino acid compositional data and limited amino-terminal amino acid sequence determination of its fragments. This domain is characterized by a wealth of glutamic acid and lysine residues (~25% each) and has a high content of proline (12%) and alanine (15%). The Ch-tail domain, isolated after chymotryptic digestion of native NF-H, can be split with CNBr into two fragments of apparent mol. wt. 110 K (CNBr I) and 45 K (CNBr II). CNBr I could be placed at the amino-terminal end and CNBr II at the carboxy-terminal end of the tail domain. The peptides appear to be rather difficult to sequence on the automatic sequencer because of a rapidly rising background of glutamic acid and lysine. However, the amino-terminal sequences obtained for the two fragments appear rather similar, except for some minor variations concerning the content of hydrophobic amino acids. They display the meanwhile well known sequence type of NF tailpiece extensions (Geisler *et al.*, 1983a, 1984). One observation shows that this array in NF-H seems to have subdomains. The amino-terminal fragment CNBr I contains a high amount of phosphoserine (~5 mol%), while the following CNBr II fragment has very little if any phosphate. However, we cannot give an absolute number of moles of phosphoserines per mole of NF-H for the following reasons. First, we do not know the true mol. wt. of CNBr I (see above). Second, CNBr cleavage performed for prolonged time at room temperature may lead to some loss of phosphate. Third, the method used by us to detect phosphorylation is not very accurate. These limitations do not detract from the conclusion that the phosphate substitutions are not evenly distributed through the tailpiece extension but rather are restricted to the amino-terminal region. This finding supports the gel electrophoretic results of Julien and Mushynski (1983), who concluded from V8 protease digests that the majority of serine phosphates is located in a fragment of 40 kd observed by gel electrophoresis. Currently it is not known if the phosphate substitution carries any functional importance, but its restriction to the amino-terminal region of the tailpiece extension indicates the presence of subdomains. So far the sequences available for the tail domains of the three triplet proteins appear similar although a true homology has not yet been detected. This raises the question of whether the cross-bridging function between neighbouring filaments is performed exclusively by the tailpiece extension of NF-H as suggested by immunoelectron microscopy (Willard and Simon, 1981; Sharp *et al.*, 1982; Debus *et al.*, 1982; Hirokawa *et al.*, 1984), or if the related domains of NF-L and NF-M also participate in the structure.

The amino-terminal sequence of the Ch-tail indicates that the chymotryptic cut has occurred either within the lysine/glutamic acid-rich sequences or directly prior to them. Thus it remains unknown whether NF-H displays a short non-α-helical tailpiece domain (domain A in Figure 3) which has been found so far in all non-epithelial IF proteins including NF-L and NF-M (Geisler *et al.*, 1983a, 1984). This domain of ~30–40 residues is in NF-L similar to the headpiece in the display of β-sheet and β-turns. In NF-L and NF-M it provides the connection between the glutamic acid/lysine-rich extensions and the rod domains. So far we have not found the link between the rod domain and the tailpiece extension in NF-H, and similarly we lack an overlap between the headpiece and the rod. The isolated headpiece is, however, unambiguously located at the amino terminus of NF-H, since it is not labelled after cleavage with ^{14}C-labelled 2-nitro-5-thiocyanobenzoic

acid. However, to place the tailpiece extension at the carboxy terminus we have to argue mainly by analogy with all other five non-epithelial IF proteins which show this feature in their sequence data. In addition, CNBr II obtained from total NF-H seems to lack a carboxy-terminal homoserine as judged by hydrazinolysis experiments. Thus it is very likely that this fragment covers the carboxy-terminal region of the whole molecule (our unpublished observation).

