Kent Academic Repository

Clarke, Charlea, Banasik, Michal, Juodeikis, Rokas, Warren, Martin J. and Pickersgill, Richard W. (2026) ***Evolutionarily divergent DUF4465 domains have a common vitamin B12-binding function.*** **FEBS Open Bio, 16 (8). pp. 1527-1549. ISSN 2211-5463.**

Downloaded from

https://kar.kent.ac.uk/113471/ The University of Kent's Academic Repository KAR

The version of record is available from

https://doi.org/10.1002/2211-5463.70231

This document version

Publisher pdf

DOI for this version

Licence for this version

CC BY (Attribution)

Additional information

Versions of research works

Versions of Record

If this version is the version of record, it is the same as the published version available on the publisher's web site. Cite as the published version.

Author Accepted Manuscripts

If this document is identified as the Author Accepted Manuscript it is the version after peer review but before type setting, copy editing or publisher branding. Cite as Surname, Initial. (Year) 'Title of article'. To be published in ***Title of Journal*** , Volume and issue numbers [peer-reviewed accepted version]. Available at: DOI or URL (Accessed: date).

Enquiries

If you have questions about this document contact ResearchSupport@kent.ac.uk. Please include the URL of the record in KAR. If you believe that your, or a third party's rights have been compromised through this document please see our Take Down policy (available from https://www.kent.ac.uk/guides/kar-the-kent-academic-repository#policies).

FEBS openbio

FEBS PRESS science publishing by scientists

RESEARCH ARTICLE

Evolutionarily divergent DUF4465 domains have a common vitamin B_{12}-binding function

Charlea Clarke[1], Michal Banasik[1], Rokas Juodeikis[2], Martin J. Warren[2,3] and Richard W. Pickersgill[1]

1 School of Biological and Behavioural Sciences, Queen Mary University of London, UK
2 Quadram Institute Biosciences, Norwich, UK
3 School of Natural Sciences, University of Kent, Canterbury, UK

Keywords
B_{12}-binding proteins; B_{12}-scavenging proteins; DUF4465 (DUF Domain of Unknown Function); IPR027828 family proteins; microbiome; β-jellyroll proteins

Correspondence
C. Clarke and R. W. Pickersgill, School of Biological and Behavioural Sciences, Queen Mary University of London, London E1 4NS, UK
E-mail: c.clarke@qmul.ac.uk; r.w.pickersgill@qmul.ac.uk

(Received 18 December 2025, revised 2 March 2026, accepted 6 March 2026)

doi:10.1002/2211-5463.70231

Edited by So Nakagawa

The DUF4465 family (DUF, domain of unknown function) contains more than 1000 members distributed across eight bacterial clades with species from diverse microenvironments including various gut microbiomes, hydrothermal vents, and soil. In the gut commensal *Bacteroides thetaiotaomicron* (*B. theta*), DUF4465 containing proteins act as high-affinity B_{12}–binding proteins that scavenge this cofactor to ensure bacterial survival. Such B_{12} capture is essential for bacteria that have lost the ability to synthesize B_{12} *de novo*. This raises the question of whether B_{12}-binding is ubiquitous across this family of proteins. Here, we show that B_{12}-binding is a recurrent function of eight distantly related members of the DUF4465 family. It is reasonable to conclude that B_{12}-binding is a common function of most DUF4465 proteins. These results establish DUF4465 as a structurally conserved family of augmented β-jellyroll B_{12}-binding proteins with widespread roles in microbial competition for this essential cofactor.

Impact statement

DUF4465 defines a widespread, structurally conserved bacterial cobalamin-binding domain and provides a promising scaffold for protein-based B_{12} capture and purification.

Nutrient acquisition is a critical aspect of bacterial survival, particularly for essential micronutrients, such as vitamin B_{12} (cobalamin or B_{12}), which are often limited in the environment [1]. This scarcity is especially evident in the human gut, where only a minority of the bacterial species present can synthesize B_{12} completely [2,3], and where over 80% of sequenced human microbial gut species nevertheless encode B_{12} dependent genes or riboswitches [4,5]. In such ecosystems, microbial communities rely heavily on both sharing and competition for B_{12} [6,7]. The importance of this cofactor stems from its central role in metabolism, acting as a cofactor for key enzymes such as methionine synthase and methyl malonyl CoA mutase [8–10]. Cobalamin biosynthesis is among the most complex and energetically demanding metabolic processes known, requiring approximately 30-enzymatic step biosynthetic pathway [11,12]. Although B_{12}-independent enzymes and pathways can substitute for survival, most bacteria that encode them also retain

Abbreviations
B_{12}, vitamin B_{12}/cobalamin; *B. theta*, *Bacteroides thetaiotaomicron*; CCP4, collaborative computation project no. 4; DUF, domain of unknown function; EFI-EST, enzyme function initiative enzyme similarity tool; ESRF, european synchrotron radiation facility; GNT, genome neighborhoodneighbourhood tool; IPR, interpro; IPTG, isopropyl β-d-1-thiogalactopyranoside; MAFFT, multiple alignment using fast fourier transform; MP, maximum parsimony; MUSCLE, multiple sequence comparison by log-expectation; NP, neighborneighbour-joining; OD, optical density; PDB, protein data bank; PFAM, protein families database; RMSD, root mean square deviation; SSN, sequence similarity network; TCEP, tris(2-CarboxyEthyl) phosphine.

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

B_{12}-dependent isoenzymes and rely on transporters to acquire complete or incomplete corrinoids through salvage pathways [13,14]. This pattern suggests that scavenging for B_{12} is not only sufficient but often preferred, as reflected in microbial communities where B_{12} producers are outnumbered by B_{12}-dependent organisms such as the gut commensal *B. theta* [15].

Recent studies have expanded our understanding of B_{12} transport and sequestration, particularly in B_{12}-limited environments such as the human gut. While the core B_{12} uptake pathway is highly conserved, scarcity exerts a strong selective pressure to diversify this protein network [16–18]. In *B. theta*, B_{12} limitation induces expression of additional outer membrane lipoproteins including those encoded by *btuG, btuH*, and *btuJ1*/*J2*, which are colocated within riboswitch-regulated loci with essential transport machinery: *btuB, btuCD*, and *btuF*, *BtuG*. It has been shown that these lipoproteins bind B_{12} and enhance uptake efficiency, conferring a competitive edge under nutrient limitation [19–21]. Such adaptations highlight how otherwise conserved B_{12} transport systems evolve diversification in response to selective pressure. We recently demonstrated that the DUF4465 proteins BtuJ1 and BtuJ2 are high-affinity B_{12}-binding proteins that localize to bacterial extracellular vesicles and promote the survival of *B. theta* [22]. Independent work has also reported B_{12}-binding to BtuJ1 aids bacterial survival [23]. Here, we show that B_{12} binding is a unifying feature of DUF4465 proteins, demonstrated across eight family members with as little as 16% sequence identity. Additionally, we show that these adaptations are not exclusive to gut bacteria, expanding our understanding of bacterial competition for micronutrients such as B_{12}.

Materials and methods

Molecular biology

DNA corresponding to the soluble domains of the DUF4465 proteins, lacking the predicted lipoprotein motif, was inserted into a modified pET3a vector resulting in constructs encoding N-terminal hexahistidine tagged proteins produced via the T7 promotor. The plasmids were validated by sequencing. The eight starred proteins in Fig. 1A were produced in this way.

