Volume 302, number 3, 239–242 FEBS 11035 May 1992

Truncation of recombinant vimentin by ompT

Identification of a short motif in the head domain necessary for assembly of type III intermediate filament proteins

Mechthild Hatzfeld, Huub Dodemont, Uwe Plessmann and Klaus Weber

Max Planck Institute for Biophysical Chemistry, Department of Biochemistry, D-3400 Goettingen, Germany

Received 23 March 1992

Recombinant vimentin expressed in *E. coli* JM101 cells is cleaved after cell lysis between arginines 11 and 12. The truncated vimentin is assembly incompetent. Expression of the same cDNA construct in BL21 cells, which lack the protease ompT, provides intact and polymerization-competent vimentin. The ompT cleavage site is contained in a short sequence motif (YRRMF) shared by the head domains of type III and IV intermediate filament (IF) proteins. We propose that a related motif present in the N-terminal 32 residues of λCII acounts for the known IF formation of a fusion protein formed with a truncated GFAP.

Intermediate filament; ompT; Protease; Sequence motif; Vimentin

1. INTRODUCTION

All intermediate filament (IF) proteins share a common subunit organization consisting of a non-α-helical N-terminal head domain, a central α-helical rod domain and a C-terminal tail domain [1–6]. The in vitro assembly process of IF seems to consist of a number of association steps that involve different protein domains. Dimer and tetramer formation depend solely on the α-helical rod domain [7]. Assembly of filaments from tetramers, however, requires the presence of the head domain [7–9], whereas the C-terminal tail domain is not essential for IF assembly in vitro [9–11]. Although the head domains are variable in size and sequence [12] they share certain features. They are positively charged, due to a series of arginine residues, and contain serine residues for phosphorylation by protein kinases A and C and the cdc2 kinase [13–15]. In order to analyze the function of the head domain in IF assembly more precisely we have expressed a vimentin cDNA clone in *E. coli* and purified the recombinant protein. We show that a proteolytic breakdown product of vimentin lacking only the 11 N-terminal residues is assembly incompetent. The truncated vimentin was obtained by specific cleavage by the bacterial protease ompT and occurs after cell lysis in certain bacterial strains.

Correspondence address: M. Hatzfeld, Max Planck Institute for Biophysical Chemistry, PO Box 2841, D-3400 Goettingen, Germany.

2. MATERIALS AND METHODS

2.1. *Cloning and characterization of vimentin cDNA*

A plasmid cDNA library representing total poly(A)$^+$ polysomal RNA from 5-day-old hamster lens was established using hte JM109/pUC18 [16] host/vector system. Procedure for mRNA isolation, cDNA synthesis, library construction and screening conditions have been described [17]. About 60,000 lysed bacterial colonies were hybridized at high stringency with a ^{32}P-labeled, nick-translated probe derived from the 5'-end of the hamster vimentin cDNA, pVim-1 [18,19], which covers amino acid residues 51–107. Of eleven positive cDNA clones obtained, three and inserts longer than 1,800 bp. The identity of the largest cDNA, used for the present study, was verified by sequence analysis [20]. The nearly full-length vimentin cDNA (1,839 bp) starts 13 bp downstream from the presumptive transcription initiation site of the gene [21] and terminates with the first adenylate residue of the poly(A) tail of the corresponding mRNA. For single strand DNA preparation and subsequent site-directed mutagenesis the clone was introduced into M13mp18. Site-directed mutagenesis was performed using a mutagenesis kit (Amersham Buchler, Braunschweig, Germany). A unique *Bam*HI site was introduced in the position coding for amino acids 4 and 5 of the original sequence. A *Hind*III site was introduced 38–43 nucleotides behind the stop codon. The *Bam*HI/*Hind*III fragment was excised and ligated into the prokaryotic expression vector pINDU [22].

