

ARTICLE

Received 13 Apr 2016 | Accepted 27 May 2016 | Published 14 Jul 2016

DOI: 10.1038/ncomms12100

OPEN

Structural basis of transcobalamin recognition by human CD320 receptor

Amer Alam¹, Jae-Sung Woo¹, Jennifer Schmitz¹, Bernadette Prinz¹, Katharina Root², Fan Chen², Joël S. Bloch¹, Renato Zenobi² & Kaspar P. Locher¹

Cellular uptake of vitamin B12 (cobalamin) requires capture of transcobalamin (TC) from the plasma by CD320, a ubiquitous cell surface receptor of the LDLR family. Here we present the crystal structure of human holo-TC in complex with the extracellular domain of CD320, visualizing the structural basis of the TC-CD320 interaction. The observed interaction chemistry can rationalize the high affinity of CD320 for TC and lack of haptocorrin binding. The *in vitro* affinity and complex stability of TC-CD320 were quantitated using a solid-phase binding assay and thermostability analysis. Stable complexes with TC were also observed for the disease-causing CD320 Δ E88 mutant and for the isolated LDLR-A2 domain. We also determined the structure of the TC-CD320 Δ E88 complex, which revealed only minor changes compared with the wild-type complex. Finally, we demonstrate significantly reduced *in vitro* affinity of TC for CD320 at low pH, recapitulating the proposed ligand release during the endocytic pathway.

¹Department of Biology, Institute of Molecular Biology and Biophysics, ETH Zurich, 8093 Zurich, Switzerland. ²Department of Chemistry and Applied Biosciences, ETH Zurich, 8093 Zurich, Switzerland. Correspondence and requests for materials should be addressed to K.P.L. (email: locher@mol.biol.ethz.ch).

Mammals depend on dietary uptake of Cobalamin (Cbl), which is a cofactor of at least two enzymes that are part of essential biochemical pathways^{1,2}. A complex uptake system involving several transport proteins and cellular receptors facilitates efficient transport of this scarce vitamin^{3–10}. Defects at any stage of this process have been found to lead to Cbl deficiency and severe developmental defects in patients^{11,12}. transcobalamin (TC)-bound Cbl is transported into cells by receptor-mediated endocytosis, which requires Ca²⁺-dependent complex formation of TC with its cognate cell surface receptor CD320 (ref. 13). CD320 is a member of the Low-density Lipoprotein Receptor (LDLR) family^{13,14}, a large group of mostly multi-ligand receptors involved in cholesterol uptake and other physiologically essential functions^{15,16}. CD320 features an N-terminal, extracellular fragment comprising two LDLR type A (LDLR-A) domains separated by an epidermal growth factor (EGF) homology domain, the exact function of which has so far not been determined. It also contains a single transmembrane helix and a C-terminal cytoplasmic region involved in complex internalization⁸. While the majority of related LDLR domain-containing receptors bind multiple and distinct ligands¹⁷, CD320 has been reported to bind TC with high specificity and affinity as evidenced by the lack of interference in TC binding from related LDLR ligands like LDL and RAP as well as high affinity for TC of both the full-length receptor and its soluble ectodomain^{8,13,18}. Even haptocorrin (HC), a high affinity Cbl-binding protein that shares significant sequence and structural similarities with TC and is also circulating in the plasma, is not recognized by CD320.

Although several of the proteins facilitating Cbl uptake and transport from saliva to the blood stream have been structurally characterized^{19–22}, the final uptake step into cells has not been visualized at a molecular level. Such insight is required not only to understand the basis of specific TC-Cbl uptake into cells, but also to rationalize disease-causing mutations that interfere with this process. Furthermore, expression of CD320 is reported to be highest in actively proliferating cells and is strongly upregulated in many cancer cells, prompting interest in its potential targeting through receptor directed drug delivery^{23,24} or its exploitation for tumour diagnostics. In light of its physiological relevance, we set out to biochemically and structurally characterize the TC-CD320 complex, revealing key molecular details of the high affinity interaction, its regulation by pH and the *in vitro* effect of mutations of CD320 on its interaction with TC.

Results

Structural and functional analysis of TC-CD320 complex. The challenges associated with the production of suitable amounts of homogenous, stable, and functional CD320 protein have so far prevented its structural analysis^{7,25,26}. To overcome these challenges, we heterologously co-expressed human TC and the external, TC-binding part of CD320 in secreted form in baculovirus-infected insect cells, taking advantage of the enhanced biochemical stability of the complex over individual components. CD320 was N-terminally fused to maltose-binding protein (MBP) that was subsequently cleaved on purification. We purified the TC-CD320 complex to homogeneity in the presence of excess cyanocobalamin (Cbl-CN) to ensure full occupancy of the Cbl-binding site. The obtained complex had a stoichiometry of 1:1 as confirmed by MALDI mass spectrometry (Supplementary Fig. 1).

The stability of the TC-CD320 complex was analysed by determining its thermostability upon incubation at increasing temperatures, followed by size exclusion chromatography (SEC) analysis. As shown in Fig. 1a, we observed a >20 °C drop in the

apparent T_m in the presence of the Ca²⁺ chelator ethylene glycol tetraacetic acid (EGTA), confirming the stabilizing effect of Ca²⁺ on the TC-CD320 complex. We developed an enzyme-linked immunosorbent assay-based solid-phase binding assay to determine the affinity of the TC-CD320 interaction by using immobilized MBP-tagged, biotinylated CD320 (CD320_{AVI}). Tag-less TC was obtained by EGTA induced dissociation of a pre-formed TC-CD320_{AVI} complex immobilized on streptavidin resin. TC binding was quantitated using a commercial monoclonal antibody (A-5) targeting a distinct C-terminal epitope of TC, away from its receptor-binding site. As shown in Fig. 1b, TC was confirmed to bind to CD320 with high affinity (K_D ~1.5 nM), in line with the high thermal stability of the complex.

To analyse whether a single LDLR-A domain of CD320 was capable of binding TC (analogous to observations with the Reelin/LDLR-A1 complex²⁷), we individually co-expressed each domain with TC. Complex formation was observed both for LDLR-A1 and LDLR-A2, which is in contrast to a previous suggestion that two LDLR-A modules are the minimal binding unit for TC (ref. 8). We found that a complex containing only LDLR-A2 had an apparent T_m only ~5 °C lower than the equivalent complex containing full-length CD320, and a dissociation constant of ~10 nM, slightly reduced compared to full-length CD320 (Fig. 1b). In contrast, the complex with LDLR-A1 appeared less stable on purification, indicating a decreased stability and precluding a detailed thermostability or solid-phase binding analysis.

Various CD320 truncations were generated and tested and a final crystallization construct comprising residues 53–199 of the extracellular domain of CD320 and TC residues 1–409 was chosen for structure determination. The TC-CD320 co-crystal structure, phased by molecular replacement using Cbl-bound TC as a search model, was refined to 2.1 Å resolution and revealed a 1:1 complex of TC-CD320 (Fig. 1c). Whereas clear electron density was observed for LDLR-A1 residues 53–89 and LDLR-A2 residues 129–171, no density was observed for the EGF-like domain (residues 90–128). Both LDLR-A domains and TC contain several disulfide bridges, the positions of the majority of which were confirmed using an anomalous difference map calculated from data collected at a wavelength of 1.6 Å (Table 1). In the same map, Ca²⁺ ions bound to the LDLR-A domains and the Cobalt of Cbl bound to TC were identified.