Protein expression

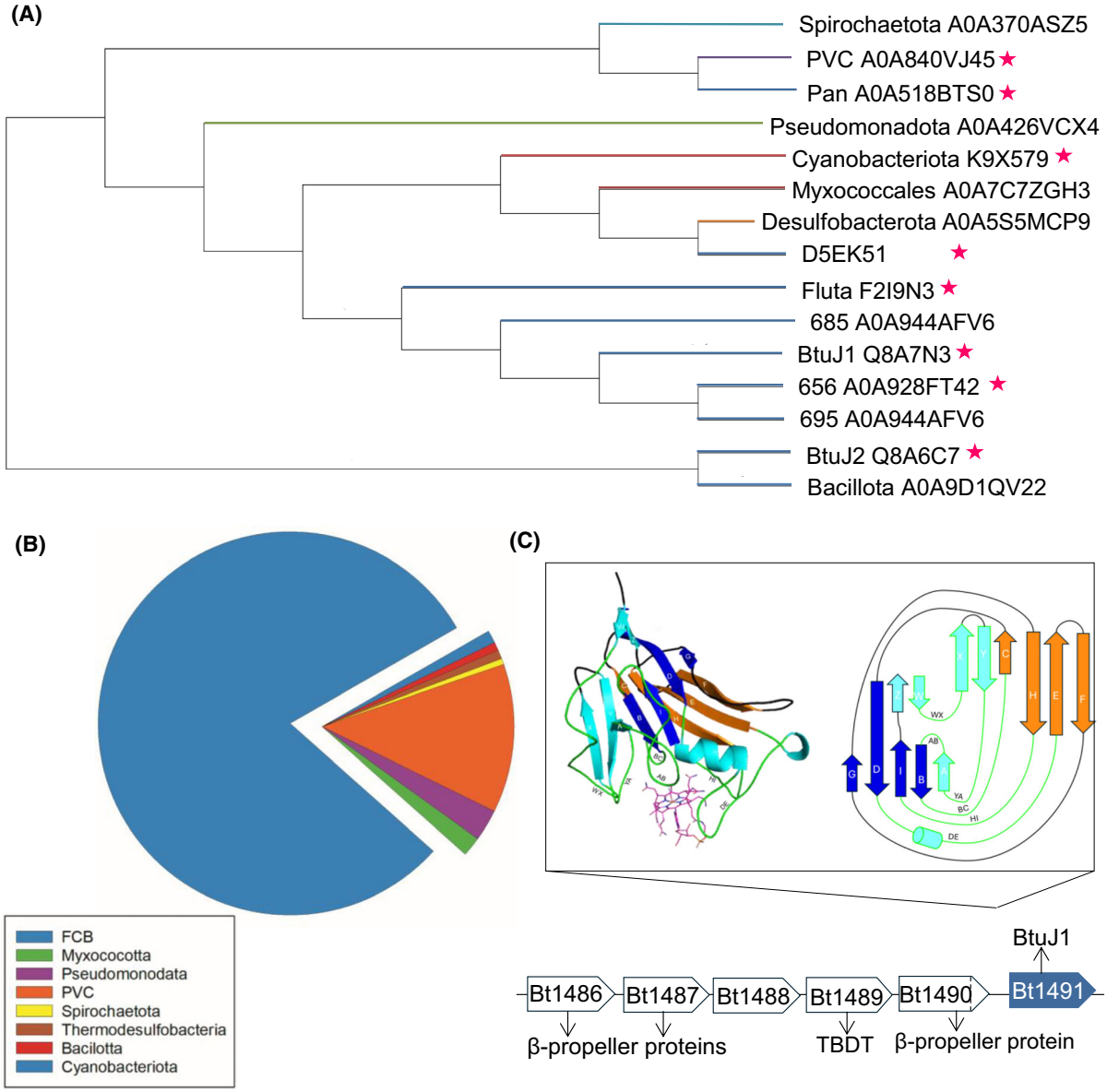
E. coli BL21(AI) cells from ThermoFisher™ were used for protein production. Transformants were selected using 100 $\mu g \cdot mL^{-1}$ ampicillin. Single colonies were used to inoculate 50 mL Luria broth (100 $\mu g \cdot mL^{-1}$ ampicillin) and cultured overnight at 37 °C with shaking. 10 mL of overnight culture was used to inoculate 800 mL fresh media and cultivated until the culture reached mid-log phase, measured with an OD_{600} value of 0.6. At this point, the culture was cooled and induced with 0.5 mM Isopropyl β-d-1-thiogalactopyranoside (IPTG). The induced culture was incubated at 20 °C overnight, while shaking. Bacteria were collected by centrifugation, and the obtained pellets kept at −20 °C until required.

Protein purification

The bacterial pellet from 400 mL culture was resuspended in 100 mL of binding buffer 20 mM Tris-Cl, 150 mM NaCl, 30 M imidazole, and 1 mM tris(2-carboxyethyl) phosphine (TCEP), and sonicated on ice. The obtained cell lysates were clarified by ultracentrifugation at 100 000×***g*** for 40 min at 4 °C. The resulting supernatant was applied to a HisTrap HP column containing 5 mL of chelating Sepharose. The His-tagged DUF4465 protein was then eluted in 5 mL fractions with elution buffer (20 mM Tris-Cl, 150 mM NaCl, 500 mM imidazole, and 1 mM TCEP). The eluted protein fractions were then concentrated and applied to a size-exclusion chromatography column (Superde×200, 16/16) in buffer 20 mM Tris–HCl (pH 7.5), 150 mM NaCl, 1 mM TCEP. The fractions corresponding to the expected elution volumes for the purified DUF domains were collected, flash-frozen in liquid nitrogen, and stored at −80 °C. To produce DUF proteins in complex with B_{12} excess cyanocobalamin was added to the cell lysate before the affinity column. Unbound B_{12} was removed during the HisTrap HP column washing step.

Fig. 1. Evolutionary distribution, structural features, and genomic context of DUF4465 family proteins. (A) Phylogenetic tree of representative DUF4465 (IPR027828 family) protein sequences from diverse bacterial phyla. Branch lengths are proportional to sequence divergence. Pink stars indicate proteins selected for study. Sequences were aligned using MUSCLE5, and the tree was constructed using the MEGA11 software. (B) Taxonomic distribution of DUF4465 homologs shown as a proportional pie chart; the FCB group represents most sequences, with smaller contributions from *Myxococcota*, *Pseudomonadota*, PVC, *Spirochaetota*, *Thermodesulfobacteria*, *Bacillota*, and *Cyanobacteriota*. (C) To the left is shown the structure of *B. theta* DUF4465 protein (gene identifier *Bt1491*, recently designated *BtuJ1*) showing cartoon representation of the *β*-jellyroll barrel with B_{12} in magenta lines (PDB code 9FCT), and right is the corresponding schematic *β*-sheet topology diagram. This panel was produced using PyMol. The inset depicts the genomic neighborhood of *Bt1491*, showing its proximity to genes encoding *btuB*, *btuC*, *btuD*, and other predicted vitamin B_{12} transport proteins. (D) Pairwise sequence identity matrix for the DUF4465 proteins shown in panel A selected for assessment of B_{12}-binding function.

22115463, 2026, 8, Downloaded from https://febs.onlinelibrary.wiley.com/doi/10.1002/2211-5463.70231 by NICE, National Institute for Health and Care Excellence, Wiley Online Library on [04/08/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License



(D)

	BtuJ2	656	BtuJ1	D5EK51	Pan	Cylst	Fluta	Halo
BtuJ2	100	19.6	20.9	21.1	21.6	19.7	22.8	23.6
656	19.6	100	19.4	19.3	16.3	23.9	20.8	21.5
BtuJ1	20.9	19.4	100	28.1	26.1	31.1	30.4	27.8
D5EK51	21.1	19.3	28.1	100	33.0	37.7	33.3	38.8
Pan	21.6	16.3	26.1	33.0	100	38.0	30.7	40.6
Cylst	19.7	23.9	31.1	37.7	38.0	100	48.5	45.0
Fluta	22.8	20.8	30.4	33.3	30.7	48.5	100	36.0
Halo	23.6	21.5	27.8	38.8	40.6	45.0	36.0	100

22115463, 2026, 8, Downloaded from https://febs.onlinelibrary.wiley.com/doi/10.1002/2211-5463.70231 by NICE, National Institute for Health and Care Excellence, Wiley Online Library on [04/08/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Protein crystallization and structure determination

The D5EK51 protein in complex with B_{12} was concentrated to 20 mg·mL^{-1} (0.7 μM) for crystallization trials using sitting drop vapor diffusion trials set up a Mosquito Crystallization robot (TTP Labtech) using the commercial screens JCSG+, Structure 1 and 2, Morpheus, MembranePlus, Proplex, LMB, MIDAS+ (Molecular Dimensions and Hampton Research). Crystallization plates were incubated at 20 °C. Crystals of D5EK51 grew in MIDAS+ and Memgold2 commercial screens. Crystallization was achieved using the MIDAS+ condition 34–1, comprising 35% poly(acrylic acid sodium salt) 2100, 0.1 M HEPES (pH 7.5), and 0.2 M ammonium sulfate. Crystals were mounted and washed in cryoprotectant augmented with 20% glycerol and then flash-frozen in liquid nitrogen in loops for data collection. Diffraction data were collected using the DECTRIS EIGER X 9 M at ESRF beamline BM30A. Structure determination, refinement, and validation used molecular replacement with an AlphaFold-generated model of D5EK51 within the CCP4 suite [24–26]. All molecular graphics were prepared using PyMOL (The PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC) [27].

