Proc. Natl. Acad. Sci. USA
Vol. 80, pp. 3548–3552, June 1983
Biochemistry

Primary and secondary structure of hamster vimentin predicted from the nucleotide sequence

(intermediate-size filament protein/helical domains)

YVONNE E. F. M. QUAX-JEUKEN, WIM J. QUAX, AND HANS BLOEMENDAL*

Department of Biochemistry, University of Nijmegen, Geert Grooteplein N21, 6525 EZ Nijmegen, The Netherlands

Communicated by L. N. M. Duysens, January 31, 1983

ABSTRACT **The nucleotide sequence of two recombinant plasmids containing hamster vimentin cDNA was determined. The sequence comprises 1,640 base pairs and reveals virtually the total primary structure of vimentin and a large part of the 3′ noncoding region. Secondary structure prediction methods allow the characterization of two distinct regions of the polypeptide chain, 135 and 145 residues long, which are able to form α helices organized in "coiled coils." Three nonhelical domains can be distinguished: a very basic NH_2-terminal domain of at least 67 residues, a nonhelical region of 45 amino acids separating the two helix domains, and a COOH-terminal region of 55 residues, which contains an excess of acidic amino acids. The meaning of each of these domains of the vimentin polypeptide for the subunit and filament formation is discussed.**

Intermediate-size (7–11 nm) filaments (IF) have been found, in addition to microtubules and microfilaments, as cytoskeletal structures in most eukaryotic cells. Immunological and biochemical studies have resulted in the recognition of five major IF protein subunit classes, each related to certain cell and tissue types: the epithelial cytokeratins, the muscle desmin, the neurofilament proteins, the astrocyte glial filament protein, and vimentin found in cells of mesenchymal origin (1, 2). The latter also is expressed in many, if not all, cultured cell lines (3). In spite of a lack of immunological crossreactivity, the structural proteins of IF must have several common properties because they constitute filamentous structures with similar morphological characteristics (4, 5), and frequently they show an end-on attachment to the plasma membrane (6). Moreover, they are all nearly insoluble in physiological buffers. Mainly because of this insolubility, little is known about the chemistry of the individual IF proteins.

Nevertheless, in some of the IF proteins, the presence of α-helical segments, which are assumed to be able to form "coiled coils," has been shown (7, 8). Steinert and co-workers (5) have proposed a general model for the IF proteins of baby hamster kidney BHK-21 cells and bovine keratinocytes, which consist of two constant α-helical domains equal in length [180 Å; 18–20 kilodaltons (kDa)] and separated by a nonhelical domain of variable size (5–17 kDa). Three adjacent polypeptides are assumed to be arranged as a three-stranded rope constituting the so-called protofilamentous unit. Limited chymotryptic cleavage and circular dichroism determination of purified chicken desmin revealed three structurally distinct domains of the protein: a very basic and highly protease-sensitive nonhelical NH_2-terminal domain of 7.5 kDa, a rod-like domain of 38 kDa (500 Å), which is very rich in helix structure (83%), and a non-α-helical COOH-terminal tailpiece of 5.5 kDa (9). These findings concerning the secondary structure of IF subunits would be more clearly understood at the molecular level if there were information available on the amino acid sequence of their proteins. The primary structures of desmin and vimentin are only known in part. Comparison of the amino acid sequence of the COOH-terminal 178 residues of porcine stomach desmin and lens vimentin showed that these molecules are closely related and differed in only 36% of their residues (10–12).

However, the primary structure of a large part of desmin and a major part of vimentin are still unknown. We choose the way of constructing plasmids containing cDNA coding for vimentin to determine the primary structure of this protein. This paper presents the virtually total primary structure of hamster lens vimentin derived from the nucleotide sequence of two cDNA clones. The predicted secondary structure is discussed in relation to IF organization.

MATERIALS AND METHODS

Isolation of Recombinant Plasmid DNA. The construction of bacterial plasmids containing sequences homologous to hamster lens vimentin mRNA inserted into the *Pst* I site of pBR322 by oligo(dG·dC)-tailing and the subsequent transformation of *Escherichia coli* HB101 have been described (13). Colony hybridization and plasmid DNA extraction were performed as reported (14, 15).

