Structural basis for mammalian vitamin B_{12} transport by transcobalamin

Jochen Wuerges*, Gianpiero Garau*†, Silvano Geremia*‡, Sergey N. Fedosov§, Torben E. Petersen§, and Lucio Randaccio*

*Centre of Excellence in Biocrystallography, Department of Chemical Sciences, University of Trieste, Via L. Giorgieri 1, I-34127 Trieste, Italy; and §Protein Chemistry Laboratory, Department of Molecular Biology, University of Aarhus, Science Park, Gustav Wieds Vey 10, 8000 Aarhus C, Denmark

Edited by Rowena G. Matthews, University of Michigan, Ann Arbor, MI, and approved January 31, 2006 (received for review October 18, 2005)

Cobalamin (Cbl, vitamin B_{12}) serves for two essential cofactors in mammals. The pathway for its intestinal absorption, plasma transport, and cellular uptake uses cell surface receptors and three Cbl-transporting proteins, haptocorrin, intrinsic factor, and transcobalamin (TC). We present the structure determination of a member of the mammalian Cbl-transporter family. The crystal structures of recombinant human and bovine holo-TCs reveal a two-domain architecture, with an N-terminal α_6-α_6 barrel and a smaller C-terminal domain. One Cbl molecule in base-on conformation is buried inside the domain interface. Structural data combined with previous binding assays indicate a domain motion in the first step of Cbl binding. In a second step, the weakly coordinated ligand H_2O at the upper axial side of added H_2O-Cbl is displaced by a histidine residue of the α_6-α_6 barrel. Analysis of amino acid conservation on TC's surface in orthologous proteins suggests the location of the TC-receptor-recognition site in an extended region on the α_6-α_6 barrel. The TC structure allows for the mapping of sites of amino acid variation due to polymorphisms of the human TC gene. Structural information is used to predict the overall fold of haptocorrin and intrinsic factor and permits a rational approach to the design of new Cbl-based bioconjugates for diagnostic or therapeutic drug delivery.

cobalamin | crystal structure | P259R polymorphism | transport protein

Cobalamin (Cbl) is an essential nutrient for mammals, because two of its derivatives, methyl-Cbl and 5′-deoxyadenosyl-Cbl (Ado-Cbl), act as coenzymes for methionine synthase and methyl-malonyl-CoA mutase, respectively (1, 2). Only microorganisms are capable of synthesizing Cbl. Consequently, animals have developed a complex pathway for gastrointestinal absorption, blood transport, and cellular uptake of dietary Cbl or vitamin B_{12} (cyano-Cbl). Cbl delivery from food to tissues involves three successive transport proteins and their cell-surface receptors (3, 4). Protein-free Cbl is first bound to haptocorrin (HC) in saliva. After proteolysis of HC–Cbl and other protein complexes of dietary Cbl in the duodenum, Cbl binds to intrinsic factor (IF) in the proximal ileum. The IF–Cbl complex enters mucosal cells in the distal ileum by receptor-mediated endocytosis, and Cbl is transferred to transcobalamin II (TC) (5), whose suggested source is the vascular endothelium (5), and which is the only nonglycosylated protein among the transporters. TC–Cbl is released to the plasma and enters cells by endocytosis via a receptor of as-yet-controversial identity (4). After dissociation of the TC–Cbl complex in lysosomes, Cbl is transformed to the coenzymes methyl- and Ado-Cbl in the cytoplasm and in mitochondria, respectively. Approximately one-quarter of Cbl in plasma is bound to TC, whereas the rest circulates bound to plasmatic HC that can enter only hepatocytes. All three proteins carry a single Cbl molecule and show high affinity to the ligand (6) but differ in specificity for Cbl and its analogues produced by microorganisms (7).

In recent years, holo-TC has been suggested to be a sensitive indicator of vitamin B_{12} deficiency (8), because it represents the physiologically useful portion of plasma Cbl. The most common cause of vitamin B_{12} deficiency, especially among the elderly population, is failure at one of the steps to internalize Cbl rather than lack of Cbl in the diet (9). The accompanying disorders include hematologic abnormalities and defects in the nervous system and metabolism (9).

The overall internalization process is very efficient and provides specificity for genuine Cbl. However, it can be impaired by mutations inside the ligand-binding site of a Cbl transporter or at the receptor–transporter interface. Regarding TC in this respect, some clinical cases of disorders have indeed been linked to defective protein in the patient (10, 11). Yet the main variant forms of this transporter stem from common polymorphisms of the TC gene (12). A structural description of the protein is a prerequisite for relating genetic polymorphisms to a potentially altered function of TC. In addition, structural information may reveal the basis for discrimination between Cbl and nonfunctional Cbl analogues during internalization of the vitamin.

Cbl contains several functional groups that are suitable for the synthesis of bioconjugates (13, 14). Because rapidly proliferating tissues take up proportionally greater amounts of Cbl, the tissue-specific delivery of both imaging agents and cytotoxic drugs via Cbl-based bioconjugates has been suggested and in part performed (15, 16). Knowledge of structural features of Cbl binding to the transporters is fundamental to a rational approach to this important application of Cbl biochemistry, which is still in its infancy.

Successful isolation (17), nucleotide sequencing (18), and expression in different recombinant systems (19, 20) have laid the ground for TC's structure determination, the first for any member of the mammalian Cbl-transporter family. We present the structures of Cbl-complexed recombinant human and bovine TC, as determined by x-ray crystallography, and a detailed description of Cbl binding.

Results

Overall Structure. Human and bovine TC, which share 73% sequence identity, adopt a very similar molecular architecture. Superposition of their backbone trace yields a rms deviation of only 1.2 Å; it is thus legitimate to describe both proteins together. TC encompasses two domains, an N-terminal "α-domain" that is entirely made up of helical secondary structure elements and a C-terminal "β-domain" showing mainly β-strand elements (Fig. 1 and Fig. 6, which is published as supporting information on the PNAS web site).