Sequence alignment and phylogenetic tree construction

Amino acid sequences of DUF4465 protein representatives D5EK51, Pan, Fluta, 656, Cylst, and Halo were collected from the UniProt database and aligned using MUSCLE5 [28]. The phylogeny was inferred using the Maximum Likelihood method and Le-Gascuel model of amino acid substitutions and the tree with the highest log likelihood is shown [29]. The initial tree for the heuristic search was selected by choosing the tree with the superior log-likelihood between a Neighbor-Joining (NJ) tree and a Maximum Parsimony (MP) tree [30]. The NJ tree was generated using a matrix of pairwise distances computed using the p-distance. The MP tree had the shortest length among 10 MP tree searches; each performed with a randomly generated starting tree. The evolutionary rate differences among sites were modeled using a discrete Gamma distribution across five categories (+G = 2.7645). The analytical procedure encompassed 15 amino acid sequences. The partial deletion option was applied to eliminate all positions with less than 95% site coverage resulting in a final data set comprising 169 positions. Evolutionary analyses were conducted in MEGA11 utilizing up to eight parallel computing threads [31].

Sequence similarity network and genome neighborhood analysis

Multiple sequence alignments were obtained using MUSCLE5 and MAFFT v7 [28,32,33]. Proteins belonging to the IPR027828 family (DUF4465) were analyzed using the Enzyme Function Initiative Enzyme Similarity Tool (EFI-EST) and Genome Neighborhood Tool (EFI-GNT) (https://efi.igb.illinois.edu) [34,35]. A Sequence Similarity Network (SSN) was generated by submitting the InterPro family identifier IPR027828, retrieving UniProtKB sequences clustered at 90% sequence identity to reduce redundancy. An alignment score threshold of 15 (corresponding to c. 40% sequence identity) was used to define edges, yielding two major clusters. For genomic context analysis, representative sequences from each connected component were submitted to EFI-GNT, examining genes within ±10 open reading frames (ORFs). Gene neighborhoods were annotated based on EFI-generated diagrams and UniProt functional descriptions. Analysis of the signal sequences was performed by SignalP 6.0 [36].

Results

Comparative analysis of the DUF4465 family

A preliminary phylogenetic tree constructed from the alignment using MEGA11 [31] illustrates the broad distribution and deep branching of DUF4465 homologs, Fig. 1A. The InterPro database [37] similarly shows that members of the DUF4465 family (IPR027828) are distributed across at least eight bacterial clades, indicating broad phylogenetic representation, Fig. 1B. To capture this diversity experimentally, representative proteins were selected to span the sequence space of the family: D5EK51 from *Coraliomargarita akajimensis*, Pan from *Mucisphaera* calidilacus (A0A518BTS0), Fluta from *Fluviicola taffensis* (F2I9N3), 656 from a *Rikenellaceae* bacterium (A0A928FT42), Cylst from *Cylindrospermum stagnale* PCC 7417 (K9X579), and Halo from *Haloferula luteola* (A0A840VJ45). All selected proteins contain predicted N-terminal signal peptides according to SignalP [36,38], consistent with localization to the periplasm or cell surface and supporting a potential role in extracellular or periplasmic B_{12} acquisition.

Within this family, the B_{12} scavenging protein from *B. theta* is encoded in B_{12} riboswitch-regulated loci, structural analyses of BtuJ1 show a β-jelly roll fold characteristic of DUF4465 proteins, Fig. 1C [22]. Notably, BtuJ1 has a close homolog in *B. theta*, BtuJ2, illustrating the presence of paralogous DUF4465 proteins within the same organism, a pattern commonly observed in nutrient scavenging systems and consistent with gene duplication within this family [23]. Initial pairwise comparisons showed that the previously characterized B_{12}-binding proteins BtuJ1 and BtuJ2 share relatively low overall sequence identity, suggesting

22115463, 2026, 8, Downloaded from https://febs.onlinelibrary.wiley.com/doi/10.1002/2211-5463.70231 by NICE, National Institute for Health and Care Excellence, Wiley Online Library on [04/08/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

substantial divergence within the family. To evaluate sequence conservation more systematically, DUF4465 sequences were retrieved from UniProt and aligned using MUSCLE5 [28,39]. The resulting multiple sequence alignment revealed a wide range of sequence identity values (approximately 15–50%), Fig. 1D. Notably, substantial sequence variability was observed within the Bacteroidetes phylum, in some cases exceeding the divergence observed between different clades. For instance, protein 656 (A0A928FT42) from a *Rikenellaceae* bacterium exhibited a particularly low sequence identity compared to BtuJ1 despite belonging to the same broader phylogenetic grouping.

Determination of B_{12}-binding to DUF4465 family proteins

To assess whether cobalamin binding is a ubiquitous function across the DUF4465 protein family, representative sequences were retrieved from the UniProt database [39]. Initially, several DUF4465 candidate sequences were analyzed and selected to maximize phylogenetic breadth and sequence divergence, ensuring coverage across major bacterial clades while avoiding over-representation of closely related homologs. Signal peptide–truncated constructs were designed to promote soluble cytoplasmic expression, and six proteins (D5EK51, Pan, Fluta, 656, Cylst, and Halo) were taken forward based on successful recombinant expression.

Among these, D5EK51 from *Coraliomargarita akajimensis* was prioritized for structural studies because it reproducibly yielded diffraction-quality crystals. All proteins were purified using a two-step chromatographic workflow consisting of immobilized metal affinity chromatography followed by size-exclusion chromatography. Purifications were performed under two parallel conditions: in the presence of excess B_{12} (cyanocobalamin) and in its absence, enabling direct comparison of ligand-bound and apo states.

Evidence of B_{12} binding was apparent during purification. For D5EK51, Pan, Fluta, Halo, Cylst and 656, nickel affinity columns developed and retained a distinct pink coloration even after extensive washing when proteins were purified in the presence of B_{12}, indicating stable retention of the protein-B_{12} complex. Eluted fractions from these purifications were similarly pink, consistent with co-elution of protein–B_{12} complexes. Protein containing fractions were confirmed by absorbance at 280 nm and SDS/PAGE analysis, while the presence of bound B_{12} was monitored via absorbance at 361 nm. Comparative UV–visible spectra of proteins purified in the presence and absence of B_{12} further supported cofactor association, Fig. S1. Proteins isolated with B_{12} displayed the characteristic cobalamin absorbance features, including a peak near 361 nm, which was absent in apo preparations. Notably, this peak exhibited a measurable shift relative to free B_{12} in solution, consistent with perturbation of the B_{12} environment upon protein binding. Such spectral changes are typical when cobalamin transitions from solvent-exposed to a protein-embedded state, where hydrogen-bonding interactions, and local polarity alter the ligand's electronic environment These data therefore provide spectroscopic evidence that B_{12} is not merely copurifying nonspecifically but is engaged in a defined binding environment within DUF4465 proteins.

Size exclusion chromatography profiles also revealed clear ligand-dependent differences in elution behavior. Chromatograms of the D5EK51–B_{12} complex showed a reproducible leftward shift in elution volume relative to the apo protein, Fig. 2. The peak is also narrower indicating a more ordered state consistent with the ordering of the flexible loops upon ligand-binding. Similar trends were observed for the other B_{12}-binding homologs, supporting the conclusion that cobalamin binding is a conserved biochemical feature within these DUF4465 proteins.

Determination of the structure of the D5EK51-B_{12} complex

The D5EK51-B_{12} complex was successfully purified and concentrated, facilitating the growth of protein crystals. The resulting crystals belonged to the space group $P3_221$, with three D5EK51-B_{12} complexes in the asymmetric unit. Structural alignment of chains 2 and 3 to chain 1 yielded root-mean-square deviations (RMSDs) of 0.154 Å and 0.084 Å, respectively, over 170 equivalent Cα atoms, indicating high structural similarity. Structure determination was performed using molecular replacement with an AlphaFold-generated [24,25] model of D5EK51 within the CCP4 suite [26]. The structure was refined to a resolution of 1.85 Å, achieving R_{work} and R_{free} values of 0.198 and 0.222, respectively. Notably, the final model exhibited no Ramachandran outliers, underscoring its stereochemical quality. Crystallographic and refinement statistics are summarized in Table S1.

