

Rat Prominin, Like Its Mouse and Human Orthologues, Is a Pentaspan Membrane Glycoprotein

Denis Corbeil, Christine A. Fargeas, and Wieland B. Huttner¹

Max Planck Institute of Molecular Cell Biology and Genetics, Pfotenhauerstrasse 108, D-01307 Dresden, Germany

Received June 26, 2001

Mouse prominin is the first characterized member of a novel family of membrane glycoproteins. It displays a characteristic membrane topology with five transmembrane segments and two large glycosylated extracellular loops. Prominin orthologues and paralogues have been identified in human, fish, fly, and worm. Recently, a cDNA sequence encoding the rat homologue of mouse prominin has been reported [Zhu *et al.* (2001) *Biochem. Biophys. Res. Commun.* 281, 951–956]. Surprisingly, due to a single nucleotide deletion that shifts the reading frame and introduces a premature stop codon, the protein predicted from this cDNA would correspond to a C-terminally truncated form of prominin with only four transmembrane segments. Here we report evidence that is in contrast to the report of Zhu *et al.* (2001). We isolated a rat prominin cDNA devoid of any frameshift mutation, demonstrate that rat prominin, like the other mammalian prominins, is a full-length 120-kDa pentaspan membrane glycoprotein, and have not been able to detect any C-terminally truncated form of rat prominin. © 2001

Academic Press

Key Words: prominin; prominin (mouse)-like 1; pentaspan membrane protein; CD133.

The prominins are an emerging family of proteins that among the multispan membrane proteins display a novel topology (1). Mouse prominin (2) and human prominin (mouse)-like 1 (PROML1; alias AC133 antigen, CD133) (3) are predicted to contain five membrane-spanning domains, with an N-terminal domain exposed to the extracellular space followed by four, alternating small cytoplasmic and large extracellular, loops and a cytoplasmic C-terminal domain. Experimentally, this pentaspan membrane topology is sup-

ported by analyses of N-glycosylation and epitope accessibility (1–4).

Prominins are widely distributed throughout the animal kingdom. Predicted proteins structurally related to prominin have been identified in *Caenorhabditis elegans* (2, 5) zebrafish *Danio rerio* and *Drosophila melanogaster* (1). Vertebrate prominins are expressed in epithelial cells (2, 6) and epithelial-derived cells such as rod photoreceptor cells (4), as well as in non-epithelial cells, notably hematopoietic stem cells (6, 7). Irrespective of the cell type, prominins are concentrated in microvilli and other plasma membrane protrusions (2, 4, 6, 8) by a molecular mechanism that involves a novel cholesterol-based lipid microdomain (9). Although the physiological role of prominin has not yet been elucidated, recent studies using morphological, biochemical and genetic approaches suggest a role of prominin in plasma membrane protrusion morphogenesis (1, 4, 9).

Recently, after identifying expressed sequence tag clones corresponding to mRNA species differentially expressed in rat skeletal muscle upon high (versus normal) concentration of blood glucose (10), Zhu and colleagues (11) have reported the molecular cloning of one of the corresponding cDNAs. The protein encoded by this cDNA (11) was proposed to be the rat homologue of mouse prominin (2) and human PROML1 (3). Surprisingly, due to a single nucleotide deletion that shifts the reading frame and introduces a premature stop codon, the rat cDNA isolated by Zhu *et al.* (11) predicts a protein that would consist of only 591 amino acid residues and that would correspond to a C-terminally truncated version of mouse prominin (858 amino acid residues). By comparison with mouse prominin, the predicted rat homologue would lack about half of the second extracellular loop, the fifth transmembrane segment and the cytoplasmic C-terminal domain (11). However, Zhu *et al.* (11) did not present any evidence eliminating the possibility that the C-terminal truncation is virtual rather than real, being due to, e.g., a sequencing error, and showing that a C-terminally truncated prominin indeed exists *in vivo*.

The sequence reported in this paper has been deposited in the GenBank database (Accession No. AF386758).

¹To whom correspondence should be addressed. Fax: +49 (0)351 210 1489. E-mail: huttner@mpi-cbg.de.