Conflict of interest statement: No conflicts declared.

This paper was submitted directly (Track II) to the PNAS office.

Abbreviations: TC, transcobalamin II; HC, haptocorrin; IF, intrinsic factor; Cbl, cobalamin.

Data deposition: The atomic coordinates and structure factors have been deposited in the Protein Data Bank, www.pdb.org (PDB ID codes 2BB5, 2BB6, and 2BBC).

†Present address: Institut de Biologie Structural, Rue Jules Horowitz, 38027 Grenoble Cedex 1, France.

‡To whom correspondence should be addressed. E-mail: geremia@univ.trieste.it.

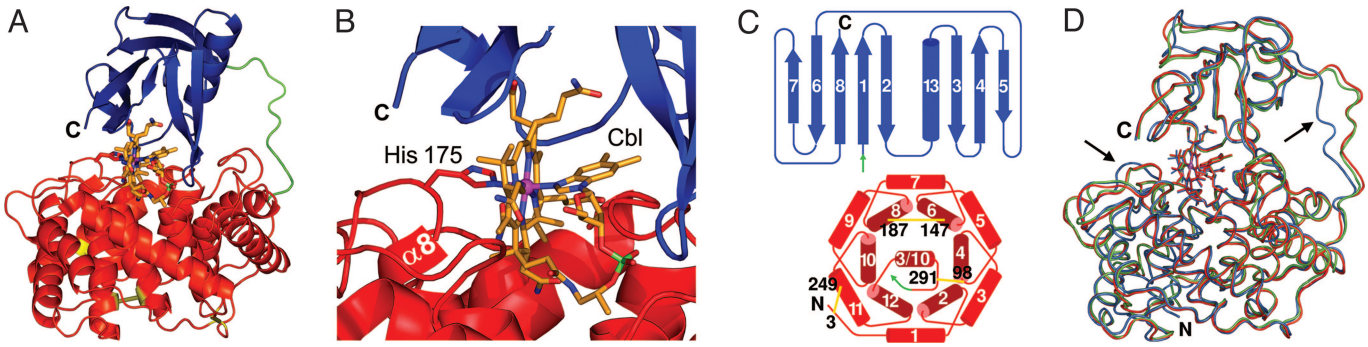

Fig. 1. Overall structure of holo-TC. (*A*) Secondary structure cartoon (here, bovine TC) showing the N-terminal α-domain (red) and C-terminal β-domain (blue) with the flexible linker (green) and the Cbl (orange) in the central domain interface. (*B*) Enlarged view to the histidine-coordinated Cbl, which is shown in ball-and-stick representation (magenta, Co ion; lime, phosphorus; red, oxygen; blue, nitrogen; orange, carbon). (*C*) Topology diagram of the two domains with a separate numbering for helices and strands and the color scheme of *A*. In the α-domain's α_6-α_6 barrel, the helix orientations run clockwise in the inner α_6 bundle and counterclockwise in the outer. A 3/10 helix at the bottom center shields the hydrophobic core of the domain. The three disulfide bridges are shown as yellow sticks and Cys residues are labeled (human TC numbering). (*D*) Superposition of the backbone trace of human holo-TC (blue) with both crystal forms of bovine TC (green and red, chain A in all cases). Arrows indicate the patches of gaps in the sequence alignment of Fig. 3. Stereo versions of *A* and *D* are shown in Fig. 6.

The α-domain spans ≈300 residues and is built of 12 α-helices and a 3/10-helix (Fig. 1*C*). Every even-numbered α-helix participates in the formation of a parallel six-helix bundle, which constitutes the core of the domain. The bundle is surrounded in an approximately hexagonal arrangement by the six other helices whose axes are strongly inclined (30–45°) with respect to that of the inner six-helix bundle. The orientation of the outer helices is parallel with respect to each other but antiparallel with respect to the inner helices. The resulting α_6-α_6 barrel is closed on the side opposite to the β-domain by the 3/10-helix as a central lid. Disulfide bridges, previously identified by chemical analysis of bovine TC (21), stabilize this globular domain. The location of the disulfide bridges suggests that two of them represent important elements for domain stability, because they involve parts of the inner six-helix bundle (Fig. 1*C*). These are the bridges Cys-98–Cys-291 and Cys-147–Cys-187 (mature human TC numbering). The bridge Cys-3–Cys-249 merely attaches the N-terminal loop to the outer α_6 barrel, explaining the finding that Cbl binding is severely impaired upon disruption of the bridges Cys-98–Cys-291 and Cys-147–Cys-187 but not of the bridge Cys-3–Cys-249 (22).

The β-domain spans ≈100 residues and consists of a five-stranded β-sheet that packs against a smaller sheet containing three β-strands and a single α-helix. A linker region of 11 residues connects the two domains in bovine TC, whereas only nine residues were found in the human form. A high degree of flexibility is attributed to the linker, because significant disorder in the electron density maps is observed for this region. Apart from the linker, other regions of elevated flexibility were identified by superposition of crystallographically independent molecules (two of human TC and five of bovine TC; see Fig. 1*D* and Fig. 7, which is published as supporting information on the PNAS web site), as well as by analysis of the thermal parameter B (Fig. 7). These regions are loops in the α-domain that encompass in human TC the stretches Gly-68–Gly-80, Gly-125–Gly-130, Pro-167–His-172, and Thr-237–Pro-241.

Although the overall fold of TC appears to be unique and unrelated to Cbl-binding proteins (discussed below), a search for structurally similar protein fragments revealed that the α-domain of TC (Fig. 1*C*) resembles domain 1 in the functionally unrelated membrane protein squalene-hopene cyclase (23).