Overall architecture of the TC-CD320 complex. In agreement with epitope mapping studies²⁸, both LDLR-A1 and LDLR-A2 of CD320 exclusively interact with the α domain of TC (Fig. 1c). The structure of CD320-bound TC is very similar to that of its free state²⁰ (r.m.s.d of all Cα atoms of 0.43 Å, Supplementary Fig. 2) with a few key differences. First, we observed electron density for a fourth disulfide bridge linking residues Cys65–Cys78, both of which are conserved in several primate and rodent species²⁰. Second, our structure revealed clear electron density for the loop connecting the α and β domains of TC. Finally, we observed electron density for the loop carrying the Cbl-binding His173.

To confirm the implications of the observed interface, we co-expressed CD320 with a truncated version of TC containing only its α domain (TCα) to mimic its apo form. We expressed the TCα-CD320 complex in insect cells to avoid denaturing conditions previously used to obtain apo-TC (ref. 29). We could indeed purify a TCα-CD320 complex with a 1:1 stoichiometry, which exhibited a decreased stability (apparent T_m of ~60 °C) compared with the complex containing full-length TC (apparent T_m of ~70 °C, Fig. 1a).

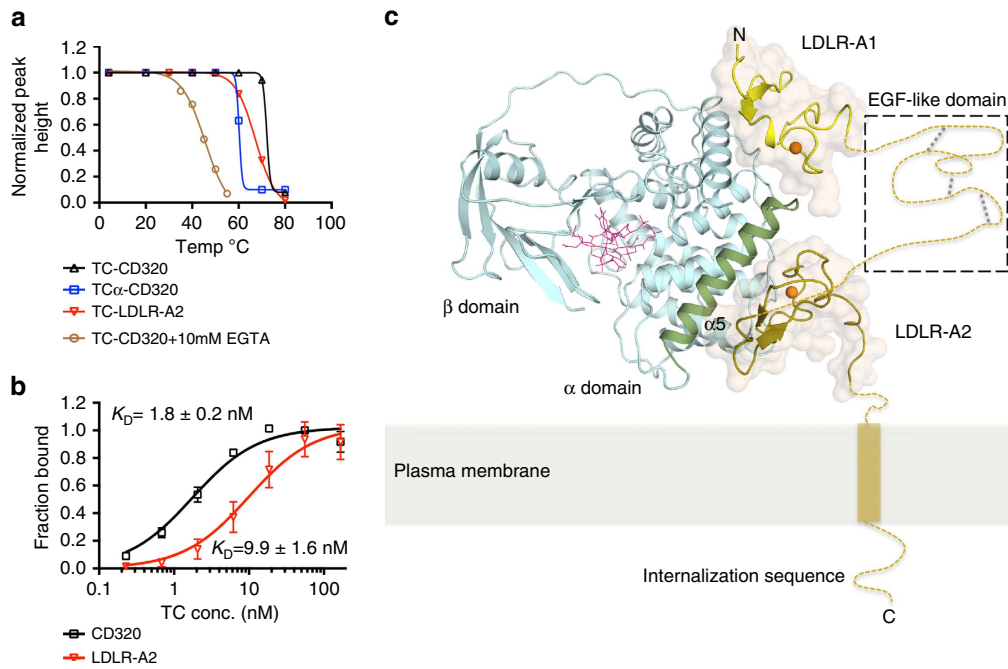


Figure 1 | Characterization of the TC-CD320 complex. (a) Thermostability curves for different TC-CD320 complexes, demonstrating decreased complex stability in the presence of EGTA as well as for the TC α -CD320 and TC-LDLR-A2 complexes. (b) Solid-phase binding assay quantitating TC binding to full-length CD320 (black) and to LDLR-A2 (red; $n = 5$, error bars indicate S.D). (c) Overall structure of TC-CD320 complex showing TC (cyan) and LDLR-A1 and LDLR-A2 of CD320 (yellow and gold, respectively). The likely position of the EGF-like module (boxed) bridging the two LDLR-A domains is shown schematically. The transmembrane domain (yellow cylinder) and internalization motif of CD320 were not part of the expression construct. Orange spheres represent bound Ca^{2+} ions. Bound Cbl, sandwiched between the α and β domains of TC, is shown in pink and TC helix $\alpha 5$ is coloured green.

Table 1 | Data collection and refinement statistics.

Crystal:	TC-CD320	TC-CD320 Δ E88	TC-CD320 _{ano}	TC-CD320 Δ E88 _{ano}
<i>Data collection</i>				
Wavelength (Å)	1.0	1.0	1.6	1.77
Space group	P43212	P43212	P43212	P43212
Cell dimensions				
a, b, c (Å)	98.10, 98.10, 356.0	98.40, 98.40, 356.3	97.94, 97.94, 356.5	97.81, 97.81, 354.7
α, β, γ (°)	90, 90, 90	90, 90, 90	90, 90, 90	90, 90, 90
Resolution (Å)	2.1	2.6	3.0	3.0
R_{merge}	0.1(0.9)*	0.121(1.62)	0.11 (0.44)	0.186(0.97)
$I/\sigma I$	12(1.0)	16.8(2.2)	19.8 (7.2)	25.4 (6.1)
Completeness (%)	95.8(62.3)	99.9(100)	98.3 (97.4)	98.4 (97.1)
Redundancy	4.8(2.5)	10.8(11.2)	9.0 (9.3)	35.7 (37.0)
<i>Refinement</i>				
Resolution (Å)	2.1	2.6	-	-
No. reflections	101389	54979	-	-
$R_{\text{work}}/R_{\text{free}}$	0.1840/0.2075	0.2001/0.2305	-	-
No. atoms				
Protein/Cbl	7515	7488	-	-
Ca^{2+} ions	6	4	-	-
Water	444	44	-	-
B-factors				
Protein/Cbl	58.6	74.4	-	-
Ca^{2+} ions	61.9	86.1	-	-
Water	55.0	62.4	-	-
r.m.s deviations				
Bond lengths (Å)	0.01	0.01	-	-
Bond angles (°)	1.93	1.89	-	-
Ramachandran plot				
Favoured (%)	96.5	96.6	-	-
Allowed (%)	3.4	3.1	-	-
Disallowed (%)	0.1	0.3	-	-

*Values in parentheses are for the highest-resolution shell.

The interface of TC and CD320 consists of $\sim 900 \text{ \AA}^2$ buried surface area, with LDLR-A2 contributing 486 \AA^2 compared with 405 \AA^2 for LDLR-A1. The fold of the LDLR-A domains of CD320 is similar to that of related family members^{30–34} (Fig. 2a–c). Both LDLR-A domains contain central Ca^{2+} ions bound by four conserved acidic residues and two backbone carbonyls in an octahedral coordination. They also contain a PLXWRCD motif that is conserved in other CD320 orthologs and related proteins¹³. The backbone carbonyl of the tryptophan residue in this motif (Trp72 and Trp150 in LDLR-A1 and LDLR-A2, respectively) coordinates the central Ca^{2+} ion, while its side chain is stabilized through a stacking interaction with the preceding proline residue. In the case of LDLR-A2, this interaction is further stabilized through stacking with Trp115 and hydrogen bonding with Glu118 of TC. These latter interactions are absent in the LDLR-A1/TC interface. In LDLR-A1, the Ca^{2+} -coordinating carboxylate groups of Asp75 and Asp79 form salt bridges with Lys105 of TC, forming an ‘acidic necklace’ previously found to be a key feature of LDLR/ligand complexes³² (Fig. 2a). The equivalent acidic necklace in LDLR-A2 is formed by Lys114 of TC and the Ca^{2+} coordinating carboxylate groups of Asp153 and Asp157 of CD320 (Fig. 2b).