In disagreement with Zhu *et al.* (11), we report here the isolation of a rat prominin cDNA devoid of any nucleotide deletion introducing a premature stop codon. By immunoblotting, rat prominin, like its mouse and human orthologues, is shown to be a full-length 120-kDa glycoprotein containing the cytoplasmic C-terminal domain.

MATERIALS AND METHODS

PCR amplification of rat prominin. To amplify by PCR the complete open reading frame of rat prominin, the gene-specific sense primer (SDF1; 5'-TTACCAGCCAGGAGGCAGAAGATT-3') and antisense primer (SDR1; 5'-ACTGTGAACCAGGACCAGTGGCT-3') were designed based on the published rat cDNA sequence (GenBank Accession No. AF263368) (11). Rat kidney Marathon-ready cDNA was used as template (Clontech, Palo Alto, CA; Cat. No. 7472-1). In addition to the amplification using the gene-specific primers, an independent PCR was performed using the SDF1 sense primer and the linker-specific adaptor primer 1 (AP1) supplied with the Marathon-ready cDNA kit. PCRs were carried out according to the manufacturer's instructions using Advantage cDNA Polymerase Mix (Clontech; Cat. No. 8417-1) for thirty cycles of denaturation (94°C, 30 s) and annealing/extension (68°C, 4 min) in a PTC-200 Peltier thermal cycler (MJ Research, Waltham, MA).

Cloning and cDNA sequencing. PCR products were subcloned in the pCR-Blunt II TOPO vector according to the manufacturer's instructions (Zero Blunt TOPO PCR cloning kit; Invitrogen BV, Groningen, Netherlands). Purified DNA plasmids were obtained from TOP10 bacterial cell cultures with a commercially available kit (Qiagen GmbH, Hilden, Germany; Cat. No. 12163). cDNA inserts were completely sequenced on both strands using the BigDye Terminator Cycle Sequencing V2.0 kit (Applied Biosystems, Foster City, CA). Sense (SDF3; 5'-CTACCTGCTCAACGATCA-3') and antisense (SDR3; 5'-CGAATGTTCTGTGCATCCA-3') primers were used for sequencing the cDNA region containing the postulated nucleotide deletion (11). DNA sequence acquisition and the initial analysis were performed using an ABI PRISM 377 DNA sequencer and the DNA Sequencing Analysis software version 3.4.1 (Applied Biosystems). Sequence alignments were performed at the National Center for Biotechnology Information using the BLAST network services (12).

Protein sequence analysis. The amino acid sequence deduced from the rat prominin cDNA was compared with its mouse and human orthologues using the CLUSTAL W program (13) with BLOSUM series matrix and a 10.0 penalty for opening up gaps and a 0.05 penalty for gap extensions and manually adjusted according to the human *PROML1* gene (4). The putative signal peptide cleavage site was determined using SignalP V1.1 software (14) at SignalP WWW Prediction Server, Center for Biological Sequence Analysis (<http://www.cbs.dtu.dk/services/SignalP/index.html>).

Membrane preparation. Kidneys from adult Sprague-Dawley rats were homogenized in buffered sucrose (300 mM sucrose, 5 mM EDTA, 10 mM HEPES-KOH, pH 7.5, 1 mM PMSF, 2 µg/ml leupeptin, 10 µg/ml aprotinin, 300 mg tissue/5 ml buffer) at 4°C by ten strokes in a glass-Teflon homogenizer at 3000 rpm, and centrifuged at 1000g for 10 min. The supernatant was centrifuged at 40,000g for 30 min, and the membrane pellet resuspended in 1% SDS, 20 mM sodium phosphate, pH 7.5, and boiled for 2 min. This material, referred to as solubilized membranes, was then used for deglycosylation and immunoblotting.

Endoglycosidase digestion and immunoblotting. Proteins (50 µg) in the solubilized rat kidney membranes were incubated overnight at 37°C in the absence or presence of 1U PNGase F according to the manufacturer's instructions (Roche Molecular Biochemicals, Mannheim, Germany). Proteins were then analyzed by SDS-PAGE and

transferred to poly(vinylidene difluoride) membranes (Millipore Corp., Bedford, MA; pore size 0.45 µm) using a semi-dry transfer cell system (Cti, Idstein, Germany). After transfer, membranes were incubated overnight at 4°C in blocking buffer (PBS containing 5% lowfat milk powder and 0.3% Tween 20). Prominin was then detected using either αE2 antiserum (1/20,000) or αI3 antiserum (1/20,000) (8) followed by horseradish peroxidase-conjugated secondary antibody, all diluted in blocking buffer. Antigen-antibody complexes were visualized using enhanced chemiluminescence (ECL system, Amersham Corp.).