Cbl Environment. Cbl in base-on conformation (i.e., Co coordinated by the dimethylbenzimidazole group) is located between the two domains with the plane of the corrin ring approximately perpendicular to the domain interface (Fig. 1*B*). The contact areas of Cbl with each of the domains are virtually equal in size. Two types of strong interaction between Cbl and TC are present. First, there is a coordination bond of the Co ion (in oxidation state +3) to a histidine side chain at the upper axial position (termed β-side) of Cbl (Fig. 2*A*). The histidine replaced the H_2O at the β-side of H_2O-Cbl used in the preparation of holo-TC. Second, a number of hydrogen bonds link Cbl to TC either directly or indirectly via ordered solvent molecules (Fig. 2*B*). Hydrogen bonds involve all side chains of Cbl's corrin macrocycle and two oxygen atoms of the phosphate group. Regarding the transport protein, six residues of the α-domain and an equal number of residues of the β-domain participate in direct hydrogen bonds. A further seven α-domain residues and six β-domain residues are involved in solvent-mediated links to Cbl. The hydrogen bond network is the same for human and bovine TC (Fig. 2*B*). The α-domain interacts mainly via side chains, whereas the β-domain exclusively uses backbone oxygen or nitrogen atoms. Hydrophobic interactions are fewer in number and in human TC involve Met-270 (at ring A of Cbl), Gly-390 (ring B), Phe-376, and Trp-409 (ring C), Tyr-137 (ring D), as well as Gly-85, Ser-359, and Tyr-362 at the nucleotide moiety. All these interactions are present also in bovine TC, with an additional interaction of Val-176 to ring D.

The histidine, forming the coordination bond to Cbl via its imidazole Nε atom, is positioned at the end of the flexible loop that precedes helix α8 (Figs. 1*B* and 3). For both TC forms, the electron density maps indicate elevated mobility of this loop. In bovine TC, continuous electron density is observed between His-175 Nε and the Co ion (Fig. 2*A*). The bond length (restrained to 2.1 Å; see *Materials and Methods*) averaged 2.13 Å over the five crystallographically independent molecules of the two crystal forms of bovine TC. In the structure of human TC, His-173 is in a position to form the analogous coordination bond. However, the clear discontinuity in the electron density map (Fig. 8, which is published as supporting information on the PNAS web site) and the greater distance (2.8 Å) between the Nε nitrogen atom of His-173 and the Co ion suggest the presence of a mix of conformations with and without His-Co coordination (within the limits of resolution of our crystallographic data; see Table 1, which is published as supporting information on the PNAS web site). Reduction of hexacoordinate Co(III) to pentacoordinate Co(II) may have been induced by x-rays during data collection, as observed in a similar case of the Cbl-dependent enzyme glutamate mutase (24), and may have led to the loss of His coordination in our case. We hypothesize that only human TC was affected, because the loop preceding His-173 in human

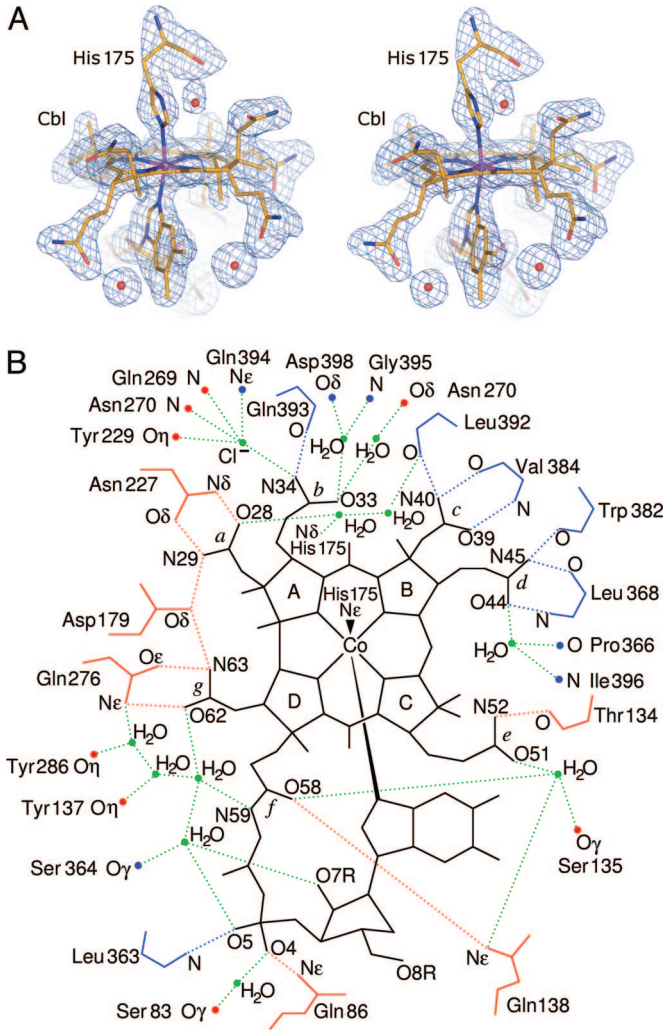

Fig. 2. Cbl interactions with bovine TC (monoclinic crystal form). (*A*) Stereoview of the $F_o - F_c$ omit electron density map at 2.0-Å resolution around the coordination of His-175 Nε to the Co ion of Cbl (contour level 3 σ). Some solvent water molecules are shown as red spheres, one of which forms a H-bond to His-175 Nδ. The Co ion (magenta) is axially coordinated by the imidazole Nε at a distance of 2.13 Å (above the corrin plane) and by the dimethylbenzimidazole nitrogen N3B at 2.09 Å. (*B*) Scheme of polar interactions. H-bonds are shown as dotted lines (red, to residues in the α-domain; blue, to residues in the β-domain; green, solvent-mediated interactions). The main or side chain is indicated for residues in direct contact with Cbl, whereas dots indicate residues linked to Cbl via solvent molecules (11 H_2O and a Cl^- ion from NaCl salt). The same scheme of direct contacts is observed in human TC (to translate from bovine to human TC numbering, see Fig. 3).