A network of disulfide linkages stabilizes the folds of both LDLR-A1 and LDLR-A2, with disulfide formation following the 1-3/2-5/4-6 pattern observed previously^{30–32,34,35}. LDLR-A1 contacts TC helices $\alpha 3$ and $\alpha 5$, whereas LDLR-A2 interacts with TC helix $\alpha 5$ and the loop between helices $\alpha 6$ and $\alpha 7$ (Fig. 2b). LDLR-A2 is involved in an extensive network of electrostatic and hydrogen-bonding interactions with TC helix $\alpha 5$ (Fig. 2b) including an additional salt bridge between TC2 residue Arg122 and Glu136 of LDLR-A2. The overall quality of electron density in LDLR-A2 is higher than that of LDLR-A1. This is especially apparent for the bound Ca^{2+} ion and Ca^{2+} -coordinating residues and disulfide bond pairs. We also observe higher average B-factors for bound Ca^{2+} ions in LDLR-A1 compared with LDLR-A2 ($\sim 80 \text{ \AA}^2$ and 44 \AA^2 , respectively). Overall these results indicate increased motion/conformational disorder in LDLR-A1 compared with LDLR-A2.

Dependence of TC-CD320 binding affinity on pH. Ligand interactions of the LDLR family members are thought to be modulated by histidine residues located at the binding interfaces and responding to changes in pH, a mechanism termed the

‘histidine switch’^{16,36–39}. An analysis of the TC-CD320 interface revealed three histidine residues from TC (His56, His104 and His154) and one from CD320 (His155) that may play a role in modulating the interaction (Fig. 3a). Among the observed interactions, the contact between His104 of TC and Asp77 of LDLR-A1 is notable because the backbone carbonyl of Asp77 coordinates the central Ca^{2+} ion. Furthermore, the backbone carbonyl of the interface residue His155 of CD320 also coordinates the central Ca^{2+} ion in LDLR-A2. This histidine also interacts with Asp157 of LDLR-A2, one of the acidic necklace forming residues. The above suggests that an intricate structural link exists between the pH-dependent interface and the Ca^{2+} -coordination in the LDLR-A2 domain. Finally, His104 of TC is in a special location where it could interact with residues from both LDLR-A1 and LDLR-A2.

We analysed the effect of changes in pH on the TC-CD320 interaction by incubating an immobilized complex of TC and biotinylated receptor (CD320_{AVI}) at pH 4.8 and pH 7.5 followed by SDS–polyacrylamide gel electrophoresis analysis. We observed a release of TC from the complex at low, but not at physiological pH (Fig. 3b). We quantitated this effect using our solid-phase binding assay by comparing binding of isolated tag-less TC to immobilized CD320_{AVI} at pH 4.8 and pH 7.5. As shown in Fig. 3c, the affinity of TC for CD320 is significantly reduced at lower pH ($K_D \sim 1.0 \text{ nM}$ at pH 7.5 and $\sim 58 \text{ nM}$ at pH 4.8).

Structure of the TC-CD320 Δ E88 complex. Deletion of a single glutamate (Δ E88) in LDLR-A1 of CD320 has been associated with Cbl deficiency observed in patients carrying the mutation⁴⁰. To investigate whether a structural effect was underpinning this defect, we generated a complex of TC with CD320 Δ E88, characterized it biochemically and determined its structure at 2.6 \AA resolution (Supplementary Fig. 3). Protein expression and purification were near identical to the wild-type complex. However, the purified complex displayed increased propensity to form non-specific multimers as analysed by SEC (Supplementary Fig. 4), requiring an additional SEC purification step for crystallization. As the mutant complex structure was of lower resolution than that of the wild type, we confirmed the positions of disulfide bonds and bound Ca^{2+} ions by using anomalous difference peaks from data collected at a wavelength of 1.77 \AA .

The TC-CD320 Δ E88 structure revealed a similar domain arrangement and binding interface as the wild-type protein

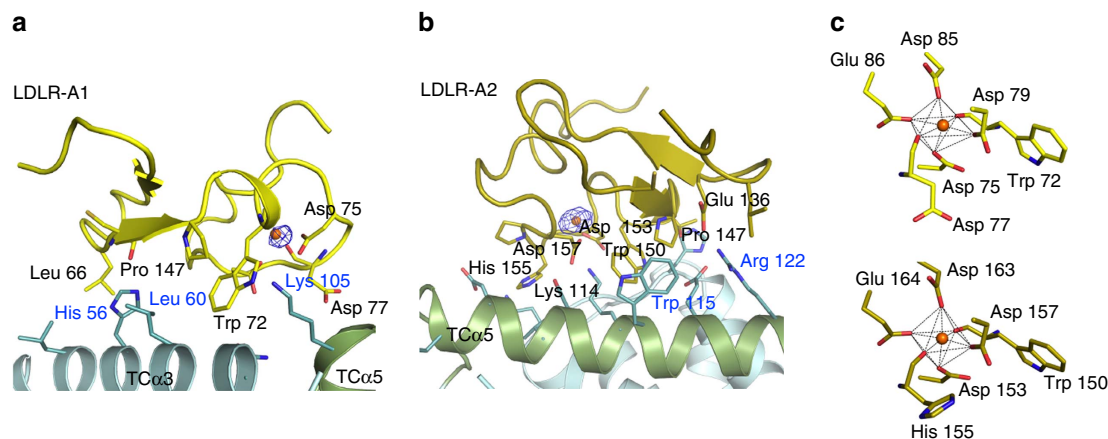


Figure 2 | Details of the TC-CD320 binding interface. (a) Interactions of LDLR-A1 and (b) LDLR-A2 with TC. TC helix $\alpha 5$ is coloured green. Ca^{2+} ions are indicated as orange spheres with corresponding anomalous difference density peaks shown (blue mesh contoured at 4σ). TC and CD320 residues involved in intermolecular contacts and Ca^{2+} coordination are shown as sticks and labelled blue and black, respectively. (c) Close-up views showing octahedral Ca^{2+} coordination in LDLR-A1 (yellow, top) and LDLR-A2 (gold, bottom).

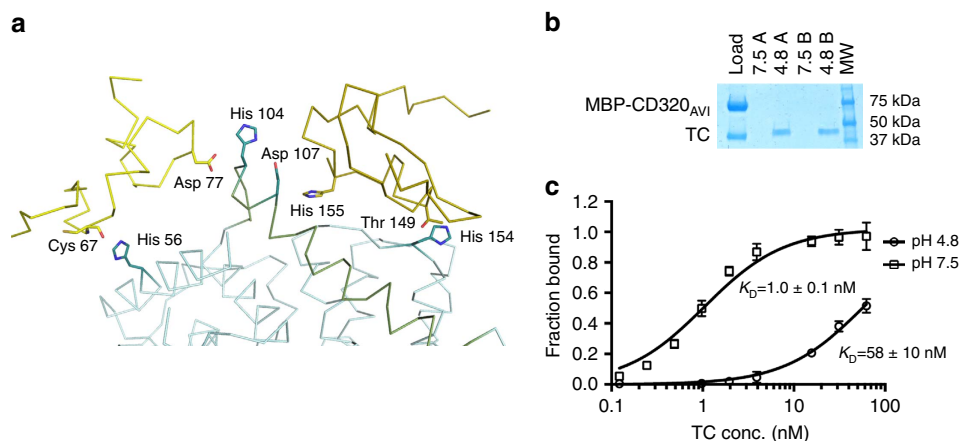


Figure 3 | Interface histidines and pH dependence of TC-CD320 affinity. (a) Histidine residues identified at the TC-CD320 interface: CD320 (yellow and gold) and TC (blue) are shown as backbone traces, with histidine residues at the binding interface shown as respectively coloured sticks. (b) SDS-polyacrylamide gel electrophoresis (SDS-PAGE) showing TC dissociation from an immobilized CD320_{AVI}-TC complex at pH 4.8 but not at pH 7.5. Two successive washes (A and B) are shown. (c) TC-CD320 binding curves at pH 4.8 and 7.5 obtained from solid-phase binding experiments ($n=5$, error bars indicate s.d.).

