Molecular Characterization of a Novel Human Hybrid-Type Receptor That Binds the $\alpha_2$-Macroglobulin Receptor-associated Protein*

(Received for publication, September 10, 1996)

Linda Jacobsens‡§, Peder Madsen‡, Søren K. Moestrup‡, Anders H. Lund‡, Niels Tommerup‡, Anders Nykjær‡, Lars Sotrup-Jensen‡, Jørgen Gliemann‡, and Claus M. Petersen‡§

From the Departments of ‡Medical Biochemistry and §Molecular and Structural Biology, University of Aarhus, DK-8000 Aarhus C, Denmark and the Department of Medical Genetics, The Panum Institute, University of Copenhagen, DK-2200 Copenhagen N, Denmark

The 39–40-kDa receptor-associated protein (RAP) binds to the members of the low density lipoprotein receptor (LDLR) gene family and functions as a specialized endoplasmic reticulum/Golgi chaperone. Using RAP affinity chromatography, we have purified a novel ~250-kDa brain protein and isolated the corresponding cDNA. The gene, designated SORLA, maps to chromosome 11q23/24 and encodes a 2214-residue type 1 receptor containing a furin cleavage site immediately preceding the N terminus determined in the purified protein. The receptor, designated sorLA-1, has a short cytoplasmic tail containing a tyrosine-based internalization signal and a large external part containing (from the N-terminal): 1) a segment homologous to domains in the yeast vacuolar protein sorting 10 protein, Vps10p, that binds carboxypeptidase Y, 2) five tandemly arrangedYWTD repeats and a cluster of 11 class A repeats characteristic of the low density lipoprotein receptor gene family receptors, and 3) six tandemly arranged fibronectin type III repeats also found in certain neural adhesion proteins. sorLA-1 may therefore be classified as a hybrid receptor. Northern blotting revealed specific mRNA transcripts in brain, spinal cord, and testis but not in several major organs. Both RAP and an antibody against a synthetic peptide derived from a sequence determined in the mature protein detected sorLA-1 in crude human brain extracts. The domain structure suggests that sorLA-1 is an endocytic receptor possibly implicated in the uptake of lipoproteins and of proteases.

The 39–40-kDa receptor associated protein (RAP) binds to the members of the low density lipoprotein receptor (LDLR) gene family (1–4) and to the recently discovered human receptor gp95/sortilin. RAP is located predominantly in the endoplasmic reticulum and Golgi compartments and is thought to prevent aggregation and premature binding of ligands to the multifunctional LDLR family receptors (5, 6). In mammals, the known members of the LDLR family comprise LDLR itself, the very low density lipoprotein receptor, the apolipoprotein receptor 2, and the giant (~600 kDa) multifunctional $\alpha_2$-macroglobulin receptor/LDL receptor-related protein (LRP) and megalin. In addition, several related receptors have been discovered in chicken (7, 8). The LDLR family members are type 1 receptors with short cytoplasmic tails containing one or a few FXNFXY internalization motifs. The extracellular domains contain 1) clusters of the ~40-residue LDLR class A repeats that display negatively charged surfaces stabilized by three disulfide bonds and bind ligands, including RAP (9, 10), 2) clusters of “spacer” regions containing the consensus peptide YWTD, and 3) single elements or pairs of growth factor repeats. In addition to the LDLR family members, a hybrid receptor with an N-terminal cluster of LDLR class A repeats and a C-terminal part similar to regions in guanine nucleotide binding protein-coupled receptors has been described in the mollusc Lymnaea stagnalis (11). The RAP binding human gp95/sortilin is a type 1 receptor whose luminal domain is homologous to that of Vps10p (12) and without the repeats that characterize the LDLR gene family members.

The aim of the present work was to search for new receptors capable of binding RAP. We report the molecular characterization of a novel human type 1 receptor, designated sorLA-1 (sorting protein-related receptor containing LDLR class A repeats), which is predominantly expressed in the central nervous system. It is a hybrid-type receptor that contains 11 LDLR class A repeats and five YWTD repeats, as well as a Vps10p/sortilin domain and six fibronectin type III repeats.