Restriction Endonuclease Mapping and DNA Sequence Analysis. The locations of restriction enzyme sites on the inserted DNA were determined by digesting plasmids containing cDNA coding for vimentin (pVim-1 and pVim-2) with one or more restriction enzymes and analyzing the fragments on 5% polyacrylamide gels containing 90 mM Tris borate buffer (pH 8.3) and 2.5 mM EDTA. Restriction fragments were recovered from the gel by electroelution, precipitated, and washed with ethanol. The purified DNA fragments were labeled, and the sequence was determined by established procedures (16, 17). The DNA sequence was recorded, edited, translated, and compared by using the computer programs of Staden (18). For the prediction of secondary structure, the heptade convention (19) was used; moreover, the rules of Chou and Fasman (20) were applied for the same goal.

RESULTS AND DISCUSSION

DNA Sequence and Predicted Polypeptide Sequence. We have reported (13) the construction of recombinant plasmids containing sequences homologous to vimentin mRNA. Two clones pVim-1 and pVim-2, which directed the synthesis of vimentin, were selected upon positive hybridization translation for sequence assay by the method of Maxam and Gilbert (16). The combined restriction map of the two clones and the se-

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: kDa, kilodaltons; bp, base pairs; IF, intermediate-size filament.

* To whom reprint requests should be addressed.

quence assay strategy are presented in Fig. 1. Both 5'- and 3'-labeling methods were used. The combined nucleotide sequence of pVim-1 and pVim-2 (Fig. 2) comprises 1,640 base pairs (bp). This is more than 80% of the mRNA information, of which the length was estimated at about 2,000 bases (13). All of the sequences shown were established by sequence analysis of both strands. In the combined sequence of Fig. 2, the G·C tails, introduced during cloning are not shown.

There is only one reading frame of sufficient length to specify a protein with the size of vimentin, which has been estimated to be 52,000 to 55,000 daltons (1). This coding region contains 1,344 bp coding for 448 amino acids. The NH_2-terminal part of the derived polypeptide sequence (residues 1–60) shows a very basic character because of the presence of nine arginine residues and only one acidic amino acid (glutamic acid). The remaining part of the polypeptide has an excess of acidic residues and shows a pronounced hydrophilic character. In spite of this hydrophilic nature of a major part of vimentin, this IF protein is almost insoluble in physiological buffers. Probably secondary structure interactions are responsible for this insolubility.

Previous findings with two-dimensional gel electrophoresis of total cell extracts or partially purified fractions containing vimentin or desmin show a staircase-type degradation pattern of decreasing isoelectric point (13, 21, 22). A specific protease (21) acts gradually at the NH_2 terminus of vimentin, possibly at the arginine residues, and causes the split-off of basic peptides (9). This explains the shift of degradation products of vimentin towards the acidic side of the gel (see Fig. 3). The functional significance of this protease activity is still unknown, but it may play a role in vimentin turnover (21).

The coding region of the cDNA sequence is followed by the stop codon TAA and a noncoding region of 296 bp. The noncoding 3' end did not show any poly(A) sequences nor the hexanucleotide sequence A-A-T-A-A-A considered to be the signal involved in polyadenylylation (23, 24). At the 5' end, the cDNA insertions did not contain the start codon. As no data are available on the NH_2-terminal amino acid sequence of vimentin, we can only estimate the number of missing residues from the reported molecular weight of vimentin (1, 25). By taking into consideration the discrepancy of the reported molecular weights of vimentin, the number of residues missing is less than 5% of the primary structure. Meanwhile, we know from sequence data obtained from the isolated chromosomal vimentin gene that the only possible sites for the start codon are either positions −3 or −16, relative to the first amino acid residue shown in Fig. 2 with some uncertainty remaining about codon 8 (unpublished data). From this data, a maximal molecular weight of 53,500 can be calculated for the monomeric vimentin molecule. Therefore, the percentage of missing amino acids is at most 3.5.

Comparison of Amino Acid Sequences of Hamster and Porcine lens Vimentin. The amino acid sequence ranging from residue 272 to 448 in Fig. 2 almost exactly fits the reported sequence of the 177 amino acids of the COOH terminus of porcine lens vimentin (11, 12). There are only two differences, most likely as a result of one-base substitutions, at positions 335 and 445, where leucine and aspartic acid of hamster lens vimentin changed into valine and asparagine, respectively, in porcine lens vimentin. These observations provide evidence for the earlier proposed evolutionary conservation of vimentin (13). A similar phenomenon has been found for amino acid sequences of desmin from different species (11).