RESULTS AND DISCUSSION

Isolation and Characterization of a Rat Prominin cDNA

To isolate, by PCR, a rat prominin cDNA containing the complete open reading frame, we designed two oligonucleotide primers, SDF1 and SDR1, based on the rat nucleotide sequence reported by Zhu *et al.* (11). The sequences of the sense (SDF1) and the antisense (SDR1) primers correspond to nucleotides (nt) 89–112 of the 5'-untranslated region and nt 3514–3536 of the 3'-untranslated region, respectively. Given the high expression of prominin in adult mouse kidney (2), a kidney cDNA pool (Marathon-ready) derived from the same rat strain (Sprague-Dawley) as used by Zhu *et al.* (11) was chosen as template. A cDNA fragment of approximately 3.5 kb was thus amplified and subcloned in the pCR-Blunt II TOPO vector. cDNA-inserts from four independent clones were completely sequenced on both strands. To reduce the probability of PCR errors, an independent PCR with SDF1 and AP1 primers (see Materials and Methods) was performed in parallel. One of the cDNA clones obtained, containing a polyadenylation consensus sequence followed by a poly-A tail (not shown), was completely sequenced on both strands. The sequence consensus obtained from the five cDNAs analyzed yielded a 3449 nt long rat prominin cDNA (GenBank Accession No. AF386758), which is identical to nucleotides 89–3536 of the rat cDNA reported by Zhu *et al.* (11), except for six nucleotides.

The first difference is the substitution of cytosine 713 (11) by a thymine (nt 625 in the present sequence), which has no consequence for the predicted rat prominin amino acid sequence. Second, thymine 1330 (11) is changed to cytosine (nt 1242 in the present sequence), which changes the phenylalanine residue 382 (11) to a serine, which is also found in the corresponding position in mouse prominin (residue 383) and human *PROML1* (residue 391) (see Fig. 3, below). The next three differences concern the region in rat prominin that has been referred to as "non-homologous" with regard to mouse prominin by Zhu *et al.* (11). The nucleotide sequence reported by these authors (11) lacks the adenine residue in position 1314 of the present sequence and contains an additional adenine residue between nucleotide 1334 and 1335 of the present sequence (Fig. 1A, shaded nucleotides marked by aster-