**MATERIALS AND METHODS**

Purification, Blotting Procedures, and Protein Sequence Analysis—Brain tissue samples were, with permission, obtained from autopsy cases 12–24 h postmortem and homogenized on ice in 250 mM sucrose. 10 mM Hepes, 5 mM EDTA, 0.1 mM phenylmethylsulfonyl fluoride, pH 8.0. A crude membrane fraction was generated, solubilized, and applied to RAP affinity chromatography in buffer with 0.25% CHAPS as described (13). After washings, proteins retained on the RAP-Sepharose column were eluted at pH 4.0 in phosphate-buffered saline with 10 mM EDTA and 0.25% CHAPS and then subjected to 4–16% gradient SDS-PAGE. For ligand blotting, the proteins were transferred to Immobilon membranes, blocked for 2 h at 20 °C in buffer containing 2% defatted milk powder and 0.05% Tween 20, washed, and incubated for 16 h in 140 mM NaCl, 2 mM CaCl$_2$, 1 mM MgCl$_2$, 10 mM Hepes, pH 7.2, with 0.5% bovine serum albumin and about 5 $\times$ 10$^5$ cpm/ml $^{35}$S-labeled RAP.

*This work was supported by grants from The Danish Medical Research Council, The Danish Science Research Council, The Danish Biotechnology Program, The Danish Cancer Society, The Novo Nordic Foundation, and The Carlsberg Foundation. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The nucleotide sequence(s) reported in this paper has been submitted to the GenBank™/EBI Data Bank with accession number(s) U69975.

§ To whom correspondence should be addressed: Dept. of Medical Biochemistry, University of Aarhus, Ole Worms Alle, Bldg. 170, DK-8000 Aarhus C, Denmark. Tel: 45-89-422880; Fax: 45-86-131160.

1 The abbreviations used are: RAP, receptor-associated protein; sorLA-1, sorting protein-receptor containing LDL receptor class A repeats; LDLR, low density lipoprotein receptor; LRP, LDL receptor-related protein/$\alpha_2$-macroglobulin receptor; aa, amino acid(s); kb, kilobase(s); PAGE, polyacrylamide gel electrophoresis; CHAPS, 3-[(3-cholamidopropyl)dimethylammonio]-1-propanesulfonic acid; FLpter, fractional length p-terminal.

2 C.M. Petersen, M.S. Nielsen, A. Nykjær, L. Jacobsen, N. Tommerup, H. H. Rasmussen, H. Reigaard, J. Gliemann, P. Madsen, and S. K. Moestrup, submitted for publication.
Binding was visualized by autoradiography of air-dried filters at -80°C. The procedure for immunoblotting was similar except that blocking was in buffer containing 2% Tween 20. Incubation with peroxidase-conjugated secondary antibody was with 0.2% defatted milk powder and 0.05% Tween 20.

For sequence analysis, the SDS-PAGE resolved proteins were transferred to Immobilon membranes by electroblotting in 50 mM Tris, 50 mM boric acid buffer, pH 8.0, stained by 0.1% amido black in 45% methanol, 9% acetic acid, and the band containing sorLA-1 was cut out and used directly for N-terminal sequencing. To obtain internal sequences, sorLA-1 was reduced and carboxymethylated prior to digestion in situ (14, 15). The Immobilon membranes were incubated for 30 min at 20°C in 100 μl of 0.1 mM ammonium bicarbonate, pH 7.8, containing 1% hydrogenated Triton X-100 followed by the addition of 0.5 μg of trypsin in 0.02 M ammonium bicarbonate, 1% Triton X-100, pH 7.8, and incubation for 18 h. The membranes were washed once in digestion buffer with 10% acetonitrile and twice in 0.1% trifluoroacetic acid, and the supernatants were combined. Separation of the released peptides was performed on an Applied Biosystems microbore high performance liquid chromatography (model 172) using a 2 × 100-mm Nucleosil C18 reverse phase column and a linear gradient of acetonitrile and 0.1% trifluoroacetic acid (flow rate, 100 μl/min). N-terminal sequence analysis of the undigested protein and of selected peptides was performed using an Applied Biosystems model 477A gas phase sequenator connected to a 120A phenylthiohydantoin analyzer and operated as described (16).