Structural features and B_{12} binding mechanism

The D5EK51 protein adopts an augmented β-jellyroll fold. This fold consists of eight β-strands arranged into two antiparallel four-stranded β-sheets, BIDG and

22115463, 2026, 8, Downloaded from https://febs.onlinelibrary.wiley.com/doi/10.1002/2211-5463.70231 by NICE, National Institute for Health and Care Excellence, Wiley Online Library on [04/08/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

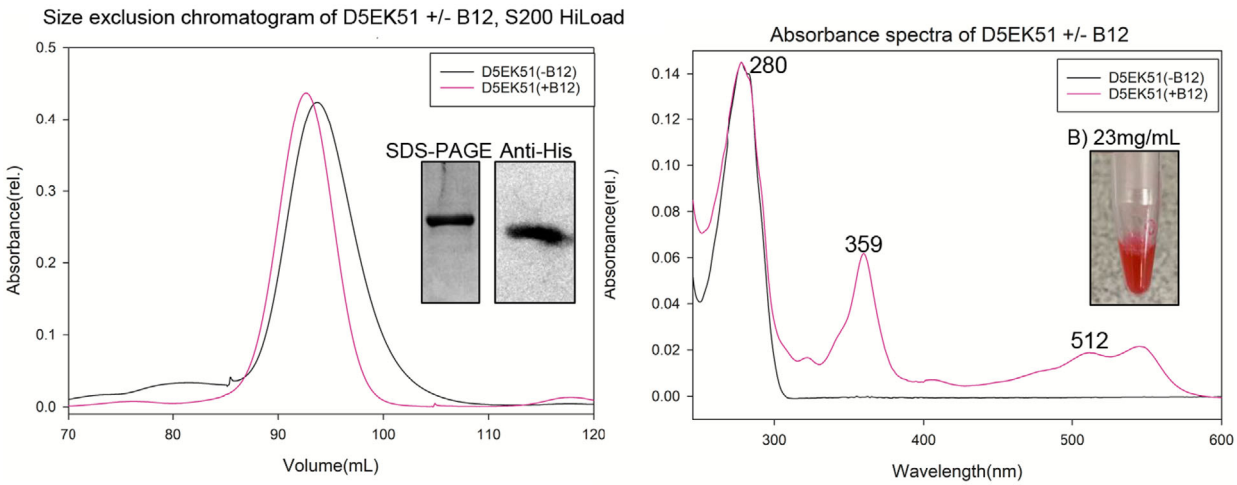


Fig. 2. Spectroscopic and chromatographic characterization of the D5EK51–B_{12} complex. (A) Size-exclusion chromatograms of D5EK51 purified with (pink line) or without (black line) B_{12} on a Superdex 200 HiLoad column. Both samples elute at ~95 mL, corresponding to the monomeric form of D5EK51 (26.9 kDa). SDS/PAGE and anti-His western blot (inset) confirm the purity and presence of D5EK51 in both conditions. (B) UV–visible absorbance spectra of purified D5EK51 alone (black) and in complex with cyanocobalamin (pink). The D5EK51–B_{12} complex exhibits distinct absorbance features at 361 nm and 550 nm, consistent with B_{12} binding, while the unbound protein shows only the characteristic 280 nm absorbance. The inset shows the deep pink coloration of the D5EK51–B_{12} complex at 23 mg·mL^{-1}, indicative of bound cobalt-containing cobalamin.

CHEF, as shown in Fig. 3A,B. In D5EK51, the β-jellyroll architecture is extended by five additional anti-parallel β-strands at the N terminus (WXYZA) and an additional α-helix inserted within loop DE. This structural arrangement is consistent with other recently solved DUF4465 family members, BtuJ1 and BtuJ2 (PDB codes: 9FCT and 9I2L [22]).

Cobalamin is well-resolved within each D5EK51 molecule of the asymmetric unit, Fig. 3C. The density for the upper ligand of cyanocobalamin is not as strong as for the rest of the molecule, which is evidence of X-ray-induced reduction of the cyanocobalamin's cobalt centre to give cob(II)alamin [40]. An upper limit to the K_d for cobalamin can be determined from the crystallization conditions as the cobalamin site is fully occupied in the electron density map. As the protein and ligand concentrations are 0.7 μM, K_d will be less than 0.2 μM and is likely much lower than this value but possibly not as low as the K_d of BtuJ1 measured at circa 10^{-10} M. The cobalamin binding pocket is formed by six loops conserved among DUF4465 proteins and is primarily lined with aromatic hydrophobic residues, including tyrosine, phenylalanine, and tryptophan. These aromatic residues not only accommodate the corrin ring but also form contacts with its side chains, facilitating tight binding, Fig. 3D.

In addition, a network of 11 hydrogen bonds is formed between D5EK51 and the side chains of B_{12}, involving both backbone and side-chain protein atoms. This dense hydrogen-bonding network is comparable to those observed in the well-characterized cobalamin-binding proteins BtuJ1 and BtuJ2, which form 16 and 13 hydrogen bonds with the B_{12}, respectively. PDBePISA analysis further defines the contribution of nonpolar interactions to complex stability, highlighting extensive hydrophobic surface complementarity between the protein and the corrin ring. All crystallized DUF4465–B_{12} complexes show substantial ligand burial, with more than 50% of the B_{12} solvent-accessible surface area occluded upon binding. For D5EK51, the buried surface area is 688Å^2 (56%), compared with 775Å^2 (62%) for BtuJ1 and 857Å^2 (69%) for BtuJ2 [41]. The magnitude of surface burial and multiple polar interactions together indicate a tightly packed, specifically coordinated binding interface. D5EK51 shares 28% and 21% sequence identity with BtuJ1 and BtuJ2 and their structures align with RMSD of 3.548 Å and 1.518 Å, respectively. The HI loop and the DE loop/helix (named after the β-strands connected) are highly conserved among the three structures, both in their conformation and in the identity of B_{12}-coordinating residues. These loops contain the conserved B_{12}-binding tyrosine's, on loop HI (Y245 in BtuJ1 numbering) and within the α-helix of loop DE (Y146), highlighting a preserved functional core of the DUF4465 family.

Surface conservation mapping further supports this conclusion, Fig. 4A. Key aromatic residues within the

22115463, 2026, 8, Downloaded from https://febs.onlinelibrary.wiley.com/doi/10.1002/2211-5463.70231 by NICE, National Institute for Health and Care Excellence, Wiley Online Library on [04/08/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