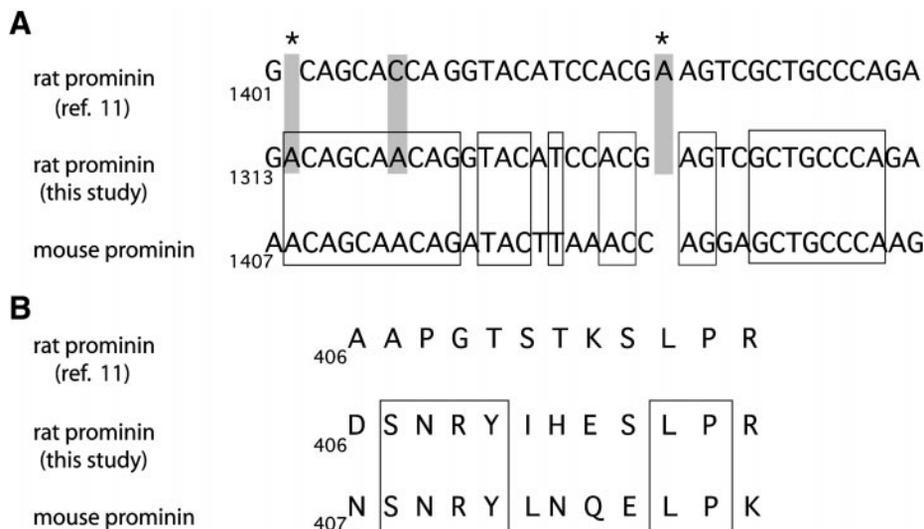


FIG. 1. The so-called "nonhomologous" region of rat prominin is, in fact, homologous to mouse prominin. (A) The nucleotide sequence of rat prominin reported by Zhu *et al.* (11) (GenBank Accession No. AF263368) was compared, from position 1401 to 1436, to the corresponding sequences of rat prominin determined in the present study (GenBank Accession No. AF386758) and mouse prominin (GenBank Accession No. AF026269); numbers refer to the respective sequence deposited in GenBank. Shaded nucleotides represent differences between the rat sequence reported by Zhu *et al.* (11) and that reported in the present study. Asterisks indicate the differences causing an apparent frameshift in the sequence reported by Zhu *et al.* (11). Nucleotides identical between the present rat prominin sequence and mouse prominin are boxed. (B) Comparison of the amino acid sequences predicted from the nucleotide sequences shown in (A). Amino acid residues identical between the present rat prominin sequence and mouse prominin are boxed.

isks). Compared to the present nucleotide sequence, which predicts, for this region of rat prominin, an amino acid sequence highly homologous to mouse prominin (Fig. 1B), this adenine deletion/addition results in reading frame shift over 21 bases, which is responsible for the apparent lack of amino acid sequence homology in the case of the rat cDNA studied by Zhu *et al.* (11). This region of the rat prominin cDNA contains an additional substitution (cytosine 1407 (11) to adenine 1320 in the present sequence) (Fig. 1A, shaded nucleotide without asterisk).

The sixth difference between the two rat prominin nucleotide sequences, but the one with the greatest consequence for the predicted protein sequence, is in the region for which Zhu *et al.* (11) reported a single nucleotide deletion (i.e., between nt 1958 and 1959) (Fig. 2A, triangle) shifting the reading frame and introducing a premature stop codon (Fig. 2A, boxed nucleotides). Aside from the fact that the position of the nucleotide missing from the sequence reported by these investigators is actually further upstream than claimed (i.e., in the triple-adenine stretch), we have found that this adenine stretch consists of four rather than three nucleotides (Figs. 2A and 2B). Consequently, no frame shift occurs, and rat prominin is predicted to be as long as the mouse and human orthologues.

Rat Prominin Is Predicted to be a Pentaspan Membrane Protein

The present rat prominin cDNA contains an open reading frame from nt 98 to 2668, which encodes a

putative 857-amino acid protein with a predicted molecular weight of 96,631. The sequence surrounding the putative translational start codon meets the Kozak consensus criteria for initiation of translation (not shown, see GenBank Accession No. AF386758). The predicted primary structure of rat prominin shows an average of 83 and 60% amino acid identity to mouse prominin and human PROML1, respectively (Fig. 3).

Hydropathy plot analysis of the rat prominin reveals six hydrophobic segments. The first, N-terminal hydrophobic segment most likely corresponds to a signal peptide (dashed line in Fig. 3). A potential signal peptidase cleavage site is found between Gly-20 and Gly-21 (Fig. 3). Cleavage at this site would generate an 837-amino acid protein with a predicted molecular weight of 94,626 and isoelectric point of 6.05. The other five hydrophobic segments constitute presumptive transmembrane domains (underlined in Fig. 3). Consequently, rat prominin, like its murine and human counterparts, would consist of an extracellular N-terminal domain, five transmembrane domains flanking two short cytoplasmic loops and two large extracellular loops, and a cytoplasmic C-terminal domain. The six cysteine residues present in the extracellular loops of rat prominin are conserved in position (Fig. 3, letters C), as is the cysteine-rich region at the transition of the first transmembrane segment and the first cytoplasmic loop (Fig. 3, asterisks). Rat prominin contains eleven potential N-glycosylation sites in its extracellular domains (Fig. 3, # symbols). Seven of these are located in the first loop (N211, N263, N281, N322, N364, N386,

