The Intrinsic Factor-Vitamin B\textsubscript{12} Receptor and Target of Teratogenic Antibodies Is a Megalin-binding Peripheral Membrane Protein with Homology to Developmental Proteins\textsuperscript{*}

(Received for publication, October 17, 1997)

Søren K. Moestrup\textsuperscript{§§}, Renata Kozyraki\textsuperscript{§§}, Mette Kristiansen\textsuperscript{§}, James H. Kaysen\textsuperscript{**}, Hanne Holm Rasmussen\textsuperscript{‡}, Didier Brault\textsuperscript{‡‡}, Françoise Pontillon, Fatime O. Goda\textsuperscript{**}, Erik I. Christensen\textsuperscript{§§}, Timothy G. Hammond\textsuperscript{**}, and Pierre J. Verroust\textsuperscript{¶}

From the \textsuperscript{§}Department of Medical Biochemistry, University of Aarhus, 8000 Aarhus C, Denmark, \textsuperscript{¶}Institut National de la Santé et de la Recherche Médicale, U489, Hôpital Tenon, 75020 Paris, France, \textsuperscript{**}Tulane University Medical and Astrobiology Centers and Veterans Affairs Medical Center, New Orleans, Louisiana 70112, \textsuperscript{‡}Département de Biochimie, Hôpital Tenon, 75020 Paris, France, and the \textsuperscript{¶}Department of Cell Biology, Institute of Anatomy, University of Aarhus, 8000 Aarhus C, Denmark

The present report shows the molecular characterization of the rat 460-kDa epithelial glycoprotein that functions as the receptor facilitating uptake of intrinsic factor-vitamin B\textsubscript{12} complexes in the intestine and kidney. The same receptor represents also the yolk sac target for teratogenic antibodies causing fetal malformations in rats. Determination of its primary structure by cDNA cloning identified a novel type of peripheral membrane receptor characterized by a cluster of eight epidermal growth factor type domains followed by a cluster of 27 CUB domains. In accordance with the absence of a hydrophobic segment, the receptor could be released from renal cortex membranes by nonenzymatic and nonsolubilizing procedures. The primary structure has no similarity to known endocytic receptors but displays homology to epidermal growth factor and CUB domain proteins involved in fetal development, e.g. the bone morphogenetic proteins. Electron microscopic immunogold double labeling of rat yolk sac and renal proximal tubules demonstrated subcellular colocalization with the endocytic receptor megalin, which is expressed in the same epithelia as the 460-kDa receptor. Furthermore, megalin affinity chromatography and surface plasmon resonance analysis revealed a calcium-dependent high affinity binding of the 460-kDa receptor to megalin, which thereby may mediate its vesicular trafficking. Due to the high number of CUB domains, accounting for 88% of the protein mass, we propose the name cubilin for the novel receptor.

\textsuperscript{*} This work was supported by the Novo Nordisk Foundation, the Danish Medical Research Council, the Danish Biomembrane Center, the Aarhus University Research Foundation, the Velux Foundation, the Foundation “Vaincre les Maladies Lysosomales,” National Institutes of Health First Award DK46117, NATO Collaborative Research Grant 94-0750, NASA Grant 9-811 Basic, a Veterans Affairs Career Development Award, a Tulane Cancer Center matching fellowship, and the Fondation “Vaincre les Maladies Lysosomales,” National Institutes of Health First Award DK46117, NATO Collaborative Research Grant 94-0750, NASA Grant 9-811 Basic, a Veterans Affairs Career Development Award, a Tulane Cancer Center matching fellowship, and the Danish Medical Research Council, the Danish Biomembrane Center, the Aarhus University Research Foundation, the Velux Foundation, the Foundation “Vaincre les Maladies Lysosomales,” National Institutes of Health First Award DK46117, NATO Collaborative Research Grant 94-0750, NASA Grant 9-811 Basic, a Veterans Affairs Career Development Award, a Tulane Cancer Center matching fellowship, and the Velux Foundation.

