Assembly of the Intrinsic Factor Domains and Oligomerization of the Protein in the Presence of Cobalamin †

ARTICLE *in* BIOCHEMISTRY · DECEMBER 2004

Impact Factor: 3.02 · DOI: 10.1021/bi048924c · Source: PubMed

6 AUTHORS, INCLUDING:

Sergey Fedosov
Odessa National Academy of Food Technolo…
65 PUBLICATIONS **1,116** CITATIONS
SEE PROFILE

Natalya U Fedosova
Aarhus University
59 PUBLICATIONS **829** CITATIONS
SEE PROFILE

Soren Moestrup
University of Southern Denmark and Universi…
244 PUBLICATIONS **12,422** CITATIONS
SEE PROFILE

Ebba Nexø
Aarhus University Hospital
411 PUBLICATIONS **9,820** CITATIONS
SEE PROFILE

Assembly of the Intrinsic Factor Domains and Oligomerization of the Protein in the Presence of Cobalamin[†]

Sergey N. Fedosov,*[,‡] Natalya U. Fedosova,[§] Lars Berglund,[||] Søren K. Moestrup,[⊥] Ebba Nexø,[#] and Torben E. Petersen[‡]

Protein Chemistry Laboratory, Department of Molecular and Structural Biology, University of Aarhus, Science Park, Gustav Wieds Vej 10, Department of Biophysics, University of Aarhus, Ole Worms Alle 185, Cobento Biotech A/S, Science Park, Gustav Wieds Vej 10, Department of Medical Biochemistry, University of Aarhus, Ole Worms Alle 170, and Department of Clinical Biochemistry, AS Aarhus University Hospital, Nørrebrogade 44, 8000 Aarhus C, Denmark

Received May 26, 2004; Revised Manuscript Received September 3, 2004

ABSTRACT: Human intrinsic factor (IF) was purified from the recombinant plant *Arabidopsis thaliana* by affinity chromatography. Cobalamin (Cbl) saturated protein was separated by gel filtration into peaks I and II, which contained according to SDS electrophoresis the 50 kDa full-length protein IF_{50} and a mixture of two fragments, respectively. Two components of peak II were identified as the 30 kDa N-terminal peptide IF_{30} and the 20 kDa C-terminal glycopeptide IF_{20}. Measurements of M_w under the nondenaturing conditions were conducted by static light scattering. They revealed 100 kDa IF dimers in peak I, whereas 50 kDa cleaved monomers were found in peak II. The protein devoid of Cbl dissociated to the elementary units incapable of association in the absence of Cbl. The individual proteolytic fragments bound Cbl at high concentration of the ligand; however, neither IF_{30}·Cbl nor IF_{20}·Cbl oligomerized. A mixture of two fragments $IF_{30} + IF_{20}$ and Cbl produced a firm complex, IF_{30+20}·Cbl, which could not associate to dimers. In contrast to IF_{30+20}·Cbl, the saturated full-length monomers IF_{50}·Cbl dimerized with $K_d \approx 1$ μM. We suggest a two-domain organization of the full-length protein, where two distant units, IF_{30} and IF_{20}, can be assembled only by Cbl. They are connected by a protease-sensitive link, whose native structure is likely to be important for dimerization. However, linkage between two domains is not compulsory for Cbl binding. Advantages of the two-domain structure of IF are discussed.

Intrinsic factor (IF)[1] is a cobalamin (Cbl, vitamin B_{12}) binding protein important for adsorption of the vitamin during digestion (*1*−*3*). Cbl of dietary source binds with high affinity to IF in the small intestine, and their complex is internalized with the help of the multiligand receptor cubilin-amnionless (*4*, *5*). Two other Cbl-specific proteins, transcobalamin and haptocorrin, carry the vitamin in plasma and other body fluids (*1*, *2*). The recent studies performed on the recombinant proteins helped to establish some fundamental properties of the Cbl transporters including details of the ligand binding (*6*−*9*), structure of the disulfide bridges (*6*, *10*), and approximate location of IF's glycosylation site (*11*).

IF stands somewhat apart from two other Cbl carriers. First, IF binds Cbl with exceptionally high selectivity and can distinguish it among different analogues (*9*, *12*). Second, only the saturated holo-IF interacts with the receptor (*11*, *13*). The mechanisms behind the above features are still unknown. The conventional scheme of interaction between Cbl, IF, and the receptor implies a stoichiometry of 1:1. Another model suggests oligomerization of gastric IF upon Cbl binding (*1*, *14*−*16*). However, the latter effect was irreproducible (*1*, *15*) which questioned its relevance.

In the present paper we describe the Cbl-induced oligomerization of recombinant IF from plants. The protein was purified as both the full-length form of 50 kDa (IF_{50}) and two cleavage products of 30 kDa and 20 kDa (IF_{30} and IF_{20}). Binding of Cbl to IF_{50} caused dimerization of the protein. Two IF fragments assembled together in a firm complex with Cbl, which did not dimerize. We suggest a novel model for the domain organization of IF and discuss its relevance for Cbl uptake.

EXPERIMENTAL PROCEDURES

Materials

All salts and media components were purchased from Merck, Roche Molecular Biochemicals, Sigma-Aldrich, and Beckton & Dickinson. All forms of Cbl were from Sigma-

[†] This work was supported by the Eureka program (CT-T2006) and Cobento Biotech A/S.

* To whom correspondence should be addressed. Tel: (+45) 89 42 50 92. Fax: (+45) 86 13 65 97. E-mail: snf@imsb.au.dk.

[‡] Protein Chemistry Laboratory, Department of Molecular and Structural Biology, University of Aarhus.

[§] Department of Biophysics, University of Aarhus.

[||] Cobento Biotech A/S.

[⊥] Department of Medical Biochemistry, University of Aarhus.

[#] Department of Clinical Biochemistry, AS Aarhus University Hospital.

[1] Abbreviations: BSA, bovine serum albumin; Cbl, cobalamin (Cbl·OH_2, aquocobalamin if not stated otherwise); Cbl·CN, cyanocobalamin; GdnHCl, guanidine hydrochloride; IF, intrinsic factor; LS, light scattering; M_{LS}, molecular mass determined by light scattering; M_w, molecular mass; P_i buffer, NaH_2PO_4/Na_2HPO_4 buffer; RI, refractive index; V_e, elution volume.

10.1021/bi048924c CCC: $27.50
Published on Web 00/00/0000 PAGE EST: 7.6

Aldrich. Sephacryl S-200 and ConA−Sepharose were obtained from Amersham Pharmacia Biotech.

Methods

Purification of Human IF. The preparation of human gastric IF (*17*) was obtained as described earlier.