cDNA Cloning and Sequencing—A degenerated probe (5'-TTTGCTGAAGCAGAAT/CTCC/GA/CTTT/GCGA-3') derived from the peptide Y177FANENPFADFeS66 and containing degeneracies at the most ambiguous positions, was used to screen a human brain cDNA library in the lambda ZAP vector (Stratagene). Following 5'-end labeling with [γ-32P]ATP, the probe was used to screen four sets of replicate 22 × 22-cm nylon filters (Amer sham), each containing 1.25 × 106 plaque-forming units. The filters were hybridized at 50°C using the tetramethylammonium technique (17). Two positive clones with identical inserts were identified, purified, and rescued into pBluescript. The 1.4-kb insert, coding for an His286-Ala750 of sorLA-1 (cf. Fig. 2) was recovered by EcoRI cleavage and random prime labeled with [α-32P]ATP for rescreening. Two clones were identified in the brain cDNA library and were shown to contain inserts of 2.2 and 5.7 kb. From these inserts, 5775 bases were obtained in addition to a 3'-untranslated region and the poly(A) tail. The 1.4-kb cDNA was then used to screen a Jurkat cDNA library in the ZAP Express vector, and a clone was identified and shown to overlap the previously determined sequence by 2004 bases.

Chromosome Mapping—Fluorescence in situ hybridization using the biotin-labeled 5.7-kb cDNA with corresponding DAPI banding and measurement of the relative distance from the short arm telomere to the signal fractional length p-terminal (FLpter) value were performed as described (18).

Northern Blotting—Tissue blots 7760–1, 7759–1, and 7767–1 (Clontech) containing 2 μg of poly(A)+ purified RNA from different human tissues were used. Hybridizations carried out with the labeled 1.4-kb cDNA probe and washings were performed at high stringency conditions.

RESULTS

Identification of sorLA-1 and Binding of RAP—Fig. 1A shows that the eluate from the RAP affinity column contained at least two proteins larger than ~200 kDa (lanes 1–8): LRP α-chain as detected by Western blotting using a specific monoclonal antibody (not shown), and a hitherto unrecognized faster migrating protein. 125I-Labeled RAP bound to both proteins (lane 9) and could be inhibited by unlabeled RAP (200 nM), EDTA (10 mM), and heparin (10 units/ml) (not shown). Fig. 1B, lanes 1 and 2, shows that the migration of sorLA-1 was retarded following reduction, indicative of the presence of multiple disulfide bonds. The size of the reduced protein was estimated at ~250 kDa. Additional results (not shown) demonstrated that treatment with PNGase-F caused a reduction in size and that antibodies (rabbit serum) directed against the LRP domain containing 10 LDLR class A repeats reacted with both LRP and sorLA-1. These data suggested that sorLA-1 might be related to the LDLR family receptors.

Proteins resolved by reducing SDS-PAGE were transferred to Immobilon membranes, Amido Black-stained bands of sorLA-1 were cut out, and sequence analysis was performed on undigested sorLA-1 and on selected peptides from the tryptic digestes. The obtained aa sequences, underlined in Fig. 2, were not identified in the GenBank and EMBL data bases. An antibody against a synthetic peptide based on the sequence Glu370, Arg388 reacted with sorLA-1 in crude brain membranes (Fig. 1B, lanes 3 and 4).