Secondary Structure Prediction. The predicted sequence of the protein was analyzed for the ability to form α helices organized in coiled coils. For this purpose we applied the heptade convention (Fig. 4), which showed its usefulness in the analysis of tropomyosin, keratins (19, 26, 27), and chicken desmin (9). The sequence of the vimentin polypeptide displayed distinct regions that could fit a helical configuration (seven residues per two turns) with the necessary alignment of apolar amino acids to form a hydrophobic backbone. The hydrophobic bands of three neighboring helix domains can link together, forming the three-stranded rope of the protofilamentous unit. The remainder of the helix surface is generally polar.

Besides helical domains, two nonhelical regions could be distinguished: an NH_2-terminal domain of at least 67 residues and a COOH-terminal region of 55 amino acids. If we look at the helix domain, we find two short interruptions at positions 131

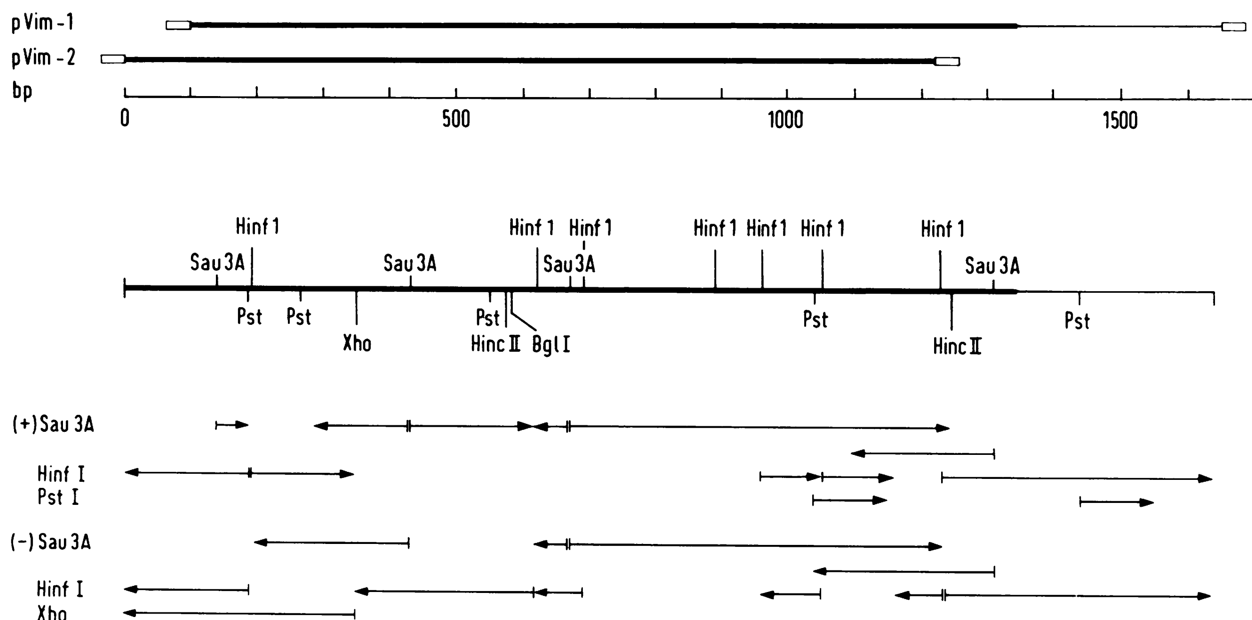

FIG. 1. Restriction map and strategy for sequence determination of the inserts of plasmids pVim-1 and pVim-2. The left and right sides of the restriction map correspond to position 3,612 of the *Pst* I site of pBR322. Restriction sites used for sequence determination are shown; arrows indicate the direction and extent of sequence analysis. The thick bars indicate the coding region, and the open boxes indicate the G·C tails introduced during the cloning procedure.