Purification of the Full-Length IF and Its Fragments from Plants. The isolation procedure corresponded to the one described earlier (*11*) with certain modifications. The full-length protein and the fragments were purified from the plant extract by affinity chromatography as the ligand-saturated preparations. They were separated into two fractions (peak I and II) by gel filtration on a 550 mL Sephacryl S-200 column equilibrated with 0.1 M Tris and 1 M NaCl, pH 7.5. The apo forms were obtained after GdnHCl treatment (*11*) followed by renaturing, when IF in 5 M GdnHCl solution was diluted 10 times with 0.2 M P_i buffer, pH 7.5. The protein was concentrated by ultr-filtration and subjected to extensive dialysis against the same buffer.

The preparation of IF_{50} was additionally purified by gel filtration on a 100 mL Sephacryl S-200 column equilibrated with 0.2 M P_i buffer, pH 7.5.

The renatured fragments IF_{30} and IF_{20} were separated from each other on ConA−Sepharose. The protein sample (4−5 mL, ≈2 mg/mL) in 0.2 M P_i buffer was applied to a 5 mL column of ConA−Sepharose at room temperature. The glycopeptide IF_{20} was effectively adsorbed on the matrix via its carbohydrate moiety whereas IF_{30} was collected in the break-through fraction. The column was washed with P_i buffer, and the adsorbed fragment was eluted in 10% glucose and 0.1 M P_i buffer, pH 7.5. The preparation of IF_{20} was slightly contaminated with IF_{50} (≈10%) and, therefore, additionally purified by gel filtration on a 100 mL Sephacryl S-200 column.

Electrophoresis, Western Blot Analysis, and Peptide Sequencing. SDS−PAGE and blotting were performed according to the standard procedure. Reduction with dithiothreitol was carried out prior to electrophoresis. The protein from the blotted samples was sequenced on the equipment from Applied Biosystems Procise.

Determination of SH Groups. Apo-IF (20−30 μM) in either 0.2 M P_i buffer, pH 7.5, or 5 M GdnHCl was mixed with 50 μM 5,5′-dithiobis(2-nitrobenzoic acid) and incubated for 30 min, whereupon absorbance at 412 nm was measured. The value of A_{412} in the control sample without IF was subtracted from the experimental data. The extinction coefficient of 13600 M^{-1} cm^{-1} was used for calculations.

Dissociation of the Protein−Ligand Complexes. The velocity of Cbl liberation was investigated by the ligand exchange method as described earlier (*9*). In short, the holoprotein saturated with Cbl·OH_2 (12 μM) was mixed with Cbl·CN (50 μM) and incubated at room temperature. At certain time intervals, the samples were taken out and mixed with charcoal. Absorbance in the supernatant was measured. The degree of the ligand displacement was measured spectroscopically from the ratio A_{361}/A_{352} and plotted versus incubation time.

Binding of IF to the Receptor. Binding of IF to cubilin was performed by surface plasmon resonance analysis on a BIAcore 2000 instrument as described elsewhere (*18*). In short, human cubilin was immobilized on the surface of the sensor chip and subjected to a flow of solution containing 10−100 nM IF. The increase of the surface plasmon resonance signal due to IF−cubilin interaction was followed in time. The flow of the IF-free buffer caused detachment of IF from cubilin and decrease of the signal.

Determination of the Molecular Mass by Laser Light Scattering. Molecular masses were determined by analysis of the static light scattering (LS), which allowed calculation of the absolute value of M_{LS} (*19*). The preparation of IF (100 μL, 100−300 μg of protein) was applied to a Superose-12 HR 30 column and run on HPLC system LC-10Advp (Shimadzu Corp.) at a flow of 0.5 mL/min in 0.1 M Tris and 0.2 M NaCl, pH 7.5. Light scattering (LS), refractive index (RI), and absorbance were measured in the elution profile in slices of 0.002 mL. The values from 10 slices were averaged and presented in the figures. The following modules were used as in-line detectors: light scattering module MiniDAWN (Wyatt Technology Corp.), differential refractometric detector RID-10A (Shimadzu Corp.), and UV−vis absorbance detector SPD-10Av (Shimadzu Corp.). When the apo forms of IF were resaturated with Cbl, the time of incubation prior to HPLC was 1−2 h. The response from RI and absorbance detectors (millivolts) corresponded to milligram per milliliter concentration and optical density at 1 cm light path, respectively. Absorbance was measured at one of the following wavelengths: 280 nm for protein detection; 355 nm for IF + Cbl·OH_2; 361 nm for IF + Cbl·CN; 570 nm for IF + Cbl·OH_2 if 50 μM Cbl·OH_2 was present in the elution buffer.

Calculation of M_{LS} was performed by a "two-detector" method (*20*, *21*) as well as using ASTRA software (Wyatt Technology Corp.). Both methods are based on the proportionality of LS to the product of the weight-average molar mass (M_w) and the concentration of the macromolecule (milligram per milliliter) (*19*).

The two-detector method is a good approximation for proteins with M_w < 500 kDa separated by gel permeation chromatography. It requires calibration of the detectors using a protein with the known molecular mass as well as assignment of the d*n*/d*c* coefficient (*20*, *21*). In this case M_{LS} can be calculated from the equation:

$$M_{LS} = \frac{LS}{RI}\frac{K_{RI/LS}}{dn/dc} \quad (1)$$

where LS and RI are measurements of the corresponding detectors and $K_{RI/LS}$ is the calibration constant. The value of $K_{RI/LS} = 32$ was obtained using bovine serum albumin (BSA) as a calibrator (M_w = 66.2 kDa, d*n*/d*c* = 0.187). The value of *dn*/d*c* for a glycoprotein can be calculated according to the formula: *dn*/d*c* = [protein fraction] × 0.187 + [carbohydrate fraction] × 0.15. The following *dn*/d*c* coefficients were assigned to IF_{30}, IF_{50}, and IF_{20}: 0.187, 0.181, and 0.172, assuming 0%, 15%, and 40% content of carbohydrates, respectively. The obtained values of M_{LS} are presented as the mean ± SD.

The multiangle analysis was performed by ASTRA software. The *dn*/d*c* coefficients mentioned above were used during calculations.

