Cubilin, the receptor for intrinsic factor-vitamin B_12, is a novel type of high molecular weight receptor consisting of a 27 CUB (complement components C1r/C1s, Uegf, and bone morphogenetic protein-1) domain cluster preceded by 8 epidermal growth factor repeats and a short N-terminal sequence. In addition to binding the vitamin B_12-cARRIER complex, cubilin also binds receptor-associated protein. To delineate the structures for membrane association and ligand binding we established a panel of stable transfected Chinese hamster ovary cells expressing overlapping segments of rat cubilin. Analysis of conditioned media and cell extracts of transfected cells revealed that the N-terminal cubilin region conveys membrane association. Helical plotting of this region demonstrated a conserved amphipathic helix pattern (Lys74–Glu109) as a candidate site for hydrophobic interactions. Ligand affinity chromatography and surface plasmon resonance analysis of the secreted cubilin fragments showed ligand binding in the CUB domain region. Further dissection of binding-active fragments localized the binding site for intrinsic factor-vitamin B_12 to CUB domains 5–8 and a receptor-associated protein-binding site to CUB domains 13–14. In conclusion, the N-terminal cubilin region seems crucial for membrane association, whereas the CUB domain cluster harbors distinct sites for ligand binding.

Uptake of dietary vitamin B_12/cyanocobalamin (B_12) depends on gastric intrinsic factor (IF), which undergoes a specific and strong complex formation with the vitamin in the intestinal lumen (1). The vitamin-carrier complex is subsequently internalized after being recognized by the high affinity membrane receptor, cubilin (2, 3). The binding of IF-B_12 is suggested to occur via the N-terminal region of IF (4) and only when B_12 is bound to the carrier protein (1–3, 5). The physiological importance of the specific B_12 carrier and its receptor is evident from the B_12 deficiency, characterized by the megaloblastic anemia and neurological symptoms related to defective functional expression of IF or cubilin (6–11). In addition to binding IF-B_12, cubilin also binds receptor-associated protein (RAP) (5), and growing evidence suggests that the receptor has other biological functions. This is in particular indicated by the fact that kidney and yolk sac epithelia exhibit a high cubilin expression (2, 5, 12), albeit IF is synthesized in the gastric epithelium and virtually only is present in the gastrointestinal lumen.

The recent determination of the primary structure of the rat and human cubilin molecule (13, 14) revealed an ~3600-aa protein with a short N-terminal segment followed by 8 epidermal growth factor (EGF) repeats and a large cluster of 27 CUB domains. No classical transmembrane segment has been predicted, and the overall structure is distinct from known endocytic receptors, including the RAP-binding members of the low density lipoprotein (LDL) receptor family (15). Two mutations, one causing a Pro/Leu substitution in CUB domain 8 and another suggested to activate an intrinsic cryptic splice site leading to truncation of the receptor in CUB domain 6, have recently been identified in the cubilin gene of patients with inherited B_12 malabsorption (11).

To link the structural and functional features of cubilin, we established a comprehensive panel of transfected Chinese hamster ovary (CHO) cells expressing various segments of the receptor. Biochemical analyses of the cells and expression products now delineate the regions important for membrane association and binding of IF-B_12 and RAP.

**EXPERIMENTAL PROCEDURES**

**Ligands and Receptors**—Porcine and human IF-B_12 were purified from gastric mucosa extract as described (16). Human RAP was produced as a recombinant protein in *Escherichia coli*. Rabbit cubilin was purified from renal cortex by IF-B_12 and RAP affinity chromatography, as described previously (5).

**Construction of Plasmids for Expression of Recombinant Rat Cubilin Fragments**—Cubilin cDNA fragments extended with enzyme restriction sites were amplified by polymerase chain reaction with the Expand™ High Fidelity PCR System (Roche Molecular Biochemicals) and purified with the QIAEX II gel extraction kit (Qiagen, CA). The polymerase chain reaction products were subcloned into the expression vector (pSecTag2B or pcDNA3.1/Zeo(−)) from Invitrogen, Groningen, The Netherlands) by use of the appropriate restriction enzymes (New England Biolabs, Beverly, MA; Amersham Pharmacia Biotech) and the T4 DNA ligase (Amersham Pharmacia Biotech). Plasmids were transformed using XL1-Blue competent cells (Stratagene, LaJolla, CA), and plasmid DNA was isolated by the Qiagen Maxiprep method (Qiagen) and sequenced before transfection as described previously (13). The following 14 constructs were subcloned and expressed.