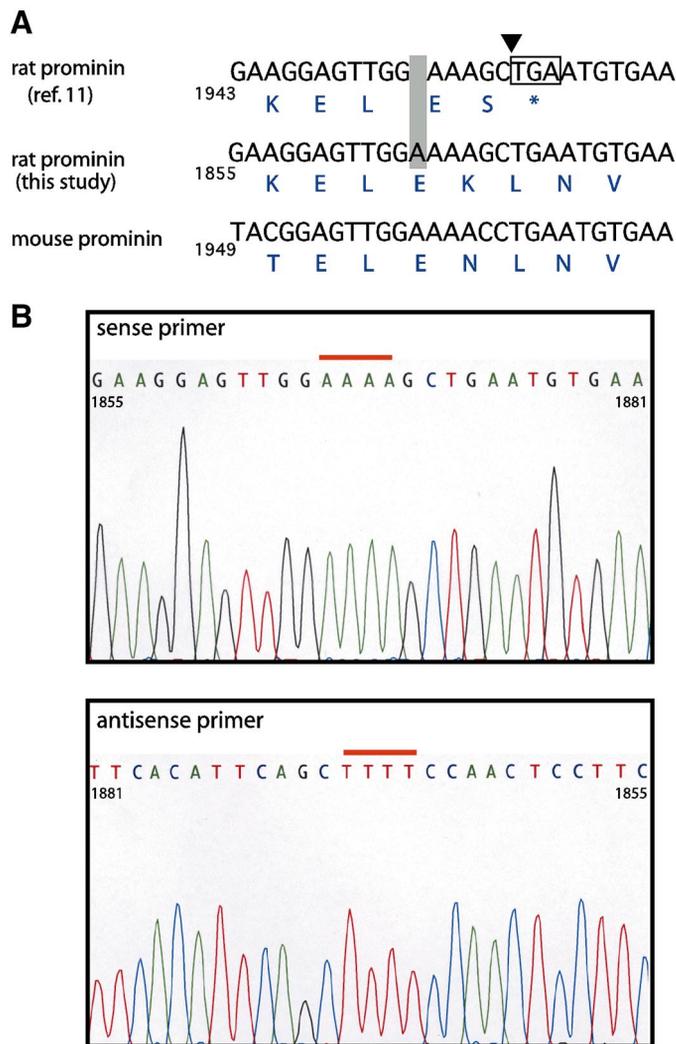


FIG. 2. The rat prominin cDNA does not contain a single nucleotide deletion leading to a shift in the reading frame. (A) The nucleotide sequence reported by Zhu *et al.* (11) (GenBank Accession No. AF263368) was aligned, from position 1943 to 1968, with the corresponding sequences of rat prominin determined in the present study (GenBank Accession No. AF386758) and mouse prominin (GenBank Accession No. AF026269); numbers refer to the respective sequence deposited in GenBank. Deduced amino acids are given in blue letters underneath the second nucleotide of the corresponding codon. The triangle indicates the position of the single nucleotide deletion as claimed by Zhu *et al.* (11); the postulated premature stop codon (nt 1959–1961) is boxed. The single nucleotide actually missing from the sequence reported by Zhu *et al.* (11) is one of the adenines of the tetra-adenine stretch (the choice of the shaded adenine is arbitrary). (B) Sequencing runs showing the stretch of four (rather than three) adenine residues in the rat prominin sequence determined in the present study. DNA sequencing reactions were performed on both strands using the SDF3 sense primer (top panel) and the SDR3 antisense primer (bottom panel). Red lines indicate the tetra-adenine stretch (nt 1866–1869) (top panel) and the tetra-thymine stretch in the complementary sequence (bottom panel).

N405) and four in the second (N544, N571, N622, N722); eight of these sites are conserved in position compared to mouse prominin (Fig. 3).

Immunological Characterization of Rat Prominin

So far, all conclusions as to whether rat prominin is C-terminally truncated (11) or full-length (this study) are based on predictions from cDNA sequences rather than experimental facts. We therefore characterized the actual rat prominin protein existing *in vivo*. Specifically, adult rat kidney membranes were analyzed by immunoblotting using the α E2 antiserum (8), which was raised against the first extracellular loop (amino acids 183–426) of mouse prominin. This region exhibits \approx 80% amino acid identity to rat prominin. Rat prominin was detected as immunoreactive band with a mean apparent molecular mass of \approx 120 kDa (Fig. 4, top panel, left lane, bracket). The same result was obtained (Fig. 4, bottom panel, left lane, bracket) using another antiserum, referred to as α I3 (8), which recognizes the highly conserved (see Fig. 3) cytoplasmic C-terminal domain of mouse prominin.