\textsuperscript{§} These authors have contributed equally to the experimental work.

\textsuperscript{¶} To whom correspondence should be addressed: Dept. Medical Biochemistry, University of Aarhus, DK-8000 Aarhus C, Denmark. Tel.: 45 89 422880; E-mail: skm@biobase.dk.

Intestinal uptake of vitamin B\textsubscript{12} (B\textsubscript{12}) is facilitated by a receptor recognizing B\textsubscript{12} in complex with gastric intrinsic factor (IF) for a review, see Ref. 1). Failure of either IF production or receptor expression leads inevitably to a B\textsubscript{12} deficiency state and disease (2, 3). Although the existence of a receptor for IF-B\textsubscript{12} was recognized nearly 3 decades ago (4), its biochemical properties and structure have remained elusive, possibly because of difficulties in purifying it in high amounts from the terminal ileum (5, 6). The observation by Seetharam et al. (7) that the receptor was expressed in considerably greater amounts in the kidney and yolk sac facilitated the production of antibodies inhibiting IF-B\textsubscript{12} binding (8, 9) and confirmed the receptor as a glycoprotein of \textgreater200 kDa. We have recently estimated a size of 460 kDa of the IF-B\textsubscript{12} affinity-purified receptor (9).

Since IF is only detected at very low levels in the circulation and in nongastrointestinal tissues (1), the high expression of the receptor in kidney and yolk sac may suggest that it has other functions/ligands. This is further indicated by several lines of evidence. First, patients with hereditary intestinal malabsorption of IF-B\textsubscript{12}, known as Imerslund-Gräsbeck syndrome (10, 11), have proteinuria, indicating that the receptor facilitating IF-B\textsubscript{12} uptake in the intestine is also important for normal kidney function. The significance of this observation is reinforced by the presence of proteinuria in a family of dogs that synthesize a nonfunctional receptor (12). Second, we have recently shown that the receptor binds RAP to a site distinct from the binding site for IF-B\textsubscript{12} (9). RAP is a chaperone-like protein that protects multiple ligand binding sites of processed low density lipoprotein receptor family proteins, in particular the two giant receptors megalin and low density lipoprotein receptor-related protein (13, 14). Finally, immunopathological studies indicate a key role of the IF-B\textsubscript{12}-binding protein in embryonic development. We have recently demonstrated that the target antigen (initially designated \textquotedblleft gp280\textquotedblright) of teratogenic antibodies (15) was associated with the endocytic pathway of yolk sac epithelial cells (16, 17) and identical to the IF-B\textsubscript{12}.

\textsuperscript{1} The abbreviations used are: B\textsubscript{12}, vitamin B\textsubscript{12}; BN cell, yolk sac–derived mouse sarcoma virus-transformed Brown Norway rat epithelial cell; EGF, epidermal growth factor; IF, intrinsic factor; RAP, receptor-associating protein; RACE, rapid amplification of cDNA ends; PBS, phosphate-buffered saline; PAGE, polyacrylamide gel electrophoresis; kb, kilobase pair(s).

\textsuperscript{2} Vitamin B\textsubscript{12} is cyanocobalamin. In the organism, cyanocobalamin is converted to the active forms of cobalamin, methyl- and 5′-deoxyadenosylcobalamin. The abbreviation B\textsubscript{12} is employed to cover all forms of cobalamin, cyanocobalamin that can be converted to the active form.

This paper is available on line at http://www.jbc.org
Molecular Characterization of Cubilin

EXPERIMENTAL PROCEDURES

Cubilin, Megalin, Ligands, and Antibodies—Purification of renal rabbit cubilin by IF-B12 affinity chromatography (9), megalin by recombinant RAP-affinity chromatography (20), and renal rat cubilin by immune-affinity chromatography (21) with monoclonal antibodies coupled to Sepharose 4B were carried out as previously reported. Human IF purified from gastric juice by B12-affinity chromatography (22) was a generous gift from Dr. Ebba Nexø. The monoclonal antibody, 5B-B11, recognizing rabbit megalin, the monoclonal antibody, 5A12, recognizing rabbit cubilin (9), the mouse monoclonal and the rabbit polyclonal antibodies against rat cubilin (18, 21), and the polyclonal sheep anti-megalin (20) have previously been described.