**N-EGF1–8**

---

* The work was financially supported by the Novo Nordisk Foundation, the Danish Medical Research Council, the Danish Bioblane Center, the Velux Foundation, and the John and Birthe Meyer Foundation, the Danish Medical Research Council, the Danish Biomembrane Center, and the John and Birthe Meyer 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.

**To whom correspondence should be addressed:** Dept. Medical Biochemistry, University of Aarhus, Ole Worms Allé Bldg 170, DK 8000 Aarhus C, Denmark. Tel.: 45 89 422880; E-mail: skm@biobase.dk.

*The abbreviations used are: B_12, vitamin B_12; CHO, Chinese hamster ovary; EGF, epidermal growth factor; IF, intrinsic factor; RAP, receptor-associated protein; SPR, surface plasmon resonance; aa, amino acids; bp, base pairs; LDL, low density lipoprotein.
Cubilin Regions Important for Membrane and Ligand Binding

**FIG. 1.** The initial strategy for mapping functional sites in rat cubilin by expression of fragments in stable transfected CHO cells. Overlapping fragments were designed according to the structural organization of the 3603-aa rat cubilin molecule, which contains a 109-aa N-terminal region (aa 1–109) including a suggested propeptide (aa 1–12), a cluster of 8 EGF repeats (aa 110–450), and 27 CUB domains (aa 454–3603). The EGF–1, CUB1–3, CUB7–14, CUB13–20 and CUB19–27 constructs were expressed by use of the murine Ig κ leader peptide. The N-EGF–1+ CUB3–1 product was expressed by use of the rat cubilin leader peptide, whereas the N-EGF product was expressed in two variants, one by using the rat cubilin leader peptide (aa 20–(−1)) and another by using the murine Ig κ leader peptide.

**TABLE 1.** Cubilin Regions Important for Membrane and Ligand Binding

| Region | Length | Membrane Binding | Ligand Binding |
|--------|--------|------------------|----------------|
| N-EGF | 1–8    | Yes              | Yes            |
| EGF   | 1–8    | Yes              | Yes            |
| CUB1–3| 1–190  | Yes              | Yes            |
| CUB4–6| 20–190 | Yes              | Yes            |
| CUB7–14| 200–2784 | Yes       | Yes            |
| CUB13–20| 2785–3603 | Yes     | Yes            |
| CUB19–27| 3604–3603 | Yes    | Yes            |

**RESULTS**

Identification of the Membrane Binding Region in Cubilin—Fig. 1 shows the structural elements of cubilin and the initial strategy for mammalian expression of cubilin fragments in stable transfected CHO cells. The EGF repeat and CUB domain products were designed to overlap to ensure that the investigated binding sites that might encompass several domains were present in at least one expression product.

Immunoblotting of the conditioned media from transfected cells showed an effective secretion of all expression products except for the products containing the N-terminal region (Fig. 2). For example, the eight-EGF repeat region (EGF1–8 product) was effectively secreted, whereas the same region added to nonreducing SDS-polyacrylamide gel electrophoresis (8–16% polyacrylamide) prior to the electrophoretic and immunodetection with a rabbit anti-rat cubilin IgG. The positions of molecular size standard markers are indicated.