Rat prominin showed a slightly slower electrophoretic mobility (120 kDa) (Fig. 4) than prominin from mouse kidney (115 kDa; (2)) (data not shown). This presumably reflects the presence of additional N-glycosylation sites in the rat (as compared to mouse) prominin (Fig. 3). Upon removal of N-linked glycans by PNGase F treatment, prominin was detected as a single immunoreactive band of \approx 90 kDa using either the α E2 (Fig. 4, top panel, right lane, arrow) or α I3 (Fig. 4, bottom panel, right lane, arrow) antibody. Collectively, our data show that there is no significant difference in length between rat prominin and its mouse and human orthologues, and that rat prominin contains the C-terminal domain previously shown (for mouse prominin) to be located in the cytoplasm (2). Importantly, we did not detect, using the α E2 antibody recognizing the first extracellular loop of prominin, any immunoreactivity band in the 90 kDa range without, and in the 65 kDa range with, PNGase F treatment, as would be expected if rat prominin were C-terminally truncated as claimed by Zhu *et al.* (11).

Conclusions

In conclusion, our results show that rat prominin is a \approx 120 kDa glycoprotein (\approx 90 kDa after N-deglycosylation) containing the cytoplasmic C-terminal domain. These data are inconsistent with the claim of Zhu *et al.* (11) that rat prominin, in contrast to its mouse and human orthologues, should not be a pentaspan membrane protein but be C-terminally truncated, lacking about half of the second extracellular loop, the fifth transmembrane segment and the cytoplasmic C-terminal domain. It should be emphasized that Zhu *et al.* (11), in contrast to the present study, did not provide any protein data on endogenously expressed rat prominin. Hence, the only discrepancy to be explained concerns the difference between the previous