90

CCC CGG CAC CTC GAA CCG GCA GGC TCC AAC CGG AGC TAT GTG ACC ACG TCC ACC CGC ACC TAC AGC CTG GGG GCA CTG CGC CCC AGC ACC

pro arg his leu glu pro ala gly ser asn arg ser tyr val thr thr ser thr arg thr tyr ser leu gly ala leu arg pro ser thr

10 20

180

AGT CGC AGC CTC TAT TCC TCA TCT CCC GGT GGC GCC TAT GTG ACC CGA TCC TCT GCG GTG CGC CTG CGG AGC AGC ATG CCC GGC GTG AGG

ser arg ser leu tyr ser ser ser pro gly gly ala tyr val thr arg ser ser ala val arg leu arg ser ser met pro gly val arg

40 50

270

CTG CTG CAG GAC TCG GTG GAC TTC TCG CTG GCC GAC GCC ATT AAC ACC GAG TTC AAG AAC ACC CGC ACC AAC GAG AAG GTA GAA CTG CAG

leu leu gln asp ser val asp phe ser leu ala asp ala ile asn thr glu phe lys asn thr arg thr asn glu lys val glu leu gln

70 80

360

GAG CTG AAT GAC CGC TTC GCC GAC TAC ATC GAC AAG GTG CGC TTC CTA GAG CAG CAG AAC AAA ATC CTG CTA GCC GAG CTC GAG CAA CTC

glu leu asn asp arg phe ala asp tyr ile asp lys val arg phe leu glu gln gln asn lys ile leu leu ala glu leu glu gln leu

100 110

450

AAG GGT CAG GGC AAG TCG CGC CTG GGC GAC CTC TAT GAG GAG GAG ATG CGG GAG CTG CGC CGG CAG GTG GAT CAG CTC ACC AAC GAC AAG

lys gly gln gly lys ser arg leu gly asp leu tyr glu glu glu met arg glu leu arg arg gln val asp gln leu thr asn asp lys

130 140

540

GCA CGC GTC GAG GTG GAG CGT GAC AAC CTA GCT GAG GAC ATC ATA CGG CTG CGA GAA AAA TTG CAG GAG GAG ATG CTC CAG AGA GAG GAA

ala arg val glu val glu arg asp asn leu ala glu asp ile ile arg leu arg glu lys leu gln glu glu met leu gln arg glu glu

160 170

630

GCG GAG AGC ACC CTG CAG TCC TTC AGA CAG GAT GTT GAC AAT GCC TCT CTG GCA CGC CTC GAC CTT GAA CGT AAA GTG GAA TCC TTG CAA

ala glu ser thr leu gln ser phe arg gln asp val asp asn ala ser leu ala arg leu asp leu glu arg lys val glu ser leu gln

190 200

720

GAA GAG ATT GCC TTT TTG AAG AAA CTG CAT GAT GAA GAG ATC CAG GAG CTA CAG GCC CAG ATT CAG GAG CAA CAT GTC CAG ATT GAC GTG

glu glu ile ala phe leu lys lys leu his asp glu glu ile gln glu leu gln ala gln ile gln glu gln his val gln ile asp val

220 230

810

GAT GTT TCT AAG CCC GAC CTC ACT GCT GCC CTG CGC GAT GTC CGC CAG CAG TAT GAA AGT GTG GCT GCC AAG AAC CTC CAG GAG GCG GAG

asp val ser lys pro asp leu thr ala ala leu arg asp val arg gln gln tyr glu ser val ala ala lys asn leu gln glu ala glu

250 260

900

GAA TGG TAC AAG TCC AAG TTT GCC GAC CTC TCT GAA GCT GCC AAC CGG AAC AAT GAT GCC CTG CGC CAG GCA AAG CAG GAG TCA AAT GAG

glu trp tyr lys ser lys phe ala asp leu ser glu ala ala asn arg asn asn asp ala leu arg gln ala lys gln glu ser asn glu

280 290

990

TAC CGG AGA CAG GTG CAG TCA CTC ACC TGC GAA GTG GAC GCA CTT AAA GGA ACT AAT GAG TCT CTG GAA CGC CAG ATG CGT GAG ATG GAA

tyr arg arg gln val gln ser leu thr cys glu val asp ala leu lys gly thr asn glu ser leu glu arg gln met arg glu met glu

310 320

1080

GAG AAT TTT GCC CTT GAA GCT GCT AAC TAC CAA GAC ACT ATT GGC CGC CTG CAG GAT GAG ATT CAG AAC ATG AAG GAA GAG ATG GCT CGT

glu asn phe ala leu glu ala ala asn tyr gln asp thr ile gly arg leu gln asp glu ile gln asn met lys glu glu met ala arg

340 350

1170

CAC CTT CGT GAA TAC CAA GAC CTG CTC AAT GTC AAG ATG GCT CTT GAC ATT GAG ATA GCC ACC TAC AGG AAG CTA CTG GAA GGC GAG GAG

his leu arg glu tyr gln asp leu leu asn val lys met ala leu asp ile glu ile ala thr tyr arg lys leu leu glu gly glu glu