Current concepts of keratin structure dating back to earlier studies on wool α-keratins (Crewther and Dowling, 1971) show that epithelial IF proteins are built as obligatory heteropolymers. The documentation of only two prototype rod sequences indicates that this complementarity most likely occurs at the level of the double-stranded coiled-coil (see for instance Crewther *et al.*, 1983; Franke *et al.*, 1983; Hanukoglu and Fuchs, 1983; Steven *et al.*, 1983; Quinlan *et al.*, 1984; Weber and Geisler, 1984). Contrary to this situation, the highly related non-epithelial IF proteins desmin, vimentin and GFA form homopolymeric filaments and are able to copolymerize *in vitro* and *in situ* (reviewed by Geisler and Weber, 1983). In the case of neurofilaments the situation seems more complex. Whereas NF-L easily forms long IF *in vitro* (Geisler and Weber, 1981a; Liem and Hutchinson, 1982) a similar ability of NF-M was only recently discovered using rather restricted experimental conditions (Gardner *et al.*, 1984). Since such experiments are necessarily based on renaturation experiments from urea solution it seems that the extensive non-α-helical tailpiece extension of NF-M (Geisler *et al.*, 1984) may have interfered in earlier assembly studies. Although NF-H provides short curly structures of a smaller diameter than NF-L and NF-H (Gardner *et al.*, 1984) long IF have so far not been obtained. Again this *in vitro* situation may be influenced by the long tailpiece since current sequence data seem to allow for a proper self assembly. NF-M and NF-H both have a non-α-helical headpiece and share, along the rod domain, 45% sequence identity. If allowance for only the most conservative exchanges (arginine *versus* lysine; aspartic *versus* glutamic acid, and exchange of hydrophobic residues) is made the sequence homology increases to 68%. We note, however, that compared with the other non-epithelial IF proteins, the NF-H rod domain displays a peculiarity. Many large hydrophobic residues such as leucine and isoleucine occurring in *a* and *d* positions of the heptades are replaced by the less hydrophobic alanine. This raises the possibility that the coiled-coils formed by NF-H alone may be less stable. This could also be related to our finding that the 40-K rod domain of NF-H, unlike the rods of NF-L and NF-M, is rather sensitive to longer exposure to chymotrypsin and can only be isolated by very short digestion (Geisler *et al.*, 1983a; Julien and Mushynski, 1983). Thus the difficulty in obtaining IF from isolated NF-H may lie in the rod structure itself. Various immuno-histochemical studies indicate that *in situ* triplet proteins present in the same neuronal compartment do not segregate into different filaments but rather show co-polymerization (Willard and Simon, 1981; Sharp *et al.*, 1982; Shaw and Weber, 1982; Shaw *et al.*, 1981; Hirokawa *et al.*, 1984). Future studies using proteolytically trimmed molecules may clarify the *in vitro* assembly properties of the high mol. wt. neurofilament proteins and also characterize their interaction patterns with NF-L.

Materials and methods

Neurofilament triplet proteins from porcine spinal cord were isolated and separated using DEAE-cellulose and gel filtration in the presence of urea (Geisler and Weber, 1981a; Geisler *et al.*, 1983a). NF-H was subjected to chemical cleavage at either cysteine or methionine using ^{14}C-labelled 2-nitro-5-thiocyanobenzoic acid (Jacobsen *et al.*, 1973; Geisler *et al.*, 1982b) or CNBr, respectively. Fragments were screened by SDS-polyacrylamide gel electrophoresis and when necessary autoradiography was also performed. The 17-K fragment derived by CNBr cleavage was purified by preparative gel electrophoresis. CNBr peptides covering the tail region were obtained either by cleavage of the whole protein or by cleavage of the Ch-tail, which was isolated after mild chymotryptic digestion of NF-H (Geisler *et al.*, 1983a) by gel filtration on Sepharose 6B followed by chromatography on DEAE-cellulose in 8 M urea. These fragments were also purified by preparative SDS-gel electrophoresis. Electro-elution, concentration and removal of SDS has been described (Geisler *et al.*, 1982b). The amino-terminal 10-K fragment derived by cysteine cleavage was purified by preparative gel electrophoresis and by chromatography on phosphocellulose. A highly related fragment was also obtained when the mixture of CNBr fragments of NF-H was subjected to CM-cellulose chromatography in the presence of 6 M urea as described for other headpiece fragments (Geisler *et al.*, 1983a; Geisler and Weber, 1983).