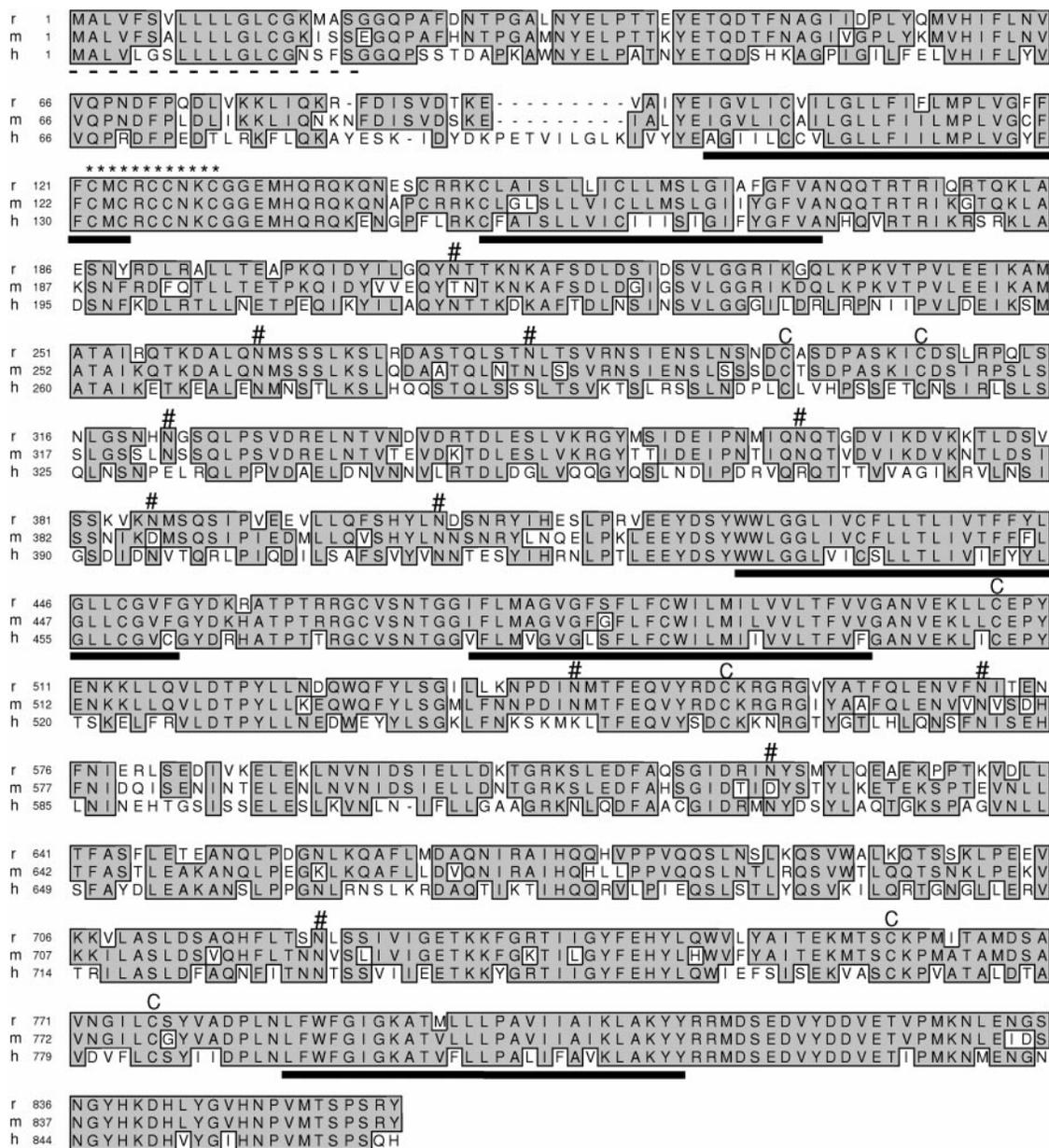


FIG. 3. Comparison of the amino acid sequence of rodent prominin and human PROML1. The rat prominin sequence determined in the present study (GenBank Accession No. AF386758) (r, top) was aligned with mouse prominin (GenBank Accession No. AF026269) (m, middle) and human PROML1 (GenBank Accession No. AF027208) (h, bottom). Shaded boxes, identical amino acid residues; dashed line, putative signal peptide; solid lines, predicted membrane-spanning segments; asterisk stretch, cysteine-rich region; C, conserved cysteine residues in the extracellular domains; #, potential N-glycosylation sites in rat prominin.

(11) and present rat prominin nucleotide sequence. The single nucleotide (adenine) deletion (Fig. 2A) present in the previously reported sequence (11) may reflect a cloning artifact or sequencing error. Clearly, this deletion is not a feature generally found in rats, as we have detected a ≈ 120 kDa protein using the $\alpha I3$ antibody in immunoblots of kidney membranes from Wistar rats (data not shown), nor is it specific for Sprague–Dawley rats, which were used by Zhu *et al.* (11), but also in the present study. Indeed, given that prominin’s membrane topology is conserved throughout the animal

kingdom (1), why should rats be different and express a C-terminally truncated form of prominin that, extrapolating from previous data (4), would not exhibit its characteristic location in plasma membrane protrusions?

Though unlikely, a point mutation in the genome of the individual rats from which the DNA analyzed by Zhu *et al.* (11) originated, cannot be ruled out at this stage. However, we feel that the claim of a “C-terminal truncated rat homologue of mouse prominin” (11) should be abandoned unless evidence is provided for the *in vivo* existence of such a protein. Even if such a

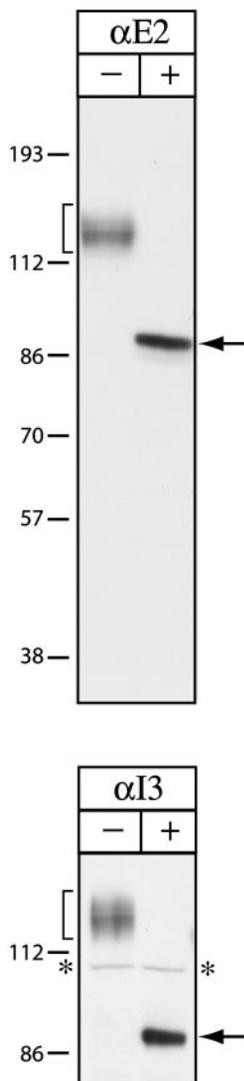


FIG. 4. Rat prominin is a full-length ≈ 120 kDa glycoprotein. Proteins ($50 \mu\text{g}$) solubilized from adult kidney membranes were incubated in the absence (-) or presence (+) of 1 U PNGase F and analyzed by immunoblotting using either αE2 antiserum (top panel) or αI3 antiserum (bottom panel). Brackets, ≈ 120 kDa form of rat prominin; arrows, N-deglycosylated ≈ 90 kDa form of rat prominin; asterisks, PNGase F-insensitive, nonspecific immunoreactive band. The position of prestained apparent molecular mass markers (in kDa) is indicated on the left.