Dissociation of the dimers $(IF{\cdot}Cbl)_2$ to monomers IF·Cbl ($d \leftrightarrow m + m$) was studied on the diluted protein samples (2−50 μM) subjected to HPLC. The apparent molecular

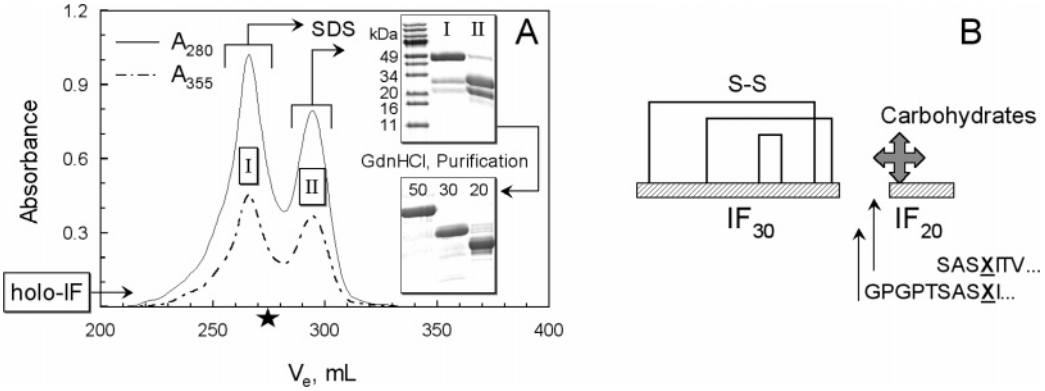

FIGURE 1: Subforms of recombinant human IF from plants. Panel A: Gel filtration profile of an IF preparation. The profile from a Sephacryl S-200 column was recorded at 280 nm (protein moiety) and 355 nm (Cbl moiety). The star indicates the elution volume of a 66.2 kDa standard BSA. Upper inset: SDS−PAGE of the material from peaks I and II. Lower inset: SDS−PAGE of the purified forms IF_{50}, IF_{30}, and IF_{20} devoid of Cbl. Panel B: Scheme of the IF's primary structure. The full-length IF_{50} is split into the fragments IF_{30} and IF_{20}. Two adjacent cleavage sites are indicated by the arrows. The plausible pattern of the disulfide bridges in IF_{30} is suggested on the basis of our previous work (*6*). The small fragment IF_{20} is composed of the 13 kDa peptide (rectangle) plus 5−7 kDa of carbohydrates (cross) (*11*). Failure to detect the Asn residue (X) during sequencing was ascribed to its glycosylation.

masses M_{LS}^{app} were determined in the eluent and plotted versus total concentration of the subunits $E_0 = m + 2d$ in the peak fractions. The values of E_0 (micromolar) were calculated according to A_{280} or RI measurements ($E_0 = A_{280}/0.063$; $E_0 = \text{RI}/0.04$). The curve M_{LS}^{app} vs E_0 was fitted to the equation:

$$M_{LS}^{app} = M_{LS}^{m} + M_{LS}^{m}\frac{K_d + 4E_0 - \sqrt{K_d^2 + 8K_dE_0}}{4E_0} \quad (2)$$

where M_{LS}^{m} is the molecular mass of a monomer based on the light scattering data, K_d is the equilibrium dissociation constant for the reaction $d \leftrightarrow m + m$, and E_0 is the total concentration of IF·Cbl subunits in the monomers and the dimers. The optimal values of M_{LS}^{m} and K_d (±SE) were found by nonlinear regression analysis using the program KyPlot 4 (Kyence Inc.).

RESULTS

Purification of the Full-Length Protein IF_{50} and Fragments IF_{30} and IF_{20}. The recombinant human IF was isolated from the plant extract by affinity chromatography and subsequent gel filtration (*11*). The elution profile (Figure 1A) revealed the presence of two peaks (I and II) containing a Cbl-saturated protein as judged from the absorbance ratio $A_{280}/A_{355} \approx 2$. According to SDS electrophoresis, the main component of peak I was the full-length IF_{50} (Figure 1A, upper inset, lane I) with the N-terminal sequence STQTQSS-(C)S.... The material from peak II revealed two major peptides with approximate molecular masses of 30 and 20 kDa (lane II). The band of IF_{30} represented the N-terminal fragment STQTQSS(C)S... and contained no carbohydrates (*11*). The band IF_{20} comprised actually two C-terminal fragments, which originated from the adjacent cleavage sites GPGPTSASXI... (≈30−50%) and SASXITVIYT... (≈50−70%). As followed from the sequence, the protein moiety of IF_{20} accounted only for 13 kDa, implying the presence of 7 kDa oligosaccharides (*11*). At least one N-glycosylation site could be assigned to the Asn residue marked as X. This residue was not detected during sequencing because of some posttranslational modification. Figure 1B schematically indicates the cleavage position and the glycosylation site in the primary structure of IF.

Notations "peak I" and "peak II" throughout the text are applied to the Cbl-saturated proteins in the corresponding gel filtration fractions. The samples under these names were not subjected to any additional treatment. The cleaved IF from peak II accounted for 30−60% of the expressed protein depending on the plant material. However, no following degradation was detected in the crude extract, at least for 24 h at room temperature.

Ligand Exchange and Receptor Affinity of Holo-IF Variants. The separated holoproteins from peaks I and II (Figure 1A) were tested with respect to velocity of Cbl exchange (Figure 2A) and interaction with the specific receptor cubilin (Figure 2B). The presented curves revealed similar characteristics for both IF variants. Their properties were comparable to IF from other sources (*9*, *11*).

Determination of M_{LS} for Holo Forms. The recombinant samples from peaks I and II [as well as a preparation of gastric holo-IF (*17*)] were subjected to size exclusion HPLC followed by measurements of laser light scattering (LS), refractive index (RI), and absorbance (Figure 3). The molecular masses M_{LS} were calculated from LS−RI data as described in the Methods section.

The full-length protein from peak I (Figure 3A) revealed characteristics similar to gastric IF (Figure 3C), $V_e \approx 12.6$ mL and $M_{LS} \approx 90-95$ kDa. The major component of peak I seemed to represent the dimer $(IF_{50}\cdot Cbl)_2$ when relating $M_{LS} = 95$ kDa of the native sample to $M_w = 50$ kDa of the elementary unit on SDS electrophoresis. On the contrary, the preparation from peak II revealed the presence of a smaller protein with $M_{LS} = 53$ kDa (Figure 3B). We assumed it to be a composite complex of two peptides with the ligand IF_{30+20}·Cbl, which lost the ability to dimerize. The low-weight complex IF_{30+20}·Cbl contaminated the dimers from peak I, which accounted for the asymmetric shape of the elution profile in Figure 3A (see also SDS−PAGE, lane I, in Figure 1A).