Amino Acid Sequencing and Estimation of N-Linked Carbohydrate of Cubilin—CNBr fragments and tryptic digests of a 100-kDa CNBr fragment of rat cubilin were purified by reverse phase high pressure liquid chromatography. Seven isolated peptides were subjected to Edman degradation using an Applied Biosystems model 477 A sequencer (data not shown). The fragmentation of the cDNA. The presence of the same mRNA in kidney and expression was confirmed, respectively, by sequencing a 3′-end RACE product and a 5′-end RACE product and a

RESULTS
dcDNA Cloning of Cubilin—By immunoscreening of a λZap cDNA library from rat yolk sac BN cells (17) we identified an initial 0.7-kb clone encoding a portion of cubilin. The sequence of this clone was used to design two nested primers to perform 5′-RACE on kidney cDNA, allowing identification of the 5′-end of cubilin. A number of clones were identified using polymerase chain reaction-generated probes for further screening of yolk sac libraries. Fig. 1 schematizes three overlapping clones used to construct the final cDNA. The last clone contained polyadenylated signal and poly(A) tail. The 5′- and 5′-ends of the 11.6-kb sequence were further confirmed, respectively, by sequencing a 3′-end RACE product and a 5′-end clone selected from a Cap Finder library. Northern blot analysis of yolk sac mRNA (Fig. 1) identified a mRNA of the same size as the cDNA. The presence of the same mRNA in kidney and intestinal mucosa, but not in liver, was confirmed by Northern blotting and reverse transcriptase polymerase chain reaction (data not shown).
Primary Structure of Cubilin—The assembled cDNA revealed an uninterrupted open reading frame of 10.8 kb encoding a 20-amino acid signal peptide and a 3603-amino acid protein with 42 potential N-glycosylation sites (Fig. 2). The molecular size of the protein backbone was calculated as 397 kDa. The seven amino acid sequences determined by N-terminal microsequencing of tryptic and CNBr peptides confirmed the identity of the sequence (boldface letters in Fig. 2). The size of the protein was confirmed by SDS-PAGE. As shown in Fig. 3, deglycosylation of the IF-B12 affinity-purified rabbit receptor by peptide N-glycosidase F increased its electrophoretic mobility corresponding to a size of 400 kDa. Compared with the 460-kDa size of the untreated glycoprotein, this indicates a carbohydrate content of \( \frac{13}{100} \)%.

Fig. 4 shows the predicted domain organization of the receptor. A stretch of approximately 110 amino acids with no apparent homology to known proteins is followed by a cluster of eight EGF type B repeats preceding 27 contiguous CUB domains accounting for 88% of the protein mass. The high degree of internal homology (overall similarity of 45%) between the CUB domains is evident from the dot plot display in Fig. 4B. A total of 76 disulfide bridges is predicted if all of the extracellular modules fold normally. The only cysteine outside of the CUB domains and EGF repeats is located in the 110-amino acid N-terminal sequence. This residue might account for the partial disulfide bond-dependent dimerization of a minor part of purified receptors (9, 17).

Except for the leader peptide, no sequence compatible with a transmembrane domain could be identified. This excludes the possibility that the protein is a type 1 membrane protein or a glycosylphosphatidylinositol-anchored protein that is synthesized with a cleavable hydrophobic C terminus. Furthermore, since almost the entire protein sequence consists of extracellular modules, it is very unlikely that the protein is a type II or III protein with a noncleaved hydrophobic signal peptide inserted in the membrane (28).

Cubilin Is a Peripheral Membrane Protein—To verify that cubilin is a peripheral membrane protein, as predicted by the lack of a transmembrane segment and cytoplasmic tail, we investigated its release from renal cortex membranes by protease digestion.