2.2. *Purification of recombinant vimentin*

Vimentin was expressed in *E. coli* JM101 and BL21, respectively. Protein purification was essentially as described for recombinant keratins 8 and 18 [23]. Vimentin, highly enriched in the inclusion body preparation, was solubilized in 8.5 M urea, 10 mM Tris-HCl, 5 mM EDTA, 1 mM DTE, pH 8.6, and purified using Mono Q anion-exchange chromatography and ss-DNA affinity chromatography. Fractions were monitored by SDS-PAGE.

2.3. *Protein sequencing*

Total bacterial lysates and partially purified fractions were separated by SDS-PAGE. Separated polypeptides were electroblotted onto a polyvinyldifluoride membrane and visualized with Amido black

staining [24]. The vimentin band was excised and subjected to automated sequencing using a Knauer model 810 sequenator.

2.4. *In vitro reconsititution studies*

Purified vimentin was diluted in 8.5 M urea buffer to a concentration of ~0.5 mg/ml as determined by the method of Bradford [25]. Aliquots of 50 μl were then dialyzed against vimentin filament buffer (10 mM Tris-HCl, pH 7.5, 170 mM NaCl, 1 mM DTE) using dialysis filters (Millipore GmbH, Eschborn, Germany). Structures formed were analyzed after negative staining with 2% uranylacetate as described [10,23,26].

3. RESULTS

A complete hamster vimentin cDNA clone was obtained from poly(A)$^+$ polysomal RNA of 5-day-old hamster lens. Unique restriction sites were introduced by site-directed mutagenesis to facilitate cloning into the expression vector pINDU [22]. Due to this cloning procedure the N-terminal sequence of authentic vimentin MSTRSVSSSSYRRMF) was slightly changed in the expected recombinant protein (MRGSVSSSSYRRMF). Expression of the cDNA clone in *E. coli* JM101 cells provided inclusion bodies highly enriched in vimentin. After solubilization in 8.5 M urea, pure vimentin was obtained by Mono Q anion- exchange chromatography and ss-DNA affinity chromatography (Fig. 1, lanes 5–7). In vitro assembly studies of JM101-vimentin revealed primarily small globular aggregates (Fig. 2b) and only very rarely short and irregularly packed IF fragments. Stepwise reconstitution by dialysis, first against 5 mM Tris-HCl, pH 8.4, and then against 10 mM Tris-HCl, pH 7.5, containing 170 mM NaCl, provided again the globular aggregates as dominant structures. Occasionally a field containing a few longer filaments was seen (Fig. 2c) but these structures had an uneven surface and were interrupted by globular aggregates.

To understand this unexpected inability of recombinant JM101 vimentin to form natural IF the purified protein was characterized by N-terminal sequencing. This showed that at least the majority of the material lacked the predicted 11 N-terminal residues and gave instead the sequence RMFGG starting at arginine 12. Blotting of the total *E. coli* extract followed by sequencing of the JM101 vimentin band provided the predicted N-terminal sequence. Thus the lack of IF forming ability of the purified protein seemed to be due to a protease cleaving the bond between arginines 11 and 12, which is contained in a sequence motif shared by type III and IV IF proteins [6]. Since an independent study on recombinant murine lamin C from XL 1-blue cells also provided a truncated protein resulting from a cleavage between two arginine residues (our unpublished results with S. Stöhr) we searched the literature for an *E. coli* protease acting on a dibasic sequence. This specificity is fulfilled by ompT, an outer membrane protease of most *E. coli* strains which is poorly inhibited by PMSF [27–29]. Since certain outer membrane proteins, such as