TC is three residues shorter than that in bovine TC (Fig. 3) and may generate a stress that facilitates detachment of the imidazole side chain. This notion is corroborated by the slightly higher rate constant of His replacement with H_2O in human TC (k_{-2} = $5{\cdot}10^{-4}{\cdot}s^{-1}$; ref. 19) compared to bovine TC (k_{-2} = $2{\cdot}10^{-4}{\cdot}s^{-1}$; ref. 21).

According to the sequence alignment (Fig. 3), the residue in human TC that corresponds to His-175 in bovine TC is expected to be His-172. However, the electron density map points to His-173 as the Cbl-binding residue (Fig. 8). Yet it cannot be ruled out, due to the observed disorder in this loop, that *in vivo* His-172 instead of His-173 may also take up a suitable position to function as Co-coordinating residue.

Cbl is tightly embedded in the protein, and the solvent-accessible surface of free Cbl is reduced upon binding to TC to only ≈7%. The portion of Cbl with contact to the bulk solvent is limited to the ribose group of the nucleotide moiety (Fig. 4*A*). The 5′-hydroxyl group is situated at the center of this small window to Cbl. In holo-TC, the oxygen atom (O8R in Fig. 2*B*) maintains ≈70% of its solvent accessible area in free Cbl. A small deformation of Cbl bound to human TC exists at the corrin ring C to avoid a steric clash between a methyl group and the phenolic oxygen of Tyr-362. Bovine TC shows a less bulky phenylalanine (Phe-367) at the same position.

Discussion

Capturing the Cbl ligand in the interface of two rigid domains is a well known binding mode. Structurally characterized Cbl-binding proteins other than TC act similarly, including the Cbl-dependent enzymes (25) methionine synthase (1), methylmalonyl CoA mutase (2), glutamate mutase (24), glycerol-dehydratase, diol-dehydratase, and Cob-adenosyltransferase, as well as the periplasmic Cbl-binding protein BtuF from *Escherichia coli* (26). Coordination of the Co ion by a histidine residue is present only in mutases and methionine synthase (25). In these cases, the histidine replaces the dimethylbenzimidazole group at the lower side (α) of Cbl, thus converting the cofactor to its base-off conformation. On the contrary, the histidine of holo-TC coordinates on the upper side (β) of Cbl, leaving Cbl in the base-on conformation. This makes the mode of axial His-Co coordination in TC unique among Cbl-related proteins.

Results from our crystal structure analysis combined with previously performed kinetics studies (6, 19) using Cbl with different upper axial groups (H_2O, CN, and N_3) allow reconstruction of the main steps during Cbl binding to TC. Cbl is almost completely buried inside the domain interface of holo-TC and thus cannot enter the binding site without an initial gap between the two domains. The width of this gap is likely to be determined by the flexibility of the linker region (Fig. 1*A*). Binding of Cbl may first occur either to the α- or the β-domain. One can speculate whether this happens without preference, because the interaction area and the number of hydrogen bonds between Cbl and either of the two domains are very similar. Once Cbl is attached to a domain, optimal shape complementarity is created between the domains, and fast encapsulation of Cbl can follow. The "enclosure mode" of binding occurs in a single step with a rate constant in the order of 10^7–10^8 $M^{-1}{\cdot}s^{-1}$ (19).

An additional slow phase (k_{+2} = 0.02 s^{-1}) was observed only for binding of H_2O-Cbl (19), i.e., Cbl having a weakly coordinated upper axial group. As follows from the crystal structures, H_2O was replaced by a histidine residue at the end of the loop connecting helices α7 and α8 (Figs. 1*B* and 2*A*). The slow phase is thus identified with the substitution of His for H_2O. The upper axial groups CN and N_3 are not expected to be displaced by His, because the corresponding slow phase is absent in the binding kinetics of CN-Cbl and N_3-Cbl (19). The absence is consistent both with the increasing affinity of Co(III) to ligands in the row H_2O < imidazole < N_3^- < CN^- (27) and with a characteristic shift in the UV-absorption spectrum of Cbl observed only upon displacement of H_2O (6, 19). The histidine residue of TC is brought into position for Co coordination by virtue of both the correct orientation of Cbl at the interface of the two domains and a moderate level of mobility present at its border location between a highly mobile loop and a firm helix α8. The activation energy of H_2O displacement in TC is quite high (120 kJ/mol), which points to some constraints during shaping of the achieved final structure (19). Nevertheless, even after substitution of the original axial ligand, high flexibility of the histidine-hosting loop remains. Consequently, no substantial structural difference can be expected between a His-off conformation, e.g., TC-(CN-Cbl) and the His-on conformation TC-(His-Cbl).

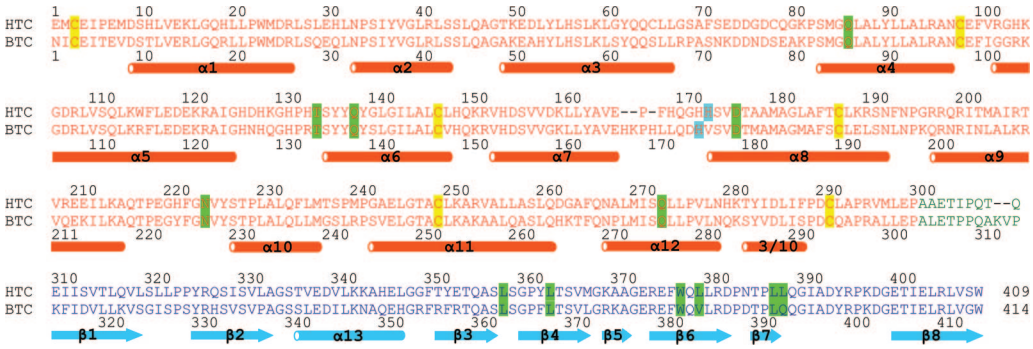

Fig. 3. Sequence alignment and secondary structure elements. The insertion of gaps in the sequence of human TC and the assignment of secondary structure elements are based on the x-ray structures. One-letter residue codes as well as helix and strand symbols are colored in correspondence to Fig. 1 *A* and *C*. Cys residues of disulfide bridges are highlighted in yellow, residues involved in direct H-bonds to Cbl in green, and the Co-coordinating His residue in blue. The sequence identity is 73%.