cDNA Library Screening, Amino Acid Sequencing, and Gene Mapping.—The overlapping 2.2- and 5.7-kb clones identified in the human brain cDNA library coded for the C-terminal 1925 aa, including two internal sequences determined in purified sorLA-1. Screening of the Jurkat library identified a clone that overlapped the previously determined sequence by 2004 bases and in addition coded for the N-terminal 289 aa, including the N terminus of the purified protein as determined by aa sequence analysis. The assembled cDNA revealed an open reading frame encoding the 2214 aa shown in Fig. 2. It should be noted that the N terminus Ser54 of the purified protein as determined by aa sequence analysis. The assembled cDNA revealed an open reading frame encoding the 2214 aa shown in Fig. 2. It should be noted that the N terminus Ser54 of the purified sorLA-1 most likely results from cleavage by furin since the preceding sequence RRKR abides by the furin consensus sequence RX(K/R)R.

Using the biotin-labeled cDNA, specific signals were only obtained distally on the long arm of chromosome 11, with 36 of 40 analyzed metaphases displaying at least one specific signal and with 79 of the 160 chromatin being labeled. The FLpter value was 0.88 ± 0.05 corresponding to localization of the gene, designated SORL1, at 11q23.2-q24.2 (Fig. 3).

Domain Structure.—The deduced aa sequence (Fig. 2) predicts a type 1 receptor with a 28-residue signal peptide (19) preceding a large luminal/extracellular part, a transmembrane domain, and a cytoplasmic tail of 56 residues. Fig. 4A provides an overview of the domain organization. The N-terminal ~700 residues constitute a single domain with homology to each of two similarly sized luminal domains of yeast Vps10p (12) and to the luminal domain of the newly identified RAP binding receptor gp95/sortilin (20). The overall homology to the N-terminal ~700-residue domain of Vps10p is weak (23% aa identity). However, as shown in Fig. 4B, a segment comprising residues 597–724 of sorLA-1 is closely related to the corresponding
segments in gp95/sortilin and Vps10p (only one Vps10p segment is shown). Based on the high degree of homology, which includes the spacing of 10 cysteine residues, we have suggested this region to constitute a novel protein module. A cleavage site for furin or homologous proteases may represent an additional feature common to the 700-residue Vps10p homology domains. Thus, in addition to sorLA-1, human gp95/sortilin also contains a potential furin cleavage sequence (RWRR), which in fact precedes the N terminus (Ser45) determined in purified gp95/sortilin, and yeast Vps10p contains a potential site (Ile28-Arg29) for cleavage by Kex2, a yeast homologue of furin.

The Vps10p domain is followed by five spacer repeats containing the tetrapeptide YWTD or closely related motifs and, notably, a cluster of 11 LDLR class A repeats (residues 1046–1525). Fig. 4 shows alignment of the class A repeats and displays consensus residues with an occurrence of more than 50% at a given position in human LRP, rat megalin, and the Caenorhabditis elegans homologue (20). Finally, the region preceding the transmembrane domain constitutes six 90-aa fibronectin type III repeats with a characteristic spacing of tryptophan, tyrosine, and leucine residues (not shown) followed by a 24-residue linker sequence.

In the tail, the hexapeptide F2144ANSHY2149 is closely similar to the putative internalization motif FENTLY of the endocytic mannose receptor (21) and in accordance with the overall internalization motif (FYXYYY(FY)). The nonapeptide Asp2162-Asp2170 has seven acidic residues and constitutes an acidic cluster similar to those in mannose-6-phosphate receptors, Vps10p, epidermal growth factor, and transferrin receptors and to the cluster that functions as a determinant in furin for trans-Golgi network localization and internalization from the cell surface (Ref. 22 and references therein).

DISCUSSION

We characterize a new receptor with structural modules described previously in different gene families: the segment containing 10 cysteine residues present in yeast Vps10p (12) and human gp95/sortilin, which precedes the signal peptide of the N-terminal sequence of the purified protein and internal sequences determined in purified sorLA-1 are underlined. The hydrophobic putative transmembrane domain as well as potential N-glycosylation sites are boldfaced. The internalization motif in the cytoplasmic tail is boxed, and the acidic cluster is indicated as a dashed underscore.