and we observed no significant changes in TC or the LDLR-A2 domain, ruling out a gross structural effect of the Δ E88 mutation. As seen for the wild-type protein, the region of LDLR-A1 carrying the deletion (EEE loop) revealed poor electron density and higher B-factors when compared with that of LDLR-A2 (average B-factors for LDLR-A1 and A2 Ca²⁺ ions ~ 112 and 56 \AA^2 , respectively). Moreover, electron density for the Ca²⁺ binding residue Glu86 as well as Glu87 and disulfide forming Cys88 (equivalent to Cys89 in wild type) was missing. However, the centrally bound Ca²⁺ as well as the LDLR-A1 acidic necklace was clearly visible in our structure indicating intact Ca²⁺ binding. Even though the apparent T_m of the complex was slightly lower than that of the wild-type protein, approaching that of the LDLR-A2 domain alone (Supplementary Fig. 3b), the mutant protein bound TC with equivalent affinity to wild type with a calculated K_D of 1.5 nM (Supplementary Fig 3c).

Role of the EGF domain of CD320 in TC binding. The exact function of the EGF-like domain in CD320 has not yet been established and its deletion in CD320 was shown to have no effect on TC binding⁵. The EGF-like domain of CD320 is not visible in our electron density maps despite chemically being present, suggesting positional disorder in the crystal. To assess what effect the domain has on CD320 complex formation with TC, we generated a mutant of CD320 in which the EGF-like domain was replaced by (Gly₄Ser)₄ sequence, which constituted a linker long enough to offer unrestricted movement and flexibility to the LDLR-A modules that the disulfide-rich EGF-like domain would not allow. The mutant protein was co-purified with TC and appeared biochemically stable and capable of TC binding. However, we found that the resulting construct yielded additional species of higher molecular weight in SEC in addition to the 1:1 stoichiometry complex peak seen for wild type (Fig. 4a). As shown in Fig. 4b, these results were confirmed by native electrospray ionization mass spectrometry analysis of the TC-CD320(Gly₄Ser)₄ complex as well as its individually purified SEC peaks. We observed single species of ~ 120 kDa mass for the larger SEC peak and 65 kDa mass for the smaller SEC peak, which we interpret as 2:1 or 2:2 TC:CD320 species and a single 1:1 complex, respectively.

Discussion

Our structure offers a basis for the selectivity of CD320 for TC over the homologous and structurally similar HC. This is especially relevant in light of HC's suggested role as a scavenger protein for corrin-like substrates in the blood stream that could otherwise bind TC and enter the cell and interfere with downstream Cbl metabolism. Unlike TC, both IF and HC are glycoproteins and it was previously suggested that N-linked glycosylation of surface asparagines of HC might prevent its binding to CD320 (ref. 21). A comparison of receptor-bound TC with the recently determined HC structure (Fig. 5a) rules out such an explanation, as the glycans are not located at the appropriate surface. Rather, TC and HC have different electrostatic potentials at the relevant surfaces (Fig. 5b), with Lys105 and Lys114 in TC replaced by Asp92 and Asn102 in HC, respectively. The resultant electrostatic mismatch is especially pronounced for regions of HC that would contact LDLR-A2, which is the major determinant of stable complex formation. Thus, HC cannot bind CD320 due to electrostatic repulsion.

The finding that the TC-CD320 interaction only involves the alpha domain of TC is in contrast to the recognition interface of IF with Cubilin, where both the α and β domains of holo-IF are involved^{19,41}. Our findings suggest that in the absence of gross structural changes in the α domain on loss of bound Cbl, both holo and apo-TC should be able to interact with CD320. Earlier studies have indeed pointed to receptor recognition of apo-TC, although a higher affinity was reported for holo-TC (refs 42–44). The TC-CD320 binding interface is only slightly larger than that of the LDLR-RAP complex (buried surface area $\sim 760 \text{ \AA}^2$) and whereas the two LDLR-A modules in RAP contribute equally to the buried surface area and use virtually identical modes of ligand recognition, LDLR-A2 in CD320 serves as the primary determinant of high affinity binding. LDLR-A1, which displays low *in vitro* affinity to TC on its own, probably has a role in enhancing the specificity and affinity of CD320 for TC. The mutant CD320 Δ E88, although more prone to forming non-specific complexes with TC (Supplementary Fig. 4), displays high thermal stability as well as tight TC binding. At least *in vitro*, these findings rule out a compromised ability of CD320 Δ E88 to bind TC. Additional studies are needed to explain the *in vivo* effect of the mutation on receptor regulation and trafficking. Our data also support a structural role of the EGF-like domain of CD320 in influencing the flexibility and the orientation of

