Volume 317, number 3, 189–194 FEBS 12091 February 1993

Variable and constant regions in the C-terminus of vinculin and metavinculin

Cloning and expression of fragments in *E. coli*

Peter Strasser[a], Mario Gimona[a], Monika Herzog[a], Benjamin Geiger[b] and J. Victor Small[a]

[a]*Institute of Molecular Biology, Austrian Academy of Sciences, Billrothstrasse 11, 5020 Salzburg, Austria* and [b]*Weizmann Institute of Science, Dept. Chemical Immunology, 76100 Rehovot, Israel*

Received 15 December 1992

Metavinculin differs from vinculin in having an additional insert of 68 to 79 amino acids in length in the C-terminal half of the molecule. Cross-species comparison of metavinculin sequences from pig, man, chicken and frog reveals a division of the insert into two parts: the first variable and the second highly conserved. The longest insert, 79 amino acids, was found in *Xenopus laevis*. Three different C-terminal constructs of vinculin and metavinculin over-expressed in *E. coli* could be purified by column chromatography. Two-dimensional gel electrophoresis and peptide analysis revealed pI values between 8.35 and 10.25 for the recombinant proteins. Biochemical and structural features of the metavinculin-specific sequence and the conserved vinculin/metavinculin carboxy-terminus are discussed.

Vinculin; Metavinculin; C-terminus; Sequence; *Xenopus*; *E. coli* expression

1. INTRODUCTION

Anchorage of the actin cytoskeleton to the cell membrane in cell contact areas is mediated by a cascade of interacting cytoplasmic and transmembrane proteins [1–3]. The biochemical composition of this complex shows some tissue variability and the number of newly identified proteins associated with it is rapidly increasing [4–6]. One protein ubiquitously found in actin-associated cell–matrix and cell–cell contacts is the cytoplasmic 117 kDa protein vinculin [7]. In smooth muscle cells an additional 124 kDa form, named metavinculin, is expressed [8,9]. Recently it has been shown that vinculin and metavinculin arise from a single gene by alternative splicing of a single exon coding for a 68 (mammalian) or 69 (avian) amino acid stretch inserted in the carboxy-terminal third of the molecule [10,11]. Structure–function analysis revealed that vinculin contains several binding sites involved in its association with cell adhesions [12]. Jones et al. [13] showed that the amino-terminal region of vinculin contains a binding site for talin [14], another member of the anchorage cascade. Recently Turner and colleagues [4] showed that paxillin, a 68 kDa focal contact protein, binds in vitro to the C-terminal 28 kDa portion of vinculin. Moreover, constructs comprising either the N- or the C-terminal regions of the molecule that were introduced into cells via transfection with partial cDNAs localized specifically to focal contacts [15]. The functional aspects of the metavinculin-specific sequence is, however, still unclear. The aim of the present work was to compare the metavinculin specific sequences from a variety of different species in an attempt to identify common structural motifs. In addition, constructs of different portions of the vinculin/metavinculin molecule have been expressed in bacteria and purified by column chromatograhy. The finding that the loss of the C-terminal half of vinculin leads to unstable constructions points to a possible role of this region in stabilizing the protein.

Correspondence address: P. Strasser, Institute of Molecular Biology, Austrian Academy of Sciences, Billrothstraße 11, A-5020 Salzburg, Austria. Fax: (43) (662) 249 6129.

Abbreviations: PCR, polymerase chain reaction; FPLC, fast-performance liquid chromatography; IPTG, isopropyl-β-D-thiogalactopyranoside; *E. coli*, *Escherichia coli*.

2. MATERIALS AND METHODS

2.1. *Bacterial strains, plasmids and enzymes*

E. coli DH 5α (GIBCO BRL, USA) was used as a standard host and the plasmid pBluescript (Stratagene) as a standard vector for cloning and sequencing of DNA fragments. The plasmid pMW 172 [16] and *E. coli* strain BL21 (DE3) [17] were a kind gift from Dr. Michael Way (MRC Cambridge, UK). All enzymes used were purchased from Boehringer Mannheim and New England Biolabs and used according to the manufacturer's recommendations.