370 380

1260

AGC AGG ATT TCT CTG CCT CTT CCC AAC TTT TCT TCC CTG AAC CTG AGA GAA ACT AAT CTG GAG TCA CTC CCT CTG GTT GAC ACC CAC TCA

ser arg ile ser leu pro leu pro asn phe ser ser leu asn leu arg glu thr asn leu glu ser leu pro leu val asp thr his ser

400 410

1350

AAA AGA ACA CTC CTG ATT AAG ACA GTG GAA ACT AGG GAT GGA CAG GTG ATC AAT GAA ACC TCT CAG CAT CAT GAT GAC CTT GAA TAA AAA

lys arg thr leu leu ile lys thr val glu thr arg asp gly gln val ile asn glu thr ser gln his his asp asp leu glu

430 440

1450

TTGCACATAC TCTGTGCAAC AACGCAGTCC AGCAAGAAGA AAAAAAAGAA ATTCGTATCT TAAGGAAACA GCTTTCAAGT GCCTTTACTG CAGTTTTTCA

1550

GGAGCGCAAG TAAGATTTGG GATAGAAATA AGCTCTAGTT TCTAACAACT GACACCCTAA AAGATTTAGA AAAGGTTTAC AACACAATCT AGTTTACGAA

GAAATCTTGT GCTAGAATAC TTTTCAAAGT ATTTTGAATA CCATTAAACG CTTTCCCAGT ATACGACCAA CTGACGCTTA TA

FIG. 2. Complete combined nucleotide sequence of the inserts of pVim-1 and pVim-2 and the predicted amino acid sequence of hamster lens vimentin.

and 344 and a more pronounced one of about 45 amino acids, between residues 217 and 263, with the "helix breaker" proline. These results are similar to the domain structure found by chymotryptic cleavage of chicken desmin (9). Apart from the application of the heptade convention, we used the rules for secondary structure prediction according to Chou and Fasman (20) (data not shown) and estimated the length of helical and nonhelical domains: a non-α-helical COOH-terminal domain of 55 residues; two helices of 145 and 135 residues, respectively, separated by a nonhelical insertion of 45 amino acids; and a nonhelical NH_2-terminal domain of at least 68 residues.

The short interruption found in each helix-domain (at positions 131 and 344) are required to let the hydrophobic backbone run down in a straight manner. The amino acids around these interruptions possibly form a loop outside the helix, as was previously proposed for the three-stranded rope of influ-

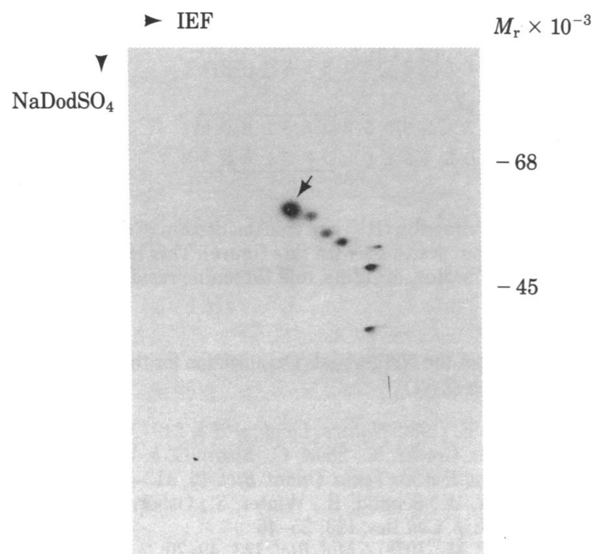

FIG. 3. Autoradiograph of translation products derived from mRNA hybridizing to pVim-1 (13). Arrow, vimentin; arrowheads, direction of electrophoretic migration under isoelectrofocusing (IEF) and in the presence of $NaDodSO_4$, respectively. A ladder of breakdown products towards the acidic side of the gel can be seen, possibly as a result of the action of a specific protease (21) at the arginine-rich NH_2-terminal part.

enza hemagglutinine (28) and for desmin (9). Such a loop structure would be exposed to the environment and is a possible site for an antigenic determinant. Following the method for locating protein antigenic determinants by analyzing amino acid sequences in order to find the point of greatest local hydrophilicity (29), we found an antigenic determinant around the interruption of the NH_2-terminal helix domain (data not shown). This may support the suggestion of a loop structure at the interruption positions.