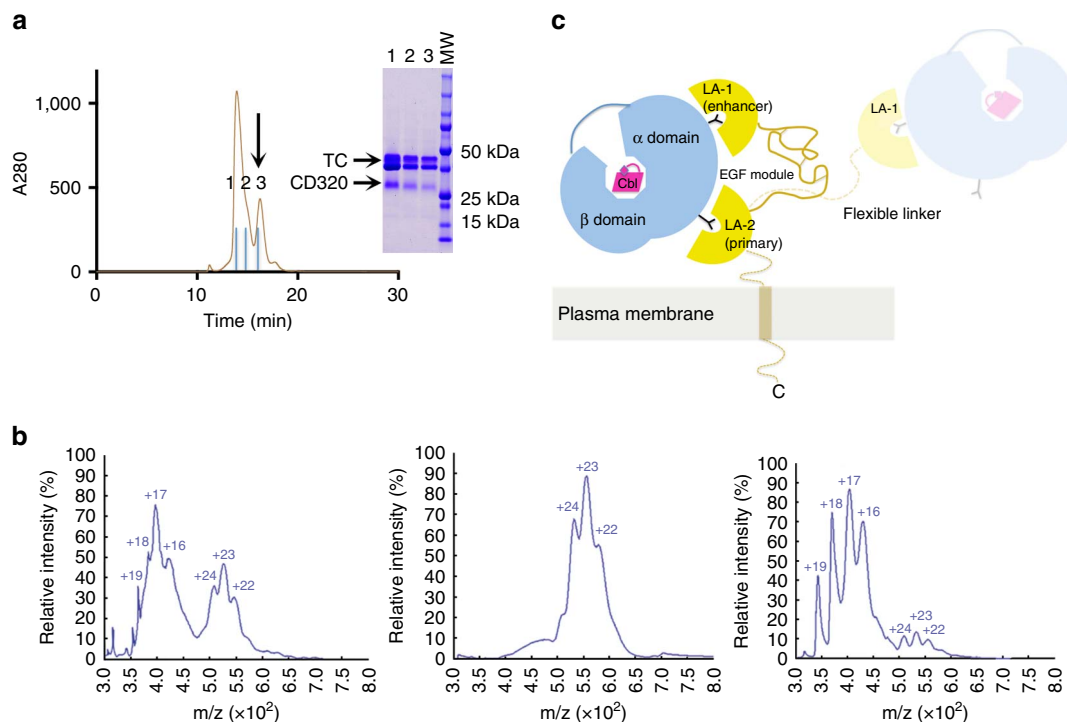


Figure 4 | Role of EGF-like domain in TC-CD320 complex formation. (a) SEC profile of TC-CD320(Gly₄Ser)₄ construct, revealing higher mass peaks indicating multimeric species. Arrow indicates position of 1:1 TC-CD320 peak. Peak fractions indicated above SEC profile correspond to lanes in SDS-polyacrylamide gel electrophoresis (SDS-PAGE). Note that TC appears as a double band depending on polyacrylamide percentage and SDS-PAGE running conditions. (b) Native electrospray ionization mass spectrometry (ESI-MS) analysis of TC-CD320(Gly₄Ser)₄ complex. Each peak cluster comprises different charged species of a single macromolecular assembly formed due to the presence of residual salt from protein purification and correspond to approximate molecular weights (MWs) of 65 and 120 for the lower and higher molecular weight species, respectively (left panel). Analysis of individually purified SEC peaks shows isolated high MW (middle) and low MW (right) species. (c) Model of properly oriented LDLR-A1 and LDLR-A2 domains (strong yellow) acting in tandem to specifically bind a single TC molecule. Dashed line indicates a flexible linker with proposed second TC molecule (faded yellow) binding.

the two LDLR-A domains to ensure that both LDLR-A domains bind to a single TC (Fig. 4c), thereby preventing formation of non-specific multimers that could potentially be detrimental to proper ligand internalization.

The combination of an acidic environment and decreased Ca²⁺ concentration has been implicated in the release of ligands of LDLRs during endocytosis (as recently reviewed⁴⁵). However, no direct demonstration of pH-mediated dissociation of the TC-CD320 complex was reported. Our *in vitro* finding that the affinity of TC for CD320 is reduced more than 50-fold at low pH is in line with the expected receptor-ligand segregation and TC delivery to the lysosomal pathway. Our *in vitro* system sets the stage for addressing more detailed questions in future studies. These include whether acidification promotes Ca²⁺ release from the LDLR-A domains, how CD320 changes at low pHs, and whether such changes alter the ability of CD320 to bind potential regulators.

Finally, the ability to produce a stable and intact TC-CD320 complex, along with knowledge of the binding interface at atomic detail, might facilitate the design and manufacture of small molecule as well as immunotherapeutic regulators that could selectively disrupt the interaction in certain tumours overexpressing CD320. As for the latter, we have put in significant effort towards and been successful in the generation of camelid nanobodies against the TC-CD320 complex targeting various epitopes. These have the potential not only to aid in capturing novel conformational states of CD320 for structure determination, but also add to list of potential tools available for diagnostic or therapeutic applications in cancer treatments.

Methods

Protein expression and purification. The gene for human TC harbouring a R209Q mutation (GenBank ID BC001176.1) was obtained from ImaGenes (Life sciences). All CD320 constructs were N-terminally fused to a nona-histidine tagged MBP construct with a tobacco etch virus (TEV) protease cleavage sequence inserted between the MBP and receptor genes. The native secretion signals for both CD320 and TC constructs were replaced with a Melittin secretion signal for enhanced secreted expression in *Spodoptera frugiperda* (Sf21 cells) using the Multibac expression system⁴⁶. CD320 and TC constructs were cloned into the dual expression pFL vector and bacmids, and viruses were generated according to recommended procedures. For CD320, residues 31–199 and 53–199 were used. LDLR-A1 and LDLR-A2 constructs comprised residues 53–95 and 125–199, respectively. The TC α construct comprised residues 1–290 of TC. Sf21 cells were grown and maintained in SF900-II medium and protein was expressed for 96 h post infection. Bulk medium was then harvested, supplemented with 1 μ M Cbl-CN, and its pH was adjusted to 6.5 before being applied to Ni-NTA resin equilibrated in 20 mM HEPES pH 7.0, 0.5 mM CaCl₂ and 150 mM NaCl. The resin was washed with 20 mM HEPES pH 7.0, 300 mM NaCl, and 40 mM imidazole and the protein eluted with 20 mM HEPES pH 7.0, 150 mM NaCl, 0.5 mM CaCl₂ and 200 mM imidazole buffer. Eluted protein was desalted in 20 mM Tris pH 7.5, 0.5 mM CaCl₂ and 150 mM NaCl (Buffer A) before overnight cleavage of the 9xHis-MBP module with His-tagged TEV protease (prepared in-house) at a w/w ratio of 1:10 (TEV:CD320). The resultant TC-CD320 complex displayed weak affinity for Ni²⁺-NTA and could be purified from excess receptor through a second Ni²⁺ affinity purification step with excess receptor separated in the flow through and the intact receptor eluting with 20 mM Imidazole. Final samples were desalted into Buffer A, concentrated to 10–20 mg ml⁻¹ and flash frozen in liquid N₂. The flexible linker mutant of CD320 was generated by replacing residues 96–135 of CD320 with the eicosapeptide (Gly₄Ser)₄. Protein expression was performed as for wild-type TC-CD320. Where required, CD320 and its variants were individually expressed without TC in the same cloning/expression system.

Biotinylated receptor was prepared by adding a C-terminal Avi Tag (GLNDIFEAQKIEWHE) to CD320 and its variants. Expression and purification was identical to the wild-type complex except that no MBP cleavage was performed. Overnight biotinylation with the biotin ligase BirA (prepared in-house) was performed by adding 5–10 μ g BirA per mg purified receptor in Buffer A