Immunoblotting of media and cells of cubilin cDNA plasmid-transfected CHO cells. Lanes 1–7 show the immunoblotting of the indicated cubilin fragments in the conditioned media. All fragments except for the constructs containing the N-terminal region were present in the medium. Lanes 8 and 9 show that the two fragments absent in the media were expressed in the cells. Twenty μl of media and 20 μg of Triton X-100 (1%) solubilized cells were subjected to nonreducing SDS-polyacrylamide gel electrophoresis (8–16% polyacrylamide) prior to the electrophoretic and immunodetection with a rabbit anti-rat cubilin IgG. The positions of molecular size standard markers are indicated.
Cubilin Regions Important for Membrane and Ligand Binding

were retained in the cells, whereas the expression products without this segment were virtually absent in the cells, thus indicating an immediate secretion upon synthesis. Immunofluorescence microscopy of the cells (not shown) expressing the N-terminal region indicated that the expression product was localized to the membrane and intracellular vesicles as previously reported for native cubilin in cultured yolk sac cells (13). Using either the cubilin signal peptide or the murine Ig kappa chain signal peptide as the leader sequence in the construct encoding the N-terminal product (NEGF1-8) led to cellular retention of the expressed protein. This indicates that the membrane association of the N-terminal region of cubilin is not because of membrane binding of a noncleaved leader peptide.

Identification of the Regions for Binding of IF-B12 and RAP—Ligand affinity precipitation of the expression products displayed in Fig. 2 with IF-B12-Sepharose beads (Fig. 4) revealed IF-B12 binding to the CUB1–8 product, whereas none of the other regions exhibited any binding. Because no binding was seen to the EGF1–8 product, shorter segments of the CUB domain regions 4–8 and 13–20 were expressed. The CUB5–8 product was effectively precipitated by IF-B12 beads, whereas CUB4–6 was not. These data combined with the absent IF-B12 binding to CUB7–14 indicate that the binding site for IF-B12 is localized within CUB domains 5–8 and is dependent on the structure of several domains within this region. Similar expression of human cubilin CUB domains 5–8 also revealed a binding site for IF-B12 (data not shown).

Identification of Recombinant Cubilin Domains—Recombinant CUB domain regions containing the distinct sites for IF-B12 (CUB1–8 and CUB5–8) and RAP (CUB7–14 and CUB13–20) were readily purified to homogeneity by IF-B12 and RAP affinity chromatography. Figs. 6 and 7 show the purification of CUB5–8 and CUB7–14 and the ligand-fragment interactions as recorded by SPR analysis. These data further confirmed that the sites for IF-B12 and RAP are distinct. The K_d for human and porcine IF-B12 binding to CUB5–8 was ~35 nM and ~45 nM, respectively, versus ~5 nM for the binding to native rabbit and human cubilin (Refs. 13 and 14 and data not shown). The K_d for binding of RAP was ~70 nM versus ~170 nM for the binding to native rabbit cubilin (13).

DISCUSSION

In the present study we used a mammalian expression approach to identify the cubilin regions important for membrane association and the binding of IF-B12 and RAP. Fig. 8 is a schematic presentation of the present mapping of functional regions in cubilin.

Membrane Association of Cubilin—Cubilin is a membrane protein although more loosely attached to the membrane compared with other membrane proteins (13). The present expression data now indicate that the N-terminal cubilin region is crucial for membrane association. The region has no classical transmembrane segment, but helical plotting demonstrated a conserved amphipathic helix motif similar to the lipid-embedded amphipathic helix structures described in apolipoproteins (17). Membrane association by means of this structure may be a hairpin-like anchoring with the extreme N-terminal end facing the noncytoplasmic environment. This will allow the

![Fig. 3. Identification of a putative amphipathic α-helix structure in cubilin.](Image 1)

![Fig. 4. Identification of IF-B12 binding recombinant cubilin fragments.](Image 2)
N-terminal trimming by furin, the Golgi proteinase suggested to recognize cubilin and induce cleavage after the furin recognition sequence preceding Arg12 in rat cubilin (14). Alternatively, the amphipathic region may be involved in the assembly of cubilin. A very recent electron microscopic study of purified bovine cubilin has shown the formation of trimers in solution without detergent and lipid (18). A region of four heptad repeats, actually a part of the putative amphipathic helix structure, has been proposed to account for the assembly. Future structural studies of membrane-associated cubilin may further elucidate how the receptor is physically attached to the membrane and how it is organized there. Maybe cubilin is functionally a monomer, but when shed from the membrane, the short hydrophobic region causes self-assembly into homodimer/homotrimers.