Primary Structure of Cubilin—The assembled cDNA revealed an uninterrupted open reading frame of 10.8 kb encoding a 20-amino acid signal peptide and a 3603-amino acid protein with 42 potential N-glycosylation sites (Fig. 2). The molecular size of the protein backbone was calculated as 397 kDa. The seven amino acid sequences determined by N-terminal microsequencing of tryptic and CNBr peptides confirmed the identity of the sequence (boldface letters in Fig. 2). The size of the protein was confirmed by SDS-PAGE. As shown in Fig. 3, deglycosylation of the IF-B12 affinity-purified rabbit receptor by peptide N-glycosidase F increased its electrophoretic mobility corresponding to a size of 400 kDa. Compared with the 460-kDa size of the untreated glycoprotein, this indicates a carbohydrate content of \( \frac{13}{100} \)%.

Fig. 4 shows the predicted domain organization of the receptor. A stretch of approximately 110 amino acids with no apparent homology to known proteins is followed by a cluster of eight EGF type B repeats preceding 27 contiguous CUB domains accounting for 88% of the protein mass. The high degree of internal homology (overall similarity of 45%) between the CUB domains is evident from the dot plot display in Fig. 4B. A total of 76 disulfide bridges is predicted if all of the extracellular modules fold normally. The only cysteine outside of the CUB domains and EGF repeats is located in the 110-amino acid N-terminal sequence. This residue might account for the partial disulfide bond-dependent dimerization of a minor part of purified receptors (9, 17).

Except for the leader peptide, no sequence compatible with a transmembrane domain could be identified. This excludes the possibility that the protein is a type 1 membrane protein or a glycosylphosphatidylinositol-anchored protein that is synthesized with a cleavable hydrophobic C terminus. Furthermore, since almost the entire protein sequence consists of extracellular modules, it is very unlikely that the protein is a type II or III protein with a noncleaved hydrophobic signal peptide inserted in the membrane (28).

Cubilin Is a Peripheral Membrane Protein—To verify that cubilin is a peripheral membrane protein, as predicted by the lack of a transmembrane segment and cytoplasmic tail, we investigated its release from renal cortex membranes by protease digestion.
cedures that do not involve solubilization of the membranes or enzymatic treatment. Fig. 6 shows the identical size of the receptor from kidney, yolk sac, and intestinal mucosa (lanes 1–3). As seen in lane 4 versus lane 3, approximately 50% of cubilin was released into the fluid phase by mechanical grinding of renal cortex in PBS, whereas megalin, the 600-kDa transmembrane protein (29) expressed in the same tissues, was released in minimal amounts. IF-B_{12} affinity chromatography performed as recently described for solubilized membranes (9), but in the present experiment, without the use of detergent in the running or elution buffer, yielded a comparable amount of purified receptor (data not shown). Cubilin, which remained membrane-associated, was tightly bound but could be released partly by EDTA, heparin, and, to a low extent, phosphorylethanolamine (Fig. 6, lanes 5–11). Heparin and phosphorylethanolamine have previously been reported to bind to the spermadhesin CUB domain (30, 31). The same treatments released virtually no megalin (Fig. 6). The size of the released cubilin, as estimated by SDS-PAGE, was not different from the membrane-associated cubilin.

Cubilin Traffics with Megalin—Previous studies have demonstrated the presence of megalin and cubilin in the same endocytic vesicles of the same absorptive epithelia in the intestine, kidney, and yolk sac (9, 16, 21). Fig. 7 shows electron microscopic examination of rat yolk sac and kidney sections subjected to double immunogold labeling using a sheep anti-megalin polyclonal antibody and a mouse anti-cubilin monoclonal antibody. The large gold particles label the megalin antibody and the small particles label the cubilin antibody. An almost identical localization of the two sizes of gold particles is seen. This close localization of the particles led us to test for the formation of cubilin-megalin complexes.