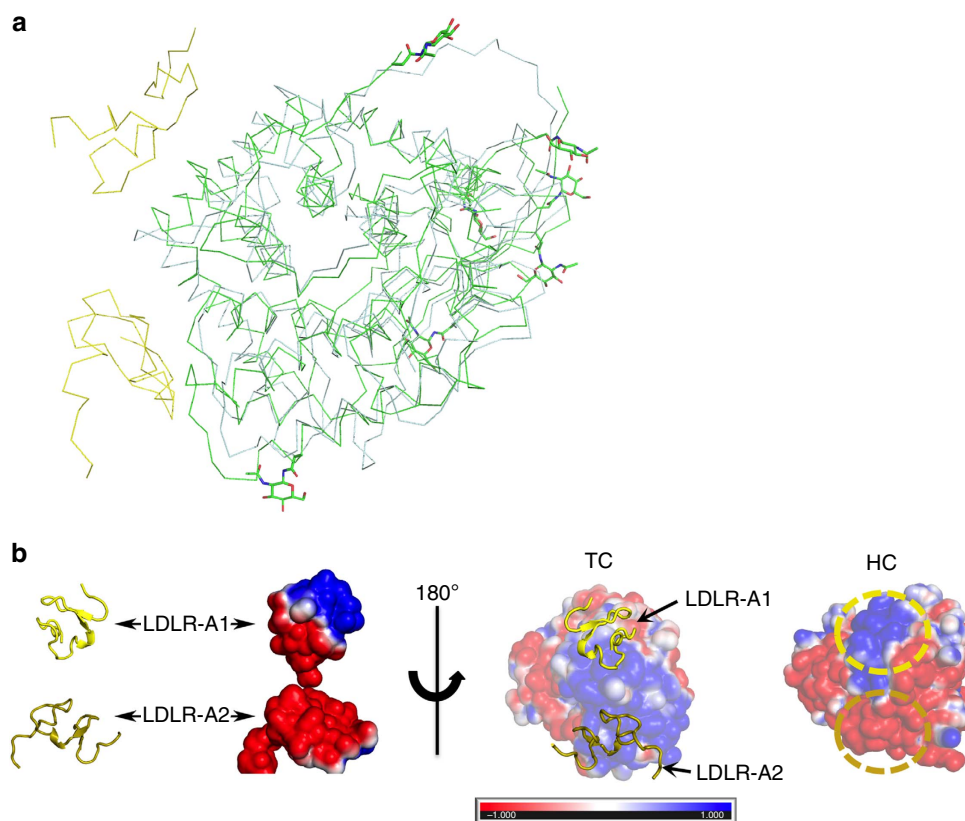


Figure 5 | Comparison of CD320-bound TC with free HC (pdb ID 4KKI). (a) Superposition of HC with CD320-bound TC revealing position of glycosylated residues (green sticks) of HC. (b) Electrostatic potential maps of CD320 (2nd from left), TC (2nd from right) and HC (right) with the latter two rotated 180° along the y axis in relation to CD320 to reveal the receptor interaction surfaces. For clarity, CD320 LDLR-A1 (yellow) and LDLR-2 (gold) are also shown in cartoon format. The regions of HC corresponding to equivalent TC binding regions for LDLR-A1 and LDLR-A2 are highlighted (dotted yellow and gold ovals, respectively).

containing 10 mM magnesium acetate, 10 mM ATP and 50 μ M biotin. Biotinylation was confirmed by gel shift analysis on mixing desalted, biotinylated samples with streptavidin before SDS-polyacrylamide gel electrophoresis.

TC binding assay. Tag-less TC was obtained by dissociating a complex of TC and biotinylated CD320 immobilized on Streptactin resin (IBA) with four column volumes of Buffer B (Buffer A lacking Ca^{2+} and containing 10 mM EGTA) at 37 °C for 30 min. The flow through containing TC was collected and the process repeated an additional three times. The pooled flow through samples were concentrated in an Amicon filter (30 kDa MWCO) and desalted in Buffer A. Protein concentrations were determined using a Bio-Rad DC Protein Assay Kit and A_{280} measurements. 1–5 pMol receptor was added to each well of pre-blocked 96-well Neutravidin plates (ThermoFischer Scientific) for 1 h at room temperature (RT). All incubations were carried out on a plate shaker at 350 r.p.m. Unbound receptor was discarded and the wells were washed three times with 200 μ l Buffer A containing 0.05% Tween-20 (Buffer AT). TC (100 μ l) serially diluted in Buffer AT was added to the wells for 1 h at RT. Unbound TC was discarded and the wells washed five times with 200 μ l Buffer AT. To analyse binding at different pHs, Buffer A was modified to contain 100 mM Tris-HCl pH 7.5 or 100 mM sodium acetate pH 4.8 instead of 20 mM Tris-HCl pH 7.5 and used for serially diluting TC. An additional wash step in TC dilution buffer was added to equilibrate the wells before TC binding as well as after discarding unbound TC. TC antibody A-5 (Santa Cruz Biotechnology, Inc, Catalogue # sc-137017) was diluted to 1 μ g ml⁻¹ (1:200) in Buffer AT and 100 μ l was added to the wells for 30 min at RT. Unbound antibody was discarded and the wells washed three times with Buffer AT. Horse radish peroxidase (100 μ l) conjugated Goat mouse Anti IgG (ThermoFischer Scientific, Catalogue # A16072) diluted to 1 μ g ml⁻¹ (1:1,500) in Buffer AT was added to the wells for 30 min. Unbound antibody was discarded and the wells washed three times with Buffer AT followed by development with the TMB Substrate Kit (ThermoFischer Scientific). Absorbance (450 nm) measurements were made with a BioTek Synergy HT plate reader. Absorbance readings were plotted against TC concentrations and curves fitted in GraphPad PRISM using a single-site specific binding equation, and normalized to the Calculated Bmax for direct comparison. All experiments were done in replicates of 5.

Thermostability assays. Frozen protein samples were thawed and diluted to 0.5–1 mg ml⁻¹ in Buffer A. A thermocycler (Bio-rad) was used to heat 80–100 μ l samples at the selected temperatures for 10 min followed by rapid cooling to 4 °C before spinning down at 120,000g for 20 min. Samples were loaded onto a G3000Sxl (TOSOH Biosciences) column pre-equilibrated in Buffer A at 4 °C. Thermostability curves were plotted in GraphPad PRISM based on complex peak heights and fitted using a sigmoidal dose response (variable slope) equation. Protein used for assays in the presence of EGTA was diluted 20 \times to 0.5–1 mg ml⁻¹ in Buffer B, which was also used for SEC.

Chemical Cross-linking. The protein complexes at 10 mg ml⁻¹ were mixed with a 10% glutaraldehyde solution in a 10/1 (v/v) ratio for one hour at RT in buffer solutions. The mixture was further diluted with the original protein buffer solution or water before mass spectrometric analysis.

Mass spectrometry. A commercial MALDI-TOF/TOF mass spectrometer (model 4800 plus, AB Sciex, Darmstadt, Germany) equipped with a high-mass detector (HM2, CovalX AG, Zurich, Switzerland) was used. All measurements were performed in the linear positive ion mode with standard settings. Ionization was achieved with an Nd:YAG laser (355 nm) with the energy just above the threshold for ion formation. Each mass spectrum was the average of 1,000 laser shots acquired at random sample positions without searching for hot spot. Sinapinic acid (20 mg ml⁻¹ in water/acetonitrile/TFA, 49.95/49.95/0.1, v/v/v) was used as the matrix. Matrix (0.5 μ l) was first spotted on a stainless steel plate and allowed to dry under ambient conditions. Afterwards, 0.5 μ l of the sample was spotted on the top of the matrix and also dried under ambient conditions. In the final step, 0.5 μ l of the matrix was spotted. All mass spectra were baseline corrected and smoothed using a Savitzky-Golay algorithm by Igor Pro 6.2, WaveMetrics, Oregon, USA.

Native ESI-mass spectrometry was carried out on a commercial hybrid quadrupole/time-of-flight mass spectrometer Synapt G2-S equipped with a Z-spray interface followed by the StepWave ion guides (Waters, Manchester, UK). The instrument was operated in the Resolution mode (resolution > 20,000) in positive ion mode. For calibration, CsI clusters (10 mg ml⁻¹) dissolved in 50% (v) isopropanol was used. The recorded spectra were averaged (500 scans), smoothed with a moving

algorithm (width of three steps) and centroid spectra were generated at 80% peak height. After polynomial fitting, the mass axis was calibrated. Typically, 3–5 μl of sample were loaded into 1 μm Au/Pd coated glass needles (Thermo Scientific, Madison, WI, USA). The source temperature was kept at 30 °C. The backing pressure was set to 0.5 bar to assist sample flow. The capillary voltage was set to 0.95 V to generate the electrospray. The sampling cone voltage and the source offset were optimized to 150 and 50 V respectively, to ensure efficient ion transmission. Further, to assure better declustering of salt adducts, the trap collision energy and the transfer collision energy were kept at 30 and 15 V, respectively. The quadrupole transmission range was adjusted for the desired m/z range. The mass spectra were recorded in a m/z window of 50–8,000 with a scan time of 2 s and an interscan delay of 0.1 s. Each native electrospray ionization mass spectrometry spectrum represents the combination of 150 individual scans. Mass spectra were acquired using Mass Lynx software (version 4.1, Waters, Manchester, UK). All mass spectra were baseline corrected, normalized and smoothed using MATLAB_R2014a.