Individual sequences were determined from tryptic, thermolytic and V8-protease peptides separated by two-dimensional fingerprint methods by paper chromatography as well as by h.p.l.c. (for details see Geisler *et al.*, 1983a). These peptides were characterized by amino acid composition and stepwise Edman degradation using the modified technique (Chang *et al.*, 1978). For further details see previous references (Geisler and Weber, 1982; Geisler *et al.*, 1983a). Using a gas phase sequenator the following regions of the 17-K fragment were also determined, i.e., residues 2–23 and residues 119– 137. The amino-terminal sequences of the fragments from the tail region were also obtained using the sequenator. The ratio of glutamic acid *versus* glutamine in the tail fragments was determined by amino acid analysis after enzymatic hydrolysis using pronase, leucine aminopeptidase and prolidase. The presence and the approximate amount of phosphoserine was obtained using hydrolysis in 6 N HCl at 100°C for 2 h. Separation was either by paper electrophoresis at pH 3.5, or by standard amino acid analysis. Appropriate phosphoserine standards were included to allow quantitation.

Acknowledgements

We thank A.Gruber and S.Gaspar for expert technical assistance. This work was supported by the Max Planck Society and a grant from the Deutsche Forschungsgemeinschaft to N.G. and K.W. J.V. is Bevoegdverklaard Navorser at the Belgian National Fund for Scientific Research (N.F.W.O.). S.F. was supported by a predoctoral fellowship from the Max Planck Society.

References

Chang,Y., Bauer,D. and Wittmanmn-Liebold,B. (1978) *FEBS Lett.*, **93**, 205-214.
Chin,T.K., Eagles,P.A.M. and Maggs,A. (1983) *Biochem. J.*, **215**, 239-252.
Crewther,W.G. and Dowling,L.M. (1971) *Appl. Polym. Symp.*, **18**, 1-20.
Crewther,W.G., Dowling,L.M. and Inglis,A.S. (1980) *Proceedings of the 6th Quinquennial International Wool Textile Research Conference,* Pretoria, Vol. **2**, pp. 79-91.
Crewther,W.G., Dowling,L.M., Steinert,P.M. and Parry,D.A.D. (1983) *Int. J. Biol. Macromol.*, **5**, 267-274.
Debus,E., Flügge,G., Weber,K. and Osborn,M. (1982) *EMBO J.*, **1**, 41-45.
Dowling,L.M., Parry,D.A.D. and Sparrow,L.G. (1983) *Biosci. Rep.*, **3**, 73-78.
Franke,W.W., Schiller,D.L., Hatzfeld,M. and Winter,S. (1983) *Proc. Natl. Acad. Sci. USA*, **80**, 7113-7117.
Gardner,E.E., Dahl,D. and Bignami,A. (1984) *J. Neurosci.*, **11**, 145-155.
Geisler,N. and Weber,K. (1981a) *J. Mol. Biol.*, **151**, 565-571.
Geisler,N. and Weber,K. (1981b) *Proc. Natl. Acad. Sci. USA*, **78**, 4120-4123.
Geisler,N. and Weber,K. (1982) *EMBO J.*, **1**, 1649-1656.
Geisler,N. and Weber,K. (1983) *EMBO J.*, **2**, 2059-2063.
Geisler,N., Kaufmann,E. and Weber,K. (1982a) *Cell*, **30**, 277-286.
Geisler,N., Plessmann,U. and Weber,K. (1982b) *Nature*, **296**, 448-450.
Geisler,N., Kaufmann,E., Fischer,S., Plessmann,U. and Weber,K. (1983a) *EMBO J.*, **2**, 1295-1302.
Geisler,N., Plessmann,U. and Weber,K. (1983b) *FEBS Lett.*, **163**, 22-24.
Geisler,N., Fischer,S., Vandekerckhove,J., Plessmann,U. and Weber,K. (1984) *EMBO J.*, **3**, 2701-2706.
Hanukoglu,I. and Fuchs,E. (1982) *Cell*, **31**, 243-252.
Hanukoglu,I. and Fuchs,E. (1983) *Cell*, **33**, 915-924.
Henderson,D., Geisler,N. and Weber,K. (1982) *J. Mol. Biol.*, **155**, 173-176.
Hirokawa,N., Glicksman,M.A. and Willard,M.B. (1984) *J. Cell Biol.*, **98**, 1523-1536.
Jacobsen,G.R., Schaffer,M.H., Stark,G.R. and Vanaman,T.C. (1973) *J. Biol. Chem.*, **248**, 6583-6591.
Julien,J.P. and Mushynski,W.E. (1982) *J. Biol. Chem.*, **257**, 10467-10470.