The profile of a standard protein (BSA) with $M_w = 66.2$ kDa is given in Figure 3D for comparison and in order to

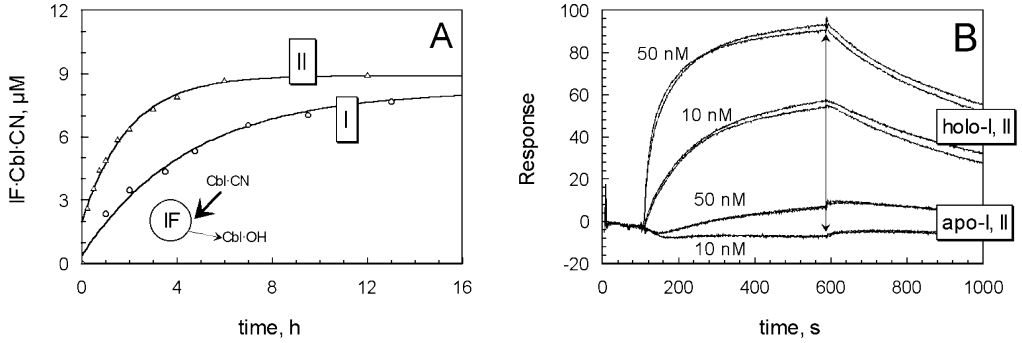

FIGURE 2: Functional properties of full-length and cleaved IF's. Panel A: Dissociation of IF·Cbl complexes. The process was followed by exchange of the endogenous ligand Cbl·OH_2 (12 μM) for the exogenous one Cbl·CN (50 μM). The curves were obtained for the full-length protein (I) and the cleaved form (II). The apparent rate constants k_{-app} of Cbl·OH_2 dissociation were estimated from the initial slope $v_0 = k_{-app}[IF{\cdot}Cbl]_0$ as 4×10^{-5} s^{-1} (I) and 1×10^{-4} s^{-1} (II). Panel B: Binding of holo-IF to the specific receptor cubilin. The binding process was followed on BIAcore equipment. The receptor was immobilized on the surface of a sensor chip and subjected to a flow of the indicated forms of IF. The relative response was measured in time. The arrow indicates the time when the protein flow was changed for the buffer and the dissociation of IF from cubilin was initiated. Interaction of holo-IF with cubilin was characterized by $k_+ = (3-6) \times 10^5$ $M^{-1}s^{-1}$ and $k_- = (1-2) \times 10^{-3}$ s^{-1}; $K_{cub} = 3-6$ nM in all cases.

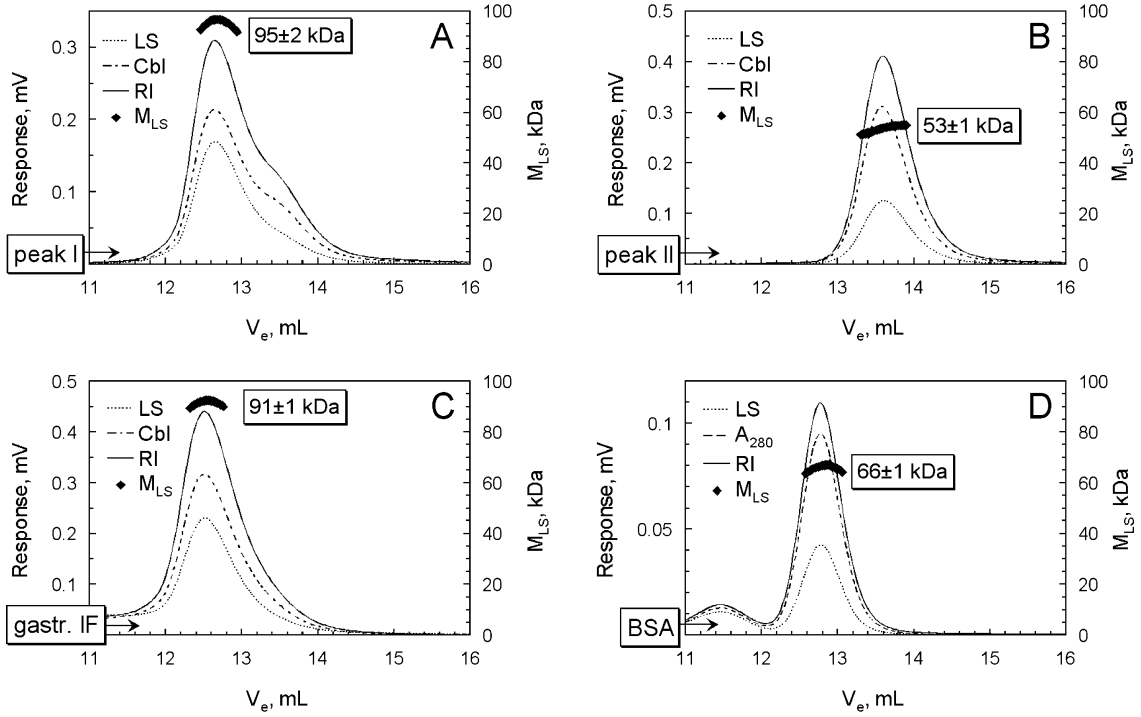

FIGURE 3: Light scattering measurements of holo-IF's and a standard. The proteins were applied to size exclusion HPLC. Measurements of light scattering (LS), refractive index (RI), and absorbance were performed in the eluent and used for M_{LS} calculations by either a two-detector method (black diamonds in the figure) or ASTRA software (subscript to the figure); see the Methods section for details. Panel A: Recombinant IF from plants (peak I), ASTRA $M_{LS} = 95 \pm 1$ kDa. Panel B: Recombinant IF from plants (peak II), ASTRA $M_{LS} = 53 \pm 1$ kDa. Panel C: Gastric IF, ASTRA $M_{LS} = 92 \pm 1$ kDa. Panel D: Standard protein, BSA, ASTRA $M_{LS} = 66 \pm 1$ kDa.

corroborate the accuracy of measurements ($M_{LS} = 66 \pm 1$ kDa).

Preparation of the Apo Forms. The apoproteins were prepared by treatment of the proteins from peaks I and II with GdnHCl and subsequent renaturing (see the Methods section). The recovered Cbl binding capacity in IF_{50} and $IF_{30} + IF_{20}$ preparations was approximately 80%. Additional purification increased this value to 95−98%. The preliminary analysis concerning Cbl binding, receptor recognition, and absorbance spectroscopy pointed to normal renaturing of all IF variants. None of the mentioned IF's contained free SH groups. The SDS−PAGE of the pure apo forms are shown in Figure 1A, lower inset. In the text below we continue to use the abbreviations IF_{50}, IF_{30}, and IF_{20}, which refer to the renatured apoproteins in contrast to the original holoproteins from peaks I and II.

Determination of M_{LS} for Apo-IF_{50} ± Cbl. When the apo form IF_{50} was subjected to HPLC, both the elution volume and the LS measurements unequivocally showed the presence of a monomeric protein with $M_{LS} = 48$ kDa (Figure 4A). No traces of the dimers were found. Partial saturation with Cbl (one-half of the binding capacity) caused visible changes in the elution profile (Figure 4B). It could now be separated in two overlapping components with $M_w \approx 80$ kDa and $M_w \approx 50$ kDa. As seen from the absorbance measurements, Cbl was associated only with the first peak.