Crystallization and structure determination. For crystallization of the wild-type TC-CD320 complex, frozen samples were thawed and diluted to 10 mg ml⁻¹ and mixed with well buffer (9.6% PEG8000, 160 mM Ca(OAc)₂, 80 mM MES pH 6.2, and 20 mM MgCl₂) in a 1:1 ratio at 20 °C in sitting drop plates. Crystals were cryo-protected by exchanging the drop solution with 9.6% PEG8000, 160 mM Ca(OAc)₂, 80 mM MES pH 6.2, and 2.0 mM MgCl₂, 20 mM Tris pH 7.5, 150 mM NaCl, 0.5 mM CaCl₂, 1 μM Cbl-CN and 25% Glycerol and flash frozen in liquid Nitrogen. An N-terminal 22 amino acid deletion of the receptor led to higher quality crystals that were used for the final structure determination.

Crystals of the ΔE88 mutant were obtained by purifying the complex by SEC in Buffer A before concentrating to 10 mg ml⁻¹ and mixing with well buffer (180 mM L-Proline, 11.7% PEG3350, 90 mM Hepes pH 7.5, and 30 mM Glycyl-Glycyl-Glycine) in a 2:1 ratio at 20 °C in sitting drop plates. Crystals were cryo-protected by exchanging the drop solution with Cryo-buffer (108 mM L-Proline, 54 mM Hepes pH 7.5, 7.2% PEG3350, 0.6 mM CaCl₂, 24 mM Tris pH 7.5, 180 mM NaCl, 1 μM Cbl-CN and 25% Glycerol) and flash frozen in liquid Nitrogen. Data were collected at the X06SA microfocuss beamline (Swiss Light Source) at a wavelength of 1 Å. A complete data set from a single wild-type crystal was collected in six discreet wedges and processed using denzo and scalepack (HKL Research). The structure was determined by molecular replacement using PHASER in PHENIX (ref. 47) using a TC monomer from the Cbl-H₂O bound human TC structure (PDB 2BB5) as a search model. Model building and refinement were carried out using COOT (ref. 48) and PHENIX. Anomalous data sets from an identical crystal collected at 1.6 Å wavelength was processed by denzo and scalepack. Two molecules of the TC-CD320 complex/asymmetric unit were found, molecule 1 comprising chains A and C and molecule 2 comprising chains B and D.

A complete data set for the ΔE88 mutant was processed by XDS (ref. 49) and aimless in CCP4 (ref. 50). The wild-type TC-CD320 structure was used as a search model for molecular replacement using PHASER in PHENIX. Model building and refinement were carried out using COOT and PHENIX. An anomalous data set collected at 1.77 Å wavelength was processed by XDS and aimless in CCP4. TC-CD320 complex molecule 2 was used to generate all figures using the PyMOL Molecular Graphics System (The PyMOL Molecular Graphics System, Version 1.8 Schrödinger, LLC). Part of the 2fo-fc electron density maps for both the wild-type and CD320 ΔE88 complexes with TC are shown in Supplementary Fig. 6.

Data availability. Atomic structure factors and coordinates for the TC-CD320 and TC-CD320 ΔE88 structures have been deposited under accession codes 4ZRP and 4ZRQ, respectively. The authors declare that all relevant data sets used in this study are available on request.

References

- Cannata, J. J., Focesi, Jr A., Mazumder, R., Warner, R. C. & Ochoa, S. Metabolism of propionic acid in animal tissues. Xii. Properties of mammalian methylmalonyl coenzyme a mutase. *J. Biol. Chem.* **240**, 3249–3257 (1965).
- Taylor, R. T. & Weissbach, H. N5-methyltetrahydrofolate-homocysteine transmethylase. Propylation characteristics with the use of a chemical reducing system and purified enzyme. *J. Biol. Chem.* **242**, 1509–1516 (1967).
- Johnston, J., Bollekens, J., Allen, R. H. & Berliner, N. Structure of the cDNA encoding transcobalamin I, a neutrophil granule protein. *J. Biol. Chem.* **264**, 15754–15757 (1989).
- Allen, R. H. Human vitamin B12 transport proteins. *Prog. Hematol.* **9**, 57–84 (1975).
- Birn, H. *et al.* Characterization of an epithelial approximately 460-kDa protein that facilitates endocytosis of intrinsic factor-vitamin B12 and binds receptor-associated protein. *J. Biol. Chem.* **272**, 26497–26504 (1997).
- Kozyraki, R. *et al.* The human intrinsic factor-vitamin B12 receptor, cubilin: molecular characterization and chromosomal mapping of the gene to 10p within the autosomal recessive megaloblastic anemia (MGA1) region. *Blood* **91**, 3593–3600 (1998).
- Quadros, E. V., Sai, P. & Rothenberg, S. P. Characterization of the human placental membrane receptor for transcobalamin II-cobalamin. *Arch. Biochem. Biophys.* **308**, 192–199 (1994).
- Jiang, W., Nakayama, Y., Sequeira, J. M. & Quadros, E. V. Mapping the functional domains of TCblR/CD320, the receptor for cellular uptake of transcobalamin-bound cobalamin. *FASEB J.* **27**, 2988–2994 (2013).
- Chanarin, I., Muir, M., Hughes, A. & Hoffbrand, A. V. Evidence for intestinal origin of transcobalamin II during vitamin B12 absorption. *Br. Med. J.* **1**, 1453–1455 (1978).
- Nielsen, M. J., Rasmussen, M. R., Andersen, C. B., Nexø, E. & Moestrup, S. K. Vitamin B12 transport from food to the body's cells—a sophisticated, multistep pathway. *Nat. Rev. Gastroenterol. Hepatol.* **9**, 345–354 (2012).
- Reynolds, E. Vitamin B12, folic acid, and the nervous system. *Lancet Neurol.* **5**, 949–960 (2006).
- Kozyraki, R. & Cases, O. Vitamin B12 absorption: mammalian physiology and acquired and inherited disorders. *Biochimie* **95**, 1002–1007 (2013).
- Quadros, E. V., Nakayama, Y. & Sequeira, J. M. The protein and the gene encoding the receptor for the cellular uptake of transcobalamin-bound cobalamin. *Blood* **113**, 186–192 (2009).
- Quadros, E. V. Advances in the understanding of cobalamin assimilation and metabolism. *Br. J. Haematol.* **148**, 195–204 (2010).
- Brown, M. S. & Goldstein, J. L. A receptor-mediated pathway for cholesterol homeostasis. *Science* **232**, 34–47 (1986).
- Blacklow, S. C. Versatility in ligand recognition by LDL receptor family proteins: advances and frontiers. *Curr. Opin. Struct. Biol.* **17**, 419–426 (2007).
- Hussain, M. M., Strickland, D. K. & Bakillah, A. The mammalian low-density lipoprotein receptor family. *Annu. Rev. Nutr.* **19**, 141–172 (1999).
- Willnow, T. E. *et al.* RAP, a specialized chaperone, prevents ligand-induced ER retention and degradation of LDL receptor-related endocytic receptors. *EMBO J.* **15**, 2632–2639 (1996).
- Andersen, C. B., Madsen, M., Storm, T., Moestrup, S. K. & Andersen, G. R. Structural basis for receptor recognition of vitamin-B(12)-intrinsic factor complexes. *Nature* **464**, 445–448 (2010).
- Wuerges, J. *et al.* Structural basis for mammalian vitamin B12 transport by transcobalamin. *Proc. Natl Acad. Sci. USA* **103**, 4386–4391 (2006).
- Wuerges, J., Geremia, S. & Randaccio, L. Structural study on ligand specificity of human vitamin B12 transporters. *Biochem. J.* **403**, 431–440 (2007).
- Furger, E., Frei, D. C., Schibli, R., Fischer, E. & Prota, A. E. Structural basis for universal corrinoid recognition by the cobalamin transport protein haptocorrin. *J. Biol. Chem.* **288**, 25466–25476 (2013).
- Quadros, E. V., Nakayama, Y. & Sequeira, J. M. Targeted delivery of saporin toxin by monoclonal antibody to the transcobalamin receptor, TCblR/CD320. *Mol. Cancer Ther.* **9**, 3033–3040 (2010).
- Quadros, E. V., Nakayama, Y. & Sequeira, J. M. Saporin conjugated monoclonal antibody to the transcobalamin receptor TCblR/320 is effective in targeting and destroying cancer cells. *J. Cancer Ther.* **4**, 1074–1081 (2013).
- Quadros, E. V., Sai, P. & Rothenberg, S. P. Functional human transcobalamin II isoproteins are secreted by insect cells using the baculovirus expression system. *Blood* **81**, 1239–1245 (1993).
- Seligman, P. A. & Allen, R. H. Characterization of the receptor for transcobalamin II isolated from human placenta. *J. Biol. Chem.* **253**, 1766–1772 (1978).
- Yasui, N., Nogi, T. & Takagi, J. Structural basis for specific recognition of reelin by its receptors. *Structure* **18**, 320–331 (2010).
- Fedosov, S. N. *et al.* Mapping the functional domains of human transcobalamin using monoclonal antibodies. *FEBS J.* **272**, 3887–3898 (2005).
- Fedosov, S. N., Fedosova, N. U., Nexø, E. & Petersen, T. E. 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 (2000).
- Fass, D., Blacklow, S., Kim, P. S. & Berger, J. M. Molecular basis of familial hypercholesterolaemia from structure of LDL receptor module. *Nature* **388**, 691–693 (1997).
- Rudenko, G. *et al.* Structure of the LDL receptor extracellular domain at endosomal pH. *Science* **298**, 2353–2358 (2002).
- Fisher, C., Beglova, N. & Blacklow, S. C. Structure of an LDLR-RAP complex reveals a general mode for ligand recognition by lipoprotein receptors. *Mol. Cell* **22**, 277–283 (2006).
- Beglov, D. *et al.* Structural insights into recognition of beta2-glycoprotein I by the lipoprotein receptors. *Proteins* **77**, 940–949 (2009).
- Verdaguer, N., Fita, I., Reithmayer, M., Moser, R. & Blaas, D. X-ray structure of a minor group human rhinovirus bound to a fragment of its cellular receptor protein. *Nat. Struct. Mol. Biol.* **11**, 429–434 (2004).
- Kurniawan, N. D. *et al.* NMR structure of a concatamer of the first and second ligand-binding modules of the human low-density lipoprotein receptor. *Protein Sci.* **9**, 1282–1293 (2000).