The IF-B12 Binding and RAP Binding Sites in Cubilin—The identification of the binding site for IF-B12 in the region encompassing CUB domains 5–8 establishes the CUB domain structure as a ligand binding domain. Although this domain has long been known as a structural entity in other membrane proteins, its biological function in these proteins is largely unknown. In view of the present mapping data, it is tempting to suggest that some other identified receptor-like membrane proteins with CUB domains, e.g. SRP-ductin (19) are orphan receptors with ligands remaining to be identified.

The recently determined three-dimensional CUB domain structure (20) agrees well with a ligand binding function. The CUB domain is a barrel-like structure containing two layers of five-stranded β-sheets with the β-turns in a surface-exposed position as in the antigen binding regions of immunoglobulins. Furthermore, structural analysis of CUB domain dimers (21) suggests that the dimers may associate via the β-sheet layers. If the CUB domains of cubilin are arranged in this way, on top of each other and with the β-sheets facing, it will leave the less conserved β-turns of the CUB domains available for interactions. In view of the existence of 27 CUB domains in cubilin, a high number of putative sites for various protein interactions may then be predicted.

Another interesting perspective arising from the recognition of the binding site for IF-B12 in cubilin relates to the recent...
identification of mutations in cubilin of patients with autosomal recessive B₁₂ malabsorption (Imerslund-Gräsbeck syndrome) (9, 10). The two mutations identified in the human cubilin gene (11) both relate to the critical region now known to harbor the IF-B₁₂ binding region. The mutation, suggested to activate a cryptic intronic splice site leading to truncation of the receptor in CUB domain 6, may lead to an inactive translation product, whereas the P→L substitution in CUB domain 8 might cause a structural change specifically impairing IF-B₁₂ binding. Further expression and mutational studies are now in progress to identify the crucial residues of the IF-B₁₂ binding site and the structural-functional implications of this mutation in human cubilin.

A site for binding of RAP was localized within CUB domains 13–14. The physiological importance of this interaction remains to be determined, but it is tempting to speculate that RAP may assist the processing/folding of cubilin in line with its suggested chaperone function and importance for normal processing of LDL receptor-related protein and megalin (22). Interestingly, the RAP binding to cubilin differs from the RAP binding to other receptors in the sense that RAP is not a general inhibitor of ligand binding in cubilin.

The estimated $K_d$ values for binding of IF-B₁₂ was 6–9-fold higher than the estimated affinities of IF-B₁₂ to native rabbit and human cubilin (12, 13), whereas the estimated $K_d$ values for binding of RAP to recombinant and native cubilin only differed about 2-fold. The lower affinity of the IF-B₁₂ binding to CUB5–8 may be explained by the small size of the fragment making it more vulnerable to immobilization compared with full-length cubilin. Alternatively, other parts of the receptor are required for complete structural integrity of the binding site. However, the existence of other independent sites for IF-B₁₂ binding seems unlikely in view of the absent IF-B₁₂ binding to other expressed regions of cubilin.

No specific function has yet been dedicated the cluster of EGF repeats in cubilin. Its position between the membrane-associated segment and the CUB domain region might indicate a structural role as a spacing segment for positioning the ligand binding region in an appropriate distance from the membrane. Interestingly, a similar cluster of EGF repeats is positioned between the transmembrane segment and the ligand binding regions of LDL receptor-related protein (23). Neither has a function of these EGF repeats, constituting the extracellular part of the LDL receptor-related protein β-subunit, been defined apart from binding the LDL receptor-related protein α-subunit.

In conclusion, the present molecular dissection of cubilin has identified the regions important for membrane association and binding of IF-B₁₂ and RAP. The high number of cubilin CUB domains suggests a high number of CUB domain interactions. To address this question we are currently investigating the existence of other cubilin ligands binding to the CUB domain region.

Acknowledgment—We thank Kirsten Lassen for excellent technical assistance.