Mutagenic studies combined with Cbl-binding assays (28) were performed to detect essential residues for Cbl binding to human TC and can now be interpreted in terms of structure–function relationships. The region Ser-174–Thr-186 in human TC was subjected to site-directed mutagenesis, because it shows the highest sequence similarity between Cbl transporters (Fig. 9*A*, which is published as supporting information on the PNAS web site). Ser-174, Asp-176, and Thr-177 were investigated in single, double, and triple mutations to alanine as well as Ala-179, Ala-181, and Ala-184 in single mutations to arginine. Only one of the above residues (Asp-176) is actually involved in a H-bond to Cbl (Fig. 2*B*). The results of substitution showed that all single mutations to alanine left the Cbl-binding capacity invariant in contrast to the double and triple mutations. This indicates that removal of the direct H-bond between Asp-176 and Cbl is still permitted but not the substitution of two or more hydrophobic Ala residues for the original polar Ser, Asp, and Thr groups in the vicinity of the Cbl-binding site. Any of the single Ala-to-Arg mutations resulted in impaired protein stability. This can be structurally rationalized by the disruption of hydrophobic interactions of helix α8 with the neighboring helices α6, α7, and α10 (Fig. 10, which is published as supporting information on the PNAS web site).

The electrostatic potential on the protein's surface reveals a contrast between the two investigated TC forms (Fig. 4). Human TC shows predominantly a negative potential, reflecting an excess of six negatively charged residues over positively charged residues in its amino acid composition. Bovine TC, with an excess of 10 positively charged residues, is mainly positive on its surface. This difference does not seem to influence the affinity of Cbl to these TC forms, because the dissociation kinetics for H_2O-Cbl are similar ($k_- \approx 10^{-6}\cdot s^{-1}$ at 37°C, pH 7.5), and the affinity of both TC forms for Cbl can be estimated as $K_a \approx 10^{14}\cdot M^{-1}$ (19, 21). This similar behavior is in agreement with the observed identity of the pattern of H-bond interactions in both forms (Fig. 2*B*).

Recent mapping of functional domains of human TC by using monoclonal antibodies (20) yielded some information about the receptor-recognition site. Several antibodies and heparin negatively affected TC–receptor interaction. The corresponding epitopes are now assignable to solvent-exposed parts of the

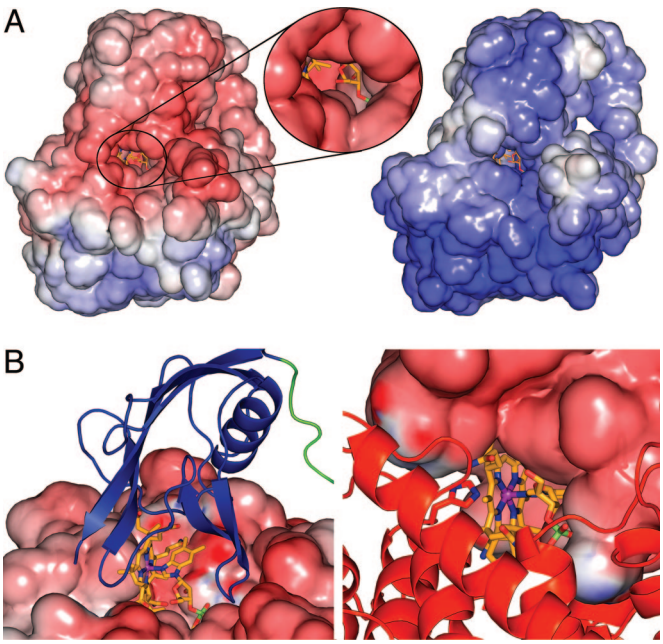

Fig. 4. Solvent-accessible surface and electrostatic potential properties of TC. (*A*) Human (*Left*) and bovine (*Right*) TC (color coding: blue surface, positive potential; red surface, negative potential). *Inset* zooms in on the solvent-accessible portion of Cbl in human TC. (*B*) A negative electrostatic potential prevails at the domain interface in human TC (*Left*, α-domain; *Right*, β-domain). Also illustrated is the hemispherical depression in both domains at their interface, each of which accommodates approximately half of the Cbl molecule.

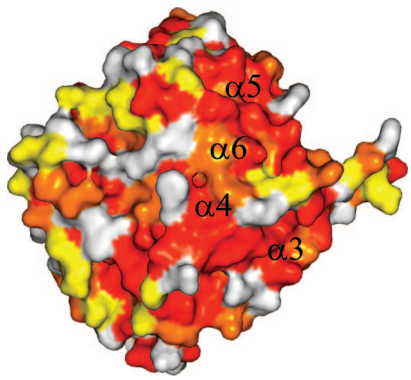

Fig. 5. Mapping the amino acid conservation among seven TCs (Fig. 9*B*) on the molecular surface of human TC. Color coding: red, identity; orange, conserved; yellow, semiconserved; white, not conserved. The view to the α-domain is as in Fig. 1C. The conserved region proposed as TC's receptor-recognition site is located on the right half and involves the surface of the labeled helices α3–α6 and their loops. A low level of conservation is present on the β-domain surface.

closely situated regions Asn-97–His-133 (helix α5 and adjacent loops) and Lys-189–Thr-207 (from the end of helix α8 to the center of helix α9). They overlap well with an extended surface region of strict amino acid conservation among seven mammalian TCs (Fig. 9*B*). This surface patch (Fig. 5) includes the region located between the two epitopes, His-149–Glu-166 (α7 and preceding loop) and Tyr-54–Leu-67 (helix α3). This conserved region on TC's surface is thus a good candidate for the composite receptor-recognition site.