Fig. 2. Deduced amino acid sequence of sorLA-1. Numbering starts from the predicted signal sequence cleavage site, and negative numbers refer to the signal peptide. The N-terminal sequence of the purified protein and internal sequences determined in purified sorLA-1 are underlined. The hydrophobic putative transmembrane domain as well as potential N-glycosylation sites are boldfaced. The internalization motif in the cytoplasmic tail is boxed, and the acidic cluster is indicated as a dashed underscore.

Fig. 3. Gene mapping of sorLA-1. Panels A and B, partial metaphases hybridized with biotin-labeled cDNA showing specific FISH signals (arrows) on a metaphase chromosome (A) and a prometaphase chromosome (B) and the simultaneously DAPI-stained chromosomes (A′ and B′). C, a chromosome ideogram with the localization of the SORL1 gene encoding sorLA-1 based on the DAPI band pattern and FLpter value. The horizontal box indicates the variation in FLpter values on individual chromosomes.

Tissue-specific Expression—Northern blotting on mRNA from human tissues showed hybridization with a transcript of 10.5 kb particularly abundant in brain, spinal cord, and testis. mRNA was also expressed in the prostate, ovary, thyroid, and spleen but was not detected in major organs as skeletal muscle, kidney, placenta, heart (Fig. 5), and liver, lung, bone marrow, and adrenals (not shown).
repeats of LDLR family receptors at the cell surface, the main function is thought to be prevention of receptor aggregation in the endoplasmic reticulum and Golgi compartments as reported for the multifunctional LRP (6, 7). RAP may similarly protect sorLA from aggregation in the synthetic pathway, whereas it is unlikely to be an external ligand. However, the structure of sorLA-1 provides some clues concerning possible extracellular ligands. Notably, apolipoprotein E-containing lipoproteins bind, like RAP, to the clusters of class A repeats of all members of the LDLR family and may therefore bind to sorLA-1 as well. Other candidate ligands include proteases since yeast Vps10p functions as receptor for carboxypeptidase Y (12) and since the class A repeats of the multifunctional LDLR family receptors bind a series of proteases and protease-inhibitor complexes (27–30). Finally, a possible functional role of the fibronectin type III repeats, e.g., in promoting neurite outgrowth as reported for the type III repeats of the neuroglia cell adhesion molecule (24), needs to be elucidated.

The tissue distribution suggests that sorLA-1 may be particularly important in the central nervous system. In fact, the distribution is quite similar to that of the recently cloned human LDLR gene family member apolipoprotein receptor 2 (4). In view of the strong association of the apolipoprotein E4 allele with late-onset Alzheimer’s disease (31), it is likely that apoE receptors abundant in brain, and among these possibly sorLA-1, are involved in the pathogenesis of neurodegenerative diseases in addition to their role in the normal function of the central nervous system.

Acknowledgments—We thank Nina Jørgensen and Lene Kristensen for expert technical assistance.

REFERENCES

1. Krieger, M., and Herz, J. (1994) Annu. Rev. Biochem. 63, 601–637
2. Farquhar, M. G., Saito, A., Kerjaschki, D., and Orlando, R. A. (1995) J. Am. Soc. Nephrol. 6, 35–47
3. Jingami, H., and Yamamoto, T. (1995) Curr. Opin. Lipidol. 6, 104–108
4. Kim, D.-H., Iijima, H., Goto, K., Sakai, J., Ishii, H., Kim, H.-J., Suzuki, H., Kondo, H., Sueki, S., and Yamamoto, T. (1996) J. Biol. Chem. 271, 8373–8380
5. Bu, G., Geuze, H., Strous, G. J., and Schwartz, A. L. (1995) EMBO J. 14, 2269–2280
6. Willnow, T. E., Kohlmann, A., Horton, J., Otani, H., Braun, J. R., Hammer, R. E., and Herz, J. (1996) EMBO J. 15, 2632–2639
7. Schneider, W. J., and Nimpf, J. (1995) Curr. Opin. Lipidol. 4, 205–209
8. Novak, S., Hiesberger, T., Schneider, W. J., and Nimpf, J. (1996) J. Biol. Chem. 271, 11732–11736
9. Russell, D. W., Brown, M. S., and Goldstein, J. L. (1989) J. Biol. Chem. 264, 21682–21686
10. Daly, N. L., Scanlon, M. J., Djordjevic, J. T., Kroon, P. A., and Smith, R. (1995) Proc. Natl. Acad. Sci. U.S.A. 92, 6334–6338
11. Tensen, P. C., van Kesteren, E. R., Planta, R. J., Cox, K. J. A., Burke, J. F., van Heerikhuizen, H., and Vreugdenhil, E. (1994) Proc. Natl. Acad. Sci. U.S.A. 91, 4816–4820
12. Marcusson, E. G., Horazdovsky, B. F., Cereghino, L. J., Bharakhanian, E., and...
New RAP Binding Hybrid Receptor