Comparison of the Structures of Hamster Vimentin and Chicken Desmin. *Nonhelical NH_2-terminal domain.* The NH_2-terminal domains of vimentin and desmin show little similarity at the sequence level of the amino acids (Fig. 5). The residues 67–74 of vimentin, which are all homologous to residues 81–88 of chicken desmin, constitute a notable exception. These residues of vimentin form the first turns of the helical domain. Obviously the homology between vimentin and desmin starts at the position where the nonhelical domain changes into the helical configuration. The amino acid composition, however, of the nonhelical NH_2-terminal domain is very striking and has great resemblance with the composition of the 7.5-kDa domain of chicken desmin. In the NH_2-terminal region of vimentin, common amino acids like aspartic acid, isoleucine, lysine, histidine, and leucine are missing or appear just once, but this region is very rich in serine, threonine, and arginine (25%, 11%, and 14.5%, respectively) just like the 7.5-kDa headpiece of chicken desmin (24.6%, 14.5%, and 14.5%, respectively). The NH_2-terminal region of vimentin contains only one acidic residue and, because of the presence of eight arginine residues, it is very basic. The same holds true for the NH_2-terminal domain of chicken desmin, which contains 10 arginine residues. Perhaps the amino acid composition and the basic character are more important for the function and properties of the NH_2 terminus [for example the weak affinity of vimentin to ribosomes and RNA (9, 30)] than is the sequence of the amino acids. The low degree of sequence homology between the nonhelical headpiece of hamster vimentin and chicken desmin is in contrast with the 66% homology of the COOH-terminal 200 residues (refs. 9, 11, and 12 and this study).

COOH-terminal domain. The last 55 residues at the COOH terminus of hamster vimentin form a nonhelical tailpiece. This domain has about the same length as the nonhelical COOH-terminal part of chicken desmin, but the homology at the sequence level is only 42% instead of the 66% homology between vimentin and desmin (9, 11, 12).

The helical domain. Two domains of almost similar size (145 and 130 residues) could easily fit an α-helical arrangement according to the heptade convention with a nonhelical insertion of 46 amino acids. As mentioned before, there are two interruptions, one in each helix domain, around positions 131 and 344, the latter occurring at the same position as in the helix region of chicken desmin (9). In view of the domain structure of vimentin, particularly the two helical domains, we searched for duplications in the nucleotide or amino acid sequence using the matrix method of Gibbs and McIntyre (31). However, because no duplications could be detected, we must conclude that the domain-structure is not the result of a duplication at the genomic level.