Julien,J.P. and Mushynski,W.E. (1983) *J. Biol. Chem.*, **258**, 4019-4025.
Kaufmann,E., Geisler,N. and Weber,K. (1984) *FEBS Lett.*, **170**, 81-84.
Lewis,S.A., Balcarek,J.M., Krek,V., Shelanski,M. and Cowan,N.J. (1984) *Proc. Natl. Acad. Sci. USA*, **81**, 2743-2746.
Liem,R.K.H. and Hutchinson,S.B. (1982) *Biochemistry (Wash.)*, **21**, 3221-3226.
Milam,L. and Erickson,H.P. (1982) *J. Cell Biol.*, **94**, 592-596.
Moll,R., Franke,W.W., Schiller,D.L., Geiger,B. and Krepler,R. (1982) *Cell*, **31**, 11-24.
Quax,W.J., Egbert,W.V., Hendriks,W., Quax-Jeuken,Y. and Bloemendal, H. (1983) *Cell*, **35**, 215-223.
Quax-Jeuken,Y.E.F.M., Quax,W.J. and Bloemendal,H. (1983) *Proc. Natl. Acad. Sci. USA*, **80**, 3548-3552.
Quinlan,R.A., Cohlberg,J.A., Schiller,D.L., Hatzfeld,M. and Franke,W.W. (1984) *J. Mol. Biol.*, **178**, 365-388.
Sharp,G., Shaw,G. and Weber,K. (1982) *Exp. Cell Res.*, **137**, 403-413.
Shaw,G., Osborn,M. and Weber,K. (1981) *Eur. J. Cell Biol.*, **26**, 68-82.
Shaw,G. and Weber,K. (1982) *Nature*, **298**, 277-279.
Sparrow,L.G. and Inglis,A.S. (1980) *Proceedings of the 6th Quinquennial International Wool Textile Research Conference*, Pretoria, Vol. **2**, pp. 237-246.
Steinert,P.M., Idler,W.W. and Goldman,R.D. (1980) *Proc. Natl. Acad. Sci. USA*, **77**, 4534-4538.
Steinert,P.M., Rice,R.H., Roop,D.R., Trus,B.L. and Steven,A.D. (1983) *Nature*, **302**, 794-800.
Steven,A.C., Hainfield,J.F., Trus,B.L., Wall,J.S. and Steinert,P.M. (1983) *J. Cell Biol.*, **97**, 1939-1944.
Traub,P. and Vorgias,C.E. (1983) *J. Cell Sci.*, **63**, 43-67.
Weber,K. and Geisler,N. (1984) *Cancer Cells (The Transformed Phenotype)*, Vol. **1**, published by Cold Spring Harbor Laboratory Press, NY, pp. 153-159.
Weber,K., Shaw,G., Osborn,M., Debus,E. and Geisler,N. (1983) *Cold Spring Harbor Symp. Quant. Biol.*, **48**, 717-729.
Willard,M.B. and Simon,C. (1981) *J. Cell Biol.*, **89**, 198-205.

Received on 22 October 1984; revised on 22 November 1984