IF and HC in Comparison to TC. IF and HC, the two other mammalian Cbl transporters, are glycosylated proteins with a total number of amino acids similar to TC. Glycosylation has posed so far a severe obstacle for the determination of their native structure. The structure of TC may serve as model for the paralogues IF and HC (28). Besides the transporters' similarity in Cbl-binding rate constants of $k_+ \approx 10^8$ $M^{-1}{\cdot}s^{-1}$ (6, 19), structural similarity exists in regions corresponding to TC's secondary structure elements (Fig. 9*A*), especially those forming the inner six-helix bundle. The six cysteine residues that form disulfide bridges in the α-domain of TC are conserved. We suggest that the overall architecture of IF and HC, and in particular the α-domain, resembles that of TC. This supposition is corroborated by recent analysis of two IF fragments (29, 30). These originated from natural proteolysis of IF from recombinant plants and closely corresponded to the α- and β-domains of TC according to their N-terminal sequencing. The separated fragments of IF were efficiently assembled by Cbl in a firm complex, which in Cbl binding and receptor recognition displayed a behavior very similar to that of uncleaved IF (30). Thus, two completely independent domains could still produce an (α-domain)–Cbl–(β-domain) sandwich complex of correct geometry. The finding that truncation of 50 amino acids from the C-terminal end of human IF abolished Cbl binding (31) is in accordance with our suggestion, because this truncation would imply for TC the loss of all H-bond interactions between Cbl and the β-domain (Figs. 2*B* and 3).

Binding of Cbl to human IF seems to occur sequentially. Preferential binding to the smaller C-terminal domain with high affinity ($k_{+1} = 10^7$–10^8 $M^{-1}{\cdot}s^{-1}$) is followed by interaction with the larger N-terminal domain with low affinity (29, 30). Assembly of the IF–Cbl complex may be accomplished via stepwise accommodation of the ligand between two protein modules ($k_{+2} = 10$–100 s^{-1}). A similar mechanism of protein–ligand interaction could be conjectured for TC. However, the approach of the two domains upon Cbl binding to TC is expected to occur faster, because no corresponding distinct phase was present in the binding kinetics (19). Another difference between IF and TC seems to be the composition of the receptor-recognition site, which in IF is formed by regions stemming from both domains (30).

The sequence alignment of all three Cbl transporters (Fig. 9*A*) indicates the lack of the Co-coordinating histidine residue in the sequences of IF and HC. This is consistent with high reactivity of IF-(H_2O-Cbl) and HC-(H_2O-Cbl) complexes with externally supplied histidine in contrast to the inert complex TC-(H_2O/His-Cbl) (6).

Assuming that IF and HC adopt the same fold as TC, the sequence alignment (Fig. 9*A*) also permits the mapping of potential Asn-linked glycosylation sites in IF and HC onto their structure. Fig. 8*A* indicates that HC possesses glycosylation sites on both domains and at the N-terminal side of its linker region, whereas IF shows such sites exclusively on the β-domain. Notably, both potential glycosylation sites on the α-domain of HC fall on the surface region above proposed as receptor-recognition site of TC. This posttranslational modification of HC may guarantee the TC-receptor's discrimination between the two plasma Cbl transporters, HC and TC.

We conclude that the protein scaffold of all of the Cbl transporters displays most likely the same two-domain architecture, with the place of Cbl binding in the domain interface. Yet the exact kinetic manifestation of binding is not identical for the three proteins. A subsequent step of Cbl binding, involving the replacement of a weak upper axial ligand of the Co ion with a histidine residue, appears to be a particular feature present solely in TC.

Polymorphisms in Human TC. Several single-nucleotide substitutions in the TC gene cause polymorphism of this protein with change of amino acids and potential alteration in the protein's ability to deliver Cbl to cells. Intracellular Cbl deficiency results in elevated plasma levels of substrates of Cbl-dependent enzymes, methylmalonic acid (MMA) and homocysteine. Increased total homocysteine (tHcy) is under intense investigation as a risk factor for vascular disease and stroke (32) as well as for neural tube defects (33). Numerous clinical studies investigated a possible association of TC polymorphisms, in particular the most common polymorphism P259R, with changes in plasma concentrations of holo-TC, total Cbl, MMA, or tHcy (12, 34–37). Some evidence relates the 259R allele to lower holo-TC and higher MMA concentrations when compared with the 259P allele (12, 35). However, the literature about an influence of this polymorphism on the plasma levels of Cbl and tHcy is still ambiguous (38).

Our structural data show that Pro-259 (i.e., Pro-241 in mature TC) is part of the solvent-exposed flexible loop between helices α10 and α11, spatially near the N terminus (Fig. 1 *C* and *D*) and suggest that the P259R polymorphism will influence neither the binding ability of Cbl to TC nor the stability of the α-domain. An involvement of this substitution in receptor recognition of holo-TC appears unlikely due to high sequence variability among seven mammalian TCs in the vicinity of Pro-241 (Fig. 9*B*), and no direct evidence exists for it.

Drug Delivery by TC. The potential of Cbl-based derivatives as targeting agents has recently been demonstrated in various fields. These include imaging of tumor cells with the help of Cbl conjugates (15) as well as cancer treatment by chemotherapy with cytotoxins linked to Cbl (16). To guarantee efficiency of the target, attachment of an active ligand to Cbl or modification of chemical groups in Cbl must not significantly diminish the affinity of the resulting analogue for TC. A rational approach to the design of successful targets is now possible with the help of the TC structure. Inspection of the Cbl environment in human TC (Fig. 4*A*) shows that only the 5′-hydroxyl group on the ribose of Cbl's nucleotide moiety can accommodate the attachment of a larger ligand molecule without disturbing the interactions between TC and Cbl (Fig. 2*B*). Attachment at the corrin side chains or the phosphate group will inevitably lead to impaired interactions. Beside steric hindrance, electrostatic interactions between the attached ligand and residues of TC should be considered. For instance, the surface of human TC around the Cbl-containing cavity is in part negatively charged (Fig. 4). The upper axial ligand position of the Co ion was successfully used as well (14) despite the absence of the His coordination.