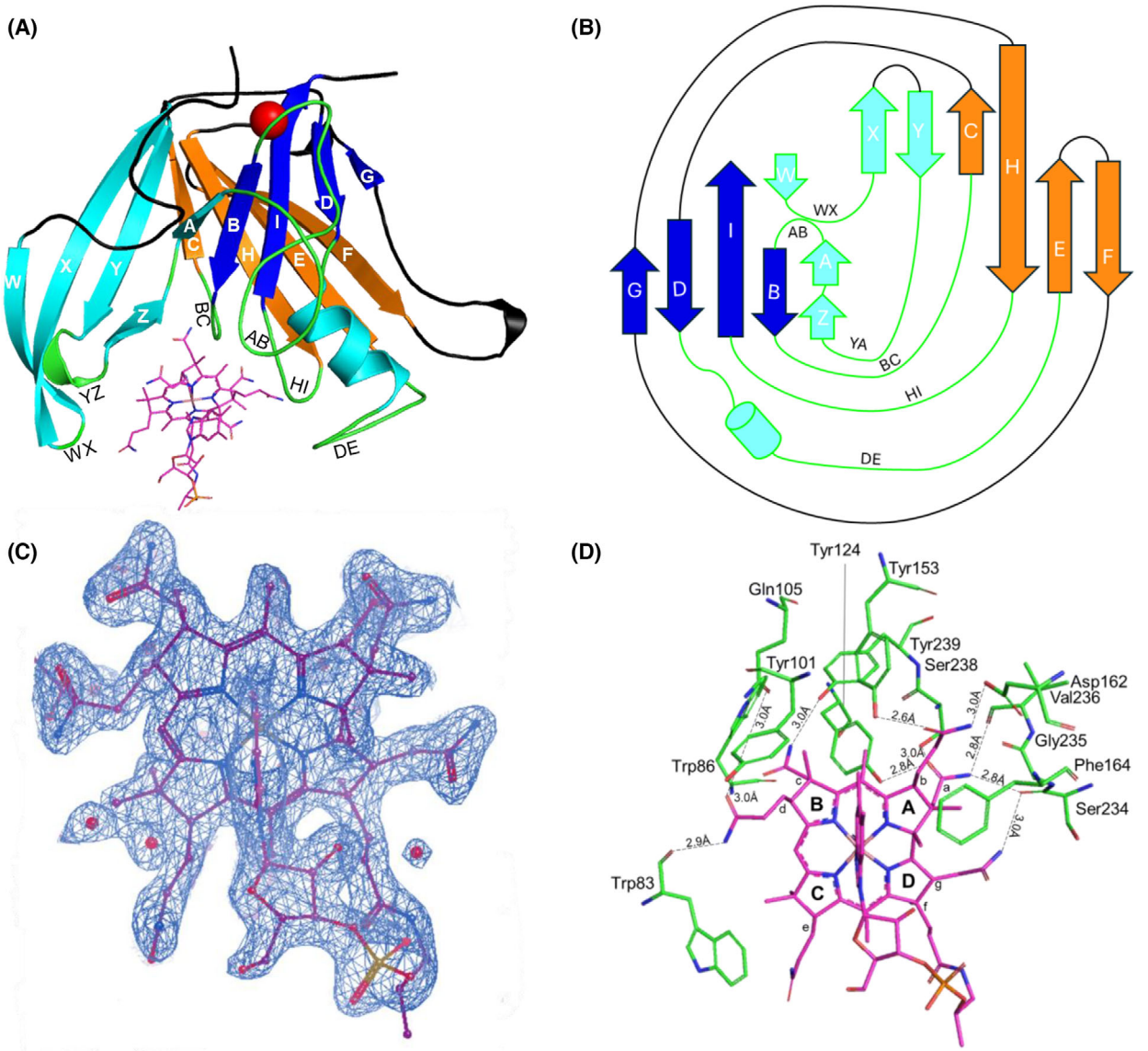


Fig. 3. Structural features of the D5EK51–B_{12} complex. (A) Overall cartoon representation of the D5EK51 structure, highlighting the augmented β-jellyroll fold characteristic of DUF4465 family proteins. The two β-sheets are labeled according to convention: BIDG (blue) and CHEF (orange). Five additional β-strands (WXYZA) at the N terminus and a loop-associated α-helix (loop DE) are also present. Loops are named according to the β-strands they link together; hence loop HI connects β-strands H and I. Loops forming the B_{12}-binding site are shown in green. Cyanocobalamin is depicted in magenta, and a stabilizing calcium ion is shown as a red sphere. (B) Topology diagram of the D5EK51 secondary structure. The β-jellyroll fold consists of two antiparallel four-stranded β-sheets with additional N-terminal β-strands (WXYZA) and an α-helix (loop DE), consistent with other DUF4465 family proteins. (C) 2Fo–Fc electron density map (blue mesh) of the B_{12} ligand contoured at 1.0σ, showing well-resolved density for the corrin ring and side chains. Ordered water molecules are also visible and contribute to ligand binding. (D) Close-up view of the B_{12}-binding pocket showing the hydrogen bonding network (dashed lines). Eleven hydrogen bonds are formed between D5EK51 and B_{12}, involving both main-chain and side-chain atoms. These interactions predominantly occur near the A-ring of the corrin ring, stabilizing B_{12} within a hydrophobic cavity composed of aromatic residues shown in green. This Figure was produced using PyMol.

binding pocket are conserved across homologs, particularly those in loops YZ, AB, BC, HI, and DE. The HI loop (GTPAYF) is especially conserved, with all residues present at greater than 70% frequency and the critical B_{12}-binding tyrosine retained in 97% of homologs [42]. Sequence motif alignment was used to create a sequence logo, Fig. 4B and extension of this analysis representatives of all eight clades strongly

22115463, 2026, 8, Downloaded from https://febs.onlinelibrary.wiley.com/doi/10.1002/2211-5463.70231 by NICE, National Institute for Health and Care Excellence, Wiley Online Library on [04/08/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

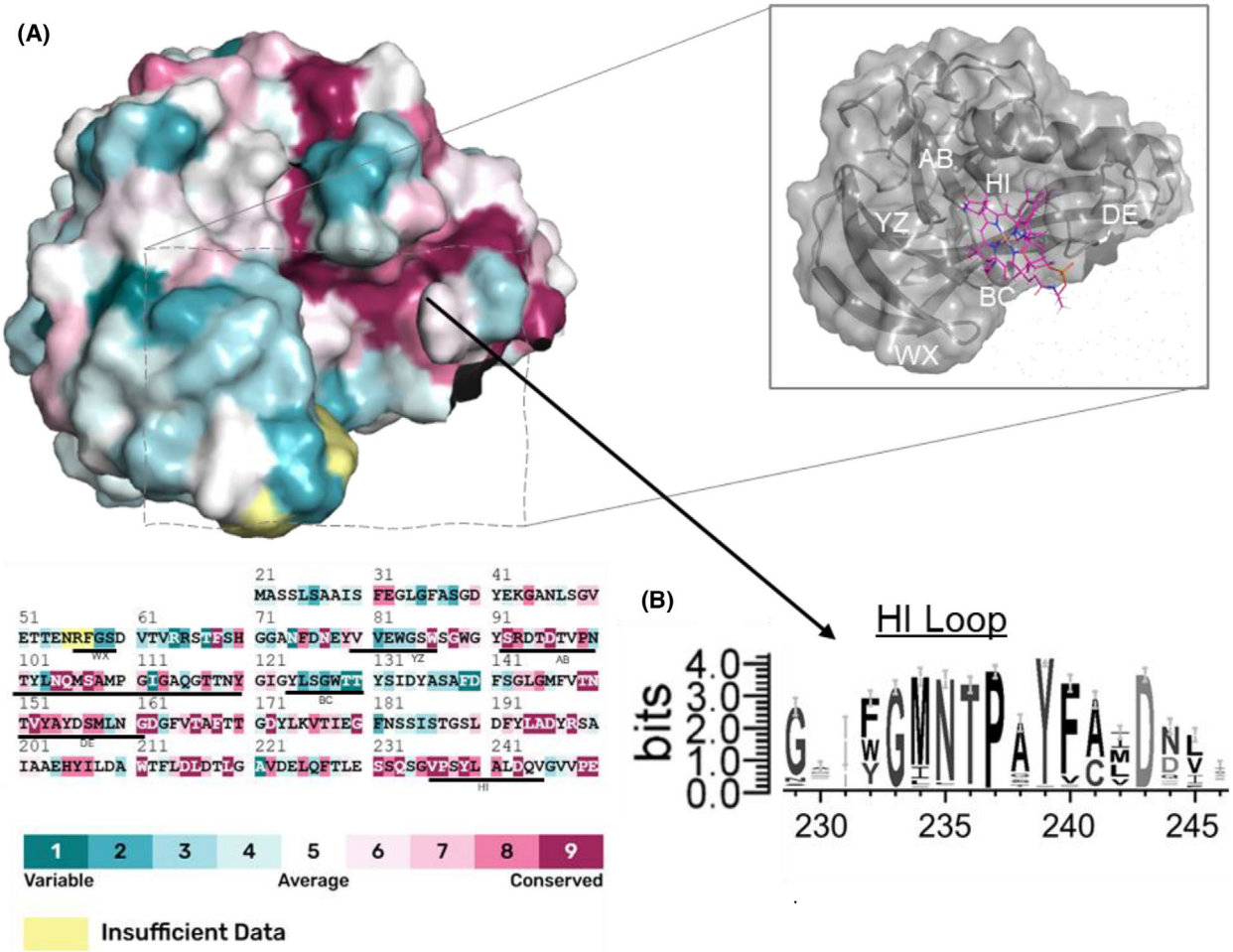


Fig. 4. ConSurf analysis of D5EK51 reveals evolutionary conservation of the B_{12}-binding pocket. (A) Surface representation of the experimentally determined structure colored by residue conservation (teal indicates variable and pink conserved residues), highlighting the conserved B_{12}-binding pocket (the boxed region indicates the positions of the loops). Below is a multiple sequence alignment of representative homologs with conservation scores (1–9) mapped to the structure. The surface conservation and corresponding sequence alignment shown in this Figure was generated using the ConSurf server. (B) Sequence logo for residues predicted on HI loop, showing a conserved NTPAY motif.

support B_{12} binding as a conserved function across the DUF4465 family.