REFERENCES

1. Nexø, E. (1998) in Vitamin B12 and B12 Proteins (Kraeutler, B., Adrigoni, D., and Golding B. T., eds) pp. 461–475, Wiley-vch Verlag, Weinheim, Germany
2. Seetharam B., Seetharam, S., Li, N., and Ramanujam, K. S. (1994) in Advances in Thomas Addison’s Diseases (Bhatt H. H., James, V. H. T., Bessey, G. M., Bottazo, G. F., and Keen, H., eds) Vol II, pp. 393–406, Journal of Endocrinology Ltd., Bristol, UK
3. Moestrup, S. K. in Vitamin B12 and B12 Proteins (Kraeutler, B., Adrigoni, D., and Golding B. T., eds) pp. 477–489, Wiley-vch Verlag, Weinheim, Germany
4. Tang, L. H., Chokshi, H., Hu, C. B., Gordon, M. M., and Alpers, D. H. (1992) J. Biol. Chem. 267, 22862–22866
5. Birn, H., Verroust, P. J., Nexø, E., Hager, H., Jacobsen, C., Christensen, E. I., and Moestrup, S. K. (1997) J. Biol. Chem. 272, 26497–26504
6. Gueant, J. L., Saunier, M., Gastin, I., Safi, A., Lamireau, T., Duclos, B., Bigard, M. A., and Grasbeck, R. (1995) Gastroenterology 108, 1622–1628
7. Fyfe, J. C., Ramanujam, K. S., Ramsamy, M., Patterson, D. F., and Seetharam B. (1991) J. Biol. Chem. 266, 4489–4494
8. Yang, Y. M., Ducos, R., Rosenberg, A. J., Catrou, P. G., Levine, J. S., Podell, E. R., and Allen, R. H. (1985) J. Clin. Invest. 76, 2057–2065
9. Imerslund, O. (1960) Acta Pathol. Scand. 1, 1–15
10. Grasbeck, G. R. (1960) Acta Med. Scand. 167, 289–296
11. Aminoff, M., Carter, J. E., Chadwick, R., Grasbeck, R., Abdelaal, M. A., Broch, H., Jenner, L. B., Verroust, P. J., Moestrup, S. K., de la Chapelle, A., and Krahe, R. (1999) Nutr. Genet. 21, 309–313
12. Sahali, D., Mulliez, N., Chatete, P., Dupuis, R., Ronco, P., and Verroust, P. (1988) J. Exp. Med. 167, 213–218
13. Moestrup, S. K., Kozyraki, R., Kristiansen, M., Kayser, J. H., Rasmussen, H. H., Braak, D., Pontillon, F., Geda, F. O., Christensen, E. I., Hammond, T. G., and Verroust, P. J. (1998) J. Biol. Chem. 273, 5235–5242
14. Kozyraki, R., Kristiansen, M., Silahlartogi, A., Hansen, C., Jacobsen, C., Tommerup, N., Verroust, P. J., and Moestrup, S. K. (1998) Blood 91, 3593–3600
15. Gismann, J. (1998) Biochim. Biophys. Acta 1391, 961–964
16. Nexø, E., and Olesen, H. (1976) Biochim. Biophys. Acta 446, 143–147
17. Segrest, J. P., De-Loof, H., Dehlin, J. G., Brouilette, C. G., and Anantharamaiah, G. M. (1996) Proteins 8, 103–117
18. Lindblom, A., Quadt, N., Marsh, T., Adrigoni, D., Morgenl, M., Mann, K., Maurer, P., and Paulsson (1999) J. Biol. Chem. 274, 6374–6380
19. Cheng, H. Bjerkmans, M., and Chen, H. (1996) Anot. Rca 244, 327–343
20. Diao, J. M., Carvalho, A. L., Koll, I., Calvette, J. T., Topfer, P. E., Varela, P. F., Romero, A., Urbanke, C., and Romao, M. J. (1997) Protein Sci. 6, 725–727
21. Varela, P. F., Romero, A., Sanz, L., Romao, M. J., Topfer, P. E., and Calvette, J. T. (1999) J. Biol. Chem. 274, 635–649
22. Willnow, T. E., Armstrong, S. A., Hammer, R. E., and Herz, J. (1995) Proc. Natl. Acad. Sci. U. S. A. 92, 4537–4541
23. Herz, J., Hamann, V., Rogus, S. and Myklebost, O., Gausepohl, M., and Stanley, K. K. (1988) EMBO J. 7, 4119–4127