Materials and Methods

Protein Expression and Purification. Expression of recombinant human and bovine TC in *Pichia pastoris* and their purification as holo-forms using H_2O-Cbl is described in refs. 19 and 21, respectively.

Crystallization. Crystallization of human TC was carried out by the hanging-drop vapor-diffusion technique, as described (39). The crystal used for structure determination grew from 20% polyethylene glycol (PEG) 4000/18% ethanol/0.2 M Tris, pH 7.6, at 293 K with the protein at 6 $mg{\cdot}ml^{-1}$ in 0.8 M NaCl/80 mM Tris, pH 7.5.

Bovine TC crystallized in two different crystal forms by using the same technique as for human TC. Both monoclinic and trigonal forms grew from 28% PEG 8000/0.2 M magnesium acetate/0.1 M Tris, pH 8.5/20% 2-methyl-2,4-pentadiol with the protein at 25 $mg \cdot ml^{-1}$ in 1 M NaCl/0.1 M Tris, pH 7.5.

X-Ray Data Collection and Processing. Diffraction data for both human and bovine TC were collected at the Elettra synchrotron radiation source (Trieste, Italy) at cryogenic temperature (100 K; see Table 1). Human TC data were processed with MOSFLM/SCALA (40) and bovine TC data with HKL2000 (41).

Structure Determination and Refinement. Solution of the crystallographic phase problem for human TC was hampered by low diffraction power of the thin-plate crystals and by purely translational noncrystallographic symmetry (NCS) (39). The phase problem was solved by using bovine TC and single-wavelength anomalous dispersion data to 2.8-Å resolution collected on the trigonal crystal form at the Co K-edge with 16-fold multiplicity. The Co coordinates were found by using SHELXD (42), and initial phases were calculated with SHARP (43). A partial protein model was obtained by interpretation of solvent-flattened electron-density maps by using the graphics program O (44) and automated procedures in ARP/WARP (45). The partial model was used in molecular replacement calculations with AMORE (46), exploiting the 2.0-Å resolution of bovine TC data for the monoclinic crystal form. Electron-density interpretation resulted in a complete model that was refined with REFMAC5 (47), applying NCS restraints and TLS refinement. This model served then as start for refinement against bovine TC data of the trigonal crystal form at 2.4-Å resolution, as well as for molecular replacement calculations to solve the phase problem for human TC. The human TC model was initially refined with CNS (48) by simulated annealing with a cartesian molecular dynamics protocol and subsequently with REFMAC5 applying NCS restraints and TLS refinement. The B_{12} monomer library of REFMAC5 was used to refine Cbl. The distance between Co and Nε of the His ligand of bovine TC was restrained to 2.1 Å with an estimated standard deviation of 0.03 but left unrestrained in human TC refinement. Table 1 reports final refinement statistics.

Protein Structure Analysis. The electrostatic potential on the solvent accessible protein surface (1.4-Å probe radius) was calculated with APBS (49), assuming neutral histidine side chains, and visualized with PYMOL (DeLano Scientific, South San Francisco, CA). Atomic accessibility areas were calculated with NACCESS (S. J. Hubbard and J. M. Thornton, University College London, London). Structurally similar proteins in the Protein Data Bank (www.rcsb.org/pdb) were searched with DALI (50).

We thank the staff of the XRD1 beamline at the Elettra synchrotron (Trieste, Italy). This work was funded by Projects FIRB2003-RBNE03PX83 and PRIN2003037580 (Ministero dell'Istruzione dell'Università e della Ricerca, Rome), Eureka (CT-T2006), and the Lundbeck Foundation.