Genomic analysis of genetic context of DUF4465 encoding genes

We identified two major genomic clusters of DUF4465 proteins [34,35]. The first, comprising approximately 900 members, showed frequent colocalization with transport-associated domains: 20% colocated with ABC-type transporters (PF00005), 40% with TonB-dependent β-barrels (PF00593), 55% with their associated plug domains (PF07715), and 40% with β-propeller domain proteins (PF16819). The second cluster, containing 195 DUF4465 members, exhibited even stronger associations, with 80–95% co-occurring with these transport systems and periplasmic binding proteins. In total, 84 DUF4465 proteins lacked neighboring Pfam-annotations, suggesting that these values may underestimate the true frequency of such associations. Together, these patterns indicate that DUF4465 proteins are frequently genomically linked to vitamin B_{12} transport systems, supporting their role in B_{12} acquisition.

Discussion

DUF4465 proteins provide valuable insight into how bacteria handle the transport of a complex and fragile cofactor such as vitamin B_{12}. Among the suite of B_{12}-

22115463, 2026, 8, Downloaded from https://febs.onlinelibrary.wiley.com/doi/10.1002/2211-5463.70231 by NICE, National Institute for Health and Care Excellence, Wiley Online Library on [04/08/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

handling proteins, there is strong selective pressure to evolve architectures that tightly bind the cofactor; both because B_{12} is scarce in most environments and because it is chemically unstable. Vitamin B_{12} is well known to be sensitive to light, heat, oxidation, and pH, all of which can lead to rapid degradation under many physiological conditions [43,44]. It is therefore unsurprising that such a vital cofactor is almost universally trafficked via dedicated protein chaperones that protect it from damage and ensure its safe delivery into the cell for enzymatic use [45,46].

Previous analyses of B_{12}-bound structures in the Protein Data Bank (PDB) have identified similar patterns described here, such as a 'hydrophobic shroud' in the B_{12} binding pocket [47]. However, many of these structures correspond to enzymes, which are designed to retain B_{12} as a cofactor rather than transfer it to other proteins. In contrast, our analysis focuses on B_{12}-binding proteins involved in scavenging and transport, providing new insight into a distinct class of B_{12} interaction mechanisms.

The DUF4465 proteins use a common binding mechanism involving loops that provide shape complementarity and a high density of hydrophobic residues, Fig. 5. Notably, loops are also often seen binding B_{12} in other bacterial B_{12} transport proteins, BtuK, BtuH, BtuG, BtuB, BtuF.

Therefore, a loop pocket architecture, supported by the conserved β-jellyroll scaffold in DUF4465 proteins, appears to be the solution for achieving high affinity and conditional release. The extensive hydrogen-bonding network from the loops consistently observed in DUF4465 and other B_{12} transport proteins further stabilizes the cofactor while still allowing partial solvent exposure; this is a structural signature of proteins specialized for transfer rather than catalysis. Interestingly, in eukaryotic B_{12} transport proteins, the B_{12} is encapsulated by two domains, which shield it from solvent; these proteins do not perform transfer and are instead internalized by the cell and digested [48].

Despite consistently conserved features in DUF4465 proteins, detailed analysis of the D5EK51 crystal structure and sequence reveals differences compared to BtuJ1 and BtuJ2, particularly in the N-terminal region of the DUF4465 fold. One notable variation is the auxiliary *β*-strands at the N terminus (*β*-strands W, X, and Y, Fig. 3). While the number of these strands varies among family members, BtuJ1 and BtuJ2 each possess four auxiliary strands, whereas D5EK51 contains five (W, X, Y, and Z, A, Fig. 3). This seemingly minor difference has a significant impact on the topology of the binding pocket, especially the lengths and conformations of loops WX and YA. Overall, the WX and YA loop architecture varies markedly across available structures and shows lowest conservation; a sequence alignment is shown in Fig. S2. BtuJ2 exhibits considerably longer WX/YA loops, whereas both BtuJ1 and D5EK51 have shorter loops. These loops, along with the auxiliary strands, shape the N-terminal side of the pocket (left-side as shown in Fig. 5) responsible for coordinating the B_{12} B and C rings; they may also be critical for lower ligand recognition or for protein–protein transfer interactions due to their position. Additionally, we observe distinct hydrogen-bonding patterns emerging at the N-terminal end of the binding site for different DUF4465 proteins. BtuJ1 employs a conserved tryptophan residue to π-stack against the adenosyl lower ligand of the B_{12} molecule. In contrast, BtuJ2/D5EK51 uses two tyrosine/tryptophan residues to achieve a similar anchoring function but via direct hydrogen bonding, avoiding contact with the lower ligand. While both mechanisms effectively coordinate side chains from C and D of the B_{12} corrin ring, they result in slightly different ligand conformations, specifically of the *d* B_{12} side chain.

Further identification of DUF4465 family members and the biochemical confirmation of their tight B_{12}-binding expands our understanding of their role in bacterial B_{12} scavenging. Particularly, the widespread distribution of DUF4465 homologs across diverse bacterial lineages suggests that this mode of B_{12} handling is not confined to gut-associated species but rather represents a general adaptation to the scarcity of B_{12} in most environments. This finding challenges the prevailing view that related dedicated B_{12} scavenging systems are a feature unique to *B. theta* and related gut microbes. Instead, the presence of DUF4465 proteins across multiple phyla suggests that bacteria from varied ecological niches have adopted similar strategies for capturing and stabilizing cobamides.

Genomic analyses further support this conclusion: DUF4465-encoding genes are frequently co-located with known B_{12} transporter loci, implying a coordinated role in uptake and trafficking. Notably, over 70% of these proteins are predicted lipoproteins, with the remainder bearing periplasmic targeting signals; a pattern consistent with localization to the outer membrane and function at the cell–environment interface. In addition, DUF4465 proteins are often genomically associated with other "DUF" protein families that share structural similarity with accessory B_{12}-scavenging proteins from *B. theta*. But due to naming convention inconsistencies, it is more difficult to group these proteins together.

22115463, 2026, 8, Downloaded from https://febs.onlinelibrary.wiley.com/doi/10.1002/2211-5463.70231 by NICE, National Institute for Health and Care Excellence, Wiley Online Library on [04/08/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