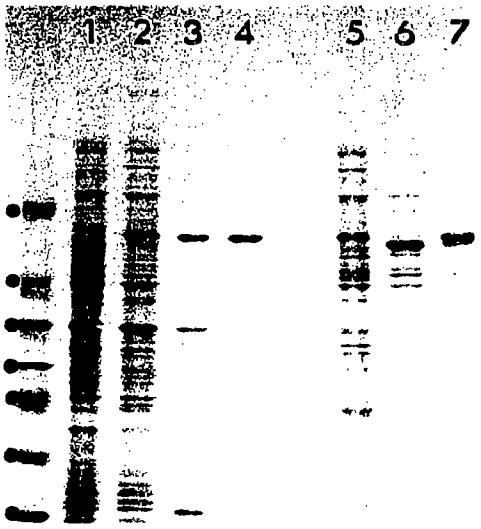

Fig. 1. SDS-polyacrylamide gel electrophoresis demonstrating the purification of recombinant vimentin from *E. coli* BL21 (lanes 1–4) and *E. coli* JM101 (lanes 5–7). Reference proteins are: BSA (66,000), ovalbumin (45,000), glyceraldehyde-3-phosphate-dehydrogenase (36,000), carbonic anhydrase (29,000), trypsinogen (24,000), trypsin inhibitor (20,100) and α-lactalbumin (14,200). Lane 1, total extract from normal BL21 cells; lane 2, total extract from BL21 cells expressing hamster vimentin; lane 3, preparation of vimentin containing inclusion bodies from a BL21 bacterial culture; lane 4, peak fraction from ss-DNA chromatography containing essentially pure vimentin; lane 5, total extract from JM101 bacteria expressing vimentin; lane 6, inclusion body preparation from the same culture as in lane 5; lane 7, JM101–vimentin in peak fraction from ss-DNA chromatography.

ompA ompF, pHoE [30] and ompC (our unpublished results with S. Stöhr) are also present in inclusion body preparations, a contamination by ompT seemed likely. Therefore we expressed the vimentin construct in BL21 cells, which lack ompT [28]. Recombinant vimentin was well expressed and readily purified as described above for vimentin from JM101 cells (Fig. 1, lanes 2–4). The pure BL21 vimentin showed the correct N-terminal sequence and readily assembled into regular IF using standard reconstitution conditions (Fig. 2a).

4. DISCUSSION

We have shown that, starting with the same cDNA construct and using the same purification procedure, vimentin is obtained in an assembly-incompetent form from *E. coli* JM101 cells, while BL21 cells provide assembly-competent protein. Amino-terminal sequencing showed that vimentin is made in JM101 cells with the correct sequence but that truncation of the first 11 residues occurs after the cells are lysed. The Arg–Arg bond opened is a known target for the outer membrane protease ompT, which is not expressed in BL21 cells [28]. The value of ompT$^-$ strains such as BL21 and B824 has been recognized in purification of T7 RNA polymerase [28] and the ω subunit of *E. coli* RNA polymerase [31]. Many recombinant proteins expressed in *E. coli* do not