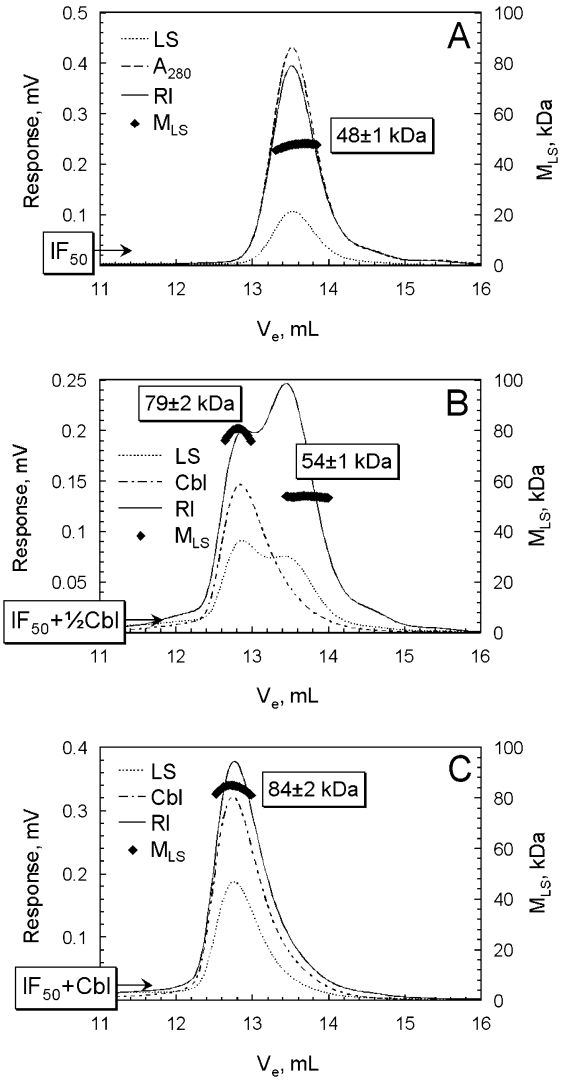

FIGURE 4: Light scattering measurements of recombinant apo-IF_{50} from plants ± Cbl. All notation as in Figure 3. Panel A: IF_{50} devoid of Cbl, ASTRA $M_{LS} = 46 \pm 1$ kDa. Panel B: IF_{50} half-saturated with Cbl, ASTRA $M_{LS1} = 76 \pm 1$ kDa, $M_{LS2} = 52 \pm 1$ kDa. Panel C: IF_{50} saturated with an excess of Cbl, ASTRA $M_{LS} = 85 \pm 1$ kDa.

When IF_{50} was exposed to an excess of the ligand (Cbl·OH_2), the prevailing amount of the protein was eluted in the lower volume, and M_{LS} increased to 84 kDa (Figure 4C). The profiles recorded for other Cbl forms (Cbl·CN and adenosyl-Cbl, not shown) did not differ from Cbl·OH_2.

Determination of M_{LS} for the Fragments ± Cbl. Two mixed fragments ($IF_{30} + IF_{20}$, 1:1) eluted separately as 28 and 20 kDa proteins (Figure 5A). No tendency to association was detected. The same was the case for the individual samples IF_{30} and IF_{20}, where the presence of Cbl (Figure 5B,C) or its absence (not shown) did not influence the measured M_{LS} of 27 and 18 kDa, respectively. The affinity of IF_{30} to Cbl was very weak, and only the presence of 50 μM Cbl in the elution buffer prevented complete dissociation of the ligand from the peptide (Figure 5B). Approximately 20−30% of IF_{30} did bind the ligand under the mentioned conditions, but no peptide association was found.

The glycopeptide IF_{20} revealed much higher affinity to the ligand. This fragment remained essentially saturated with Cbl·OH_2 (Cbl·CN) even without any additives to the elution buffer (Figure 5C). The only molecular form on the HPLC profile was an 18 kDa protein.

The mixture of two peptides, $IF_{20} + IF_{30}$, with Cbl (1:1:2) showed a shift in V_e and an increase of M_{LS} to 50 kDa, which corresponded to the assembled complex IF_{30+20}·Cbl (Figure 5D).

Stability of the IF Oligomers. We analyzed the stability of the produced protein associates in terms of their M_{LS}^{app}. A number of the diluted samples were used in this experiment. The lowest concentration prior to HPLC was ≈2 μM, which gave the level of ≈0.3 μM in the eluted peak (Figure 6A,B). The averaged records of LS and RI were sufficiently accurate to perform M_{LS} calculations. A noticeable change was found for the diluted full-length protein (Figure 6A) when compared with its concentrated sample (Figure 4C). Thus, the diluted sample contained a protein of lower Stokes radius and molecular mass according to the determined values of V_e and M_{LS}^{app}. On the contrary, the peptide complex IF_{30+20}·Cbl revealed no visible change in its molecular mass (compare Figure 5D with Figure 6B). The presence or absence of 10 mM Ca^{2+} in the elution buffer did not influence M_{LS}^{app} of any sample and neither did the time of incubation from 2 min to 24 h (not shown).

Several HPLC profiles, analogous to those in Figure 6A,B, were recorded. They were used to produce dependencies of the apparent molecular masses versus total concentration of the subunits (IF_{sub}) in the peak fractions (Figure 6C). The Cbl-saturated dimeric forms (both freshly purified and GdnHCl-treated) revealed a clear tendency to dissociate to monomers with $K_d \approx 1$ μM at the decreasing protein concentrations (Figure 6C, upper curves). However, dissociation to monomers did not influence the degree of IF saturation with Cbl. At $IF_{sub} \rightarrow 0$, the molecular masses of the monomers were $M_{LS}^{m} = 53$ kDa (peak I) and 47 kDa (IF_{50} + Cbl). The slight discrepancy could be ascribed to loss of a native or artificial peptide, which was originally bound to holo-IF (I) but dissociated during GdnHCl treatment.

In contrast to the unstable dimers, the cleaved construction IF_{30+20}·Cbl did not essentially change its M_{LS}^{app} upon dilution or concentration (Figure 6C, lower line). There was a potential drift from 46 kDa ($IF_{sub} = 0$) to 50 kDa ($IF_{sub} = 12$ μM) which could be ascribed, with certain reservations, to weak dimerization ($K_d \approx 300$ μM).

DISCUSSION

In the present work we investigated the influence of Cbl on the oligomeric state of IF. Oligomerization of gastric IF upon Cbl binding was occasionally observed before (*1*, *14*, *16*). At the same time, this phenomenon was difficult to reproduce (*15*). Confirmation or rejection of the mentioned effect would be important to interpret the interaction between Cbl, IF, and the specific receptor cubilin. Therefore, we have conducted a systematic analysis of IF oligomers using static light scattering (LS) coupled to size exclusion chromatography. Determination of the molecular mass by LS is independent of the Stokes radius, protein conformation, or branching (*19*−*21*) in contrast to gel filtration and sedimentation methods.