31383

Emr, S. D. (1994) Cell 77, 579–586

13. Petersen, C. M., Ellgaard, L., Nykjær, A., Vilhardt, F., Vorum, H., Thøgersen, H. C., Nielsen, M. S., Jacobsen, C., Moestrup, S. K., and Gliemann, J. (1996) EMBO J. 15, 4165–4173

14. Fernandez, J., deMott, M., Atherton, D., and Mische, S. M. (1992) Anal. Biochem. 210, 255–264

15. Atherton, D., Fernandez, J., and Mische, S. M. (1993) Anal. Biochem. 212, 98–105

16. Sottrup-Jensen, L. (1995) Anal. Biochem. 225, 187–188

17. Honore, B., Madsen, P., and Leffers, H. (1993) J. Biochem. Biophys. Methods 27, 39–48

18. Tommerup, N., and Vissing, H. (1995) Genomics 27, 259–264

19. von Heijne, G. (1983) Eur. J. Biochem. 133, 17–21

20. Saito, A., Pietromonaco, S., Loo, A. K.-C., and Farquhar, M. G. (1994) Proc. Natl. Acad. Sci. U. S. A. 91, 9725–9729

21. Johnson, K. F., Chan, W., and Kornfeld, S. (1990) Proc. Natl. Acad. Sci. U. S. A. 87, 10010–10014

22. Voorhees, P., Deignan, E., van Donselaar, E., Humphry, J., Marks, M. S., Peters, J. P., and Bonifacino, J. S. (1995) EMBO J. 14, 4961–4975

23. Kornblitt, A. R., Umezawa, K., Vibe-Pedersen, K., and Baralle, F. E. (1985) EMBO J. 4, 1755–1759

24. Burgoon, M. P., Hazan, R. B., Phillips, G. R., Crossin, K. L., Edelman, G. M., and Cunningham, B. A. (1995) J. Cell Biol. 130, 733–744

25. Prince, J. T., Alberti, L., Healy, P. A., Nauman, S. J., and Stallcup, W. B. (1991) J. Neurosci. Res. 30, 567–581

26. Ruoslahti, E., and Pierschbacher, M. D. (1987) Science 238, 491–497

27. Moestrup, S. K., Holtet T. L., Etzerodt, M., Thøgersen, H. C., Nykjær, A., Andreasen, P. A., Rasmussen, H. H., Sottrup-Jensen, L., and Gliemann, J. (1993) J. Biol. Chem. 268, 13891–13896

28. Willnow, T. E., Orth, K., and Herz, J. (1994) J. Biol. Chem. 269, 15827–15832

29. Bu, G., Williams, S., Strickland, D. K., and Schwartz, A. L. (1992) Proc. Natl. Acad. Sci. U. S. A. 89, 7427–7431

30. Nielsen, M. S., Nykjær, A., Warshawsky, I., Schwartz, A. L., and Gliemann, J. (1995) J. Biol. Chem. 270, 23713–23719

31. Strittmatter, W. J., and Roses, A. D. (1995) Proc. Natl. Acad. Sci. U. S. A. 92, 4725–4727