2.2. *Cloning, sequencing and expression of metavinculin-specific fragments*

RNA was isolated from porcine stomach and chicken gizzard smooth muscle according to Chirgwin et al. [18] and subsequently

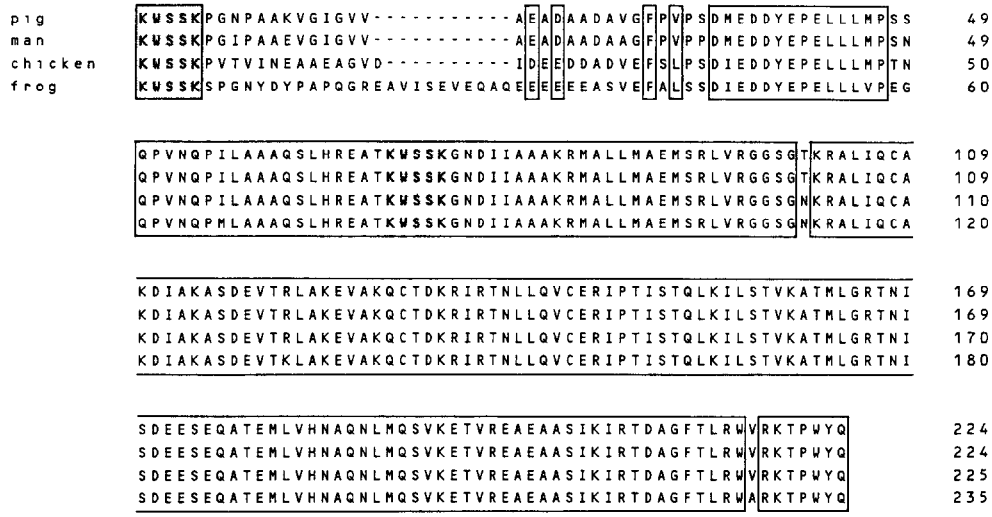

Fig. 1. Cross-species comparison of the C-terminal region of the metavinculin molecule including the metavinculin-specific sequences reveals high homology (boxed) throughout the C-terminus outside the metavinculin-specific piece. The metavinculin-specific sequences show a striking separation into a variable and a conserved domain (open and boxed regions). The grey areas indicate the KWSSK motifs flanking the metavinculin inserts. Sequences for chicken and human vinculin and metavinculin were taken from [10,11,21,29]. The nucleotide sequence encoding the *Xenopus* fragment was reconstructed from metavinculin-specific (PCR) and vinculin-specific (cDNA) sequences.

purified by oligo(dT)-cellulose affinity chromatography (Pharmacia, Sweden). Poly A^+ RNA from *Xenopus laevis* intestinal tract was kindly provided by Mr. C. Wechselberger (Chemistry Dept., Institute of Molecular Biology, Salzburg). Amplification, cloning and sequencing of chicken and porcine metavinculin-specific fragments have been reported [10]. PCR-generated clones of porcine C-terminal vinculin- and metavinculin-specific cDNA fragments were obtained using degenerated primers according to the upstream amino acids **KDEEFPE** [19] and the C-terminal sequence **VRKTPWYQ** [20,21]. For designing specific PCR primers for amplification of the amphibian metavinculin insert fragment, the C-terminal region of a *Xenopus* vinculin cDNA clone was sequenced starting upstream of the KWSSK motif and extending across the stop codon. The two PCR primers were derived from the cDNA sequence flanking the KWSSK motif. Forty cycles of PCR amplification were performed as recommended by United States Biochemicals Corp. (USB). PCR fragments of correct size were eluted from agarose gels and cloned into linearized pBluescript vectors. DNA sequencing was carried out using the Sequenase version 2.0 DNA sequencing kit (USB Corp.). Each fragment described was sequenced from 2 independent PCR amplifications.

Fragments were cloned into the expression vector pMW 172 and expressed in *E. coli* BL21 (DE3) as described [16]. All constructs were sequenced prior to expression.

2.3. *Electrophoresis*

Analytical gel electrophoresis and second dimension electrophoresis were carried out on 8–22% gradient acrylamide minislab gels according to the procedure of Matsudaira and Burgess [22] in the buffer system of Laemmli [23]. Two-dimensional gels were run according to the method of O'Farrell [24] with modifications as described elsewhere [25]. Gels were stained in Coomassie brilliant blue R-250 and destained in 10% acetic acid.