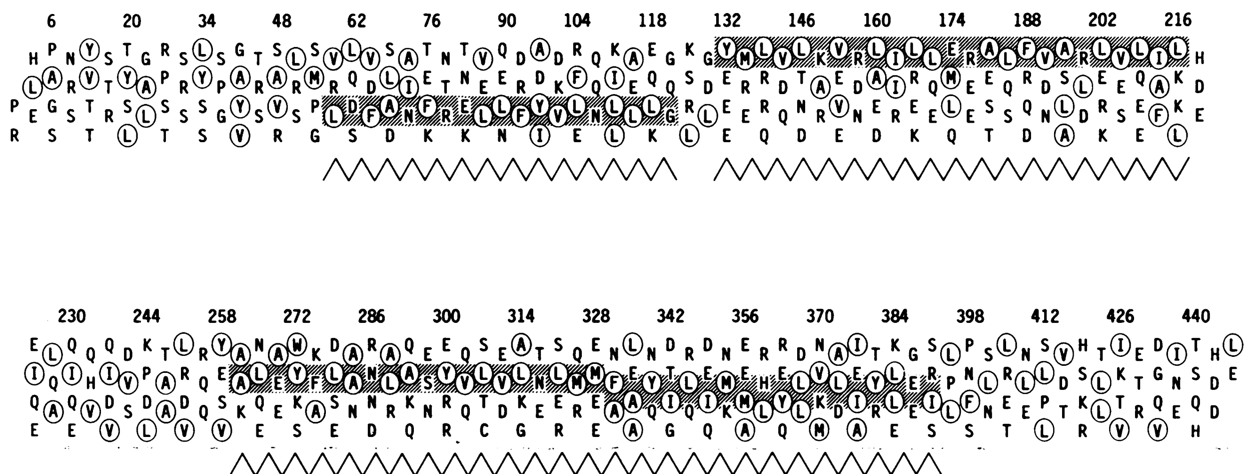

FIG. 4. Amino acid sequence of vimentin drawn in a helical net with 3.5 residues per turn. Nonpolar amino acids [alanine (A), valine (V), methionine (M), isoleucine (I), leucine (L), tyrosine (Y), phenylalanine (F), and tryptophan (W)] are encircled. Note that these hydrophobic residues are aligned in a major part of the polypeptide. Proposed helical regions are shaded and indicated by zig-zag lines. The helical arrangement is probably stabilized by the many salt-bridges between basic [lysine (K) and arginine (R)] and acidic [aspartic acid (D) and glutamic acid (E)] residues. Two distinct helical domains can be distinguished separated by a nonhelical proline (P)-containing region of about 45 amino acids.

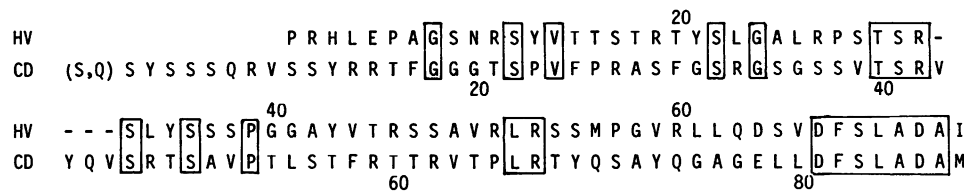

FIG. 5. Comparison of the amino acid sequences of the NH_2-terminal domains of hamster vimentin (HV) and chicken desmin (CD) (9). Identical residues in vimentin and desmin are placed in boxes. Sequence homology appears clearly from position 67 (in this figure). This is the site where the helical configuration starts in the vimentin polypeptide. Note the striking abundance of serine, arginine, and threonine residues in the non-homologous part of both proteins.

Previous experiments with the isolated α-helical 38-kDa domain of chicken desmin (9) have shown the ability of this part of the polypeptide chain to form higher aggregates and long twisted ribbons with a variable diameter (10–20 nm), indicating that this helical part plays an important role in 10-nm filament formation. Another argument for the important role of the helical part is the fact that there is a sequence homology of more than 80% between the helix domain of hamster vimentin and the known part of chicken desmin. On the other hand, the COOH-terminal regions of vimentin and desmin show only 42% homology and the NH_2-terminal domains have no significant homology at all at the sequence level. Probably evolutionary conservation of the primary structure of the helical part of the IF proteins is important for filament formation. Our findings of the similarity of the topographical models of vimentin and desmin explains the ability of vimentin to copolymerize with desmin (32).

CONCLUDING REMARKS

Our model of the vimentin molecule derived from the primary structure shows two distinct regions that fit a helical conformation. The helical arrangement is stabilized by salt bridges between basic (lysine and arginine) and acidic (glutamic acid and aspartic acid) residues (Fig. 4).

Moreover, we postulated an antigenic determinant in the NH_2-terminal α-helical domain at the position with the highest hydrophilicity. The primary structure of that part of other IF proteins is still unknown. It may well be possible that differences in the structure of that part of IF proteins are responsible for the differences in immunological properties.

The meaning of the basic character of the NH_2 terminus of vimentin is not quite clear yet. Because both vimentin and desmin conserve the same composition of amino acids and the same distribution of basic residues, it might be a necessary feature, probably for all IF proteins. The NH_2-terminal part could play a role in the attachment to the membrane, for it displays a more hydrophobic character than the rest of the molecule, or it may be involved in the head-to-tail aggregation of IF proteins as long filaments. In this context it is noteworthy that the COOH terminus of vimentin contains an excess of acidic residues.

Note Added in Proof. After submission of this paper, the amino acid sequence of desmin was reported (33).

We thank Johan den Dunnen and Ben Dekker for assistance with the nucleotide sequence analysis and Jack Leunissen for help with computer analyses of vimentin nucleotide and amino acid sequences. The present investigations have been carried out partly under the auspices of the Netherlands Foundation for Chemical Research (SON) and with the financial aid of the Netherlands Organization for the Advancement of Pure Research (ZWO).