The work was performed primarily on the recombinant protein from plants. The main protein species of plant IF

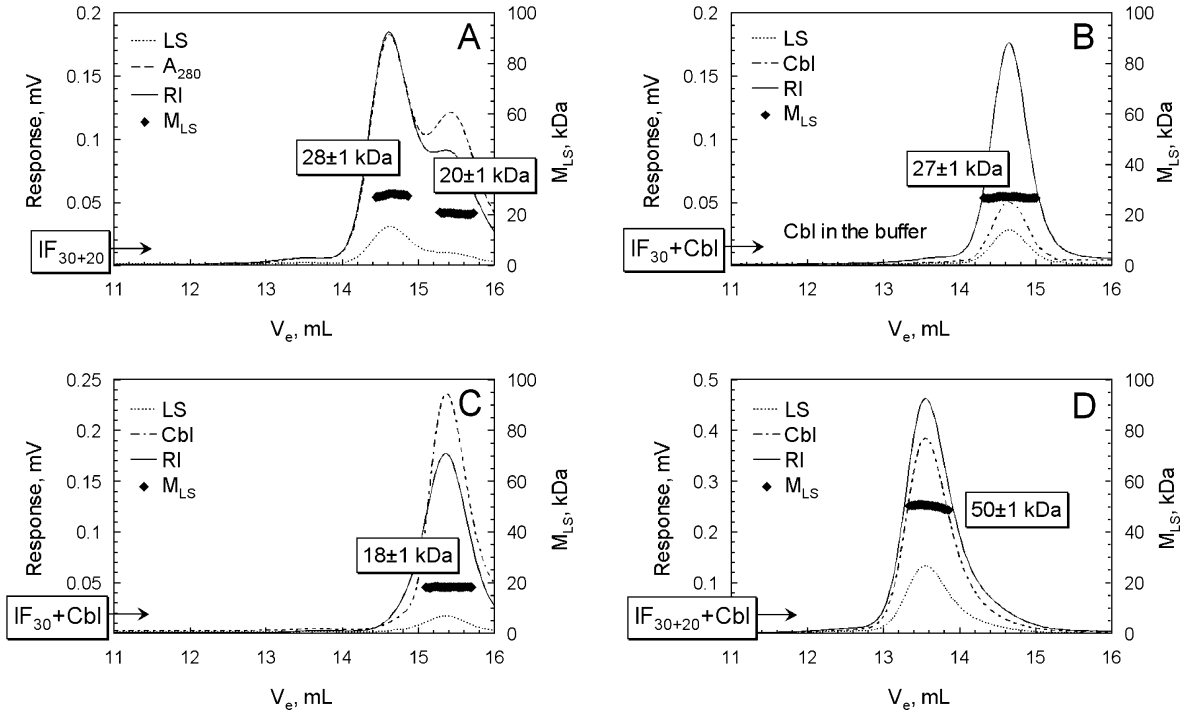

FIGURE 5: Light scattering measurements of the IF fragments ± Cbl. All notation as in Figure 3. Panel A: $IF_{30} + IF_{20}$, ASTRA $M_{LS1} = 26 \pm 1$ kDa, $M_{LS2} = 22 \pm 3$ kDa. Panel B: IF_{30} in the presence of 50 μM Cbl in the elution buffer, ASTRA $M_{LS} = 27 \pm 1$ kDa. Panel C: IF_{20} + Cbl, ASTRA $M_{LS} = 18 \pm 1$ kDa. Panel D: $IF_{30} + IF_{20}$ + Cbl, ASTRA $M_{LS} = 50 \pm 1$ kDa.

observed on SDS−PAGE were the 50 kDa full-length form IF_{50} and the fragments IF_{30} and IF_{20} (Figure 1A). The latter two corresponded to the N-terminal peptide of ≈30 kDa and the C-terminal glycopeptide of ≈20 kDa, respectively (Figure 1B). Cleavage of IF at the indicated site (Figure 1B) seems to be a ubiquitous phenomenon, because varying amounts of kindred fragments were also observed for gastric IF (*16*) and recombinant IF from yeast (*8*, *9*).

Gel filtration analysis showed that the plant protein contained two subforms, which were separated during gel filtration into peaks I and II (Figure 1A). LS measurements established different molecular masses of these samples: (I) $M_{LS} = 95$ kDa and (II) $M_{LS} = 53$ kDa; see Figure 3A,B. Comparison of M_{LS} to the SDS−PAGE bands of (I) 50 kDa and (II) 30 + 20 kDa indicated the presence of the dimers and the nicked monomers in the above fractions. The sample of gastric holo-IF (≈55 kDa on SDS−PAGE) demonstrated the presence of the dimers (Figure 3C); i.e., the behavior of the recombinant protein from plants matched the natural protein in that respect. Other physiological properties of the dimer and the nicked monomer were surprisingly similar, at least concerning the velocities of Cbl exchange (Figure 2A) and the receptor binding (Figure 2B). None of the mentioned features differed from the plant protein from IF's of other sources (*9*, *11*, *13*).

Analysis of the holo forms I and II from the gel filtration profile revealed the presence of IF oligomers, yet the effect could be artificial and caused by the purification method. Therefore, all IF forms were devoid of Cbl, renatured, and isolated as the individual samples IF_{50}, IF_{30}, and IF_{20} (Figure 1A, lower inset). When the full-length IF_{50} was mixed with increasing amounts of Cbl (1:0, 1:0.5, 1:2), a clear transition from a 48 kDa monomer to a 84 kDa "dimer" was found (Figure 4). Only Cbl-saturated protein dimerized, as was particularly noticeable in a half-saturated sample (Figure 4B). The molecular mass of the obtained dimer was, however, reproducibly lower than could have been expected. The asymmetric shape of the elution peak (Figure 4C) pointed to an existing dimer−monomer equilibrium, which shifted rapidly upon dilution during gel filtration.

Analysis of the apparent molecular masses in a broad range of the protein concentration demonstrated that the dimers $(IF_{50}{\cdot}Cbl)_2$, indeed, had a tendency to dissociate to the monomers $IF_{50}{\cdot}Cbl$ at decreasing protein concentrations (Figure 6C). The calculated dissociation constant $K_d \approx 1$ μM for the dimer−monomer equilibrium (Figure 6C) indicates that the dimers can be observed only at a relatively high concentration of the protein, >5 μM. In other words, dimerization does not seem to be important under physiological concentrations of IF ≈50 nM (*14*−*16*), unless there is a stabilizing agent to fasten the dimer. The presence of such an agent can be hypothesized from the 2:1 ratio reported for the $IF_{50}{\cdot}Cbl$ + receptor interaction (*23*). Our negative result with Ca^{2+} eliminates this ion from the list of potential fasteners. Despite uncertainty about the physiological role of the dimers, oligomerization to $(IF_{50}{\cdot}Cbl)_2$ appeared to be essential for separation of the full-length protein from the fragments.