2.4. *Purification of recombinant polypeptides*

500-ml *E. coli* cultures expressing the various constructs were centrifuged for 10 min at 7,000 rpm in a Sorvall RC-5 refrigerated centrifuge using a GS-A rotor and the pellets were suspended in 20 ml final volume of QS buffer (20 mM KCl, 20 mM imidazole, 0.1 mM EGTA, 0.1 mM EDTA, pH 7.3) on ice. Cells were lysed by passing the suspension through a french press (LSM AMINCO, USA) at 1,000 bar and the cell debris was removed by centrifugation in an SS-34 rotor at 18,000 rpm for 15 min. The supernatant was immediately applied onto a Q-Sepharose anion exchanger column (Pharmacia, Sweden) and bound proteins were eluted with a linear gradient ranging from 20 to 300 mM KCl. Vinculin-containing fractions were pooled and further purified on a Sephacryl S-300 gel filtration column equilibrated in QS-buffer (for MW-2 and MW-13) or Sephacryl S-100 (for MW-13c). Fractions enriched in the respective vinculin- and metavinculin fragments were pooled and purified by gel filtration on an FPLC Superose 6 column equilibrated in QS buffer. All gel filtration media used were from Pharmacia (Sweden).

2.5. *Structural and biophysical analysis*

The pI values and molecular weights were calculated using the UW gcg 7 program PEPSTRUCTURE (University of Wisconsin, USA). Predictions of α-helix, β-turn, β-sheet or random coil secondary structures were made using the Garnier–Osguthorp–Robson algorithms [26].

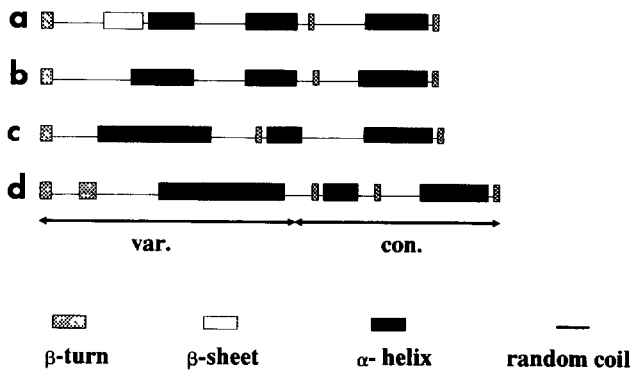

Fig. 2. Comparison of secondary structure predictions according to Garnier–Osguthorpe–Robson analysis of the metavinculin inserts reveals structural similarity in the conserved region (con.) but divergence in the variable portion (var.). a, porcine; b, human; c, chicken and d, *Xenopus*.

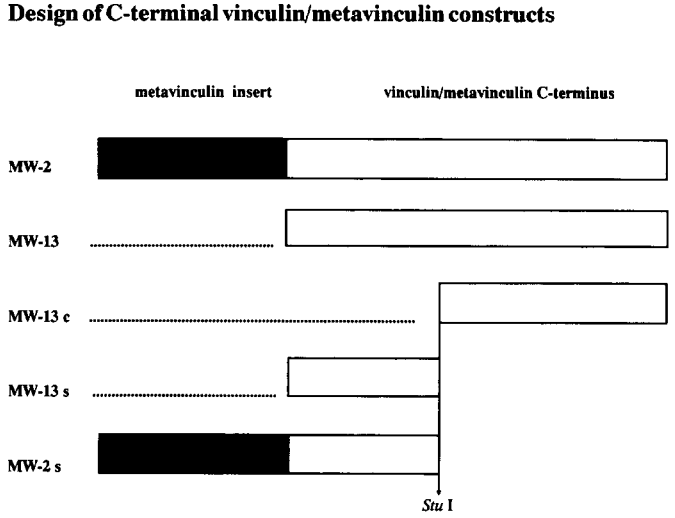

Fig. 3. MW-2, MW-13 and MW-13c: constructs encoding different segments of the vinculin/metavinculin C terminus that were expressed in *E. coli*. Constructs MW-13s and MW-2s, both carrying a stop codon at the *Stu*I site could not be expressed in *E. coli*.