36. Beglova, N., Jeon, H., Fisher, C. & Blacklow, S. C. Cooperation between fixed and low pH-inducible interfaces controls lipoprotein release by the LDL receptor. *Mol. Cell* **16**, 281–292 (2004).
37. Beglova, N. & Blacklow, S. C. The LDL receptor: how acid pulls the trigger. *Trends Biochem. Sci.* **30**, 309–317 (2005).
38. Chung, N. S. & Wasan, K. M. Potential role of the low-density lipoprotein receptor family as mediators of cellular drug uptake. *Adv. Drug Deliv. Rev.* **56**, 1315–1334 (2004).
39. Yamamoto, T., Chen, H. C., Guigard, E., Kay, C. M. & Ryan, R. O. Molecular studies of pH-dependent ligand interactions with the low-density lipoprotein receptor. *Biochemistry* **47**, 11647–11652 (2008).
40. Quadros, E. V. *et al.* Positive newborn screen for methylmalonic aciduria identifies the first mutation in TCblR/CD320, the gene for cellular uptake of transcobalamin-bound vitamin B(12). *Hum. Mutat.* **31**, 924–929 (2010).
41. Fedosov, S. N. *et al.* Composite organization of the cobalamin binding and cubilin recognition sites of intrinsic factor. *Biochemistry* **44**, 3604–3614 (2005).
42. Rothenberg, S. P. & Quadros, E.V. Transcobalamin II and the membrane receptor for the transcobalamin II-cobalamin complex. *Baillieres Clin. Haematol.* **8**, 499–514 (1995).
43. Quadros, E. V., Nakayama, Y. & Sequeira, J. M. The binding properties of the human receptor for the cellular uptake of vitamin B12. *Biochem. Biophys. Res. Commun.* **327**, 1006–1010 (2005).
44. Orning, L. *et al.* Characterization of a monoclonal antibody with specificity for holo-transcobalamin. *Nutr. Metab. (Lond)* **3**, 3 (2006).
45. Andersen, C. B. & Moestrup, S. K. How calcium makes endocytic receptors attractive. *Trends Biochem. Sci.* **39**, 82–90 (2014).
46. Bieniossek, C., Richmond, T. J. & Berger, I. MultiBac: multigene baculovirus-based eukaryotic protein complex production. *Curr. Protoc. Protein Sci.* Chapter 5, Unit 5 20 (2008).
47. Adams, P. D. *et al.* PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 213–221 (2010).
48. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 486–501 (2010).
49. Kabsch, W. Xds. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 125–132 (2010).
50. Winn, M. D. *et al.* Overview of the CCP4 suite and current developments. *Acta Crystallogr. D Biol. Crystallogr.* **67**, 235–242 (2011).

Acknowledgements

We gratefully acknowledge the assistance of staff scientists at the PX beamline of the Swiss Light Source, Villigen, Switzerland. This work was supported by the Swiss Cancer League (KFS 3004-08-2012 to K.P.L.), NCCR TransCure (grant to K.P.L.), the Swiss National Science Foundation (SNF 200020-124663 to R.Z. and SNF 31003A_146191 and SNF 310030B_166672 to K.P.L.) and the European Molecular Biology Organization (EMBO ALTF-770-2010 to A.A.).

Author contributions

A.A. carried out the structure determination of TC-CD320 complexes as well as all other experimental procedures excluding Mass Spectrometry analyses. J.-S.W. cloned and crystallized the initial TC-CD320 complexes. J.S. contributed to the cloning and purification of TC α -CD320 and TC-LDLR-A2 complexes. B.P. helped with protein expression and purification experiments. J.S.B. generated expression plasmids and viruses for some constructs. F.C., K.R. and R.Z. performed mass spectrometry analyses. K.P.L., A.A., and J.-S.W. conceived the project. K.P.L. and A.A. analysed the data and wrote the manuscript.

Additional information

Supplementary Information accompanies this paper at <http://www.nature.com/naturecommunications>

Competing financial interests: The authors declare no competing financial interests.

Reprints and permission information is available online at <http://npg.nature.com/reprintsandpermissions/>

How to cite this article: Alam, A. *et al.* Structural Basis of Transcobalamin Recognition by human CD320 Receptor. *Nat. Commun.* 7:12100 doi: 10.1038/ncomms12100 (2016).



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>

© The Author(s) 2016