protein were shown to exist, it should be referred to as prominin₁₋₅₉₁ rather than called by the new name “fudenine,” which bears a (perhaps coincidental) resemblance to the first name of the last author of the Zhu *et al.* paper (11).

ACKNOWLEDGMENTS

We thank Glenis Wiebe for the excellent sequencing service. W.B.H. was supported by grants from the DFG (HU 275/7-1), the EC

(ERB-FMRX-CT96-0023 and ERBBI04CT960058), the German-Israeli Foundation for Scientific Research and Development, and the Fonds der Chemischen Industrie.

REFERENCES

- Corbeil, D., Röper, K., Fargeas, C. A., Joester, A., and Huttner, W. B. (2001) Prominin: A story of cholesterol, plasma membrane protrusions and human pathology. *Traffic* **2**, 82–91.
- Weigmann, A., Corbeil, D., Hellwig, A., and Huttner, W. B. (1997) Prominin, a novel microvilli-specific polytopic membrane protein of the apical surface of epithelial cells, is targeted to plasmalemmal protrusions of non-epithelial cells. *Proc. Natl. Acad. Sci. USA* **94**, 12425–12430.
- Miraglia, S., Godfrey, W., and Yin, A. H., *et al.* (1997) A novel five-transmembrane hematopoietic stem cell antigen: Isolation, characterization, and molecular cloning. *Blood* **90**, 5013–5021.
- Maw, M. A., Corbeil, D., and Koch, J., *et al.* (2000) A frameshift mutation in prominin (mouse)-like 1 causes human retinal degeneration. *Hum. Mol. Genet.* **9**, 27–34.
- Corbeil, D., Röper, K., Weigmann, A., and Huttner, W. B. (1998) AC133 hematopoietic stem cell antigen: Human homologue of mouse kidney prominin or distinct member of a novel protein family? *Blood* **91**, 2625–2626.
- Corbeil, D., Röper, K., and Hellwig, A., *et al.* (2000) The human AC133 hematopoietic stem cell antigen is also expressed in epithelial cells and targeted to plasma membrane protrusions. *J. Biol. Chem.* **275**, 5512–5520.
- Yin, A. H., Miraglia, S., and Zanjani, E. D., *et al.* (1997) AC133, a novel marker for human hematopoietic stem and progenitor cells. *Blood* **90**, 5002–5012.
- Corbeil, D., Röper, K., Hannah, M. J., Hellwig, A., and Huttner, W. B. (1999) Selective localization of the polytopic membrane protein prominin in microvilli of epithelial cells—A combination of apical sorting and retention in plasma membrane protrusions. *J. Cell Sci.* **112**, 1023–1033.
- Röper, K., Corbeil, D., and Huttner, W. B. (2000) Retention of prominin in microvilli reveals distinct cholesterol-based lipid microdomains in the apical plasma membrane. *Nat. Cell Biol.* **2**, 582–592.
- Zhu, G., Zuo, J., and Fang, F. (2000) Isolation of genes related to blood glucose-control in rat skeletal muscle. *Chin. Med. J.* **113**, 894–898.
- Zhu, G., Chang, Y., Zuo, J., Dong, X., Zhang, M., Hu, G., and Fang, F. (2001) Fudenine, a C-terminal truncated rat homologue of mouse prominin, is blood glucose-regulated and can up-regulate the expression of GAPDH. *Biochem. Biophys. Res. Commun.* **281**, 951–956.
- Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990) Basic local alignment search tool. *J. Mol. Biol.* **215**, 403–410.
- Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994) CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* **22**, 4673–4680.
- Nielsen, H., Engelbrecht, J., Brunak, S., and von Heijne, G. (1997) Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. *Protein Eng.* **10**, 1–6.