3. RESULTS AND DISCUSSION

3.1. *Inter-species comparison of vinculin and metavinculin sequences*

Koteliansky et al. [10] and Byrne et al. [11] have shown that metavinculin and vinculin arise by alternative splicing of a single exon encoding the metavinculin-specific insert in chicken and human. The splice sites were identified at amino acid positions 915–916 of the vinculin sequence. The data shown here for the porcine and the *Xenopus* cDNAs are in agreement with these findings. For both of the latter species the nucleotide sequences encoding the single **KWSSK** motif of vinculin and the first upstream **KWSSK** motif in metavinculin

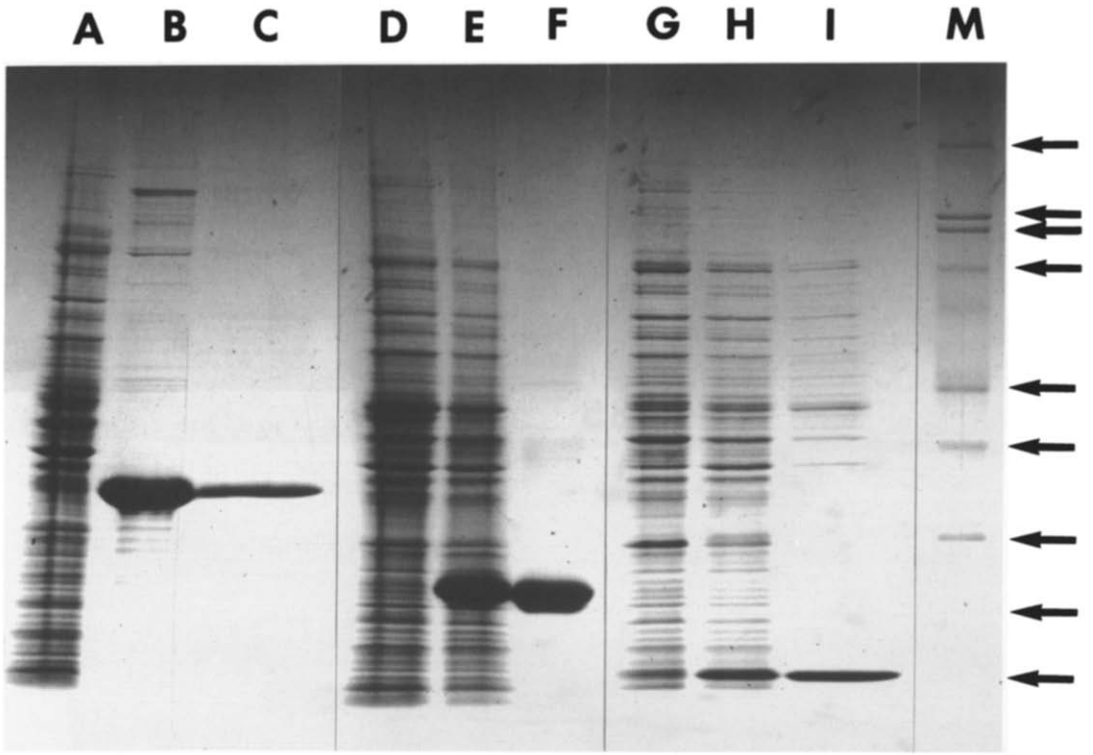

Fig. 4. Expression and purification of three different recombinant metavinculin fragments. a–c, MW-2; d–f, MW-13; g–i, MW-13c; M, molecular weight markers. a, d and g show *E. coli* before and b, e and h after induction with IPTG. c, f and i show the purified constructs. Molecular weight markers from top to bottom: myosin heavy chain (200), metavinculin (150), vinculin (130), phosphorylase *a* (92.5), hexokinase (51), actin (42), carbonic anhydrase (29.5), myosin light chain (20) and cytochrome *c* (11.7) kDa.

Table I

Calculated molecular weights and pI values of expressed fragments

Construct	Mol. mass (Da)	Isoelectric point (pI)
MW-2	28,822	8.35
MW-13	21,755	10.25
MW-13c	12,795	9.35

Note the extremely basic value of the MW-13 construct, representing the complete vinculin C-terminus.

were identical, whereas sequences coding for the second downstream motif in metavinculin were different (not shown). Furthermore, comparison of the porcine vinculin and metavinculin PCR clones showed nucleotide sequence identity outside the insert region (not shown). Taken together, these data suggest that the splice site is conserved in all tissues expressing metavinculin.