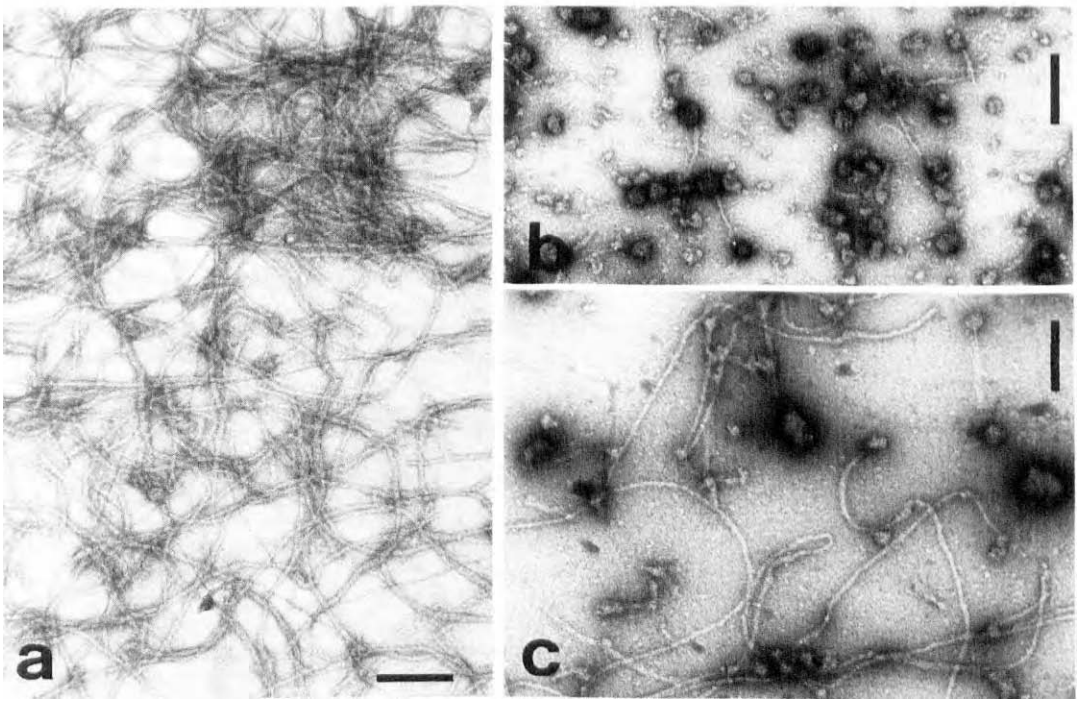

Fig. 2. Electron micrographs of structures assembled from recombinant vimentin in standard vimentin filament buffer (10 mM Tris-HCl, pH 7.5, 170 mM NaCl, 1 mM DTE). (a) Typical IF obtained after dialysis of vimentin purified from BL21 cells. (b) Globular aggregates obtained after dialysis of vimentin purified from JM101 cells. Only very few IF-like fragments were detected. (c) Structures detected after dialysis of vimentin from JM101 cells first against 5 mM Tris-HCl, pH 8.4, and then against filament buffer. Aside from some globular aggregates only short and irregular IF were formed. Bars=0.2 μm.

contain cleavage sites readily accessible to ompT (see for instance [10,23]). Since vimentin displays several dibasic sequences in addition to the Arg–Arg bond at residues 11 and 12 the specific cleavage of the latter bond in JM101 extracts indicates that the head domain is not protected against ompT. Similarly, in recombinant murine lamin C, an Arg–Arg bond is opened in the head domain. The inaccessibility of dibasic sequences in the rod domains possibly suggests a proper coiled-coil formation of these proteins in the inclusion bodies.

The function of the head domain in IF assembly has been known for several years. Proteolytic removal of the desmin head domain rendered the resulting fragment incompetent for in vitro assembly [7]. cDNA cloning and expression allowed more subtle modifications. Deletion mutants of keratins were assembly incompetent when expressed in eukaryotic cells [11], although co-polymerization with an intact complementary keratin was observed in some cases ([22,33] see, however, [11]). More subtle deletion analysis using a desmin cDNA clone revealed that removal of amino acids 5–17 was sufficient to interfere with desmin IF network formation in eukaryotic cells [8], although co-polymerization with vimentin of this mutant as well as mutants containing larger deletions was still observed [8]. Although immunofluorescence analysis of transfected cells clearly demonstrates that the deletion mutant of desmin had an effect on assembly, it cannot distinguish between formation of modified IF that aggregate into 'dots' in vivo and complete inhibition of IF formation.

The 11 N-terminal residues absent from the assembly incompetent JM101 vimentin emphasize the sequence motif S(S,A)(Y,A)RRXF, which is conserved in the head domain of most type III and type IV IF proteins [6]. Given the functional importance of this short basic motif, two problems are posed by mouse glial fibrillary acidic protein (GFAP). First, the sequence of this type III IF protein deduced from a genomic DNA has an unusually small head domain and lacks the motif [34] which is, however, present in the longer head domain of human GFAP [37] and porcine GFAP [6]. Second, a fusion protein containing the first 32 amino acids of the λCII protein coupled to the information contained in a murine cDNA provides a recombinant GFAP protein able to form IF [35]. Inspection of the short CII sequence [36] shows that the N-terminal 7 residues are MVRANKR and thus are a variation of the dibasic motif. Therefore we also checked the published DNA sequence for murine GFAP [34]. The motif is clearly present upstream of the presumptive initiation codon (Fig. 3). Choice of the first upstream methionine codon and assumption of a small sequence mistake to correct for the reading frame predicts a head domain containing the motif sequence. The corrected murine GFAP sequence is highly homologous to the corresponding human GFAP (Fig. 3) sequence.