Cubilin Binds Megalin—As shown in Fig. 8, \(^{125}\text{I}-\text{cubilin}\) binds to megalin covalently linked to Sepharose 4B. Bound radiolabel was released from the column by EDTA. Surface plasmon analysis (Fig. 9A) confirmed this binding. No difference in the rate of dissociation of cubilin from megalin was seen in the pH interval 4–8 (not shown). Binding of cubilin to megalin was reduced partially (~75%) when RAP was pre-bound to megalin, indicating that cubilin binds to the extracellular domain of megalin. Megalin-bound cubilin was still capable of binding IF-B_{12} as shown by subjecting the megalin chip to flow with IF-B_{12} after the binding of cubilin (Fig. 9B). Thus, the response after adding IF-B_{12} represents the formation of a megalin-cubilin-IF-B_{12} complex. Control experiments showed no binding of IF-B_{12} to megalin (Ref. 9 and data not shown).

**DISCUSSION**

Cubilin: A Novel Type of Multifunctional Giant Receptor—The present study provides novel molecular information on cubilin, previously known as the yolk sac target antigen of teratogenic antibodies and the intestinal receptor for IF-B_{12}. The primary structure predicts 35 extracellular modules uniquely organized in a cluster of eight EGF repeats followed from a molecular point of view, a huge cluster of 27 CUB domains accounting for 88% of the protein. Northern and Western blotting of kidney, yolk sac, and intestine did not indicate differences in size of the receptor in these organs.

The EGF type B repeats are similar to the carboxyl-terminal extracellular modules of megalin and low density lipoprotein receptor-related protein. Cubilin has otherwise very little homology to these two giant receptors, which also bind RAP and...
mediate endocytosis of a variety of ligands. Also, cubin does not display homology to sortilin, the 95-kDa putative vesicular sorting receptor, which also binds RAP (32).

The CUB domains conform to the description of Bork and Beckmann (27) based on the analysis of 31 copies of a module initially identified in the C1r and C1s components of complement and subsequently in a variety of proteins associated with fetal development. They consist of 110 amino acids defining a characteristic hydrophobicity pattern predicted to form antiparallel \( \beta \)-barrels (Refs. 33–35 and see “Note Added in Proof”).

The four conserved cysteines, generally thought to form two S-S bridges (Cys1–Cys2, Cys3–Cys4), are found in all but domain 13 of cubilin that lacks the first two cysteines as already described in the first CUB domains of C1r/s and the homologues MASP1/2 (see Ref. 36 and references therein). When analyzed individually, the CUB domains of cubilin are more closely related to those seen in developmental control proteins.

On the functional level, there is compelling evidence that the CUB domains are involved in the binding of proteins, as described for the \( \text{Ca}^{2+} \)-dependent formation of the C1 complex (37), as well as for binding of phospholipids and carbohydrates, as demonstrated for spermadhesins (30, 31, 38). In addition to the CUB domains, the EGF repeats might also account for some of the binding properties of cubilin. EGF repeats are widely expressed and participate in a number of receptor-ligand interactions (for a review, see Ref. 39). Two of the EGF repeats in cubilin have the consensus sequence for calcium binding (26, 40) and may be involved in the calcium-dependent binding of e.g. RAP or IF-B12 (9).

Membrane Binding and Internalization of Cubilin—The lack of a transmembrane segment was surprising because cubilin is internalized via clathrin-coated organelles (18) and recycles to the membrane (19). However, early studies have indicated that both an intrinsic factor-B12-binding protein (6) and the target protein of teratogenic antibodies (41) could be released, at least in part, from intestinal or renal tissue using mechanical dissociation in the absence of detergents. We further showed that some cubilin was released by heparin, phosphorylethanol-
amine, and EDTA, whereas the membrane association was resistant to acid conditions. These observations are in line with the membrane binding properties of spermadhesins. These proteins consist of a single CUB domain and lack a transmembrane segment yet are tightly bound to the surface of sperm cells via nonionic interactions with phospholipids (31), whereas another region of the CUB domain is free to bind to carbohydrates of the zona pellucida surrounding the mammalian egg.