Species comparisons of vinculin sequences, C-terminal to the vinculin/metavinculin splice site revealed complete identity apart from three amino acid exchanges, one of them being conservative (**R→K**) at position '132' in the *Xenopus* sequence (Fig. 1). The cDNA fragments coding for the metavinculin-specific insert piece from porcine and avian smooth muscle tissue were described in a previous part [10]. The metavinculin C-termini including the metavinculin-specific fragment from *Xenopus laevis* and porcine smooth muscle were obtained by PCR as described in section 2. Cross-species comparison of the metavinculin insert regions of the four species investigated (Fig. 1) show that the respective region from *Xenopus laevis* is longer (79 amino acids) than that from chicken (69 amino acids) or those from human and pig (68 amino acids). Furthermore, alignment of the peptide sequences reveals a variable sequence stretch in the N-terminal portion of the insert, but strong similarity throughout the C-terminal half. The demonstrated tissue-specific expression of metavinculin [9,18,27] together with the observed variability in the amounts of metavinculin expressed in smooth muscle tissue of different origin point to a distinct function of this vinculin variant.

3.2. *Structural and physicochemical analysis of the insert and C-terminal sequences*

Structural analysis of the four metavinculin insert sequences obtained by applying the Garnier–Osguthorp–Robson algorithms showed a common, conserved C-terminal half and a variable N-terminal region. As suggested earlier [19], the insert is predicted to contain 3 helices interspersed by β-turns (Fig. 2).

The pI values calculated for the short metavinculin-specific peptides demonstrate a strong acidic nature with values ranging from 3.86 (chicken) to 4.21 (human). In contrast, the C-terminus of vinculin is seen to be extremely basic (pI 10.6) and has a predicted high helical content (>52%).

3.3. *Expression of constructs in* E. coli

To provide tools for probing for specific functions of the metavinculin insert piece we have constructed several vinculin- and metavinculin C-terminal fragments

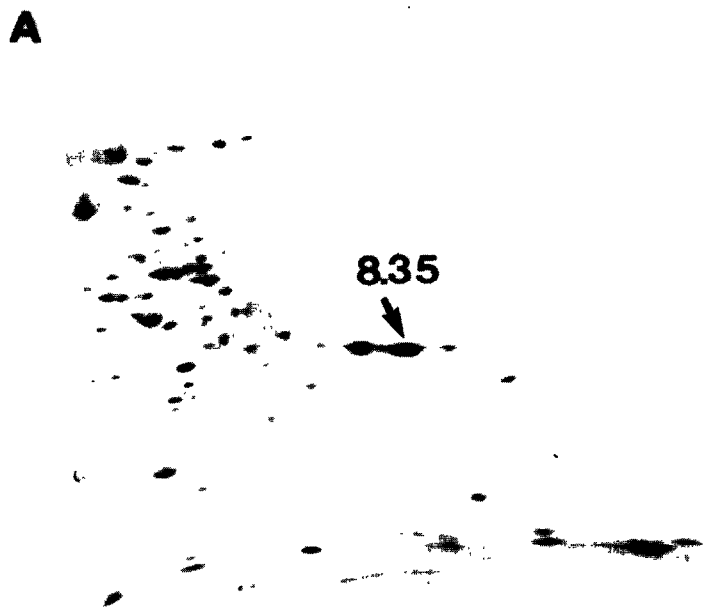

Fig. 5A.

(Fig. 3). Construct MW-2 spans the region from the beginning of the metavinculin insert throughout the complete C-terminus. MW-13 comprises the C terminus lacking the metavinculin difference piece and MW-13c encodes a C-terminal peptide starting at the *Stu*I restriction site. These three constructs were over-expressed in *E. coli* and could be purified by column chromatography (Fig. 4). The calculated, basic isoelectric points (see