1. Lazarides, E. (1980) *Nature (London)* **283,** 249–256.
2. Osborn, M., Geisler, N., Shaw, G., Sharp, G. & Weber, K. (1982) *Cold Spring Harbor Symp. Quant. Biol.* **46,** 413–429.
3. Franke, W. W., Schmid, E., Winter, S., Osborn, M. & Weber, K. (1979) *Exp. Cell Res.* **123,** 25–46.
4. Steinert, P. M. (1978) *J. Mol. Biol.* **123,** 49–70.
5. Steinert, P. M., Idler, W. W. & Goldman, R. D. (1980) *Proc. Natl. Acad. Sci. USA* **77,** 4534–4538.
6. Benedetti, E. L., Dunia, I., Ramaekers, F. C. S. & Kibbelaar, M. A. (1981) in *Molecular and Cellular Biology of the Eye Lens,* ed. Bloemendal, H. (Wiley, New York), pp. 137–188.
7. Steinert, P. M., Zimmermann, S. B., Starger, J. M. & Goldman, R. D. (1978) *Proc. Natl. Acad. Sci. USA* **75,** 6098–6101.
8. Renner, W., Franke, W. W., Schmid, E., Geisler, N., Weber, K. & Mandelkow, E. (1981) *J. Mol. Biol.* **149,** 285–306.
9. Geisler, N., Kaufmann, E. & Weber, K. (1982) *Cell* **30,** 277–286.
10. Geisler, N. & Weber, K. (1981) *FEBS Lett.* **125,** 253–256.
11. Geisler, N. & Weber, K. (1981) *Proc. Natl. Acad. Sci. USA* **79,** 4120–4123.
12. Geisler, N., Plessmann, U. & Weber, K. (1982) *Nature (London)* **296,** 448–450.
13. Dodemont, H. J., Soriano, P., Quax, W. J., Ramaekers, F., Lenstra, J. A., Groenen, M. A. M., Bernardi, G. & Bloemendal, H. (1982) *EMBO J.* **1,** 167–171.
14. Grunstein, M. & Hogness, D. S. (1975) *Proc. Natl. Acad. Sci. USA* **72,** 3961–3965.
15. Birnboim, H. C. & Doly, J. (1979) *Nucleic Acids Res.* **7,** 1513–1523.
16. Maxam, A. M. & Gilbert, W. (1980) *Methods Enzymol.* **65,** 499–560.
17. Smith, M., Leung, D. W., Gillam, S., Astell, C. R., Montgomery, D. L. & Hall, B. D. (1979) *Cell* **16,** 753–761.
18. Staden, R. (1977) *Nucleic Acids Res.* **4,** 4037–4051.
19. MacLachlan, A. D. & Stewart, M. (1975) *J. Mol. Biol.* **98,** 293–304.
20. Chou, P. Y. & Fasman, G. D. (1978) *Annu. Rev. Biochem.* **47,** 251–276.
21. Nelson, W. J. & Traub, P. (1981) *Eur. J. Biochem.* **116,** 51–57.
22. Granger, B. L. & Lazarides, E. (1979) *Cell* **18,** 1053–1063.
23. Proudfoot, N. J. & Brownlee, G. G. (1976) *Nature (London)* **263,** 211–214.
24. Fitzgerald, M. & Shenk, T. (1981) *Cell* **24,** 251–260.
25. Franke, W. W., Schmid, E., Osborn, M. & Weber, K. (1978) *Proc. Natl. Acad. Sci. USA* **75,** 5034–5038.
26. Doolittle, R. F., Goldbaum, D. M. & Doolittle, L. R. (1978) *J. Mol. Biol.* **120,** 311–325.
27. MacLachlan, A. D. (1978) *J. Mol. Biol.* **124,** 297–304.
28. Ward, C. W. & Dopheide, T. A. (1980) *J. Biol. Sci.* **33,** 441–447.
29. Hopp, T. P. & Woods, K. R. (1981) *Proc. Natl. Acad. Sci. USA* **78,** 3824–3828.
30. Nelson, W. J. & Traub, P. (1981) *J. Cell Biol.* **91,** 232.
31. Gibbs, A. J. & McIntyre, G. A. (1970) *Eur. J. Biochem.* **16,** 1–11.
32. Steinert, P. M., Idler, W. W., Cabral, F., Gottesman, M. M. & Goldman, R. D. (1981) *Proc. Natl. Acad. Sci. USA* **78,** 3692–3696.
33. Geisler, N. & Weber, K. (1982) *EMBO J.* **1,** 1649–1656.