The lectin binding characteristics of the spermadhesins are not fully characterized but include heparin and a variety of carbohydrates (30, 38). In view of its 27 CUB domains, cubilin may link to the membrane in a heterogeneous manner via multiple sites, which might account for our inability to release it entirely from the membrane.

The identification of the membrane components interacting with cubilin is also essential for explaining its internalization and recycling. The present data show that the cubilin is capable of binding to the endocytic receptor megalin, which colocalizes at the subcellular compartments of yolk sac, kidney, and probably also the intestine (9). Co-internalization of a receptor that lacks internalization signal(s) by means of another receptor has previously been shown. Thus, the glycosylphosphatidylinositol-anchored urokinase receptor is endocytosed by coupling of urokinase receptor-bound urokinase–inhibitor complex to low density lipoprotein receptor-related protein (42, 43). Megalin, which also binds the urokinase–inhibitor complex (20), may perform a similar function in regard to both the urokinase receptor- and cubilin-ligand complexes. Once internalized, IF-B12 is segregated from the receptor and directed to lysosomes for degradation of IF, whereas the receptor is recycled to the membrane (9, 19, 44). Since the in vitro cubilin-megalin complex is stable at pH 5, the two receptors might remain in complex during the entire recycling pathway. In contrast, the urokinase receptor recycles to the plasma membrane without being linked to the low density lipoprotein receptor-related protein (42). When we analyzed the effect of polyclonal megalin antibodies and RAP on the endocytosis of 125I-IF-B12 in uptake in cultured yolk sac cells, we only found a reduction of 10–15%. This modest effect might be accounted for by a short cell surface expression of cubilin and megalin due to a rapid recycling of the two proteins and thereby a too short time for the cubilin-megalin to dissociate, a prerequisite for RAP to block binding. Furthermore, a continuous incubation with RAP will probably have no effect on intracellular receptors, since exter-

4 S. K. Moestrup and P. J. Verroust, unpublished data.
addition of cholesterol to the Sonic hedgehog protein. It is therefore intriguing to speculate that the anti-cubilin antibodies could interfere with cholesterol uptake either directly or indirectly via binding of cubilin to megalin in the yolk sac.

In conclusion, the present data establish cubilin as a novel type of peripheral membrane receptor with multiple potential sites for interaction with other proteins and membrane components. It is at present known that cubilin can bind IF-B₁₂, RAP, megalin, and most likely also calcium, phospholipids, and carbohydrates. However, a number of ligands may remain to be identified to explain the role of the receptor in kidney function and its importance in fetal development. Our future studies on this protein will concentrate on its multiple interactions to define its role as a receptor for ligands in the fluids lining the yolk sac, ileum, and kidney epithelium.

Acknowledgments—Technical assistance by Merete Greffe, Kirsten Lassen, Marthe Galleran, and Hanne Sidelmann is gratefully acknowledged. We thank Christian Jacobsen for helping with the analysis of the BLAcore data; Ebbi Nees and Jørgen Gliemann for continuous support and helpful suggestions; and Pierre Fouret and D. Sahali for assistance in the initiation of the project.


REFERENCES
Molecular Characterization of Cubilin

The Intrinsic Factor-Vitamin B_{12} Receptor and Target of Teratogenic Antibodies Is a Megalin-binding Peripheral Membrane Protein with Homology to Developmental Proteins

Søren K. Moestrup, Renata Kozyraki, Mette Kristiansen, James H. Kaysen, Hanne Holm Rasmussen, Didier Brault, Françoise Pontillon, Fatime O. Goda, Erik I. Christensen, Timothy G. Hammond and Pierre J. Verroust

*doi: 10.1074/jbc.273.9.5235*

Access the most updated version of this article at [http://www.jbc.org/content/273/9/5235](http://www.jbc.org/content/273/9/5235)

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

This article cites 51 references, 18 of which can be accessed free at [http://www.jbc.org/content/273/9/5235.full.html#ref-list-1](http://www.jbc.org/content/273/9/5235.full.html#ref-list-1)