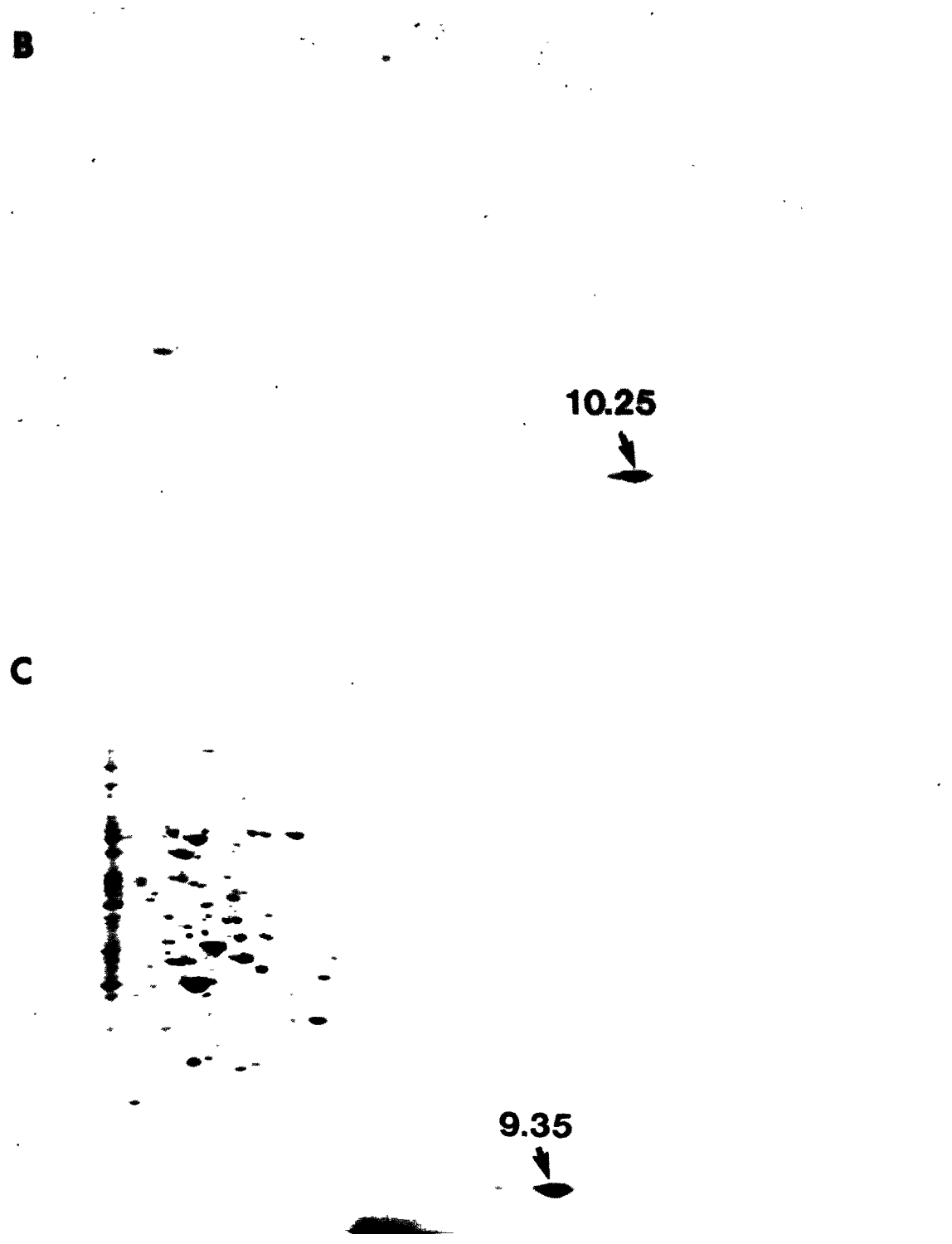

Fig. 5. Two-dimensional nonequilibrium electrophoresis gels of expressed constructs. (A) MW-2; (B) MW-13; (C) MW-13c. The mobility is consistent with their predicted pI values (indicated).

Table I) are consistent with the observed migration of these fragments on two-dimensional nonequilibrium electrophoresis gels (Fig. 5).

While we were able to express the entire C-terminal domain of vinculin we failed, however, to express the metavinculin-specific insert alone in *E. coli*. Furthermore, the insertion of a stop codon at the *StuI* site (constructs MW-2s and MW-13s) did not yield the corresponding peptides in *E. coli*. We therefore conclude that C-terminal sequences play a crucial role in stabilizing either the mRNA or the expressed protein. As described earlier [15], the C terminus of vinculin contains a region that can bind autonomously to focal adhesions. This site might overlap with the binding site for paxillin [4]. It is intended to use the purified constructs described in this communication, covering different lengths of metavinculin and vinculin, to define the binding motifs of these proteins in more detail as well as to identify as yet unknown binding partners.

It is noteworthy that the vinculin C-terminal sequences downstream of the KWSSK motif show 98.7% homology even between widely separated vertebrate species. In contrast, the KWSSK motif occurs as QWSSQ in nematode vinculin [28] and the nematode/vertebrate sequence homology for the C-terminus is only 68% [19] indicative of a branch point in the evolution of these molecules.

Acknowledgements: We are grateful to Mr. A. Weber for photography. This project was supported in part by a short term fellowship from the EMBO (to M.G.) and by grants from the Austrian Research Council (to P.S.) and the Austrian National Bank (to B.G. and J.V.S.). B.G. is E. Neter Professor in Cell and Tumor Biology.