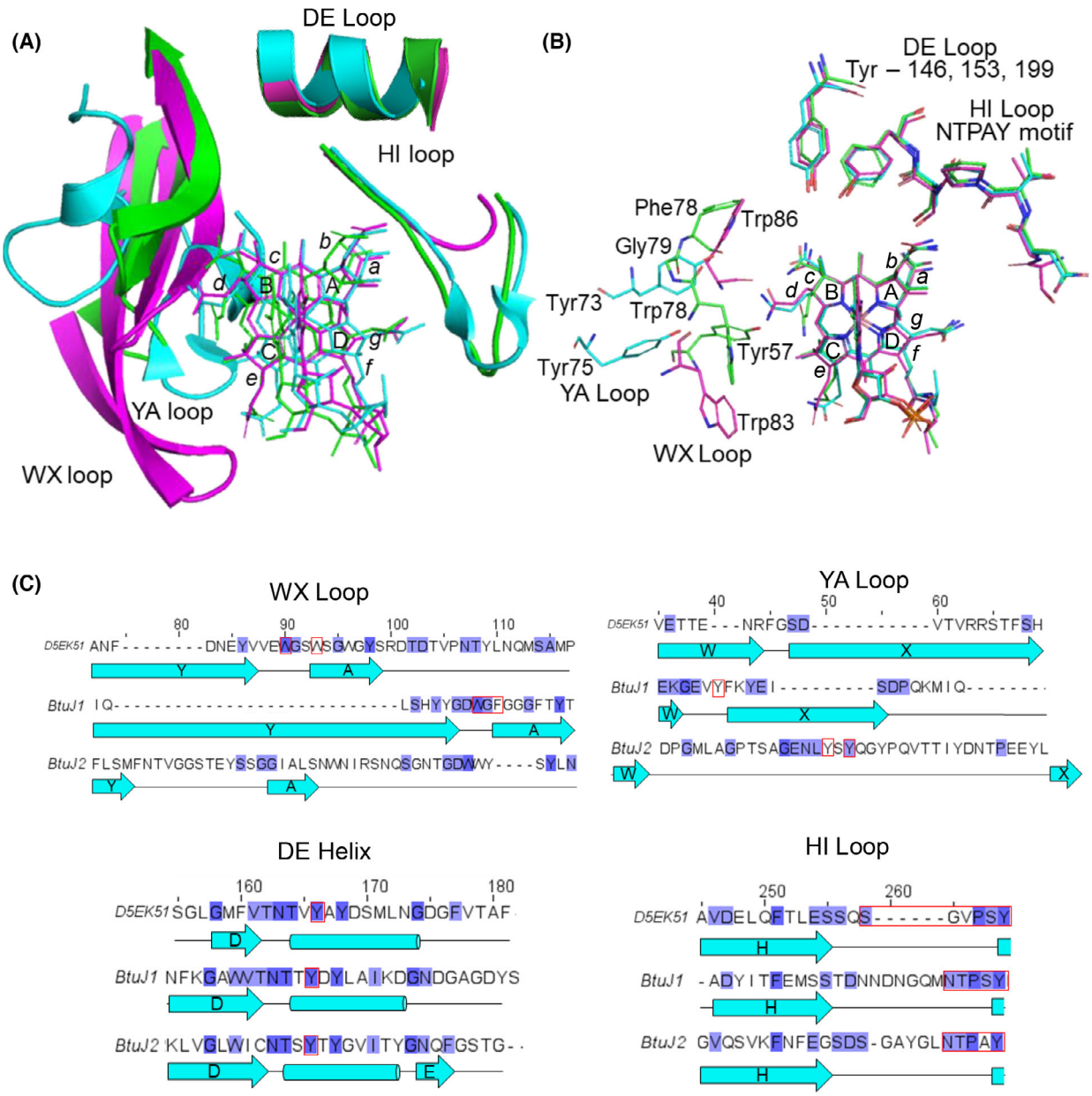


Fig. 5. Comparison of the B_{12}-binding loops of D5EK51, BtuJ1 and BtuJ2. (A) Cartoon representation of the HI, DE, WX and YA B_{12}-binding loops of D5EK51 (magenta), BtuJ1 (green) and BtuJ2 (cyan) together with their bound B_{12} molecules after superimposition (PDB codes: 9HMU, 9FCT, and 9I2L respectively). The N-terminal WX and YA loops vary significantly, but the C-terminal DE and HI loops are highly conserved. (B) Close-up view of the B_{12} binding site reveals minimal conservation of aromatic and hydrogen bonding residues responsible for ligand coordination on the YA, WX side of the pocket. However, side chains from all three structures show near-identical orientation around the A ring of the B_{12} molecule (DE and HI loops), highlighting a shared ligand recognition strategy. (C) Structure-based sequence alignment of the loops from BtuJ1, BtuJ2, and D5EK51. Conserved residues involved in B_{12}-binding are boxed in red, and secondary structure elements are shown as arrows (β-strands) and cylinders (α-helices). This Figure was produced using PyMol.

These results suggest that while not all bacteria may invest in such specialized machinery, a substantial proportion do, likely reflecting ecological or metabolic trade-offs in cobamide acquisition. Interestingly, this same selective pressure is mirrored in laboratory evolution studies; Mok et al.[49] showed that *E. coli*, when forced to grow with a suboptimal B_{12} analog, rapidly accumulates mutations that enhance B_{12} transport and activation [49]. Although such experiments cannot fully capture the complexity of natural

22115463, 2026, 8, Downloaded from https://febs.onlinelibrary.wiley.com/doi/10.1002/2211-5463.70231 by NICE, National Institute for Health and Care Excellence, Wiley Online Library on [04/08/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

ecosystems, they reinforce the idea that strengthening cobamide uptake and handling is a universal adaptive solution to cofactor limitation.

Acknowledgements

This work was supported by the Biotechnology and Biological Sciences Research Council (BBSRC) under Grant BB/X001946/1. We acknowledge the European Synchrotron Radiation Facility (ESRF) for provision of synchrotron radiation facilities under proposal number mx2606/mx2698 and we would like to thank beamline staff for assistance and support in using beamline BM30A. The authors would also like to thank Diamond Light Source for beamtime (proposal mx38316), and the staff of beamlines I02, I03, and I24 for assistance with assessing the suitability of crystals for data collection and for data collection.

Conflict of interest

The authors declare no conflict of interest.

Peer review

The peer review history for this article is available at https://www.webofscience.com/api/gateway/wos/peer-review/10.1002/2211-5463.70231.

Data accessibility

The structure and structure factors for DUF4465 protein D5EK51 are deposited in the protein databank (PDB) with code 9HMU.

Author contributions

CC was involved in conceptualization, methodology, formal analysis, investigation, writing – original draft, writing – review and editing. MB was involved in methodology, formal analysis. RJ was involved in methodology. MW was involved in conceptualization, funding acquisition. RWP was involved in conceptualization, writing – original draft, writing – review and editing, supervision, funding acquisition.

References

1 Sokolovskaya OM, Shelton AN and Taga ME (2020) Sharing vitamins: Cobamides unveil microbial interactions. *Science* **369**, eaba0165.
2 Bassford PJ, Bradbeer C, Kadner RJ and Schnaitman CA (1976) Transport of vitamin B_{12} in tonB mutants of *Escherichia coli*. *J Bacteriol* **128**, 242–247.
3 Shelton AN, Seth EC, Mok KC, Han AW, Jackson SN, Haft DR and Taga ME (2019) Uneven distribution of cobamide biosynthesis and dependence in bacteria predicted by comparative genomics. *ISME J* **13**, 789–804.
4 Degnan PH, Barry NA, Mok KC, Taga ME and Goodman AL (2014) Human gut microbes use multiple transporters to distinguish vitamin B_{12} analogs and compete in the gut. *Cell Host Microbe* **15**, 47–57.
5 Magnúsdóttir S, Ravcheev D, de Crécy-Lagard V and Thiele I (2015) Systematic genome assessment of B-vitamin biosynthesis suggests co-operation among gut microbes. *Front Genet* **6**, 148.
6 Croft MT, Lawrence AD, Raux-Deery E, Warren MJ and Smith AG (2005) Algae acquire vitamin B_{12} through a symbiotic relationship with bacteria. *Nature* **438**, 90–93.
7 Degnan PH, Taga ME and Goodman AL (2014) Vitamin B_{12} as a modulator of gut microbial ecology. *Cell Metab* **20**, 769–778.
8 Kräutler B (2012) Biochemistry of B_{12}-Cofactors in Human Metabolism. *Subcell Biochem* **56**, 323–346.
9 Roth J, Lawrence J and Bobik T (1996) Cobalamin (coenzyme B_{12}): synthesis and biological significance. *Annu Rev Microbiol* **50**, 137–181.
10 Smith AD, Warren MJ and Refsum H (2018) Vitamin B_{12}. *Adv Food Nutr Res* **83**, 215–279.
11 Rodionov DA, Vitreschak AG, Mironov AA and Gelfand MS (2003) Comparative genomics of the vitamin B_{12} metabolism and regulation in prokaryotes. *J Biol Chem* **278**, 41148–41159.
12 Zhang Y, Rodionov DA, Gelfand MS and Gladyshev VN (2009) Comparative genomic analyses of nickel, cobalt and vitamin B_{12} utilization. *BMC Genomics* **10**, 78.
13 Villa EA and Escalante-Semerena JC (2024) Corrinoid salvaging and cobamide remodeling in bacteria and archaea. *J Bacteriol* **206**, e0028624.
14 Helliwell KE, Lawrence AD, Holzer A, Kudahl UJ, Sasso S, Kräutler B, Scanlan DJ, Warren MJ and Smith AG (2016) Cyanobacteria and eukaryotic algae use different chemical variants of vitamin B_{12}. *Curr Biol* **26**, 999–1008.
15 Helliwell KE (2017) The roles of B vitamins in phytoplankton nutrition: new perspectives and prospects. *New Phytol* **216**, 62–68.
16 Shultis DD, Purdy MD, Banchs CN and Wiener MC (2006) Outer membrane active transport: structure of the BtuB:TonB complex. *Science* **312**, 1396–1399.
17 Korkhov VM, Mireku SA, Veprintsev DB and Locher KP (2014) Structure of AMP-PNP–bound BtuCD and mechanism of ATP-powered vitamin B_{12} transport by BtuCD–F. *Nat Struct Mol Biol* **21**, 1097–1099.
18 Cadieux N, Bradbeer C, Reeger-Schneider E, Köster W, Mohanty AK, Wiener MC and Kadner RJ (2002)