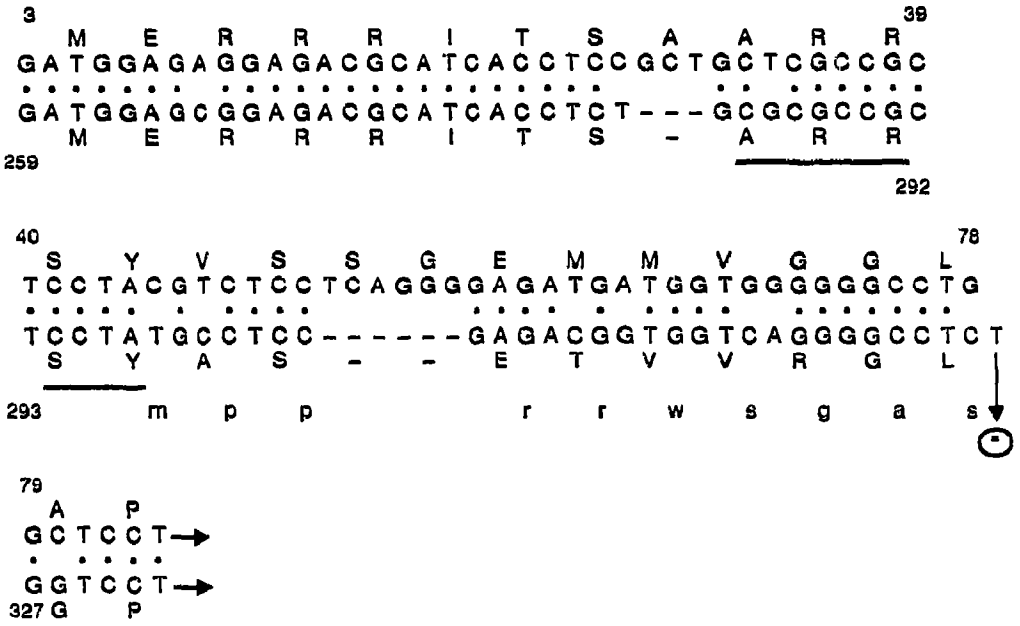

Fig. 3. Alignment of the published nucleotide and amino acid sequences of human (upper lane) and murine (lower lane) GFAP [34,37]. Dots denote identical nucleotides. Lower case letters in the lower lane represent the published amino acid sequence of mouse GFAP; upper case letters represent the corrected amino acid sequence prediction. Note the high homology between the human sequence and the corrected murine sequence. The arrow in nucleotide position 326 of mouse GFAP [34] indicates the nucleotide that has to be removed in order to correct the reading frame. The sequence motif ARRSY that is conserved in the head domain of type III and type IV IF proteins is underlined in the corrected murine sequence.