Isolation of two IF fragments, IF_{30} and IF_{20}, in the individual form opened a convenient way to explore the structure of the protein and establish the functional role of different domains. The frequent occurrence of these fragments, when working with natural and recombinant IF's from different sources, points to a two-domain organization of this Cbl binder. The N-terminal "head" domain IF_{30} accounts for 28−30 kDa of the protein sequence and contains all S−S bridges, as well as most of the conservative residues (*6*, *22*). The C-terminal "tail" domain includes only a minor part of the protein sequence, ≈13 kDa, while the rest of its mass originates from the oligosaccharides. These two domains

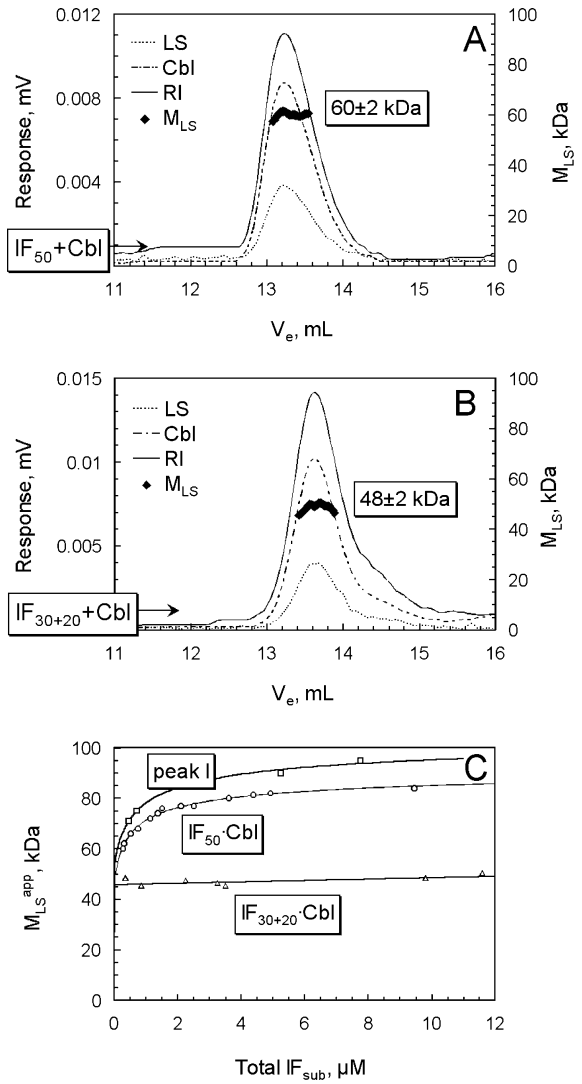

FIGURE 6: Light scattering measurements of the diluted samples. All notation as in Figure 3. Panel A: IF_{50} + Cbl, ASTRA processing of the unaveraged raw data was unreliable. Panel B: IF_{30} + IF_{20} + Cbl, ASTRA processing of the unaveraged raw data was unreliable. Panel C: Dependence of the apparent molecular mass on the total concentration of IF. The data were fitted by eq 2 with the following parameters: (1) peak I, $M_{LS}^{m} = 53 \pm 1$ kDa, $K_d = 1.3 \pm 0.4$ μM; (2) IF_{50}·Cbl, $M_{LS}^{m} = 47 \pm 1$ kDa, $K_d = 0.9 \pm 0.1$ μM; (3) IF_{30+20}·Cbl, $M_{LS}^{m} = 46 \pm 1$ kDa, $K_d = 300 \pm 200$ μM.

seem to be relatively independent from each other and resist proteolysis. On the contrary, the connecting interdomain part of ≈2 kDa is quite susceptible to cleavage.

Each isolated fragment (IF_{30} or IF_{20}) did not oligomerize, when taken either alone (not shown) or in the presence of Cbl (Figure 5B,C). Curiously enough, the affinity of the small tail fragment IF_{20} to Cbl was significantly higher than that of the bigger peptide IF_{30}. Thus, IF_{20} retained Cbl during HPLC, whereas the IF_{30}·Cbl complex easily liberated the ligand in the absence of Cbl in the elution buffer.

When the fragments were mixed together at a ratio of 1:1 (IF_{30} + IF_{20}), they did not form any aggregates without Cbl (Figure 5A). However, they produced a firm associate of 50 kDa as soon as Cbl was added (Figure 5D). Despite the proper molecular weight, the cleaved construction IF_{30+20}·Cbl was not completely identical to the full-length IF_{50}·Cbl because the ability to form dimers was completely lost. At

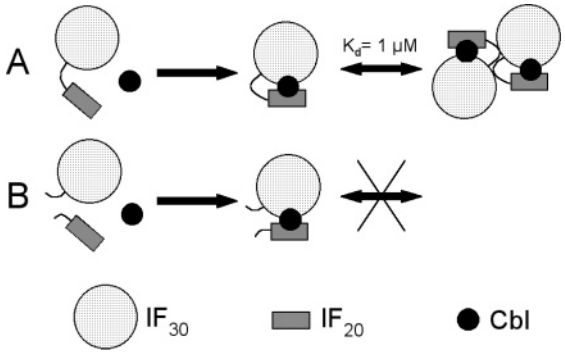

FIGURE 7: Scheme of Cbl-induced oligomerization of IF species. (A) Dimerization of the full-length protein IF_{50} after attachment of Cbl. (B) Association of the fragments IF_{30} and IF_{20} into a cleaved monomer where no following dimerization is possible. See the main text for details.

the same time, the cleaved monomers IF_{30+20}·Cbl were quite stable. Neither dilution nor concentration caused any significant change in the measured molecular mass of $M_{LS}^{m} = 46 \pm 1$ kDa.

Self-assembly of two disjointed IF fragments into a stable structure, IF_{30+20}·Cbl, upon Cbl binding supports our hypothesis about a two-domain organization of IF. These domains are expected to be sufficiently distant and operate almost independently from each other. They can even be disconnected but still retain the ability to bind the ligand. As for the lack of dimerization observed for the cleaved monomer IF_{30+20}·Cbl, this effect can be interpreted in different ways. The most straightforward explanation implies cleavage inside the dimerization site, where, in addition, a small peptide can be eliminated from the sequence. That makes the stretch of the sequence SPDHE... ...PTSAS (connecting the domains IF_{30} and IF_{20}) the most legitimate candidate for the dimerization site.

The simplest uncontroversial scheme of interaction between Cbl and IF domains is depicted in Figure 7. The upper sketch (A) shows the reaction for the full-length protein IF_{50}. Binding of Cbl brings together the distant domains, whereupon two monomers associate into a dimer. An antiparallel association might be conjectured, since the alternative parallel scheme presumes contacts between two homologous domains IF_{30}−IF_{30} and IF_{20}−IF_{20}. The experimental data (Figure 5B,C) do not support the latter hypothesis.