22115463, 2026, 8, Downloaded from https://febs.onlinelibrary.wiley.com/doi/10.1002/2211-5463.70231 by NICE, National Institute for Health and Care Excellence, Wiley Online Library on [04/08/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Identification of the periplasmic cobalamin-binding protein BtuF of *Escherichia coli*. *J Bacteriol* **184**, 706–717.
19 Wexler AG, Schofield WB, Degnan PH, Folta-Stogniew E, Barry NA and Goodman AL (2018) Human gut *Bacteroides* capture vitamin B_{12} via cell surface-exposed lipoproteins. *Elife* **7**, e37138.
20 Putnam EE, Abellon-Ruiz J, Killinger BJ, Rosnow JJ, Wexler AG, Folta-Stogniew E, Wright AT, van den Berg B and Goodman AL (2022) Gut commensal Bacteroidetes encode a novel class of vitamin B_{12}-binding proteins. *MBio* **13**, e0284521.
21 Abellon-Ruiz J, Jana K, Silale A, Frey AM, Baslé A, Trost M, Kleinekathöfer U and van den Berg B (2023) BtuB TonB-dependent transporters and BtuG surface lipoproteins form stable complexes for vitamin B_{12} uptake in gut *Bacteroides*. *Nat Commun* **14**, 4714.
22 Juodeikis R, Ulrich R, Clarke C, Banasik M, Deery E, Saalbach G, Kräutler B, Carding SR, Geeves MA, Pickersgill RW *et al.* (2025) Extracellular vesicle-linked vitamin B_{12} acquisition via novel binding proteins in *Bacteroides thetaiotaomicron*. *Biochem J* **482**, 1793–1812.
23 Abellon-Ruiz J, Pacheco-Gomez R, Watts J, Hart A, Hirt RH, Baslé A and van den Berg B (2025) BtuJ1, a novel surface-exposed B_{12}-binding protein in Bacteroidetes, functions as an extracellular vitamin reservoir that enhances fitness.
24 Varadi M, Bertoni D, Magana P, Paramval U, Pidruchna I, Radhakrishnan M, Tsenkov M, Nair S, Mirdita M, Yeo J *et al.* (2024) AlphaFold protein structure database in 2024: providing structure coverage for over 214 million protein sequences. *Nucleic Acids Res* **52**, D368–D375.
25 Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, Tunyasuvunakool K, Bates R, Žídek A, Potapenko A *et al.* (2021) Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589.
26 Agirre J, Atanasova M, Bagdonas H, Ballard CB, Baslé A, Beilsten-Edmands J, Borges RJ, Brown DG, Burgos-Mármol JJ, Berrisford JM *et al.* (2023) The CCP4 suite: integrative software for macromolecular crystallography. *Acta Crystallogr Sect D Struct Biol* **79**, 449–461.
27 Schrödinger L (2025) The PyMOL Molecular Graphics System, Version 2.0.
28 Edgar RC (2021) High-accuracy alignment ensembles enable unbiased assessments of sequence homology and phylogeny.
29 Le SQ and Gascuel O (2008) An improved general amino acid replacement matrix. *Mol Biol Evol* **25**, 1307–1320.
30 Saitou N and Nei M (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol* **4**, 406–425.
31 Tamura K, Stecher G and Kumar S (2021) MEGA11: molecular evolutionary genetics analysis version 11. *Mol Biol Evol* **38**, 3022–3027.
32 Kuraku S, Zmasek CM, Nishimura O and Katoh K (2013) aLeaves facilitates on-demand exploration of metazoan gene family trees on MAFFT sequence alignment server with enhanced interactivity. *Nucleic Acids Res* **41**, W22–W28.
33 Katoh K, Rozewicki J and Yamada KD (2019) MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization. *Brief Bioinform* **20**, 1160–1166.
34 Oberg N, Zallot R and Gerlt JA (2023) EFI-EST, EFI-GNT, and EFI-CGFP: enzyme function initiative (EFI) web resource for genomic enzymology tools. *J Mol Biol* **435**, 168018.
35 Zallot R, Oberg N and Gerlt JA (2019) The EFI web resource for genomic enzymology tools: leveraging protein, genome, and metagenome databases to discover novel enzymes and metabolic pathways. *Biochemistry* **58**, 4169–4182.
36 Almagro Armenteros JJ, Tsirigos KD, Sønderby CK, Petersen TN, Winther O, Brunak S, von Heijne G and Nielsen H (2019) SignalP 5.0 improves signal peptide predictions using deep neural networks. *Nat Biotechnol* **37**, 420–423.
37 Blum M, Andreeva A, Florentino LC, Chuguransky SR, Grego T, Hobbs E, Pinto BL, Orr A, Paysan-Lafosse T, Ponamareva I *et al.* (2025) InterPro: the protein sequence classification resource in 2025. *Nucleic Acids Res* **53**, D444–D456.
38 Teufel F, Almagro Armenteros JJ, Johansen AR, Gíslason MH, Pihl SI, Tsirigos KD, Winther O, Brunak S, von Heijne G and Nielsen H (2022) SignalP 6.0 predicts all five types of signal peptides using protein language models. *Nat Biotechnol* **40**, 1023–1025.
39 Ahmad S, da Jose Costa Gonzales L, Bowler-Barnett EH, Rice DL, Kim M, Wijerathne S, Luciani A, Kandasaamy S, Luo J, Watkins X *et al.* (2025) The UniProt website API: facilitating programmatic access to protein knowledge. *Nucleic Acids Res* **53**, W547–W553.
40 Champloy F, Gruber K, Jogl G and Kratky C (2000) XAS spectroscopy reveals X-ray-induced photoreduction of free and protein-bound B12 cofactors. *J Synchrotron Radiat* **7**, 267–273.
41 Krissinel E and Henrick K (2007) Inference of macromolecular assemblies from crystalline state. *J Mol Biol* **372**, 774–797.
42 Ashkenazy H, Abadi S, Martz E, Chay O, Mayrose I, Pupko T and Ben-Tal N (2016) ConSurf 2016: an improved methodology to estimate and visualize evolutionary conservation in macromolecules. *Nucleic Acids Res* **44**, W344–W350.

22115463, 2026, 8, Downloaded from https://febs.onlinelibrary.wiley.com/doi/10.1002/2211-5463.70231 by NICE, National Institute for Health and Care Excellence, Wiley Online Library on [04/08/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

43 Schnellbaecher A, Binder D, Bellmaine S and Zimmer A (2019) Vitamins in cell culture media: stability and stabilization strategies. *Biotechnol Bioeng* **116**, 1537–1555.
44 Yamada K, Shimodaira M, Chida S, Yamada N, Matsushima N, Fukuda M and Yamada S (2008) Degradation of vitamin B_{12} in dietary supplements. *Int J Vitam Nutr Res* **78**, 195–203.
45 Gherasim C, Lofgren M and Banerjee R (2013) Navigating the B_{12} road: assimilation, delivery, and disorders of cobalamin. *J Biol Chem* **288**, 13186–13193.
46 Sañudo-Wilhelmy SA, Gómez-Consarnau L, Suffridge C and Webb EA (2014) The role of B vitamins in marine biogeochemistry. *Ann Rev Mar Sci* **6**, 339–367.
47 Pradeep P and Nagarajan D (2023) Conserved structural motifs across diverse vitamin B_{12} binding proteins. *J Bioinform Syst Biol* **6**, 312–326.
48 Mathews FS, Gordon MM, Chen Z, Rajashankar KR, Ealick SE, Alpers DH and Sukumar N (2007) Crystal structure of human intrinsic factor: cobalamin complex at 2.6-Å resolution. *Proc Natl Acad Sci* **104**, 17311–17316.
49 Mok KC, Hallberg ZF, Procknow RR and Taga ME (2025) Laboratory evolution of *E. coli* with a natural vitamin B12 analog reveals roles for cobamide uptake and adenosylation in methionine synthase-dependent growth. *J Bacteriol* **207**, e0028424.

Supporting information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Table S1. Crystallographic and refinement statistics for D5EK51 (PDB code: 9HMU).

Fig. S1. Spectroscopic and chromatographic characterization of the Pan, Halo, 656, Fluta, Cylst–B_{12} complexes.

Fig. S2. Sequence alignment showing conservation and the positions of secondary structures of D5EK51, BtuJ1 and BtuJ2.

22115463, 2026, 8, Downloaded from https://febs.onlinelibrary.wiley.com/doi/10.1002/2211-5463.70231 by NICE, National Institute for Health and Care Excellence, Wiley Online Library on [04/08/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License