1. Drennan, C. L., Huang, S., Drummond, J. T., Matthews, R. G. & Ludwig, M. L. (1994) *Science* **266,** 1669–1674.
2. Mancia, F., Keep, N. H., Nakagawa, A., Leadlay, P. F., McSweeney, S., Rasmussen, B., Bösecke, P., Diat, O. & Evans, P. R. (1996) *Structure (London)* **4,** 339–350.
3. Seetharam, B., Bose, S. & Li, N. (1999) *J. Nutr.* **129,** 1761–1764.
4. Quadros, E. V., Nakayama, Y. & Sequeira, J. M. (2005) *Biochem. Biophys. Res. Commun.* **327,** 1006–1010.
5. Quadros, E. V., Regec, A. L., Kahn, K. M., Quadros, E. & Rothenberg, S. P. (1999) *Am. J. Physiol.* **277,** G161–G166.
6. Fedosov, S. N., Berglund, L., Fedosova, N. U., Nexø, E. & Petersen, T. E. (2002) *J. Biol. Chem.* **277,** 9989–9996.
7. Kolhouse, J. F. & Allen, R. H. (1977) *J. Clin. Invest.* **60,** 1381–1392.
8. Lloyd-Wright, Z., Hvas, A. M., Møller, J., Sanders, T. A. & Nexø, E. (2003) *Clin. Chem.* **49,** 2076–2078.
9. Wolters, M., Strohle, A. & Hahn, A. (2004) *Prev. Med.* **39,** 1256–1266.
10. Haurani, F. I., Hall, C. A. & Rubin, R. (1979) *J. Clin. Invest.* **64,** 1253–1259.
11. Seligman, P. A., Steiner, L. L. & Allen, R. H. (1980) *N. Engl. J. Med.* **303,** 1209–1212.
12. von Castel-Dunwoody, K. M., Kauwell, G. P. A., Shelnutt, K. P., Vaughn, J. D., Griffin, E. R., Maneval, D. R., Theriaque, D. W. & Bailey, L. B. (2005) *Am. J. Clin. Nutr.* **81,** 1436–1441.
13. Hogenkamp, H. P., Collins, D. A., Grissom, C. B. & West, F. G. (1999) in *Chemistry and Biochemistry of B12*, ed. Banerjee, R. (Wiley, New York), pp. 385–410.
14. Pathare, P. M., Wilbur, D. S., Heusser, S., Quadros, E. V., McLoughlin, P. & Morgan, A. C. (1996) *Bioconjugate Chem.* **7,** 217–232.
15. McGreevy, J. M., Cannon, M. J. & Grissom, C. B. (2003) *J. Surg. Res.* **111,** 38–44.
16. Russell-Jones, G., McTavish, K., McEwan, J., Rice, J. & Nowotnik, D. (2004) *J. Inorg. Biochem.* **98,** 1625–1633.
17. Allen, R. H. & Majerus, P. W. (1972) *J. Biol. Chem.* **247,** 7709–7717.
18. Platica, O., Janeczko, R., Quadros, E. V., Regec, A., Romain, R. & Rothenberg, S. P. (1991) *J. Biol. Chem.* **266,** 7860–7863.
19. Fedosov, S. N., Fedosova, N. U., Nexø, E. & Petersen, T. E. (2000) *J. Biol. Chem.* **275,** 11791–11798.
20. Fedosov, S. N., Örning, L., Løvli, T., Quadros, E. V., Thompson, K., Berglund, L. & Petersen, T. E. (2005) *FEBS J.* **272,** 3887–3898.
21. Fedosov, S. N., Berglund, L., Nexø, E. & Petersen, T. E. (1999) *J. Biol. Chem.* **274,** 26015–26020.
22. Kalra, S., Li, N., Seetharam, S., Alpers, D. H. & Seetharam, B. (2003) *Am. J. Physiol.* **285,** C150–C160.
23. Wendt, K. U., Poralla, K. & Schulz, G. E. (1997) *Science* **277,** 1811–1815.
24. Champloy, F., Gruber, K., Jogl, G. & Kratky, C. (2000) *J. Synchrotron Rad.* **7,** 267–273.
25. Banerjee, R. & Ragsdale, S. W. (2003) *Annu. Rev. Biochem.* **72,** 209–247.
26. Borths, E. L., Locher, K. P., Lee, A. T. & Rees, D. C. (2002) *Proc. Natl. Acad. Sci. USA* **99,** 16642–16647.
27. Pratt, J. M. (1972) *Inorganic Chemistry of Vitamin B12* (Academic, London).
28. Kalra, S., Li, N., Yammani, R. R., Seetharam, S. & Seetharam, B. (2004) *Arch. Biochem. Biophys.* **431,** 189–196.
29. Fedosov, S. N., Fedosova, N. U., Berglund, L., Moestrup, S. K., Nexø, E. & Petersen, T. E. (2004) *Biochemistry* **43,** 15095–15102.
30. Fedosov, S. N., Fedosova, N. U., Berglund, L., Moestrup, S. K., Nexø, E. & Petersen, T. E. (2005) *Biochemistry* **44,** 3604–3614.
31. Tang, L. H., Chokshi, H., Hu, C. B., Gordon, M. M. & Alpers, D. H. (1992) *J. Biol. Chem.* **267,** 22982–22986.
32. Wald, D. S., Law, M. & Morris, J. K. (2002) *BMJ* **325,** 1202–1206.
33. Guéant-Rodriguez, R. M., Rendeli, C., Namour, B., Venuti, L., Romano, A., Anello, G., Bosco, P., Debard, R., Gérard, P., Viola, M., *et al.* (2003) *Neurosci. Lett.* **344,** 189–192.
34. Lievers, K. J. A., Afman, L. A., Kluijtmans, L. A. J., Boers, G. H. J., Verhoef, P., den Heijer, M., Trijbels, F. J. M. & Blom, H. J. (2002) *Clin. Chem.* **48,** 1383–1389.
35. Miller, J. W., Ramos, M. I., Garrod, M. G., Flynn, M. A. & Green, R. (2002) *Blood* **100,** 718–720.
36. Zetterberg, H., Nexø, E., Regland, B., Minthon, L., Boson, R., Palmér, M., Rymo, L. & Blennow, K. (2003) *Clin. Chem.* **49,** 1195–1198.
37. Swanson, D. A., Pangilinan, F., Mills, J. L., Kirke, P. N., Conley, M., Weiler, A., Frey, T., Parle-McDermott, A., *et al.* (2005) *Birth Defects Res.* **73**, 239–244.
38. Sunder-Plassmann, G. & Födinger, M. (2003) *Kidney Int.* **63,** S141–S144.
39. Garau, G., Fedosov, S. N., Petersen, T. E., Geremia, S. & Randaccio, L. (2001) *Acta Crystallogr. D* **57,** 1890–1892.
40. Collaborative Computational Project 4 (1994) *Acta Crystallogr. D* **50,** 760–763.
41. Otwinowski, Z. & Minor, W. (1997) *Methods Enzymol.* **276,** 307–326.
42. Schneider, T. R. & Sheldrick, G. M. (2002) *Acta Crystallogr. D* **58,** 1772–1779.
43. Bricogne, G., Vonrhein, C., Flensburg, C., Schiltz, M. & Paciorek, W. (2003) *Acta Crystallogr. D* **59,** 2023–2030.
44. Jones, T. A., Zou, J. Y., Cowan, S. W. & Kjeldgaard., M. (1991) *Acta Crystallogr. A* **47**, 110–119.
45. Perrakis, A., Morris, R. & Lamzin, V. S. (1999) *Nat. Struct. Biol.* **6,** 458–463.
46. Navaza, J. (2001) *Acta Crystallogr. D* **57,** 1367–1372.
47. Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997) *Acta Crystallogr. D* **53,** 240–255.
48. Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., *et al.* (1998) *Acta Crystallogr. D* **54**, 905–921.
49. Baker, N. A., Sept, D., Joseph, S., Holst, M. J. & McCammon, J. A. (2001) *Proc. Natl. Acad. Sci. USA* **98,** 10037–10041.
50. Holm, L. & Sander, C. (1996) *Science* **273,** 595–603.