The lower sketch (B) in Figure 7 depicts the interaction between two separated units, IF_{30} and IF_{20}. The presence of Cbl "glues" them into an almost native construction. However, the following association to dimers is essentially impaired because the responsible region of the sequence is damaged or even lost.

A two-domain organization of IF may be advantageous for Cbl recognition. Thus, examination of the ligand by two practically independent units ensures a "double proof" during the ligand binding. Furthermore, connection of two distant domains by Cbl may build a distinctive signal to the receptor. In that way the receptor can discriminate the Cbl-saturated binder from the apoprotein. These issues are investigated in our next publication.

ACKNOWLEDGMENT

We acknowledge gratefully the excellent technical assistance of A. L. Christensen and C. Jacobsen.

REFERENCES

1. Gräsbeck, R. (1984) Biochemistry and clinical chemistry of vitamin B12 transport and the related diseases, *Clin. Biochem. 17*, 99−107.
2. Nexø, E. (1998) Cobalamin binding proteins, in *Vitamin B_{12} and B_{12}-Proteins* (Kräutler, B., Arigoni, D., and Golding, T., Eds.) pp 461−475, Wiley-VCH, Weinheim, Germany,
3. Okuda, K. (1999) Discovery of vitamin B12 in the liver and its absorption factor in the stomach: a historical review, *J. Gastroenterol. Hepatol. 14*, 301−308.
4. Moestrup, S. K. (1998) Cellular surface receptors important for vitamin B_{12} nutrition, in *Vitamin B_{12} and B_{12}-Proteins* (Kräutler, B., Arigoni, D., and Golding, T., Eds.) pp 477−489, Wiley-VCH, Weinheim, Germany.
5. Fyfe, J. C., Madsen, M., Højrup, P., Christensen, E. I., Tanner, S. M., Chapelle, A., Quianchuan, H., and Moestrup S. K. (2004) The functional cobalamin (vitamin B_{12})-intrinsic factor receptor is a novel complex of cubilin and amnionless, *Blood 103*, 1573−1579.
6. Fedosov, S. N., Berglund, L., Nexø, E., and Petersen, T. E. (1999) Sequence, S-S bridges, and spectra of bovine transcobalamin expressed in *Pichia pastoris*, *J. Biol. Chem. 274*, 26015−26020.
7. Fedosov, S. N., Fedosova, N. U., Nexø, E., and Petersen, T. E. (2000) Conformational changes of transcobalamin induced by aquocobalamin binding. Mechanism of substitution of the cobalt-coordinated group in the bound ligand, *J. Biol. Chem. 275*, 11791−11798.
8. Wen, J., Kinnear, M. B., Richardson, M. A., Willetts, N. S., Russel-Jones, G. J., Gordon, M. M., and Alpers, D. H. (2000) Functional expression in Pichia pastoris of human and rat intrinsic factor, *Biochim. Boiphys. Acta 1490*, 43−53.
9. Fedosov, S. N., Berglund, L., Fedosova, N. U., Nexø, E., and Petersen, T. E. (2002) Comparative analysis of cobalamin binding kinetics and ligand protection for intrinsic factor, transcobalamin, and haptocorrin, *J. Biol. Chem. 274*, 26015−26020.
10. Kalra, S., Li, N., Seetharam, S., Alpers, D. H., and Seetharam, B. (2003) Function and stability of human transcobalamin II: role of intramolecular disulfide bonds C98−C291 and C147−C187, *Am. J. Physiol. Cell Physiol. 285*, C150−C160.
11. Fedosov, S. N., Laursen, N. B., Nexø, E., Moestrup, S. K., Petersen, T. E., Erik Ø., Jensen, E. Ø., and Berglund, L. (2003) Human intrinsic factor expressed in the plant *Arabidopsis thaliana*, *Eur. J. Biochem. 270*, 3362−3367.
12. Stupperich, E., and Nexø, E. (1991) Effect of the cobalt-N coordination on the cobamide recognition by the human vitamin B12 binding proteins intrinsic factor, transcobalamin and haptocorrin, *Eur. J. Biochem. 199*, 299−303.
13. Birn, H., Verroust, P. J., Nexø, E., Hager, H., Jacobsen, C., Christensen, E. I., and Moestyrup, S. K. (1997) Characterization of an epithelial ≈460 kDa protein that facilitates endocytosis of intrinsic factor-vitamin B_{12} and binds receptor associated protein, *J. Biol. Chem. 272*, 26497−26504.
14. Allen, R. H., and Mehlman, C. S. (1973) Isolation of gastric vitamin B12-binding proteins using affinity chromatography. II. Purification and properties of hog intrinsic factor and hog nonintrinsic factor, *J. Biol. Chem. 248*, 3660−3669.
15. Visuri, K., and Gräsbeck, R. (1973) Human intrinsic factor, isolation by improved conventional methods and properties of the preparation, *Biochim. Biophys. Acta 310*, 508−517.
16. Christensen, J. M., Hippe, E., Olesen, H., Rye, M., Haber, E., Lee, L., and Thomsen, J. (1973) Purification of human intrinsic factor by affinity chromatography, *Biochim. Biophys. Acta 303*, 319−332.
17. Nexø, E. (1975) A new principle in biospecific affinity chromatography used for purification of cobalamin-binding proteins, *Biochim. Biophys. Acta 379*, 189−192.
18. Kristiansen, M., Kozyraki, R., Jacobsen, C., Nexø, E., Verroust, P. J., and Moestrup, S. K. (1999) Molecular dissection of the intrinsic factor-vitamin B12 receptor, cubilin, discloses regions important for membrane association and ligand binding, *J. Biol. Chem. 274*, 20540−20544.
19. Wyatt, P. J. (1993) Light scattering and the absolute characterization of macromolecules, *Anal. Chim. Acta 272*, 1−40.
20. Takagi, T. (1990) Application of low-angle laser light scattering detection in the field of biochemistry, *J. Chromatogr. 506*, 409−416.
21. Wen, J., Arakawa, T., and Philo, J. S. (1996) Size-exclusion chromatography with on-line light-scattering, absorbance, and refractive index detectors for studying proteins and their interactions, *Anal. Biochem. 240*, 155−166.
22. Li, N., Seetharam, S., Lindemans, J., Alpers, D. H., Arwert, F., Seetharam, B. (1993) Isolation and sequence analysis of variant forms of human transcobalamin II, *Biochim. Biophys. Acta 1172*, 21−30.
23. Yammani, R. R., Seetharam, S., and Seetharam, B. (2001) Identification and characterization of two distinct ligand binding regions of cubilin, *J. Biol. Chem. 276*, 44777−44784.

BI048924C