REFERENCES

[1] Turner, Ch.E. and Burridge, K. (1991) Curr. Opinion Cell Biol. 3, 849–853.
[2] Geiger, B. and Ginsberg, D. (1991) Cell Motil. Cytoskel. 20, 1–6.
[3] Small, J.V., Fürst, D.O. and Thornell, L.-E. (1992) Eur. J. Biochem. 208, 559–572.
[4] Turner, Ch.E., Glenney Jr., J.R. and Burridge, K. (1990) J. Cell Biol. 111, 1059–1068.
[5] Crawford, A.W. and Beckerle, M.C. (1991) J. Biol. Chem. 266, 5847–5853.
[6] Funayama, N., Nagafuchi, A., Sato, N., Tsukita, S. and Tsukita, S. (1991) J. Cell Biol. 115, 1039–1048.
[7] Geiger, B. (1979) Cell 18, 193–205.
[8] D'Angelo-Siliciano, J. and Craig, S.W. (1982) Nature 300, 533–535.
[9] Glukhova, M.A., Kabakov, A.E., Frid, M.G., Ornatsky, O.I., Zhidkova, N.I. and Koteliansky, V.E. (1986) FEBS Lett. 207, 139–141.
[10] Koteliansky, V.E., Ogryzko, E.P., Zhidkova, N.I., Weller, P.A., Critchley, D.R., Vancompernolle, K., Vandekerckhove, J., Strasser, P., Way, M.P., Gimona, M. and Small, J.V. (1992) Eur. J. Biochem. 204, 767–772.
[11] Byrne, B.J., Kaczorowsky, Y.J., Coutu, M.D. and Craig, S.W. (1992) J. Biol. Chem. 267, 12845–12850.
[12] Bendori, R., Salomon, D. and Geiger, B. (1987) EMBO J. 6, 2897–2905.
[13] Jones, P., Jackson, P., Price, G.J., Patel, B., Ohanion, V., Lear, A.L. and Critchley, D.R. (1989) J. Cell Biol. 109, 2917–2927.
[14] Molony, L., McCaslin, D., Abernethy, J., Paschal, B. and Burridge, K. (1987) J. Biol. Chem. 262, 7790–7795.
[15] Bendori, R., Salomon, D. and Geiger, B. (1989) J. Cell Biol. 108, 2383–2393.
[16] Way, M., Pope, B., Gooch, J., Hawkins, M. and Weeds, A.G. (1990) EMBO J. 9, 4103–4109.
[17] Studier, F.W. and Moffatt, B.A. (1986) J. Mol. Biol. 189, 113–130.
[18] Chirgwin, J.M., Przbyla, A.E., MacDonald, J.R. and Rutter, W. (1979) Biochemistry 18, 5914–5922.
[19] Gimona, M., Small, J.V., Moeremans, M., Van Damme, J., Puype, M. and Vandekerckhove, J. (1988) EMBO J. 7, 2329–2334.
[20] Price, G.J., Jones, P., Davison, M.D., Patel, B., Bendori, R., Geiger, B. and Critchley, D.R. (1989) Biochem. J. 159, 453–461.
[21] Weller, P.A., Ogryzko, E.P., Corben, E.B., Zhidkova, N.I., Patel, B., Price, G.J., Spurr, N.K., Koteliansky, V.E. and Critchley, D.R. (1990) Proc. Natl. Acad. Sci. USA, 87, 5667–5671.
[22] Matsudaira, P. and Burgess, D.R. (1978) Anal. Biochem. 87, 286–396.
[23] Laemmli, U.K. (1970) Nature 227, 680–685.
[24] O'Farrell, P.H. (1975) J. Biol. Chem. 250, 4007–4021.
[25] Gimona, M., Sparrow, M.P., Strasser, P., Herzog, M. and Small, J.V. (1992) Eur. J. Biochem. 205, 1067–1075.
[26] Garnier, J., Osguthorpe, D.J. and Robson, B. (1978) J. Mol. Biol. 120, 97–120.
[27] Gimona, M., Fürst, D.O. and Small, J.V. (1987) J. Muscle Res. Cell Motil. 8, 329–341.
[28] Barstead, R.J. and Waterston, R.H. (1989) J. Biol. Chem. 264, 10177–10185.
[29] Price, G.J., Jones, P., Davison, M.D., Patel, B., Bendori, R., Geiger, B. and Critchley, D.R. (1989) Biochem. J. 259, 453–461.