REFERENCES

[1] Geisler, N. and Weber, W. (1982) EMBO J. 1, 1649–1656.
[2] Geisler, N., Kaufmann, E. and Weber, K. (1982) Cell 30, 277–286.
[3] Steinert, P.M., Rice, R.H., Roop, D.R., Trus, B.L. and Steven, A.C. (1983) Nature 302, 794–800.
[4] Steinert, P.M., Steven, A.C. and Roop, D.R. (1985) Cell 42, 411–419.
[5] Hanukoglu, I. and Fuchs, E. (1983) Cell 33, 915–924.
[6] Weber, K. and Geisler, N. (1986) in: Cell and Molecular Biology of the Cytoskeleton (Shay, J.W., ed.) pp. 41–68, Plenum Press, New York.
[7] Kaufmann, E., Weber, K. and Geisler, N. (1985) J. Mol. Biol. 185, 733–742.
[8] Raats, J.M.H., Pieper, F.R., Vree Egberts, W.T.M., Verrijp, K.N., Ramaekers, F.C.S. and Bloemendal, H. (1990) J. Cell Biol. 111, 1971–1985.
[9] Traub, P. and Vorgias, C.E. (1983) J. Cell Sci. 63, 43–67.
[10] Hatzfeld, M. and Weber, K. (1990) J. Cell Sci. 97, 317–324.
[11] Bader, B.L., Magin, T.M., Freudemann, M., Stumpp, S. and Franke, W.W. (1991) J. Cell Biol. 115, 1293–1307.
[12] Steinert, P.M. and Roop, D.R. (1988) Annu. Rev. Biochem. 57, 593–625.
[13] Inagaki, M., Gonda, Y., Matsuyama, M., Nishizawa, K., Nishi, Y. and Sato, C. (1988) J. Biol. Chem. 263, 5970–5978.
[14] Geisler, N., Hatzfeld, M. and Weber, K. (1989) Eur. J. Biochem. 183, 441–447.
[15] Peter, M., Nakagawa, J., Dorée, M., Labbé, J.C. and Nigg, E.A. (1990) Cell 61, 591–602.
[16] Yanisch-Perron, C., Vieira, J. and Messing, J. (1985) Gene 33, 103–119.
[17] Dodemont, H., Riemer, D. and Weber, K. (1990) EMBO J. 9, 4083–4094.
[18] Dodemont, H.J., Soriano, P., Quax, W.J., Ramaekers, F., Lenstra, J.A., Groenen, M.A.M., Bernardi, G. and Bloemendal, H. (1982) EMBO J. 1, 167–171.
[19] Quax-Jeuken, Y., Quax, W. and Bloemendal, H. (1983) Proc. Natl. Acad. Sci. USA 80, 3548–3552.
[20] Sanger, F., Nicklen, S. and Coulson, A.R. (1977) Proc. Natl. Acad. Sci. USA 74, 5463–5467.
[21] Quax, W., Vree Egberts, W., Hendriks, W., Quax-Jeuken, Y. and Bloemendal, H. (1983) Cell 35, 215–223.
[22] Bujard, H., Gentz, R., Lanzer, M., Stueber, D., Mueller, M., Ibrahimi, I., Haeuptle, M.-T. and Dobberstein, B. (1987) Methods Enzymol. 155, 416–433.
[23] Hatzfeld, M. and Weber, K. (1990) J. Cell Biol. 110, 1199–1210.
[24] Bauw, G., De Loose, M., Inze, D., Van Montagu, M. and Vandekerckhove, J. (1987) Proc. Natl. Acad. Sci. USA 84, 4806–4810.
[25] Bradford, M.M. (1976) Anal. Biochem. 72, 248–254.
[26] Hatzfeld, M. and Franke, W.W. (1985) J. Cell Biol. 101, 1826–1841.
[27] Rupprecht, K.R., Gordon, G., Lundrigan, M., Gayda, R.C., Markovitz, A. and Earhart, C. (1983) J. Bacteriol. 153, 1104–1106.
[28] Grodberg, J. and Dunn, J.J. (1988) J. Bacteriol. 170, 1245–1253.
[29] Sugimura, K. and Nishihara, T. (1988) J. Bacteriol. 170, 5625–5632.
[30] Veeraragavan, K. (1989) FEMS Microbiol. Lett. 61, 149–152.
[31] Gentry, D.R. and Burgess, R.R. (1990) Protein Expression Purification 1, 81–86.
[32] Lu, X. and Lane, E.B. (1990) Cell 62, 681–696.
[33] Albers, K. and Fuchs, E. (1989) J. Cell Biol. 108, 1477–1493.
[34] Lewis, S.A., Balcarek, J.M., Krek, V., Shelanski, M. and Cowan, N.J. (1984) Proc. Natl. Acad. Sci. USA 81, 2743–2746.
[35] Quinlan, R.A., Moir, R.D. and Stewart, M. (1989) J. Cell Sci. 93, 71–83.
[36] Schwarz, E., Scherer, G., Hobom, G. and Kössel, H. (1978) Nature 272, 410–414.
[37] Reeves, S.A., Helman, L.J., Allison, A. and Israel, M.A. (1989) Proc. Natl. Acad. Sci. USA 86, 5178–5182.